# Robotic drumming: synchronization in social tasks

Christopher Crick Matthew Munz Tomislav Nad
Brian Scassellati
Yale University
New Haven CT 06511
christopher.crick@yale.edu, mmm92@pantheon.yale.edu,
tomislav.nad@yale.edu, scaz@cs.yale.edu

September 23, 2005

**Abstract**

Development of a rhythmic sense comes early in childhood development, and appears crucial for learning speech as well as music. Synchronized, rhythmic movements governed by oscillators play significant roles in industrial robots, but only in tightly controlled, precisely defined environments. Prior research focuses on the synchronization of the robot's control system to either proprioceptive signals or external physical oscillators. Social tasks, however, require participants to detect, interpret and attune to the activity of humans, and by extension, to their intentions and perceptions. The synthesis of multiple sensory perceptions may be a fruitful approach to this problem. In order to evaluate this approach, we programmed a humanoid robot, Nico, to play a drum in concert with human drummers and at the direction of a human conductor. Our results show that sensory integration can enable precise synchronization in social tasks even when perceptual data is imperfect, misleading and subject to extensive processing delay. Through sensory integration, Nico can attune to a tempo that is set by a human conductor, in concert with human performers.

## 1 Introduction

Humans are particularly good at predicting and harmonizing with a rapidly changing environment in real-time. For example, many activities involving coordinated movement, such as dancing and playing catch, require fast decision-making, based on partial information. This behavior may best be understood under the general framework of synchronization [1], which has been used to model cognition in social tasks [2].

Synchronization tasks present significant difficulty to existing robot architectures. Musical performance, and drumming in particular, exemplifies the kind of synchronization challenge at which humans excel, and at which robots often fail. The musical environment, however, is relatively well-structured and constrained. This provides an excellent opportunity to explore solutions to synchronization problems, in a social setting, without oversimplifying the task or environment. As Breazeal's group [3] suggests, this supportive scaffolding may be as necessary for humanoid robots in their development, as it is for infants in theirs.

Figure 1: Nico, drumming in concert with human performers

While oscillator control provides an elegant solution to physical synchronization problems, its application in uncertain, complex sensory environments presents significant challenges. A robot attempting to drum in synchrony with human performers must perceive, classify and predict human actions, compensating for incomplete and uncertain sensory input in real time. The range of visual and auditory processing tasks involved contrasts sharply with the immediate, reliable input on which force feedback-based oscillator control architectures rely.

A successful robot drummer must spend time processing the perceptual signals it receives in order to recognize a rhythm. Given this information, it then needs to predict when the next beat should occur, taking into account variations in sensing and processing time. Furthermore, its arm motion must begin exactly far enough in advance of the predicted beat so that the drum strike happens at the correct time, perfectly in phase with the humans. Finally, it must detect its own mistakes and use them to refine its model of the delays involved in sensing, processing and actuation.

We programmed a humanoid robot, Nico (shown in Figure 1), to play a drum in concert with human drummers and at the direction of a human conductor. This process involved no preconceived idea of Nico's own physical dimensions or performance characteristics; the robot had to work out its internal dynamics for itself. Nico's dimensions match those of the average one-year-old male, and it has a finely articulated arm, head and shoulder [4][5][6]. In our experiment, Nico's ability to attune to the beat of human performers provides a clear, practical measure of the efficacy of our approach.

### 1.1 Related Research

Musical perception and performance has been recognized as a fundamental aspect of human development, the study of which provides insight into speech development, enculturation, emotional development, and group social interaction [7]. Musical tasks have also proved a rewarding testbed for humanoid robotics. Robot drumming was previously explored by Hajian [8] and Williamson [9]. The drumming robot described by Hajian [8] exploited the dynamics of the task by using a low impedance drumstick holder, taking advantage of the drumstick's bounce to produce high frequency drumming. Robotic musicianship was also explored by the AI research group at Waseda University [10]. Their Wabot-2 robot had 50 degrees of freedom and was able to read a musical score and reproduce it on the piano.

The history of entertainment robotics includes many examples of musical robots, the exemplar of which is the player piano [11]. In recent years, music performance has been a key element in the demonstration and marketing of commercial robots. Notably, Toyota's Partner robot has a large number of actuators for lips and breath control, enabling it to play wind instruments (e.g. the trumpet) with great precision [12].

From the humanoid robotics literature, our work can be best understood as an extension of the work of Williamson [9], who used oscillators (implemented as neural networks) to exploit the natural dynamics of rhythmic action such as drumming. Appropriate to non-social tasks like crank turning, this research focuses on the synchronization of the robot's control system to proprioceptive signals and external physical oscillators. In a social context, however, it becomes necessary to detect and synchronize with the rhythmic activity of humans, and by extension, their intentions and perceptions. Accordingly, our focus is on the attentional dynamics of the system.

In contrast to the performance of modern commercial robots, which can be sophisticated, yet tightly scripted, Nico's performance is musically rudimentary, but highly adaptable. The robot learns and adjusts its behavior by observing and reasoning about itself and its partners in real time. Our work attempts to model the psychophysical integration of auditory and visual information, using this biological principle to enhance robotic proprioception, as suggested by Williamson [9]. Likewise, our multi-oscillator architecture and approach are informed by Dynamic Attending Theory [2]. This theory proposes that humans, when listening to a complex audio sequence, begin by focusing their attention based on an internal *referent rate*, and dynamically shift this attentive rate to form a comfortable internal rhythmic representation of the sounds they hear. It includes the concept of *hierarchical levels* of rhythmic attention, which may provide a framework for the development of further sophistication in Nico's perceptual capabilities.

Furthermore, our approach reflects recent work in developmental studies of rhythm learning in infants [13]. By the age of one, an infant can discriminate between sounds on the basis of their implied meter and rhythmic structure. Our developmental formulation of Nico's behaviors and musicianship finds parallels in early human development.

The architecture of Nico and the design decisions behind it are extensively discussed in several papers [4][5][6]. Gold [6] investigates the use of correlation in timing between the actions that Nico takes and the movements that it sees, to develop the capacity for self-recognition. Our research extends this idea, using the self-knowledge thus obtained to generate precisely-timed actions in a complex environment.

# 2 Methodology

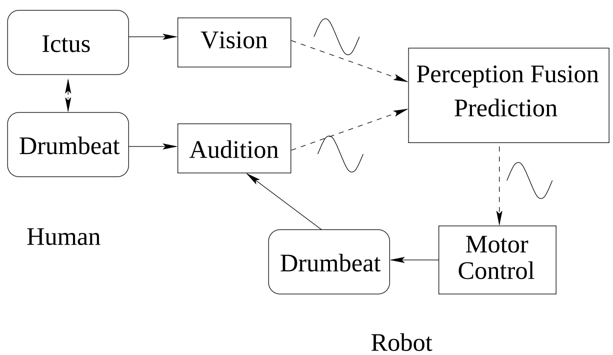


Figure 2: High-level system design

The overall architecture of the system is summarized in Figure 2. The vision and audition subsystems each produce a stream of event detections, while the motor control subsystem consumes a stream of arm-swing commands. The vision subsystem recognizes the arm motions of a human drummer or conductor, while the audition subsystem detects the drumbeats of both human and robotic performers. A high-level control subsystem (marked "Perception Fusion and Prediction" in Figure 2) ties the others together, using learning, analysis and prediction, to generate high-level arm commands. Finally, those commands are translated into low-level motor directives, which cause Nico to tap a snare drum with its end effector.

## 2.1 Visual Perception

The problem of visual perception is a difficult one, in a robotic context or indeed in a human one. A musician in a band or orchestra must, while playing an instrument and listening to his surroundings, obtain precise timing information from a conductor who may be standing at a distance, in difficult lighting. Thus, conducting gestures are designed to communicate timings as clearly as possible. This has the happy effect of simplifying Nico's recognition problem as well.

Our visual perception routine looks for the *ictus* of the conductor's beat, the point in time at which the conductor's hand "bounces" off an imaginary line, indicating the beat. In order to be useful, however, the ictus must be detected quickly. Our algorithm, as described below, performs very simple threshold-based filtering, and then computes the average vertical position of the moving, skin-colored objects in the visual field, as influenced by its previous position estimate. It uses a history of these positions to generate a trajectory, which it then analyzes to determine the proper timing of detected beats.

1. For each pixel of a 320x240 frame,
   (a) *Skin detection*: If the red component of the pixel is higher than the blue and green components, and
   (b) *Motion detection*: If the difference between the combined RGB intensity of this pixel and the one in the same position in the previous frame is more than 100 (on a 0-255 scale), then
   (c) Mark this pixel.
2. Average the vertical positions of every marked pixel within 50 pixels of the last computed vertical centroid to form the current vertical centroid.
3. If the current vertical centroid is higher than the previous centroid, and
4. If the previous centroid was lower than any of the previous 10 centroids, then
5. The ictus of a beat has been detected.

This algorithm takes advantage of the fact that a conductor's arm movements are smooth and regular, and that human skin, specifically the conductor's hand, is relatively easy to detect through simple color filtering.

## 2.2 Audition

Nico has two small microphones for ears, which are connected to an off-the-shelf sound card. A simple intensity analysis is sufficient to accurately detect the sound of a drum beat in the audio stream. First the absolute value of the data is taken to find its intensity. Then the intensity values are "smoothed", by taking the average of a window of a fixed number of samples. If a value is lower than a minimum threshold, it is replaced with 0, otherwise it is replaced with 1.

This is defined by the analysis function $a = t \circ s$ where $w$ is the smoothing window size, $m$ is the minimum intensity threshold, $r(x)$ is the audio sample at time $x$, $s$ is the smoothing function and $t$ is the threshold function.

$$s(x) = \frac{\sum_{n=(x-w)}^{x} |r(n)|}{w}$$

$$t(x) = \begin{cases} 0 & \text{where } x \leq m \\ 1 & \text{where } x > m \end{cases}$$

Finally, a search for edges within this binary stream is performed. Any change in value that lasts for longer than a minimum time threshold is taken as either the beginning of a new drum hit (when switching from 0 to 1) or the end of one (when switching from 1 to 0).

Some limitations of this scheme are that 1) all drums sound the same to Nico, and 2) any sufficiently loud noise may produce a false positive in drum beat detection. The second limitation is not too severe, at least in the friendly environment of the lab. By quieting the audience and by adjusting the microphone gains and the thresholds used in the audio analysis algorithms, false positives can be minimized. The first limitation, however, requires extra, high-level processing to distinguish the drum beats that are detected from Nico's action from those played by the musicians accompanying Nico. This classification is discussed further in section 2.4.

## 2.3 Motor Control

Nico's arm contains six revolute joints, which roughly correspond to the joints in a human infant's arm. Basic motor control for Nico's drumming motion is implemented in software, whereby a "strike" command is translated into low-level physical motor directives. The robot calibrates its arm position against the placement of the drum – the drum need not be in exactly the same position for each trial.

The arm placement routine raises the arm to a fixed position over the drumhead, relative to the arm's resting position. It then manipulates its wrist motor until the robot's end effector comes into contact with the drum head. Fixing its wrist in this position, it produces a drum beat by raising and lowering its forearm from the elbow, so that the drum is struck each time the forearm is at its lowest position.

This motion is simplistic – human drumming motions are generally more efficient and employ more joint movement – yet adequate for the task. The entire drumming motion takes about 800 ms, limiting Nico's drumming speed to about 75 bpm. Some of the energy from the impact of the arm end-point against the drum head is absorbed in the arm, resulting in minor changes in arm position over time. In our experiments, these disturbances did not significantly affect Nico's performance.

## 2.4 Sensory Integration and Prediction

As Nico's control program runs, it collects a history of its inputs and outputs as a set of arrays of timestamps – one for visual ictus detections, one for audible drumbeat detections, and one for arm motion commands. These three streams are sufficient for the learning required for the robot to attune its drumming performance with its human partners. Streams of events can, of course, be re-cast as sets of oscillator measurements, and we will use these terms interchangeably in this section.

The system has several sources of error and indeterminacy, for which Nico gradually learns to compensate. First of all, the arm's tapping motion takes a significant and variable amount of time from the moment the controller sends the arm command to the time when Nico's end effector actually strikes the drum head. Nico learns to adjust for this delay by examining the time difference between when it swings its arm and when it detects a corresponding drumbeat in its audio stream.

In an ensemble performance, and in the absence of proprioceptive input, this task is made difficult by the need to classify sounds as self-generated or external. Accordingly, we provide a warmup period during which Nico is allowed to beat its drum by itself. After it has learned its own internal timings, it can use this self-knowledge to classify the drumbeat events it encounters during a performance. The audio stream should contain a pattern which differs from the arm-motion command history only by phase. This difference is, roughly, the amount of time that it takes to swing the arm, and it is precisely this value that Nico learns in the warmup period.

Nico must also learn to associate the streams of visual and auditory events that it encounters. The visual stream contains only inputs from others – Nico's gaze is directed toward the conductor or drummer, and away from itself. The auditory stream, on the other hand, contains the effects of both human and robot action. Only by looking at both streams at the same time can Nico properly classify its audio input, and thereby estimate the tempo (both frequency and phase) to which it must attune.

The warmup period is followed by a period of observation where Nico only listens and watches its counterparts, so as to learn the association between ictus and drumbeat events. After a few seconds, Nico's attentional oscillators are attuned to the ensemble's tempo. Only then is its arm engaged, allowing the robot to join the performance.

This causal relationship between stimulus and action does not mean that Nico's attentive oscillators merely act in resonance to external stimulus. Nico is programmed with a referent period (60 bpm) and its oscillators are active and self-sustained. While an initial "push" is required to start Nico's arm moving, it will continue to play its drum after the external stimulus terminates, at the tempo to which it has attuned. Likewise, during performance, sudden, brief changes in tempo do not disturb the robot's playing as they would if it were programmed to passively resonate to stimulus.

The following algorithm outlines the processing task.

1. *Sensory Integration*: Ictus events are associated with corresponding audio events. If events from two different sensory systems correspond, then this likely indicates a valid beat perception.

2. *Self-Awareness and Classification*: Arm motion events are associated with

corresponding audio events. This pattern is then compared against the beats found in the prior step, to determine the degree to which Nico is synchronized with its partners.

3. *Attunement*: An approximation of the external tempo is determined from the median of the intervals between the beats that have been deemed reliable in the first step.

4. *Prediction*: This tempo is used to calculate the next timepoint at which a beat will occur.

5. *Action*: Nico's self-knowledge is then used to predict the best time to next initiate an arm motion. If the perceived tempo is too fast, given its learned physical limitations, then the controller places the motion initiation time two beats (or more, if necessary) in the future.

The final step of this algorithm assures that robot performs effectively even when its physical limits are surpassed. When the tempo is too fast, Nico will still keep time, albeit on alternate beats. In the terms of Dynamic Attending Theory, it finds a lower *hierarchical level* to which it may comfortably attune. This behavior nicely parallels the results from Drake's synchronized tapping experiments [2].

## 2.5 Performance

We developed two types of trials, with some variations, which we then repeated several times, noting behaviors that seemed consistent in each. The first type of trial involves an interaction between a human musician and Nico. Both the human and Nico play the same drum. In the second type of trial, there are two humans – a conductor and a drummer who plays a different drum than Nico does. The conductor makes traditional conducting gestures, and the human drummer follows this direction as closely as he can. In order to ensure that Nico followed the humans' lead, rather than vice versa, the conductor and drummer used a silent, flashing-light metronome (invisible to Nico) to enforce tempo accuracy.

Each trial begins with the "warmup" period discussed in section 2.4, where Nico has a chance to independently explore its arm motion. Then the human drummer (and conductor, if there is one) begin(s) playing at a steady tempo. After an "observational" period during which Nico quietly watches and listens to its human companions, it begins to play. In the the most forgiving variation, the humans continue playing at a fixed tempo for an extended period of time, after which they stop playing. In another variation, the humans shift to higher and lower tempos at various points during the trial, spending tens of seconds at each tempo. The third variation on the testing procedure involves gradual tempo shifts between low and high values, over an extended period of time.

# 3 Results

We collected our performance data from the event streams that Nico uses for learning and prediction, giving us precise timing relationships for every action Nico perceived or initiated. The results shown here come from two representative test runs, out of 14 for which we recorded program data. Both tests involve one person drumming with Nico, and no conductor. In this test configuration, the “ictus” was generated by the motion of the drummer’s arm, rather than the conductor’s. In the figures below, ”Arm Swing” indicates the timepoints at which Nico’s control software generated a command to initiate its arm swing, ”Drum Detect” indicates the timepoints at which the audio processing subsystem detected a drumbeat, and ”Ictus” indicates the timepoints at which the visual processing subsystem detected the change in the direction of motion of the human drummer’s arm (corresponding to the drum being struck). The horizontal axis of each graph is measured in seconds.

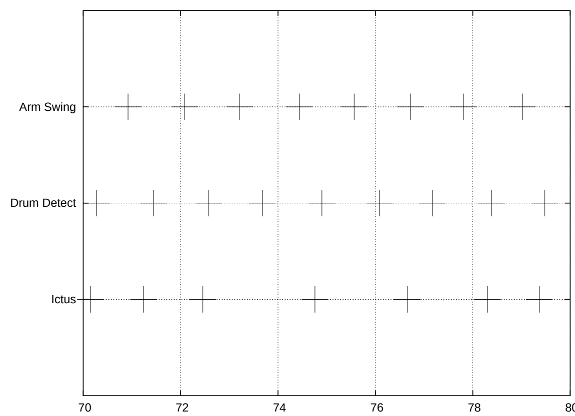


Figure 3: Test run #1, seconds 70-80. Nico performs in perfect synchrony with the human musician. Nico’s synchronization with the human drummer is so exact that their drumstrokes sound as one. Accordingly, Nico’s audition subsystem detects only one drum beat every 5/6 second. The interval between arm swing initiation, in software, and the corresponding drumbeat detection, can also be seen by comparing the pattern in the ”Arm Swing” and ”Drum Detect” event streams.

In the first test, the human drummer played the drum at a fixed tempo, 50 bpm, for an extended duration. A silent, flashing metronome was used to keep the beat accurate. 12 seconds later, at the 66-second mark of the recording, Nico makes its first arm swing, and achieves the correct beat immediately. The data from the 70 - 80 second interval shows the high accuracy of Nico’s performance (Figure 3). In several places during this interval, the ictus signal is lost, but the performance is not disturbed.

In the second test run, a human drummer started with a low tempo and then switched to a higher tempo after about 50 seconds had elapsed. In a subsequent interval (Figure 4), Nico loses the beat momentarily, and then recovers it.

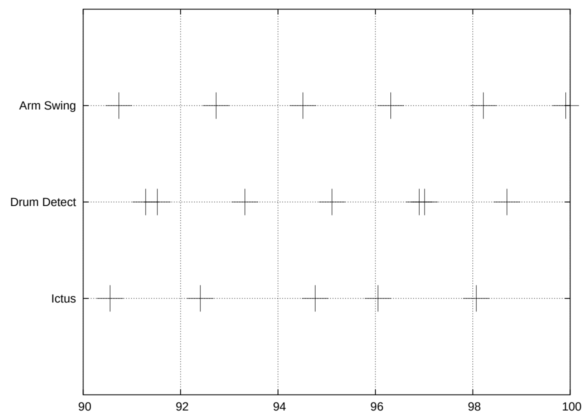


Figure 4: Test run #2, seconds 90-100. 30 seconds after the human drummer started playing, the ensemble is still at a low tempo. In intervals 90-92 and 96-98, Nico is not in perfect synchrony with the human drummer. Each of these intervals contains one arm swing initiation, but *two* drum beat detections. In these instances, Nico has struck its drum either before or after its human counterpart, by a noticeable amount, thereby generating the "extra" drumbeat.

Not long after, however, the drummer changes to a new, faster tempo, and Nico loses the beat for a second time (Figure 5). This state of discord continues until the drummer slows a bit, and Nico begins to regain synchrony (Figure 6). A few seconds later, Nico returns to a perfect unison performance with the human drummer.

# 4 Discussion

In the simplest scenario, when the oscillating patterns presented to the robot's eyes and ears maintain a constant tempo, Nico's entrainment usually occurs within a few seconds. When the tempo shifts, however, Nico may lose the beat temporarily, as it struggles to account for the varying intervals between recent reliable beats. Nico makes all of its calculations based on actual beat detections, so it has no way of knowing that an accelerando is happening until it detects a beat earlier than it expects. Human performers, by contrast, note the accelerating downstroke of a conductor's arm *before* the actual beat occurs, and thus can follow the indicated tempo shift without the lag time that Nico shows, even in its best performances. Drake [2] found that the synchronization performance of humans degrades in an impoverished sensory environment – for example, synchronization with a metronome is much more difficult than with the beats of a real musical performance. Likewise, Nico has access to very little

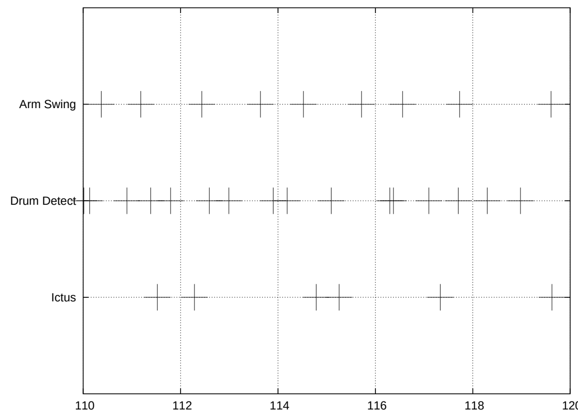


Figure 5: Test run #2, seconds 110-120. Although the human drummer is drumming regularly in this period, Nico's drumming is out of phase and frequency with its human counterpart. As a result, the pattern of drumbeat detections is irregular. The irregular arm swing data from this section confirms that Nico has yet to converge on a stable pattern. This behavior continues for 30 seconds after the interval shown here.

of the context of the oscillatory input it receives.

Interplay between the auditory and visual inputs helps provide some of the necessary context. As the tempo accelerates, ictus detection becomes less and less reliable, and the beneficial effect of sensory integration lessens. When Nico's oscillator is well-entrained to the beat provided by its human partners, the loss of reliable visual perceptions doesn't much matter. Even without support from the visual system, Nico can hear that drumbeats occur exactly when it expects them, and therefore maintains the proper frequency and phase. In this difficult environment, once Nico loses the beat, it has a hard time finding it again – especially if it manages to correlate an erratic ictus indication with one of its own erroneous drumbeats. It will tend to play erratically until it chances to get several solid perceptually-fused beats in a row, which it can then use to fix its own behavior. This accounts for the pattern we see in performance, where Nico plays very badly for half a minute or so, and then suddenly catches itself and plays perfectly for a while.

When Nico is having trouble finding the beat, it often settles into the same error conditions that Dynamic Attending Theory [2] would predict. Nico will often entrain a harmonically-related frequency to the one which its human partners present. For example, if the tempo shifts suddenly from a fast one to a slow one, Nico often entrains a frequency exactly double the true one. The robot's auditory input provides a steady, accurate-sounding oscillation (even though half the beats are entirely self-generated), and the video input confirms every other beat. Since the robot assumes that it will sometimes fail to detect an ictus, it assumes that it is playing correctly, even though fully half of its

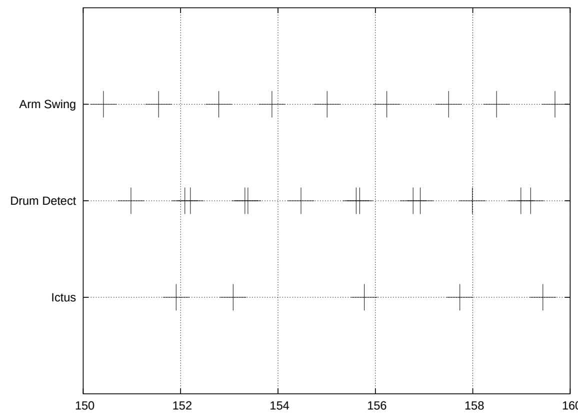


Figure 6: Test run #2, seconds 150-160. This interval shows Nico in the process of re-converging on a stable pattern. The consistent intervals between arm swings indicate that the robot has established the basic frequency, and is now experimenting to lock in the precise phase shift. Nico's beats occur alternately just ahead, just behind or exactly in time with the human drummer's. A few seconds after the interval shown here, Nico works out the correct timings and enters perfect synchrony with the human drummer. No figure is shown for that period since it is very similar to the interval shown in Figure 3.

drumbeats are off.

Finally, Nico encounters grave difficulties when the tempo is close to its maximum physical ability to beat. If the tempo is *very* fast, Nico reliably chooses to play every other beat, and problems don't emerge. If Nico tries to play two beats in quick succession, however, the arm sometimes fails to respond in time, and the drumbeat is delayed. To Nico, this beat does not look like a self-generated one, because the timing is wrong. Thus it begins to credit this delayed beat to its drumming partner, and tries to adjust the tempo accordingly.

Musicmaking is a cooperative, social task, and as such everyone involved wants it to succeed. As mentioned previously, we used a metronome to enforce a strict tempo among the human performers, but to do so challenged our ingrained propensity to adapt our drumming whenever Nico started to have difficulty. Human ensembles do this all the time – when a performer makes a mistake, the ensemble tries to accommodate it as smoothly as possible, adjusting their own playing to compensate. This mutual behavioral reinforcement is a characteristic of teaching with play [3]. Nico's active, self-sustained behavior presents, to the human participants, an opportunity for attunement and play. Interestingly, this engagement is evoked by the relatively simple, oscillator-based mechanism of Nico's control system.

# 5 Conclusion

Our research demonstrates an effective method for fusing diverse sources of oscillatory input of varying accuracy and phase shift in order to produce a reliable response in a social robot. Drumming, as a social task, proved to be well-suited for this experiment, as it requires sensory integration to compensate for the robot's limited perceptual ability. Our work also highlights what we believe are the essential aspects of competence in music making – perception and prediction, in contrast to the many examples of robotic musicianship that focus on mechanics. In addition, we have reconfirmed the benefit of developmental and behavioral approaches in the solution of real-time social tasks.

Functionally, it would be useful to extend the architecture described here with simple abstract models for music representation such that Nico may play a piece of music more sophisticated than simply striking the drum on every beat. Learning may be used to associate tempo values with sections of music, to provide better prediction in the performance of a given score. A richer representation would also provide more context for Nico's entrainment, hopefully replicating the improved performance that such context allows in humans [2].

Currently, our design is well-informed by current psychological models, but the implementation is only vaguely biologically-inspired. Specifically, we may be able to increase performance and reliability by refining our architecture for sensory integration based on more sophisticated psychophysical models. A formalized feedback control system, such as employed by Mukherjee [14], would allow much more sophisticated drumming motions and would improve Nico's ability to characterize different performance regimes.

Finally, the architecture explored here may be applied to other social tasks. For further research involving Nico, it would be relevant to explore the synchronization tasks in which a typical one-year-old might engage. This might include simple object manipulation, reaching, and game playing.

# 6 Acknowledgment

Support for this work was provided by a National Science Foundation CAREER award (#0238334) and the Doris Duke Charitable Foundation.

# References

[1] A. Rosenblum, Michael; Pikovsky, "Synchronization: from pendulum clocks to chaotic lasers and chemical oscillators," *Contemporary Physics*, 2003.

[2] C. Drake, M. R. Jones, and C. Baruch, "The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending." *Cognition*, vol. 77, no. 3, pp. 251–88, 2000.

[3] A. G. Brooks, J. Gray, G. Hoffman, A. Lockerd, H. Lee, and C. Breazeal, "Robot's play: interactive games with sociable machines," *Comput. Entertain.*, vol. 2, no. 3, pp. 10–10, 2004.

[4] P. Michel, K. Gold, and B. Scassellati, "Motion-based robotic self-recognition," in *IEEE-RAS/RSJ International Conference on Humanoid Robots*. Sendai, Japan: IEEE, 2004.

[5] G. Sun and B. Scassellati, "Reaching through learned forward model," in *IEEE-RAS/RSJ International Conference on Humanoid Robots*. Santa Monica, CA: IEEE, 2004.

[6] K. Gold and B. Scassellati, "Learning about the self and others through contingency," in *AAAI Spring Symposium on Developmental Robotics*. Palo Alto, CA: AAAI, 2005.

[7] S. E. Trehub, "The developmental origins of musicality." pp. 669–73, 2003.

[8] A. Z. Hajian, D. S. Sanchez, and R. D. Howe, "Drum roll: Increasing bandwidth through passive impedance modulation," *Proceedings IEEE International Conference on Robotics and Automation*, pp. 2294–2299, 1997.

[9] M. Williamson, "Robot arm control exploiting natural dynamics," Ph.D. dissertation, MIT Department of Electrical Engineering and Computer Science, 1999.

[10] I. Kato, S. Ohteru, K. Shirai, S. Narita, S. S. an T Matsushima, T. Kobayashi, and E. Fujisawa, "The robot musician wabot-2 (waseda robot-2)," *ROBOTICS*, vol. 3, no. 2, pp. 143–155, 1987.

[11] A. W. Ord-Hume, *Pianola : the history of the self-playing piano*. George Allen and Unwin, 1984.

[12] Toyota, "Toyota partner robot." [Online]. Available: http://www.toyota.co.jp/en/special/robot/

[13] E. E. Hannon and S. P. Johnson, "Infants use meter to categorize rhythms and melodies: Implications for musical structure learning," *Cognitive Psychology*, no. 50, pp. 354–377, 2005.

[14] R. Mukherjee, "Feedback control strategies for a nonholonomic mobile robot using a nonlinear oscillator," *Journal of Robotic Systems*, vol. 16, no. 4, pp. 237–48, April 1999.