Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, Technische Universität München, Munich, Germany, August 1-3, 2008

Emergent Dynamics of Turn-Taking Interaction in Drumming Games with a Humanoid Robot

Hatice Kose-Bagci, Kerstin Dautenhahn, and Chrystopher L. Nehaniv

Abstract— **We present results from an empirical study investigating emergent turn-taking in a drumming experience involving Kaspar, a humanoid child-sized robot, and adult participants. In this work, our aim is to have turn-taking and role switching which is not deterministic but emerging from the social interaction between the human and the humanoid. Therefore the robot is not just 'following' and imitating the human, but could be the leader in the game and being imitated by the human. Data from the first implementation of a human-robot interaction experiment are presented and analysed qualitatively (in terms of participants' subjective experiences) and quantitatively (concerning the drumming performance of the human-robot pair). Results are analysed statistically and show significant differences for the three games (with different probabilistic models) where the models enabling more interaction and more 'natural' turn-taking were preferred by the human participants.**

I. Introduction

TURN-TAKING is an important ingredient of human-human interaction and communication whereby role switch ('leader' and 'follower') is not determined by external sources but emerges from the interaction. Human beings typically 'know' when to start and stop their turns in the social interactions, based on various factors including the context and purpose of the interaction, feedback from the social interaction partners, emotional and motivational factors etc. People use different criteria for making these decisions. Our work proposes a novel framework which enables emergent turn-taking, and role-switching between a human and a humanoid in an imitation game.

There are several example works that studied turn-taking in games and conversations in the literature, focusing on different aspects. An example from developmental psychology describes the case of the emergent turn-taking between a mother and a baby without any explicit control mechanism is described [10]. The mother starts jiggling in response to her baby's sucking to encourage her baby to resume sucking, which results in emergent turn-taking between the jiggling and sucking actions.

In the field of robot assisted play and therapy, one of the most difficult issues in teaching and education of children with autism is to teach children the concept of 'turn-taking'. Turn-taking games have been used in several studies to engage children with autism in social interactions [8, 17]. Another example of turn-taking games is given from a cognitive robotics view [6]. In this work, a ball game between a humanoid robot Cog, and the human experimenter is described. Cog and the human were reaching out and grasping a ball in alteration. But here the turn-taking behaviour was led by the human experimenter in reaction to the robot's visually driven actions.

Ito and Tani studied joint attention and turn-taking in an imitation game played with the humanoid robot QRIO, where the human participants try to find the action patterns, which were learned by QRIO previously, by moving synchronously with the robot [11].

From a linguistics point of view, Sacks et al. identify some of the important features of turn-taking in human conversation as follows [18]:

- Speaker-change recurs, or at least occurs.
- Mostly, one party talks at a time.
- Occurrences of more than one party speaking at the same time are common but brief.
- Transitions (from one turn to the next) with no gap and no overlap are common (slight gap or slight overlap is accepted).
- Turn order is not fixed, but varies.
- Turn size is not fixed, but varies.
- Length of conversation is not specified in advance.
- What parties say is not specified in advance.
- Relative distribution of turns if not specified in advance.
- Number of parties can vary.
- Talk can be continuous or discontinuous.

Built on these features, Thorisson developed a turn-taking mechanism for conversations based on his previous work on the *Ymir mind model* for communicative creatures and humanoids [19]. The expressive humanoid robot KISMET [4, 5] used social

Manuscript received January 28, 2008. This work was conducted within the EU Integrated Project RobotCub ("Robotic Open-architecture Technology for Cognition, Understanding, and Behaviours"), funded by the EC through the E5 Unit (Cognition) of FP6-IST under Contract FP6-004370.

H. Kose-Bagci, K. Dautenhahn and C. L. Nehaniv are with the Adaptive Systems Research Group, University of Hertfordshire, School of Computer Science, Hatfield, Herts, AL 10 9AB, U.K.
(e-mail: {h.kose-bagci, k.dautenhahn, c.l.nehaniv}@herts.ac.uk).

cues for regulating turn-taking in non-verbal interactions with people. Here, a sophisticated robot control architecture modeling motivations, emotions and drives was used to satisfy KISMET's internal "needs". Turn-taking between KISMET and humans emerged from the robot's internal needs and goals and its perceptions of cues from its interaction partner. Similarly, in our work we study emergent turn-taking, but based on minimal, probabilistic control models.

Our particular test bed for studying emergent turn-taking here is human-robot drumming games. We used imitation games involving drumming as a test bed since they seem a suitable tool for studying the interaction between humans and robots in terms of social aspects including imitation, turn-taking and synchronization. Also, different from the above-mentioned work with KISMET, where the interaction was the goal in itself, we wanted to include a certain (enjoyable) task that needs to be achieved jointly by the human-robot pair, to provide the overall context. Drumming is relatively straightforward to implement and test, and can be implemented technically without special actuators like fingers or special skills or abilities specific to drumming [12]. There are several works concerning drumming in human-robot interaction. Robotic percussionists play drums in collaboration with human partners [7, 20]. These artifacts use robotic arms that are specially designed to play drums. An approach based on the movement generation using dynamical systems was tested on a Hoap-2 humanoid robot using drumming as a test case. [9], Similarly, humanoid drumming is used as a test bed for exploring synchronization [14].

In our study, the humanoid robot Kaspar plays drums autonomously with a human 'partner' (interactant), trying to imitate the rhythms produced by the human (as a follower) and trying to motivate (as a leader in the game) the human to respond. With a simple, but novel probabilistic method Kaspar decides when to start and stop its turn. It observes the human playing and uses its observations as parameters to decide whether to listen to the human or to take the turn actively in the game. This is different from our previous work [12] where we tested deterministic turn-taking involving gestures performed by the robot. In the current work Kaspar does not use any gestures, but only drumming to interact with the human. We found in our previous work that different robot nonverbal gestures influence people's responses in the drumming game, and thus decided to carry out this experiment without any gestures in order to be able to focus our analysis on the turn-taking behaviour.

The emergence of turn-taking as such is not the primary aim of the paper, instead the particular *turn-taking dynamics* (e.g. when and how long the robot and the human play for each turn) emerged from the interaction. To clarify, in our experiments, the *pattern and timing* of the turns is emerging: we cannot beforehand predict the pattern of timing of turns in a concrete interaction. This is the key ingredient of emergence. Kaspar's internal decision mechanism is constructed from several models and functions and the dynamics and pattern of turn-taking emerge in the interaction with the human. Even if playing the game many times with same robot, model and human participant, the outcome would be different, depending on probabilistic responses from the humanoid and current drumming behaviour of the human, -- hence, we speak of 'the emergent dynamics of turn-taking interactions'.

The rest of this paper is organized as follows: the next section describes the methodology. Section 3 presents the research questions and expectations. The experiments, results and analysis are described in section 4. Section 5 includes a conclusion on what was learned from this work, and presents ideas for future work.

II. BACKGROUND & METHODOLOGY

In our previous study [12] the human partner played a rhythm, which Kaspar tried to replicate in a simple form of imitation (mirroring). Kaspar had two modes: listening and playing. In the listening mode, it recorded and analysed the human's rhythm, and in the playing mode, it played the rhythm back by hitting the drum positioned on its lap. Then the human partner played again. This (deterministic) turn-taking in this game continued for the fixed. Kaspar did not imitate the strength of the beats but only the number of beats and durations between beats, due to its limited motor skills. For beats beyond its skill, it used instead minimum values allowed by its capabilities: Kaspar needed at least 0.3 seconds between beats to get its joints 'ready'; so that, even if the human plays faster, Kaspar's imitations would still require minimum durations of at least 0.3 seconds between beats. It also needed to wait for a few seconds before playing any rhythm in order to get its joints into correct reference positions.

One of the fundamental problems addressed by this scenario is the timing of the interaction, as timing plays a fundamental role in the regulation of human interaction (cf. [16]). It is not always clear when the robot or human partner should initiate interaction in taking a turn. Therefore, in the previous work, some predefined fixed time duration heuristics were used for synchronization. Kaspar started playing if the human partner was silent for a few seconds, and would also try to motivate the human partner with simple nonverbal gestures.

In the new work reported here, we instead used a novel, probability-based for timing and turn-taking. The temporal dynamics of turn-taking thus emerge from the interaction between the human and the humanoid. To begin to gain insight into possible interaction dynamics we selected three different simple models, to control the starting and stopping of the robot's regular drumming beats. This response is based on the duration time of the previous turn and on the number of beats played in the previous turn by the interaction partners. We denote the models *model1, model2* and *model3*. *Model1* uses a step function, *model2* a simple triangular function, and *model3* a hyperbolic function generate probabilities for starting or stopping the robot's drumming based on these inputs from previous interaction (see Figure 1). The output is bounded by maximum and minimum limits to ensure that Kaspar and the human have time to play at least once in every turn. For every turn, Kaspar looks up the probability of start or stop, and takes action accordingly. For the start, Kaspar uses the time duration of its last bout of playing, and for the stop, the number of beats of the human participant from the previous turn. The minimum number of beats Kaspar will play is 1 even if the resulting number of the beats recommended by the models is below 1.

The human starts the game with Kaspar using its turn-taking strategy when the human participant is silent for two seconds (only for the first turn). After the first turn, the turn-taking strategy is always determined by Kaspar's probabilistic models. The probability functions for the three computational models are presented in (1), (2), and (3).

$$p(x) = \begin{cases} 0 & x < Th \\ 1 & x \geq Th \end{cases} \qquad \text{(step: } \textit{model1}\text{)} \qquad (1)$$

$$p(x) = x / Th \qquad \text{(linear: } \textit{model2}\text{)} \qquad (2)$$

$$p(x) = 1 - 1/x \qquad \text{(hyperbolic: } \textit{model3}\text{)} \qquad (3)$$

where *Th* represents the respective threshold of time for starting and of number of beats for stopping (Figure 1). Regarding the probabilistic algorithm ((2), and (3)), a random value *r* in [0,1] is generated and if *r* is not less than the function output, then the model returns 1 (otherwise 0) in the conditionals (IF-statements) of the robot control (see pseudocode below). [Note: we had also tried to *start* using beats and *stop* using time with simulated data, but the current combination resulted in more drumming time and a higher number of beats for both human and Kaspar, so this combination was preferred in the current implementation.] Thus depending on the previous duration and number of beats in the interaction, according to their respective probability functions (1), (2), (3), the three models may return value 1, which triggers starting or stopping in the turn-taking algorithm (Algorithm 1). In future, other models could also easily be assessed.

Algorithm 1 The turn-taking algorithm

```
Human plays (turn # i=1)
Kaspar plays after  waiting 2 seconds when human stops
FOR i=2 to n DO
  ThTime_i= KasparPlayingTime_i-1
        IF model_j (HumanPlayingTime_i, ThTime_i) = 1
            THEN KASPAR STARTS PLAYING
               ThBeat_i= # of HumanBeats_i
              IF model_j (# of  KasparBeats_i, ThBeat_i) = 1
                 THEN KASPAR STOPS PLAYING
END FOR (end of the game)
```

So at every turn, Kaspar decides when to start and stop according to the performances of both the human participant and itself. Thus, the game and its dynamics are not deterministic but emerge from the moment-to-moment status of both Kaspar and the human interactant.

III. Research Questions & Expectations

In this paper we study the effect of the different computational models on emergent turn-taking in an imitation game. A simple drumming game enriched with different models determining the turn-taking strategy of the humanoid robot was used as a test bed, and the subjective evaluations of the participants were analysed. Our primary research questions were:

1) How do different robot turn-taking strategies based on computational, probabilistic models impact the drumming performance of the human-robot pair?

2) How do the different robot turn-taking strategies impact the participants' subjective evaluation of the drumming experience?

We expect to have 'successful' games in terms of turn-taking emerging from the interaction between human and the humanoid. Our 'success' criteria are be the number of turns with no or slight overlaps and gaps. Also the number of human beats detected by the robot and number of beats played by the robot itself will give us hints about the quality of the games.

We set up simulated experiments before the real experiments, to define the maximum and minimum limits and thresholds for the real experiments with humanoid and human participants.

We studied three models with different parameters. Each model is used both for starting and stopping the robot's play. For *start* the time duration of the previous turn is used, and for *stop* the number of beats of the previous turn is used as threshold. As described in the previous section in detail, *model1* was a step function, where the new value of could not be smaller than the threshold, thus we expect this model to give more play time and a

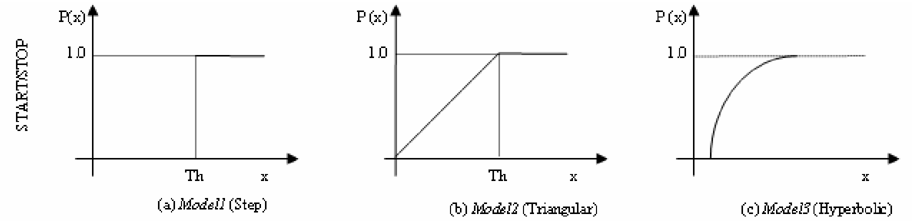

Figure 1 Computational Models for START/STOP actions. For START actions, Th = ThTime., since the x axis variable is time (t). For STOP actions, Th = ThBeat. The x axis variable is number of beats (b). For START, Th is the duration of Kaspar's previous drumming bout, and for the STOP action, Th is the number of beats in the human's previous drumming bout; except that the minimum value for Th is 1.5 sec (experimentally determined) for START and 1 beat for STOP actions. The only model which does not have threshold limitations is model3 due to its hyperbolic nature. The y axis gives the probability of START/STOP as a function of time/number of beats based on previous interaction.

Figure 2 A screen shot from the experiments showing a person playing a drumming game with Kaspar.

higher number of beats than the other models. Ideally, if the human beats long sequences, this model would reach very high values so we put a maximum time limitation (both interactants cannot play longer than 10 seconds per turn). Unlike *model1*, *model2* has a triangular shape which has the threshold as an upper bound. Since we have a probabilistic approach we can have values smaller than the threshold. In fact, we expect this model to give the least play time and lowest resulting number of beats for human participants, so we foresee that the model would not be as popular as the other two models among the participants. The last condition is *model3*, a hyperbolic model, which cannot be bounded by the thresholds. It reaches high values (close to 1) very fast compared to *model2*. Therefore we predict that it would give more play time and enables to play more beats than *model2*. Also, in our simulations we noticed that it could enable 'good games' (i.e. with a very low number of overlaps and conflicts between the human's and robot's drumming) if we played short sequences, but since the model is not bounded by thresholds, it 'reacts' to the human but does not exactly 'imitate' the games, which might not be accepted by participants.

IV. Experiments, Results & Analysis

A. Kaspar

The experiments were carried out with the humanoid robot called Kaspar. Kaspar is a child-like humanoid robot which was designed and built by the members of the Adaptive Systems Research Group at the University of Hertfordshire to study human-robot interactions with a minimal set of expressive robot features. Kaspar has 8 degrees of freedom in the head and neck and 6 in the arms and hands. The face is a silicon-rubber mask, which is supported on an aluminum frame. It has 2 DOF eyes fitted with video cameras, eyelids capable of blinking, and a mouth capable of opening and smiling, see description in [2].

B. Experimental Setup

The experiments were carried out in a separate room isolated from other people and noises which could affect the drumming experiment. Kaspar was seated on a table with the drum on its lap. The human partner was seated in front of the robot using another drum that was fixed on the table (Figure 2). The human participants used a pencil, or their bare hands to hit the drum. Although we suggested to the participants to use one pencil and hit on the top of the drum, sometimes they used two pencils with a single hand or with both hands, and several times used the tambourine-style bells around the drum's sides.

C. Software Features

The implementation of robot perception and motor control used the YARP environment [15]. YARP is an open-source framework used in the project RobotCub that supports distributed computation that emphasizes robot control and efficiency. It enables the development of software for robots, without considering a specific hardware or software environment. Portaudio [1] software was used to grab audio from the audio device, within the YARP framework.

The acoustic sound waves recorded by the sound grabber module are converted to digital music samples, which allows using mathematical computations and sample based techniques. To detect the patterns of a sound wave, a filter based method is used, based on the work of [13] originally used to detect visual patterns.

D. Participants

Twelve participants in the age range of 23-32 (4 female and 8 male) took part in the study. All participants were right-handed and worked in computer science or similar disciplines at the University. Only two of them had interacted with Kaspar prior to the experiment, and they were overall not familiar with robots. Three of them had children aged 1-3 years.

E. Interaction Game Setup

We used a one minute demo of the robot without any drumming game involved where participants were shown how to interact with Kaspar. This was followed by three games reflecting the three experimental conditions described above each lasting three minutes, without indicating to the participants anything about the differences between the conditions. Participants were simply instructed that they could play drumming games with Kaspar. We used all six possible different presentation orders of the games, to analyze the effect of the order of the games on the humans. To account for possible fatigue, habituation, or learning by the participants, in the *sequential order* section below, we analyse the games according to their order number in the sequence experienced by the participants (independent of the particular experimental condition), as being the first game, second or third, disregarding their game types, e.g. for one participant the first game (order 1) would be the *model1* game, and for another participant, *model1* would be the third game (order 3).

F. Evaluation of Questionnaire Data

After the experiment the participants were asked to complete a questionnaire investigating their preferences and opinions on the three experimental conditions.

1) Most and least preferred games according to type:

The number of participants which rated each game as most preferred and least preferred can be seen below in Table 1. It shows that both the *model1* and *model3* games were preferred by the same amount of participants, while no participant most preferred *model2*.

Table 1 also shows that most of the participants considered the *model2* game as the least preferred, while the *model1* and *model2* games had a small number of participants which considered them the least preferred. The *model3* game was slightly more popular than the *model1* game.

TABLE 1
MOST AND LEAST PREFERRED GAMES ACCORDING TO TYPE

Game type	Participants	
	Most preferred game	Least preferred game
model1	6	3
model2	0	8
model3	6	1

2) Most and least preferred games according to sequential order

The number of participants which rated each game as most preferred and least preferred according to the sequential order can be seen below in Table 2. It is shown that the most popular game type was the third game, while first and second games were less preferred.

TABLE 2.
MOST AND LEAST PREFERRED GAMES ACCORDING TO ORDER

Order	Participants	
	Most preferred game	Least preferred game
1	3	4
2	2	3
3	7	4

According to Table 2, all ordinal positions of occurrence in the sequence of the games had a similar number of participants which considered them the least preferred.

3) Reasoning behind preferences

While an exhaustive description of the qualitative analysis of the participants' responses concerning their impressions and preferences about the drumming games is beyond the scope of this brief paper, a short summary will be given below:

The order of the games had an impact on the participants. Their liking of the games increased significantly between the first and third trials (for drumming, $F(2,22)=3.29$, $p=0.069$; for sociality, $F(2,22)=4.904$, $p<0.05$, with ANOVA). They preferred the last game more, which could be because they got used to the scenario as they played more, so they had more successful plays as they spent more time; this is consistent with our previous findings [10]. According to the game types there appeared also to be an impact in terms of drumming, ($F(2,22)=2.444$, $p=0.110$ with ANOVA); but no significant difference in terms of sociality, ($F(2,22)=2.895$, $p=0.77$, with ANOVA).

G. Behavioural Data

1) Sequential order

There is no significant difference between the games according to the order (e.g. for number of turns, $F(2,22)=0.007$, $p=0.99$, with ANOVA). Only the human's total number of beats per game increased with the order of the games as they got used to the scenario while they played more (Tables 3 and 4).

TABLE 3
OBSERVED BEHAVIOUR OF KASPAR ACCORDING TO ORDER

Order	Avg. # of beats per turn	Max/ Min # of beats	Total # of beats	Avg. time per turn	Max/ Min time per turn	Total time
1	1.7±0.8	6/1	136±32	1.08±0.1	3/1	97.8±41
2	1.74±0.8	6/1	136±29	1.07±0.1	3/1	95.5±40
3	1.81±0.7	7/1	139±23	1.07±0.1	4/1	94.8±41

TABLE 4

OBSERVED BEHAVIOUR OF HUMAN ACCORDING TO ORDER

Order	#of turns	#of nonzero turns	Max # of beats	Total # of beats (Kaspar's view)	Total # of beats (real)	Avg. time per turn	Max/Min time per turn	Total time
1	93±45.08	27.83±14.3	5	44.33±25.8	104.3±27.5	0.99±0.567	3.11/0.01	70±27.53
2	91.1±43	29±12	4	47.8±27	114.8±34.5	0.99±0.6	2.06/0.01	69.1±27
3	90.4±44.27	32.3±15.4	5	55.1±32.69	122.8±23.8	1±0.57	3.11/0.01	68±24.8

TABLE 5

OBSERVED BEHAVIOUR OF HUMAN ACCORDING TO GAME TYPE

Game type	#of turns	#of nonzero turns	max # of beats per turn	Sum of beats (Kaspar's view)	Total # of beats (real)	Avg. time per turn	Max/Min time per turn	Total time
model1	65.1±4.03	37.3±15	5	72.1± 27.8	113±29.223	1.53±0.02	3.11/1.5	99.3±5.31
model2	151±3.46	21.1±7.8	3	25.6±9.67	116.7±24.69	0.25±0.01	0.61/0.01	37.4±1.7
model3	59±1.5	31±13	5	50±22	112.08±34.9	1.2±0.01	1.8/1	70±1.8

TABLE 6

OBSERVED BEHAVIOUR OF KASPAR ACCORDING TO GAME TYPE

Game type	Avg. # of beats per turn	Max/ Min # of beats	Total # of beats	Avg. time per turn	Max/ Min time per turn	Total time
model1	1.6±0.3	5/1	99±9	1±0.04	3/1	67±3.1
model2	1±0.01	3/1	153±4	1±0.004	3/1	151±3.2
model3	2.7±0.1	7/2	158±3	1.2±0.04	4/1	70±1.7

2) Interaction game type

The game types are compared in detail in Table 5 (Human's perspective) and Table 6 (Kaspar's perspective).

According to the game types, *model1* and *model3* show more similarities than *model2*. In *model1*, the total number of beats of Kaspar was higher than the total number of beats of the human participants (100/65), whereas, the total game duration is higher for human participants than for Kaspar (70/100), as well as the average time per turn. In *model3*, the total number of beats is lower than for *model1*. Although the total play durations for Kaspar and the humans were almost identical, the total number of beats for Kaspar was almost three times as high as that of human participants. In terms of maximum and minimum durations per turn, for Kaspar there is no significant change, but in the case of the human player both are significantly longer for *model1* than for the other models. The *model2* enabled the least play time and number of beats for human participants by far.

In *model1* and *model3* almost half of the turns were nonzero (i.e. the human played at least one beat). These two models showed almost similar behaviour, with *model1*'s number of turns and number of nonzero turns slightly larger compared to *model3*. The *model2* game has the largest number of turns which is almost twice as high as the other two, but the number of nonzero turns is the smallest by far (14%). Also, here the robot beats much more than the human (the number of human beats is 17% of Kaspar's beats).

Note that although, as observed from Table 5, human participants appeared to play similarly in all three games, Kaspar only detected human participants' beats, and recorded them, when it decided that the humans play a turn according to its computational model. Kaspar discarded the beats played by human participants at other times, namely during Kaspar's own play times.

V. DISCUSSION OF RESULTS

We analysed the humanoid-human drumming games in terms of sequence of order, and according to game type. While the sample size makes it difficult to make any strong inferences, as such the following analysis and the statistical analysis in the previous sections are only descriptive.

In terms of sequence there is an impact on the participants drumming behaviour and evaluation of the games while they played the three games. They tend to beat more, in fewer turns, and, in terms of the questionnaire data, they liked the games more as they played more.

As stated in the previous section, the *model2* game, due to its nature, gives the least play time to the human and Kaspar. So Kaspar does not seem to imitate the human participants' game at all, but rather 'plays on its own' (Kaspar plays at least one beat even when it does not detect a response from the human participant). As a consequence, Kaspar was a leader in the game most of the time. There were also many overlaps between Kaspar's play turns and human participants' play turns in *model2*. So either Kaspar or the human participants interrupted the other which was found 'annoying' by the human participants, some of them even called this action "rude". and caused the loss of detection of human participants' beats for Kaspar (as described before, Kaspar did not 'listen' when it played itself). So when compared to the questionnaire results, it is logical that humans did not like the *model2* game.

As stated in the previous sections, since *model1* uses the previous play's play time as a threshold, it ensures that the current play time is at least as long as the previous play time for human participants. But since they were given more play time than the other games, there were time gaps between their turns and Kaspar's turns, so they felt the tempo of the game was slower than the others. These participants preferred *model3* since the tempo of the game was faster than the *model1* for them. It is observed that this game gives them play time shorter than *model1* but long enough to have a coordinated game, so could be viewed as more 'natural'. In this game both human and Kaspar had 3-4 beats every turn (its probabilities are increasing fast, so it does not give small values very often), there were less gaps than when using *model1,* and less overlaps compared to *model2* between two turns.

But *model3* was not bounded by thresholds by nature, so seems to be independent of the human participants' performance, which annoyed some of the participants. Still one participant found this like "teaching her son to play drum". Another participant asked if she should consider Kaspar as a professional drummer or a child while she commented on the games, since it "looks like a child drumming rather than a professional".

In *model1* the human was given more time than Kaspar, but Kaspar played more beats than the human participants. Whereas in *model3,* Kaspar and the human participant where given almost equal durations and opportunities to play. So in the *model3*, Kaspar is given a chance to be a follower and leader almost equally. Kaspar had more impact on the play and played longer rhythms.

Note, there is a big amount of the zero turns (where the human could not do any action, but Kaspar played at least one beat, and passed the turn to the human) in all of the three models. However, only in *model2* is their amount high enough to affect the whole game. When these turns are distributed among normal turns as in *model1* and *model3*, they do not dominate the behaviour but can be compensated for by non-zero turns. But for *model2*, zero turns dominate the whole game and are disliked by participants.

Although there were gaps between the humans' and the robot's turns in *model1*, and *model3* did not seem to imitate the human participants in every turn, both models were successful in terms of emergent turn-taking. As a consequence according to the explanations of the human participants in the questionnaires, they liked *model1 and model3* more than *model2*.

It is important to note that while Kaspar's drum playing changed in terms of timing based on simple models, some human participants commented that Kaspar behaved 'intelligently', e.g. they thought that the robot interrupted them in a structured way, in order "to tell them something". In our computational models, we aimed not to imitate the human participants' drumming exactly, but tried to get some emergent effects from the interaction between human and humanoid instead. Although some of the participants found this "annoying" since Kaspar did "not imitate them well", surprisingly, another group of the participants thought Kaspar played like a small child, and they enjoyed the games.

Also over time, the participants learned the limits of Kaspar and the rules of the game, and adapted themselves to the game better, so they had better games, in terms of turn-taking and synchronization. We could observe long sequences of plays without any overlaps or gaps between the turns, and human participants were really enthusiastic about the games. Humans, as shown here, were not passive subjects in this game, but adapted themselves unconsciously to the capabilities of the robot. This finding is consistent with the notion of 'recipient design', a concept from ethnomethodology, where we find that natural speech is always designed for its recipient, i.e. the interaction partner, and interpreted as having been so designed. Here, the speaker creates his or her turn "with recipients in mind, and listeners are motivated to 'hear' a turn that is for them and all participants closely and constantly track the trajectory of the talk to hear 'their' turn" ([3], p.71). According to conversation analysis, this turn-taking is integral to the formation of any interpersonal exchange ([3], p. 66). While in our study the robot's behaviour was controlled based on simple computational models, the human participants used their recipient design skills in the interaction.

The issue of recipient design will be explored further in our future research. Also, we plan to add robot gestures to our future games (using head movements and facial expressions), since most of the participants commented in the questionnaires that gestures might improve Kaspar's social interaction skills, and we observed the same result in our previous work [12].

VI. CONCLUSIONS

In this work, we introduced probabilistic computational models in an imitative rhythmic interaction game that facilitates emergent turn-taking between a robot and a human partner. We based our test bed on drumming, which is a very suitable task for testing human-robot interaction. It is intended as more than a simple drumming synchronization task. Our long-term agenda is to develop rich social interaction between the robot and the human partner, which would not simply focus on synchronization to produce the same tempo, but result in producing a joyous and fruitful experience, emerging from human-robot interaction.

We used drumming interaction games enriched with different probabilistic computational models which enables Kaspar to start and stop its turns using its observations on the human participant's play. According to the play time per turn and number of beats played during a turn, Kaspar starts and stops its own turn, and therefore influences the human participant's turn. So each turn is emerging from the current play status of Kaspar and the human participants. This is more similar to natural human-human conversation, where human beings start and stop their turns in conversations and also in non-verbal communication according to criteria of their own without an external or internal rigid 'clock'.

This work was conducted within the EU Project RobotCub to carry out basic research into the regulation of interaction dynamics during social/playful human-robot interaction. The importance of turn-taking in conversations and interactions has been highlighted above. Our previous work [12] used deterministic turn-taking, simply mirroring the human's playing, causing problems in terms of timing and negatively affecting human participants' enjoyment. In this new work we developed novel turn-taking methods which appear more natural and engage the human participants more positively in the interaction games. Although we used very simple models, and this work is a first step in this domain, we were able to observe some very 'natural' games in terms of coordinated turn-taking, and some of the participants even compared the game to a normal game you may play with your children.

These methods and results will be used in other human-robot interaction studies, and are relevant for a wide area of applications that involve dynamic interactions between people and robots, including service, as well as entertainment and therapy robotics.

REFERENCES

[1] R. Bencina, and P. Burk, *Portaudio*, 2007 [cited; Available from: http://www.portaudio.com/trac/wiki/.

[2] M.P. Blow, K. Dautenhahn, A. Appleby, C. Nehaniv, D. Lee, Perception of robot smiles and dimensions for human-robot interaction design, *Proc. 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN06),* 2006, pp.469-474.

[3] D. Boden, The Business of Talk: Organizations in Action. 1994: Polity Press.

[4] C. Breazeal, *Toward sociable robots.* Robotics and Autonomous Systems, 2003. **42**(3-4): p. 167-175.

[5] C. Breazeal, *Designing Sociable Robots*. 2002: The MIT Press.

[6] R.A. Brooks, Personal Communication at *2nd International Conference on Cognitive Technology (CT'97),August 28*. 1997. Aizu, Japan.

[7] C. Crick, M. Munz, and B. Scassellati, Synchronization in social tasks: Robotic drumming, *Proceedings of IEEE RO-MAN 2006*, 2006, pp. 97-102.

[8] K. Dautenhahn, and A. Billard, *Games children with autism can play with robota, a humanoid robotic doll*. Universal Access and Assistive Technology, ed. S. Keates, et al. 2002, London: Springer-Verlag. 179–190.

[9] S. Degallier, C. P. Santos, L. Righetti and A. Ijspeert, Movement generation using dynamical systems: a humanoid robot performing a drumming task, In *Proceedings of the IEEE-RAS International Conference on Humanoid Robots (HUMANOIDS06)*, 2006, pp. 512 – 517.

[10] H. Hendriks-Jansen, *Catching ourselves in the act: Situated activity, interactive emergence, and human thought*. 1996, MIT Press, Cambridge, MA.

[11] M. Ito, and J. Tani, Joint attention between a humanoid robot and users in imitation game, *Proc. of the Int. Conf. on Development and Learning (ICDL)*, 2004.

[12] H. Kose-Bagci, K. Dautenhahn, D. S. Syrdal, and C. L. Nehaniv, Drum-mate: A Human-Humanoid Drumming Experience, *In IEEE-RAS International Conference on Humanoid Robots (Humanoids2007)*, 2007, Pittsburgh, Pennsylvania, USA.

[13] H. Kose, and H. L. Akin, Object Recognition in Robot Football Using a one Dimensional Image, *The Tenth Turkish Symposium on Artificial Intelligence and Neural Networks (TAINN)*, June, 2001: p. 291-300.

[14] S. Kotosaka, and S. Schaal, Synchronized robot drumming by neural oscillator,. *The International Symposium on Adaptive Motion of Animals and Machines*, 2000.

[15] G. Metta, P. Fitzpatrick, and L. Natale, *YARP: yet another robot platform.* International Journal on Advanced Robotics Systems Special Issue on Software Development and Integration in Robotics, 2005.

[16] B. Robins, K. Dautenhahn, C. L. Nehaniv, N. A. Mirza, D. Francois, and Olsson, L. Sustaining interaction dynamics and engagement in dyadic child-robot interaction kinesics: Lessons learnt from an exploratory study. In *Proc. of the 14th IEEE International Workshop on Robot and Human Interactive Communication, ROMAN2005*, 2005, pp. 716-722.

[17] B. Robins, K. Dautenhahn, R. te Boekhorst, and Billard, A., *Effects of repeated exposure to a humanoid robot on children with autism*. Designing a More Inclusive World, ed. S. Keates, et al. 2004, London: Springer Verlag. 225–236.

[18] H. Sacks, E.A. Schegloff, and G. Jefferson, *A Simplest Systematics for the Organization of Turn-Taking for Conversation.* Language, 1974. **50**(4): p. 696-735.

[19] K.R. Thórisson, *Natural turn-taking needs no manual: Computational theory and model, from perception to action.* Multimodality in Language and Speech Systems, 2002: p. 173-207.

[20] G. Weinberg, and Driscoll, S. Robot-human interaction with an anthropomorphic percussionist. *Proceedings of International ACM Computer Human Interaction Conference (CHI 2006).* Montreal, Canada, 2006, pp.1229-1232.