SHARED UNDERSTANDING AND SYNCHRONY EMERGENCE
Synchrony as an Indice of the Exchange of Meaning between Dialog Partners

Ken Prepin and Catherine Pelachaud
LTCI/TSI, Telecom-ParisTech/CNRS, 37-39 rue Dareau, 75014, Paris, France

Keywords: Synchrony, Shared understanding, Coupled oscillators, Dynamical systems, Inter-subjectivity.

Abstract: Synchrony is claimed by psychology as a crucial parameter of any social interaction. In dialog interactions, the synchrony between non-verbal behaviours of interactants is claimed to account for the quality of the interaction: to give to human a feeling of natural interaction, an agent must be able to synchronise on appropriate time. The synchronisation occurring during non-verbal iteractions has recently been modelised as a phenonomenon emerging from the coupling between interactants. We propose here, and test in simulation, a dynamical model of verbal communication which links the emergence of synchrony between non-verbal behaviours to the level of meaning exchanged through words by interactants: if partners of a dyad understand each other, synchrony emerges, whereas if they do not understand, synchrony is disrupted. In addition to retrieve the fact that synchrony emergence within a dyad of agents depends on their level of shared understanding, our tests pointed two noteworthy properties of synchronisation phenomenons: first, as well as synchrony accounts for mutual understanding and good interaction, di-synchrony accounts for misunderstanding; second, synchronisation and di-synchronisation emerging from mutual understanding are very quick phenomenons.

1 INTRODUCTION

When we design agents capable of being involved in verbal exchange, with humans or with other agents, it is clear that the interaction cannot be reduced to speech. When an interaction takes place between two partners, it comes with many non-verbal behaviours that are often described by their type such as smiles, gaze at the other, speech pauses, head nod, head shake, raise eyebrows, mimicry of posture and so on (Kendon, 1990; Yngve, 1970). But another aspect of these non-verbal behaviours is their timing according to the partner's behaviours.

In 1966, Condon and Ogston's annotations of interactions have suggested that there are temporal correlations between the behaviours of two person engaged in a discussion (Condon and Ogston, 1966): micro analysis of discussion videotaped conduces Condon to define in 1976 the notions of auto-synchrony (synchrony between the different modalities of an individual) and hetero-synchrony (synchrony between partners).

Since Condon et al.'s findings, synchronisation between interactants has been investigated in both behavioural studies and cerebral activity studies. These studies tend to show that when people interact together, their synchronisation is tightly linked to the quality of their communication: they synchronise if they managed to exchange and share information; synchronisation is directly linked to their friendship, affiliation and mutual satisfaction of expectations.

In developmental psychology, generations of protocols have been created, from the "still face" (Tronick et al., 1978) to the "double video" (Murray and Trevarthen, 1985; Nadel and Tremblay-Leveau, 1999), in order to stress the crucial role of synchronisation during mother-infant interactions.

Behavioural and cerebral imaging studies show that oblivious synchrony and mimics of facial expressions (Chammat et al., 2010; Dubal et al., 2010) are involved in the emergence of a shared emotion as in emotion contagion (Hatfield et al., 1993).

In social psychology, in teacher-student interaction or in group interactions, synchrony between behaviours occurring during verbal communication has been shown to reflect the rapports (relationship and intersubjectivity) within the groups or the dyads (Bernieri, 1988; LaFrance, 1979).

The very same results have been found for human-machine interactions: on one hand synchrony of non-verbal behaviour improves the comfort of the human and her/his feeling of sharing with the machine (ei-

Prepin K. and Pelachaud C..
SHARED UNDERSTANDING AND SYNCHRONY EMERGENCE - Synchrony as an Indice of the Exchange of Meaning between Dialog Partners.
DOI: 10.5220/0003140600250034
In *Proceedings of the 3rd International Conference on Agents and Artificial Intelligence* (ICAART-2011), pages 25-34
ISBN: 978-989-8425-41-6

ther a robot or a virtual agent) (Poggi and Pelachaud, 2000) and on the other hand, the human spontaneously synchronises during interaction with a machine when her/his expectations are satisfied by the machine (Prepin and Gaussier, 2010).

In the case of non-verbal interactions, the phenomenon of synchronisation between two partners has recently been investigated as a phenomenon emerging from the dynamical coupling of interactants: that is to say a phenomenon whose description and dynamics are not explicited in each of the partners but appear when the interactants are put together and when the new dynamical system they form is more complex and richer than the simple sum of partners dynamics.

In mother-infant interactions via the “double-video” design cited above, synchrony is shown to emerge from the mutual engagement of mother and infant in the interaction (Mertan et al., 1993; Nadel and Tremblay-Leveau, 1999). In adult-adult interactions mediated by a technological device, synchrony and coupling between partners has been shown to emerge from the mutual attempt to interact with the other in both behavioral studies (Auvray et al., 2009) and cerebral activity studies (Dumas et al., 2010).

These descriptions of synchrony as emerging from the coupling between interactants, are consistent with the fact cited before, that synchrony reflects the quality of the interaction. Given interactants, both the quality of their interaction and the degree of their coupling are tightly linked to the amount of information they exchange and share: high coupling involves both synchrony and good quality interaction; synchrony and quality of the interaction are covarying indices of the interaction. That makes the synchrony parameter particularly crucial: on one hand it carries dyadic information, concerning the quality of the ongoing interaction; on the other hand it can be retrieved by each partner of the interaction, comparing its own actions to its perceptions of the other (Prepin and Gaussier, 2010).

The emergence of synchrony during non-verbal interaction has been modelled by both robotics implementation (Prepin and Revel, 2007) and virtual agent coupling (Paolo et al., 2008).

In the robotic experiment, two robots controlled by neural oscillators are coupled together by the way of their mutual influence: turn-taking and synchrony emerge (Prepin and Revel, 2007).

In the virtual agent experiment, Evolutionary Robotics was used to design a dyad of agents able to favour cross-perception situation; the result obtained is a dyad of agents with oscillatory behaviours which share a stable state of both cross perception and synchrony (Paolo et al., 2008).

The stability of these states of cross-perception and synchrony is a direct consequence of the reciprocal influence between the agents.

We have seen there that literature stresses two main results concerning synchrony. First, synchrony of non-verbal behaviours during verbal-interactions is a necessary element for a good interaction to take place: synchrony reflects the quality of the interaction. Second, synchrony has been described and modelled as a phenomenon emerging from the dynamical coupling between agents during non-verbal interactions. In this paper, we propose to conciliate these two results in a model of synchrony emergence during verbal interactions.

We propose and test in simulation a model of verbal communication which links the emergence of synchrony of non-verbal behaviours to the level of shared information between interactants: if partners understand each other, synchrony will arise, and conversely if they do not understand each other enough, synchrony could not arise. By constructing this model of agents able to interact as humans do, on the basis of psychology, neuro-imaging and modelisation results, that are both the understanding of humans and the believability of artifacts (e.g. virtual humans) which are assessed.

In Section 2 we describe the architecture principle and show how a level of understanding can be linked to non-verbal behaviours. In Section 3, we test this architecture, i.e. we test in simulation a dyad of architectures which interact together. We characterise the conditions of emergence of coupling and synchrony between the two virtual agents. Finally, in Section 4, we discuss these results and their outcomes.

2 MODEL PRINCIPLE

We propose a model accounting for the emergence of synchrony depending directly on a shared level of understanding between agents. This model is based on the four next properties of humans’ interactions:

P1. To emit or receive a discourse modify the internal state of the agent (Scherer and Delplanque, 2009).

P2. Non-verbal behaviours reflect the internal states (Matsumoto and Willingham, 2009).

P3. Humans are particularly sensitive to synchrony, as a cue of the interaction quality and and the mutual understanding between participants (Ducan, 1972; Poggi and Pelachaud, 2000; Prepin and Gaussier, 2010).

P4. Synchrony can be modelled as a phenomenon emerging from the dynamical coupling of agents (Prepin and Revel, 2007; Paolo et al., 2008; Auvray et al., 2009)

The model of agent we propose in the present section is implemented in Section 3 as a Neural Network (NN). Groups of neurons are vectors of variables represented by capital letters (e.g. $V_{Input} \in [-1,1]^n$ and $S \in [-1,1]^m$) and the weights matrices which modulate the links between these groups are represented by lower case letters (e.g. $u \in [-1,1]^{m \times n}$): we obtain equations such as $u \cdot V_{Input} = S$. For sake of simplicity, in both the description of the model principle (this section) and in its implementation and tests (Section 3) groups of neurons and weights matrices are reduced to single numerical variables ($\in [-1,1]$).

In the next two subsections, we model the two first properties, P1 and P2. We describe how the non-verbal behaviour can be linked to a level of mutual understanding. Then, in the subsections 2.3 and 2.4, we describe how this will give to a dyad of agents coupling capabilities. That constitute the modelling of the third and fourth properties, P3 and P4.

2.1 Speak and Listen Modifies Internal State

Let us consider a dyad of agents, Agent1 and Agent2. Each agent's state is represented by one single variable, S_1 for Agent1 and S_2 for Agent2 ($\in [-1,1]$). Now, let us consider the speech produced by each agent, the verbal signal $V_{Act\ i}$ ($\in \{0,1\}$), and the speech heard by each agent, the perceived signal $V_{Per\ i}$ ($\in \{0,1\}$).

P1 claims that each agent, either listener or speaker, has its internal state S_i modified by verbal signals: the listener's internal state is modified by what it hears, and the speaker's internal state is modified by what it says. Two "level of understanding", the weights u_i and u'_i, are defined for each agent of the dyad. u_i modulates the perceived verbal signal $V_{Per\ i}$, and u'_i modulates the produced verbal signal $V_{Act\ i}$ (see fig.1). To model interaction in more natural settings these u_i parameters should be influenced by many variables, such as the context of the interaction (discussion topic, relation-ship between interactants), the agents moods and personalities. However in the present model we combine all these parameters in the single variable u_i ($\in [-1,1]$). The choice of the values of u_1 and u_2 is arbitrary near 0.01: it enables a well balanced sampling of the oscillators' activations, the period last around 100 time steps; the other parameters of the architecture are chosen depending on this one so as not to modify the whole systems dynamics.

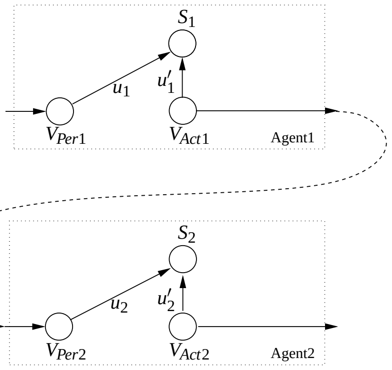

Figure 1: Verbal perception, $V_{Per\ i}$, and verbal action, $V_{Act\ i}$, both influence the internal state S_i. These influences depend respectively on the level of understanding u_i and u'_i.

If t is the time we have the following equations:

$$\begin{cases} S_1(t+1) = S_1(t) + u_1 V_{Per1}(t+1) + u'_1 V_{Act1}(t+1) \\ S_2(t+1) = S_2(t) + u_2 V_{Per2}(t+1) + u'_2 V_{Act2}(t+1) \end{cases} \quad (1)$$

Assuming that communication is ideal, i.e. $V_{Per i} = V_{Act j}$, and that Agent1 is the only one to speak, i.e. $V_{Act2} = V_{Per1} = 0$,the system of equations 1 gives:

$$\begin{cases} S_1(t+1) = S_1(t) + u'_1 V_{Act1}(t+1) \\ S_2(t+1) = S_2(t) + u_2 V_{Act1}(t+1) \end{cases} \quad (2)$$

This first property P1 is crucial in our model, as it links together the agents' internal states: each one is modified by speech depending on its own parameter u_i. In the present model, we assume that for a given agent, understanding of its productions and of its perceptions are similar: for Agent i, $u_i = u'_i$.

2.2 Non-verbal Behaviours Reflect Internal State

The second property P2, claims that "non-verbal behaviours reflect internal state". That is to say, agent's arousal, mood, satisfaction, awareness, are made visible thanks to facial expressions, gaze, phatics, backchannel, prosody, gestures, speech pauses. To make visible the internal properties of Agent i, a non-verbal signal, $NV_{Act\ i}$, is triggered depending on its internal state, S_i. When S_i reaches the threshold β, the agent produces non-verbal behaviours with th_β the threshold function (see fig. 2):

$$NV_{Act\ i}(t) = th_\beta(S_i(t)) \quad (3)$$

We suggest here that pitch accents, pauses, head nods, changes of facial expressions and other non-verbal cues are, for a certain part, produced by agents when a particularly important idea arises, when the explanation reach a certain point, when an idea or a

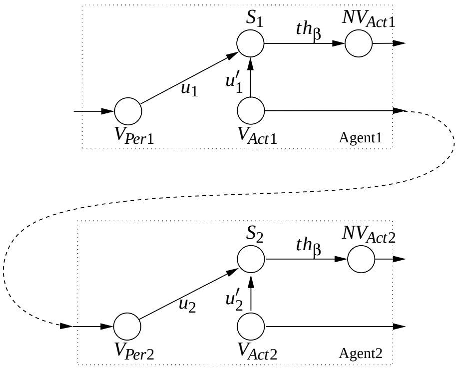

Figure 2: Each agent produces non-verbal behaviours $NV_{Act\ i}$ when S_i reaches the threshold β. $NV_{Act\ i}$ depends on how much the internal state S_i has been influenced by what has been said.

concept starts to be outlined. We assume that the phenomenon is similar in both speaker and listener, it is driven by the evolution of what is wanted to be expressed in one case and it is driven by what is heard in the other case. If speaker and listener understand each other, these peaks of arousal and understanding should co-occur: they appear to be temporally linked. These peaks will be the bases of entrainment for intentional coordination between partners. And then this coordination could be seen as a marker of interaction quality.

Considering these two first points, that is to say, equations 2 and 3 we have the following system of equations :

$$\begin{cases} NV_{Act1}(t_1) = th_\beta(\sum_{t_0}^{t_1} u_1 V_{Act1}(t)) \\ NV_{Act2}(t_1) = th_\beta(\sum_{t_0}^{t_1} u_2 V_{Act1}(t)) \end{cases} \quad (4)$$

If an agent is enough influenced by what is said, it produces non-verbal signals. And if $u_1 = u_2$ then $NV_{Act1} = NV_{Act2}$, agents' non-verbal behaviours may be synchronised, where as if u_1 and u_2 are too different, agents will not be able to synchronise.

2.3 Sensitivity to Synchrony

To account for the property P3, "sensitivity of human to synchrony", we use the fact that sensitivity to synchrony can be modelled by simple model of mutual reinforcement of the perception-action coupling (Auvray et al., 2009; Paolo et al., 2008). In addition to the influence from speech (either during its perception or its production), each agent's internal state S_i is influenced by the non-verbal behaviour it perceives from the other $NV_{Act\ j}$, modulated by sensitivity to non-verbal signal σ (see fig.3).

The internal state of each agent is modified by both what it understand of the speech and what it sees from the non-verbal behaviour of the other:

$$\begin{cases} S_1(t+1) = S_1(t) + u_1 V_{Act1}(t+1) + \sigma NV_{Act2}(t) \\ S_2(t+1) = S_2(t) + u_2 V_{Act1}(t+1) + \sigma NV_{Act1}(t) \end{cases} \quad (5)$$

This last equation will favour the synchronisation by increasing the reciprocal influence when agents' internal state reach together a high level.

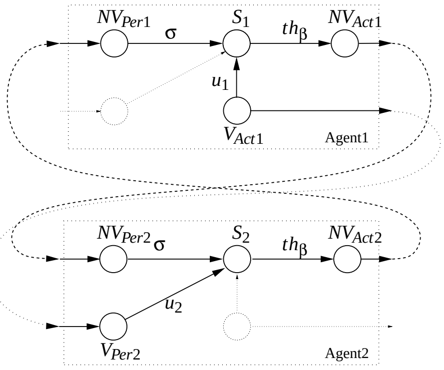

Figure 3: Agent1's internal state, S_1, is influenced by both its own understanding of what it is saying $u_1 \cdot V_{Act1}$ and the non-verbal behaviour of Agent2, $\sigma \cdot NV_{Act2}$. Agent2's internal state, ${}_2$, is influenced by its own understanding of what Agent1 says $u_2 \cdot V_{Act1}$ and the non-verbal behaviour of Agent1, $\sigma \cdot NV_{Act1}$

2.4 Coupling between Dynamical Systems

How to enable agents involved in a verbal interaction, to be as much synchronised as they share information? To enable synchrony to emerge between the two agents, we used the fact that synchronisation can be modelled as a phenomenon emerging from the dynamical coupling within the dyad (Prepin and Revel, 2007): on one hand agents must have internal dynamics which control their behaviour; on the other hand, they must be influenced by the other's behaviours.

In the previous subsections, we proposed a dyad of agent which mutually influence. If we replace the non-verbal behaviours of agents by their internal states in the system of equations 5, it gives:

$$\begin{cases} S_1(t+1) = S_1(t) + u_1 V_{Act1}(t+1) + \sigma th_\beta(S_2(t)) \\ S_2(t+1) = S_2(t) + u_2 V_{Act1}(t+1) + \sigma th_\beta(S_1(t)) \end{cases} \quad (6)$$

To enable coupling to occur, the agents should also be dynamical systems: systems which state evolves along time by themselves. The internal state of the agents S_i produces behaviours and is influenced by the other agent's behaviour. To ensure internal dynamics, we made this internal state a relaxation oscillator, which increases linearly and decreases rapidly when it reaches the threshold 0.95 (fig. 5 shows an example of the signals obtained). By oscillating , the internal states agents will not only influence each other but also be able to correlate one with the other (Prepin and Revel, 2007).

Here, two cases are interesting.

When the internal states of both agents are under the threshold triggering non-verbal behaviours, β, the system of equation 6 becomes:

$$\begin{cases} S_1(t+1) = S_1(t) + u_1 V_{Act1}(t+1) \\ S_2(t+1) = S_2(t) + u_2 V_{Act1}(t+1) \end{cases} \tag{7}$$

The two agents are almost independent, they are only influenced by the speech of Agent1 and each one produces its own oscillating dynamic. That could be the case if two tired people (high β) speak about a not so interesting subject (u_i are low): they are made apathic by the conversation, they do not express anything.

The second interesting case is when both agents' internal states are above the threshold β. The system of equation 6 becomes:

$$\begin{cases} S_1(t+1) = S_1(t) + u_1 V_{Act1}(t+1) + \sigma S_2(t) \\ S_2(t+1) = S_2(t) + u_2 V_{Act1}(t+1) + \sigma S_1(t) \end{cases} \tag{8}$$

In this case agents are not anymore independent, they influence each other depending on the way they understand speech. If we push the recursivity of these equations one step further we obtain:

$$\begin{cases} S_1(t+1) = S_1(t) + u_1 V_{Act1}(t+1) + \\ \quad \sigma(S_2(t-1) + u_2 V_{Act1}(t) + \sigma S_1(t-1)) \\ S_2(t+1) = S_2(t) + u_1 V_{Act1}(t+1) + \\ \quad \sigma(S_1(t-1) + u_1 V_{Act1}(t) + \sigma S_2(t-1)) \end{cases} \tag{9}$$

And now we see the effect of coupling, that is to say that agents are not only influenced by the state of the other but they are influenced by their own state, mediated by the other: the non-verbal behaviours of the other becomes their own biofeedback (Nadel, 2002). When the threshold β is overtaken, the reciprocal influence is recursive and becomes exponential: the dynamics of S_1 and S_2 are not any more independent, they are influenced in their phases and frequencies (Pikovsky et al., 2001; Prepin and Revel, 2007).

3 TEST OF THE MODEL

We tested this model by implementing a dyad of agent as a neuronal network in the neuronal network simulator Leto/Prometheus (developed in the ETIS lab. by Gaussier et al. (Gaussier and Cocquerez, 1992; Gaussier and Zrehen, 1994)), and by studying its emerging dynamics with different sets of parameters.

3.1 Implementation

We implemented the model on the neural networks simulator Leto/Prometheus. Leto/Prometheus simulates the dynamics of neural networks by an update of the whole network at each time step. We use groups of neurons with one neuron, and non-modifiable links between groups. The schema of fig. 4 show this implementation.

The internal states of agents, S_i, are relaxation oscillators: the re-entering link of weight 1 makes the neuron behave as a capacity, and the *Relax* neuron which fires when a 0.95 threshold is reached, inhibits S_i and makes it relax (see fig. 5 for an example of the activation obtained).

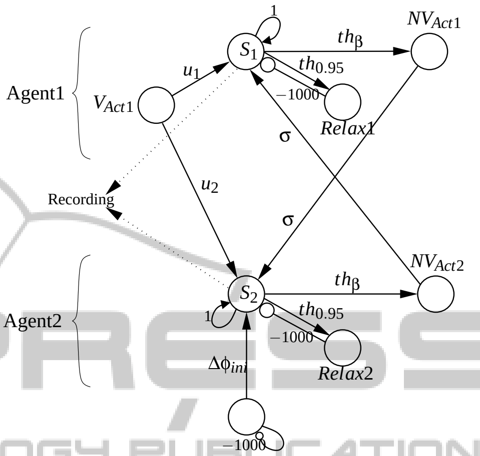

Figure 4: Implementation of the two agents. The couples $(S_1; Relax1)$ and $(S_2; Relax2)$ are relaxation oscillators. The parameters which will be tested are the following: β, the threshold which controls the non-verbal production; u_1 and u_2 which control the agents' level of sharing; $\Delta\phi_{ini}$, the initial phase-shift between agents.

V_{Act1}, Agent1's verbal production, is a neuron of constant activity 1. This neuron feeds the oscillators of both agents, weighted by their level of understanding u_1 and u_2. The values of u_1 and u_2 are near 0.01: it enables a well balanced sampling of the oscillators' activations, the period last around 100 time steps.

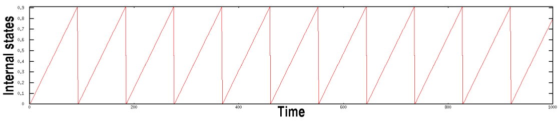

Figure 5: Activations of the internal state $S_1(t)$ for $u_1 = 0.01$.

In addition to agent understanding u_1 and u_2, three other parameters are modifiable in this implementation:

- The threshold β which controls the triggering of non-verbal signal.

- The sensitivity of agent's internal state to non-verbal signal σ which weights $NV_{Act\ i}$. These two parameters β and σ directly control the amount of non-verbal influence between the agents: they must be high enough to enable coupling, for instance reducing initial phase-shift between oscillators or compensating phase deviation when $u_1 \neq u_2$.
- The initial phase shift $\Delta\phi_{ini}$, which makes agents start with a phase shift between $S_1(t_{ini})$ and $S_2(t_{ini})$ at the beginning of each test of the architecture.

Finally, the variables recorded during these tests are the internal states of both agents, $S_1(t)$ and $S_2(t)$ (see fig. 6 for an example).

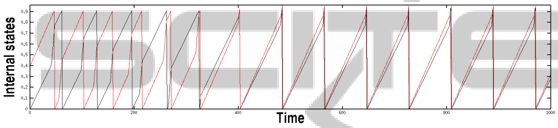

Figure 6: Activations recorded for $u_1 = 0.01$, $u_2 = 0.011$, $\beta = 0.85$, $\sigma = 0.05$ and $\Delta\phi_{ini} = 0.4$. Despite the initial phase shift and the phase deviation, the two agents synchronise. This is a stable state of the dyad, it remains until the end of the experiment (5000 time steps).

3.2 Test of Synchrony Emergence

For a given set of parameters, to determine if in-phase synchronisation occurred between agents, we used a procedure described by Pikovsky, Rosenblum and Kurths in their reference book "Synchronisation" (Pikovsky et al., 2001). This procedure consists in comparing the phases of two signals to determine if they are synchronous or not.

First we used the fact that relaxation oscillators can be characterised by their peaks. There is a peak at time t_k when $S_i(t_k) \geq 0.9\beta$ and $S_i(t_k+1) = 0$. Then, we used the fact that phase can be rebuilt from these peaks (Pikovsky et al., 2001). We assign to the time t_k the values of the phase $\phi(t_k) = 2\pi k$, and for every instants of time $t_k < t < t_k + 1$ determine the phase as a linear interpolation between these values (see fig.7):

$$\phi(t) = 2\pi k + 2\pi \frac{t - t_k}{t_{k+1} - t_k} \tag{10}$$

After that, when the phases of signals are obtained, we consider their difference modulo 2π (see fig.8). Horizontal plateaus in this graph reflect periods of constant phase-shift between signals, i.e. synchronisation. Horizontal plate aux near zero reflect periods of synchronisation and co-occurrence of non-verbal signals.

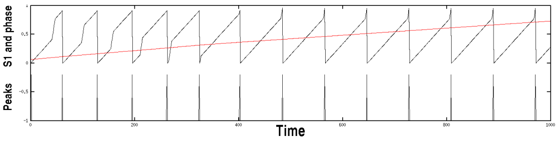

Figure 7: Signal, Peaks and Phase. In the upper part of the graph, there is the original signal S_1 (shown in fig.6) and the associated re-built phase (we can notice the change of phase slope when synchronisation occurs). In the lower part of the graph, there are the peaks extracted from S_1 in order to rebuild the phase.

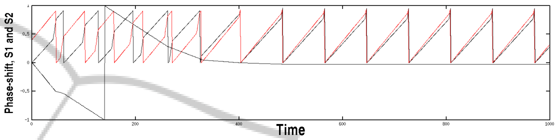

Figure 8: Signals of two agents and their associated phase-shift $\Delta_{\phi_1,\phi_2}(t)$. When agents synchronise with each other, their phase-shift remains constant and near zero.

Finally, for each 5000 time steps simulation, we define that in-phase synchronisation occurs if the phase-shift becomes near zero at a time t_{synch}, smaller than 3000, and remains constant until the end. We defined the synchronisation speed as $SynchSpeed = (3000 - t_{synch})/3000$. If in-phase synchronisation is immediate $SynchSpeed = 1$; if in-phase synchronisation occurs at time step 3000 $SynchSpeed = 0$; and if in-phase synchronisation do not occurs $SynchSpeed < 0$.

3.3 Test of Architecture Parameters

We tested different parameters of this model, first to show the direct link existing between emergence of synchrony and level of sharing between interactants, and second to characterise the different properties of this model.

To show the direct link existing between emergence of synchrony and level of sharing between interactants, we fixed u_1 to 0.01 and made u_2 vary between 0 and 0.02, that is to say the shared understanding of the two agents differs between 0 and 100%. Notice here the importance to test synchronisation when $u_2 = 0$: if synchronisation occurs when $u_2 = 0$, i.e. when Agent2 does not perceived the speech of Agent1, that means that agents synchronise every time just thank to non-verbal signal of Agent1; in that case, synchrony is not any more an in dice of the interaction quality, the influence of non-verbal signals (linked to β and σ) is too high.

To evaluate the influence of the amount of non-verbal signal exchanged, we made the threshold β vary between 0 and 0.95.

To evaluate the influence of the sensitivity to non-verbal signal, we made the sensitivity σ vary between 0 and 0.09.

Finally, to evaluate the abilities of such a dyad of agents to re-synchronise after an induced phase-shift or after a misunderstanding, we made the initial phase shift $\Delta\phi_{ini}$ vary between 0 and π.

Shared Understanding Influence. When the two agents are synchronous in phase ($\Delta\phi_{ini} = 0$), we tested which of the u_2 values keep agents synchronised or make them disynchronise. For fixed $\beta = 0.7$, $\sigma = 0.05$ and $\Delta\phi_{ini} = 0$, u_2 varies between 0 and 0.02. The following graph of fig. 9 shows the associated disynchronisation speed.

When the difference between u_1 and u_2 is to high, no synchronisation can occur since even when synchrony is forced at the beginning of the experiment, agent disynchronise.

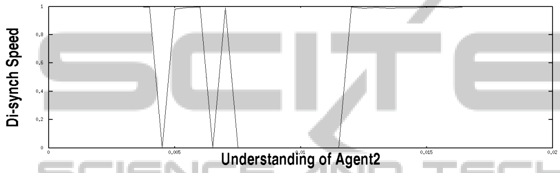

Figure 9: Di-synchronisation speed of the dyad, depending on the Agent2 understanding u_2. u_2 varies from left to right between 0 and 0.02. A null disynchronisation speed means that synchronisation has been maintained until the end of the experiment. A disynchronisation speed 1 is for a dis-synchronisation occurring at the very beginning of the experiment.

Influence of Amount of Non-verbal Signals. The coupling and synchronisation capabilities of the dyad of agents, may directly depend on the amount of non-verbal signals they exchange: among other, the ability to compensate a difference of understanding may be improved by an increase of non-verbal signals exchanged. We tested this effect by calculating disynchronisation speeds as just above, making u_2 vary between 0 and 0.02 and the threshold β varying between 0 and 0.9 ($\sigma = 0.05$). We obtained the 3D graph of fig. 10.

When $\beta = 0.9$, that is to say when very few non-verbal signals are exchanged, synchrony maintains only when the two agents have equal level of understanding, $u_1 = u_2 = 0.01$. For other values, the influence of the threshold β is not so clear: the dyad does not resist better to disynchronisation when $\beta < 0.5$ than when $6 \leq \beta \leq 8$. This effect, or this absence of effect, may be due to the fact that the more β decreases, the less accurate in time the non-verbal signals are: if β is low, non-verbal signals are emit earlier before the peaks of S_i activation and on a larger time window, they are not enough precise in time to maintain synchrony. We chosen $\beta = 0.7$, i.e. the mean of its best performances values.

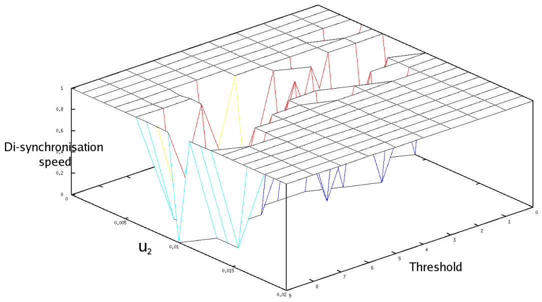

Figure 10: Di-synchronisation speed of the dyad, depending on the Agent2 understanding u_2 and the threshold β ($\sigma = 0.5$). u_2 varies between 0 and 0.02. β varies from 0.9 to 0, in the sens of non-verbal signals increase. When the d i-synchronisation speed value is null, synchronisation has been maintained until the end of the experiment. A disynchronisation speed 1 is for a disynchronisation occurring at very beginning of the experiment.

Sensitivity to Non-verbal Signals. Another way to modify the influence of non-verbal signals on coupling and synchronisation properties of the dyad, is to modify the sensitivity to the perceived non-verbal signal, σ. We tested this effect by calculating disynchronisation speeds as previously, making u_2 vary between 0 and 0.02 and the sensitivity σ varying between 0 and 0.09 ($\beta = 0.07$). We obtained the 3D graph of fig. 11.

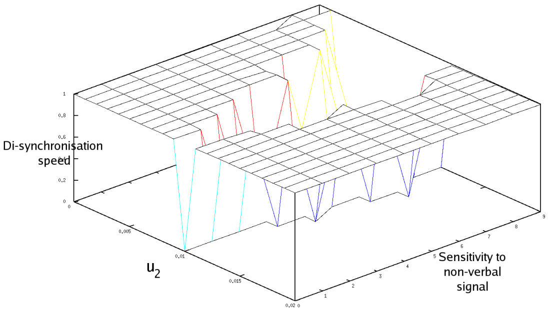

Figure 11: Di-synchronisation speed of the dyad, depending on the Agent2 understanding u_2 and the sensitivity σ ($\beta = 0.7$). u_2 varies between 0 and 0.02. σ varies from 0 to 0.09. When the d i-synchronisation speed value is null, synchronisation has been maintained until the end of the experiment. A disynchronisation speed 1 is for a disynchronisation occurring at the very beginning of the experiment.

Sensitivity to non-verbal signal σ have a direct effect on agents to stay synchronous even with different understandings: the higher is sensitivity σ, the more resistant to difference between u_i the synchronisation capability of the dyad is. The effect of σ is important despite its low value ($\sigma < 0.1$) due to the high number of non-verbal signal exchanged: when Agent i's internal state S_i reaches the threshold β, it produces the non-verbal signals $NV_{Act\ i}$ at every time step until S_i relaxes. That can last between 0 and 20 time steps for each oscillation period. The effect of σ is multiplied by this number of steps.

It is important to notice here that the σ effect on the dyad resistance to u_i differences, has a counterpart. This counter-part is the fact that when σ increase and make the dyad more resistant to disynchronisation, it also makes the synchronisation of the dyad less related to mutual understanding. For instance, when $\sigma \geq 0.7$, agents stay synchronous even when Agent2 do not understand anything, $u_2 = 0$. To balance these two effects, facilitation of synchronisation and decrease of synchrony significance, we chosen a default value of $\sigma = 0.05$.

Re-synchronisation Capability. Given a value of Agent2 understanding u_2, we tested the ability of the dyad Agent1-Agent2 to re-synchronise after a phase shift. We made the initial phase-shift $\Delta\phi_{ini}$ vary between 0 and π for every values of u_2 and calculated the speed of synchronisation if any. The 3D graph of fig. 12 shows the synchronisation speed for each couple $(u_2; \Delta\phi_{ini})$.

The initial phase-shift between S_1 and S_2 does not appear to affect the synchronisation capacities of the dyad. With the chosen $\sigma = 0.05$ and $\beta = 0.7$, when the agents' levels of understanding u_1 and u_2 do not differ more than 15% of each other, they synchronise systematically and very quickly: for instance they synchronise even when they start in anti-phase ($\Delta\phi_{ini} = \pi$). And conversely, when the levels of understanding u_1 and u_2 are more than 15% different, synchronisation is no more immediate.

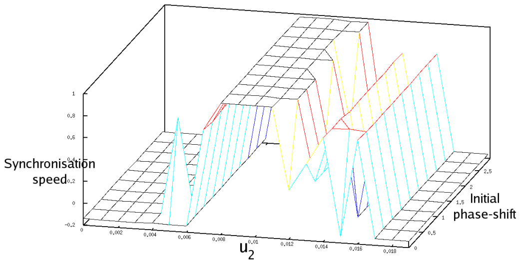

Figure 12: Synchronisation speed of the dyad, depending on the Agent2 understanding u_2 and initial phase-shift $\Delta\phi_{ini}$ ($\sigma = 0.05$ and $\beta = 0.7$). u_2 varies between 0 and 0.02. $\Delta\phi_{ini}$ varies from 0 to π. When the synchronisation speed value is null, the dyad did not synchronised until the end of the experiment. A synchronisation speed 1 is for a synchronisation occurring at the very beginning of the experiment.

4 DISCUSSION

We proposed and tested a model which links emergence of synchrony between dialogue partners to their level of shared understanding. This model assesses both the understanding of humans and the believability of artifacts (e.g. virtual humans). When two interactants have similar understanding of what the speaker says, their non-verbal behaviours appear synchronous. Conversely, when the two partners have different understanding of what is being said, they disynchronise. This model is implemented as a dynamical coupling between two talking agents: on one hand, each agent proposes its own dynamics; on the other hand, each agent is influenced by its perception of the other. These are the two minimal conditions enabling coupling. What makes this model particular is that the internal dynamics of agents are generated by the meaning exchanged through speech. It links the dynamical side of interaction to the formal side of speech.

We tested this model in simulation, and showed that synchrony effectively emerges between agents when they have close level of understanding. We noticed a clear effect of the level of understanding on the capacity of the agents to both remain synchronous and re-synchronise: agents disynchronise if the level of shared understanding is lower than 85% (with our parameters) and conversely agents synchronise if the level of shared understanding is higher than 85%. These results tend to prove that, considering that synchrony between agents is an indice of good interaction and shared understanding, the reciprocal property is true too; that is disynchrony accounts for misunderstanding.

We have shown that agents remain synchronous depends on both their shared understanding (the ratio between u_1 and u_2) and their sensitivity to non-verbal behaviour (σ in our implementation). The more sensitive to non-verbal behaviours are the agents, the more resistant to disynchronisation is the dyad and the easier is the synchronisation. An important counter-part of this easier synchronisation is that it makes synchrony less representative of shared understanding: agents or people with very different levels of understanding will be able to synchronise; if sensitivity to non-verbal behaviour is too high, the dyadic parameter of synchrony is not a cue of shared understanding. By contrast, the facility agents trigger non-verbal behaviours when their internal states are high (thresh-

Figure 13: Greta, Obadia, Poppy and Prudence. They are four agents implemented on the open source system Greta. Each one has its own personality and level of understanding. When interacting together, different levels of non-verbal synchrony should appear between the agents of this group.

old β) does not appear to change the synchronisation properties of the dyad: the higher number of exchanged non-verbal signals seems to be compensated by their associated decrease of precision.

In addition to the effect of shared understanding on the stability of synchrony between agents, we have tested the effect of shared understanding on the capacity of the dyad to re-synchronise. For instance, during a dialogue, synchrony can be broken by the use of new concept by the speaker. That may result in lowering the level of shared understanding below the 85% necessary for remaining synchronous. Synchrony can also be disrupted by an external event which can introduce a phase-shift between interactants. Given fixed sensitivity to non-verbal behaviour (σ) and facility to trigger non-verbal behaviours (β), we tested how quickly the dyad can re-synchronise after a phase-shift. The shared level of understanding necessary to enable re-synchronisation appeared to be the same as the one under which agents disynchronise.

Two crucial points must be noticed here. First, when agents' understanding do not differ more than 15% (shared understanding higher than 85%), agents synchronise systematically whatever the phase-shift is, and when agent's understanding differ more than 15% they disynchronise. Second, both synchronisation and disynchronisation of agents are very quick, lasting about one oscillation of the agents' internal states. Synchronisation and disynchronisation are very quick effects of respectively misunderstanding and shared understanding: agents involved in an interaction do not have to wait to see synchrony appears when they understand each other, they have a fast answer to whether they understand each other or not.

The 5000 time steps length of our tests allowed us to test the stability of synchrony or disynchrony after their occurrence; however it is clearly not a natural situation. Synchrony in natural interaction is a varying phenomenon involving multiple synchronisation and disynchronisation phases: the level of shared understanding varies along the interaction. In fact disynchrony may be quite informative for the dyad as its detection enables agents to adapt one another. In natural interactions, synchrony occurring after disynchrony shows that agents share understanding whereas they did not before: they have benefited from the interaction and exchanged information.

As a consequence, the mean level of shared understanding necessary for good interaction to take place between persons in natural context would be much more reasonable: the 85% of shared understanding occurs in phases of particularly good interaction and its is not a hard constraint on the whole dialogue; this very high level necessary for synchronisation should be divided by the ratio of synchrony vs disynchrony phases present in natural interaction. For instance we can imagine that a level of shared understanding higher than 85% would occur when people involved in a discussion have just reached an agreement. By contrast, when the level of shared understanding stays all along the dialogue far under 85%, the dyad would be more like two strangers trying to talk together, or a professional talking with technical words to a naive listener.

Our model has been tested and its principle has been validated in agent-agent context. To go a step farther, in "wild world" situations involving humans, two elements must be added: Understanding of language during interaction with human; Recognition of non-verbal behaviours of human users. In the near future, we will adapt the present neural architecture to the open source virtual agent Greta (Pelachaud, 2009). The system Greta enables one to generate multi-modal (verbal and non-verbal) behaviours online and with accurate timing. The verbal signals will be modelled as elements of "small-talk" and the non-verbal signal will be modelled as, pitch accents, pauses, head nods, head shakes and facial expressions. To test the real impact of such a model on human perception of interaction, we will perform perceptive evaluation: we aim to simulate a group of virtual agents dialoguing with each other (see fig.13). Each agent will have its own personality and level of understanding of what being said. This will lead to pattern of synchronisation and disynchronisation. Among other, agents which share understanding should display inter-synchrony pattern (Condon, 1976). Finally, human observers should clearly fill which agent is sharing understanding with which other agent.

In conclusion, we can notice that, in addition to the two main results of this study –"disynchrony accounts for misunderstanding" and "synchronisation and disynchronisation are very quick phenomenons"– another result is the model itself. It proposes a link between synchrony and intersubjectivity by the use of dynamical system coupling: synchrony and dynamical coupling emerge together when agents mutually understand each other; as a consequence synchrony account for good interaction.

We believe, this model is a start to answer the issues of what is the part of dynamical coupling between agents involved in verbal interaction? What is the part of emerging dynamics in the communication of meanings and intentions? And moreover, how these two parts can co-exist and feed each other?

ACKNOWLEDGEMENTS

This work has been partially financed by the European Project NoE SSPNet (Social Signal Processing Network). Nothing could have been done without the Leto/Prometheus NN simulator, lent by the Philippe Gaussier's team (ETIS lab, Cergy-Pontoise, France).

REFERENCES

Auvray, M., Lenay, C., and Stewart, J. (2009). Perceptual interactions in a minimalist virtual environment. *New ideas in psychology*, 27:32–47.

Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal Behavior*, 12(2):120–138.

Chammat, M., Foucher, A., Nadel, J., and Dubal, S. (2010). Reading sadness beyond human faces. *Brain Research*, In Press, Accepted Manuscript:–.

Condon, W. S. (1976). An analysis of behavioral organisation. *Sign Language Studies*, 13:285–318.

Condon, W. S. and Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease*, 143:338–347.

Dubal, S., Jouvent, A. F. R., and Nadel, J. (2010). Human brain spots emotion in non humanoid robots. *Social Cognitive and Affective Neuroscience*, in press:–.

Ducan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of Personality and Social Psychology*, 23(2):283–292.

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchonization during social interaction. *PLoS One*, 5(8):e12166.

Gaussier, P. and Cocquerez, J. (1992). Neural networks for complex scene recognition : simulation of a visual system with several cortical areas. In *IJCNN Baltimore*, pages 233–259.

Gaussier, P. and Zrehen, S. (1994). Avoiding the world model trap: An acting robot does not need to be so smart! *Journal of Robotics and Computer-Integrated Manufacturing*, 11(4):279–286.

Hatfield, E., Cacioppo, J. L., and Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Sciences*, 2:96–99.

Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge University Press, Cambridge, UK.

LaFrance, M. (1979). Nonverbal synchrony and rapport: Analysis by the cross-lag panel technique. *Social Psychology Quarterly*, 42(1):66–70.

Matsumoto, D. and Willingham, B. (2009). Spontaneous facial expressions of emotion in congenitally and non-congenitally blind individuals. *Journal of Personality and Social Psychology*, 96(1):1–10.

Mertan, B., Nadel, J., and Leveau, H. (1993). *New perspective in early communicative development*, chapter The effect of adult presence on communicative behaviour among toddlers. Routledge, London, UK.

Murray, L. and Trevarthen, C. (1985). Emotional regulation of interactions vetween two-month-olds and their mothers. *Social perception in infants*, pages 101–125.

Nadel, J. (2002). *Imitation and imitation recognition: their functional role in preverbal infants and nonverbal children with autism*, pages 42–62. UK: Cambridge University Press.

Nadel, J. and Tremblay-Leveau, H. (1999). *Early social cognition*, chapter Early perception of social contingencies and interpersonal intentionality: dyadic and triadic paradigms, pages 189–212. Lawrence Erlbaum Associates.

Paolo, E. A. D., Rohde, M., and Iizuka, H. (2008). Sensitivity to social contingency or stability of interaction? modelling the dynamics of perceptual crossing. *New ideas in psychology*, 26:278–294.

Pelachaud, C. (2009). Modelling multimodal expression of emotion in a virtual agent. *Philosophical Transactions of Royal Society. Biological Science*, 364:3539–3548.

Pikovsky, A., Rosenblum, M., and Kurths, J. (2001). *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge University Press, Cambridge, UK.

Poggi, I. and Pelachaud, C. (2000). Emotional meaning and expression in animated faces. *Lecture Notes in Computer Science*, pages 182–195.

Prepin, K. and Gaussier, P. (2010). How an agent can detect and use synchrony parameter of its own interaction with a human? In et al., A. E., editor, *COST Action2102, Int. Traing School 2009, Active Listening and Synchrony. LNCS 5967*, pages 50–65. Springer-Verlag, Berlin Heidelberg.

Prepin, K. and Revel, A. (2007). Human-machine interaction as a model of machine-machine interaction: how to make machines interact as humans do. *Advanced Robotics*, 21(15):1709–1723.

Scherer, K. and Delplanque, S. (2009). Emotions, signal processing, and behaviour. In *Chemosensory Perception Symposium*, Geneva. Firmenich.

Tronick, E., Als, H., Adamson, L., Wise, S., and Brazelton, T. (1978). The infants' response to entrapment between contradictory messages in face-to-face interactions. *Journal of the American Academy of Child Psychiatry (Psychiatrics).*, 17:1–13.

Yngve, V. H. (1970). On getting a word in edgewise. In Society, C. L., editor, *Papers from the 6th regional meeting*, pages 567–578.