Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009) Jyväskylä, Finland
Jukka Louhivuori, Tuomas Eerola, Suvi Saarikallio, Tommi Himberg, Päivi-Sisko Eerola (Editors)

Considerations Concerning a Methodology for Musical Robotics and Human-Robot Interaction

Birgitta Burger,[*1] Lüder Schmidt[*2]

[*]*Institute of Musicology, University of Cologne, Germany*
[1]b.burger@uni-koeln.de, [2]lueder.schmidt@uni-koeln.de

ABSTRACT

Robot technology is increasingly employed in artistic (musical) applications and as modeling tool for the investigation of general cognitive abilities and music related behavior in particular. Apart from the specifications of required system behavior and technological aspects of system design/implementation, problems occur concerning the evaluation of the systems. In recent approaches, techniques such as collecting informal reports, perceptual tests, video-based observational studies, or rating scales have been employed. Questions arise, however, as to the reliability and validity of these measures, and the lack of standardization diminishes the comparability of different studies. To attack these problems, we regard the paradigm of systematic observation as one suitable approach to study musical human-robot behavior. In a pilot study, observers were asked to assign labels to sequences of robot movements exhibiting features derived from expressive movements of musicians. Taking into account the implications of these remarks for the design of investigations, the generality of results may still be limited, e.g. by the choice of participants. As long as results of this kind cannot be integrated into a coherent theoretical framework, these considerations may run counter to recent attempts to use musical behavior as a testing field for principles underlying more general cognitive abilities/processes.

I. INTRODUCTION

Robot technology is increasingly employed in artistic (musical) applications and as a modeling tool for the investigation of general cognitive abilities and music related behavior in particular. Elsewhere (e.g. Schmidt 2005, 2007, 2008), we have argued for the inclusion of robotic modeling into an approach labeled *embodied cognitive science of music* as a means to integrate previous research in cognitive science of music in order to specify systems exhibiting a capacity for adequate behavior in realistic, music-related contexts. Taking up ideas extensively discussed by Pfeifer and Scheier (1999), but dating back to the work of psychologist M. Toda in the 1960s (cf. Toda 1982), such a system would ultimately need to comply with the requirements characterizing a *complete agent*, namely autonomy, self-sufficiency, situatedness, and embodiment (Pfeifer / Scheier 1999, Ch. 4). Current systems, however, seem to be far from achieving this goal.

Specifying a system will involve highly interdisciplinary work, drawing on disciplines such as psychology / cognitive science, (bio)mechanics, engineering, signal processing, and computer science; for a short sample of music-related robotic systems, subsumed under the heading of *musical robotics*, see e.g. Schmidt 2008, Ch. 5 or Kim et al. 2008.

Apart from the specification of required system behavior and technological aspects of system design/implementation, problems arise concerning the evaluation of the actually produced system behaviors. These problems may be encountered on different levels. Even in the absence of humans, robot behavior emerges from a complex interplay of system, environment, and task characteristics, which may be difficult to control and to estimate with regard to their consequences. Nehmzow (2008) therefore points out the need to develop adequate methods to quantify and compare observed robot behaviors in order to arrive at sound conclusions about the factors determining the behaviors and to be able to relate to underlying theoretical assumptions. While Nehmzow strives to adapt standard statistical methods to the quantification and evaluation of (mobile) robot behavior, Dautenhahn (2007) points out the need to include more informal approaches such as small-scale investigations or case studies within the field of *human-robot interaction* (HRI).

Regardless of the level of the approach – a more technically oriented evaluation of system behavior or considering system performance in a common environment with humans – there seems to be general agreement that the development of methodology is in its beginning, as yet exhibiting a lack of reliability and validity of the measures employed and thus leading to little comparability of different studies (e.g. Bartneck et al. 2008, 2009, Dautenhahn 2007, Nehmzow 2008). In recent work on music-related robotic applications, techniques such as collecting informal reports (e.g. Weinberg/ Driscoll 2006), perceptual tests (ibid.), video-based observational studies (Michalowski et al. 2007; Tanaka et al. 2005), or rating scales (Burger 2007) have been employed, liable to suffer from the problems described above.

Taking into account the implications of these remarks for the design of investigations, the generality of results may still be limited, e.g. by the choice of participants: The goal of assessing an application in a musically relevant context may lead to a focus on expert subjects, whose responses may not be representative of a general population (cf. the “user study” by Weinberg / Driscoll 2006, 38-42). As long as results of this kind cannot be integrated into a comprehensive theoretical framework, these considerations may run counter to recent attempts (e.g. Crick et al. 2006) to use musical behavior as a testing field for principles underlying more general cognitive abilities/processes.

In the following sections, we will take a closer look at some examples of previous work, touching upon problems arising from the fact that music-related robotic applications have to be considered within a broad interdisciplinary field, before returning to more specific methodological considerations in the context of a pilot study that was performed with participants of the International Summer School in Systematic Musicology 2007 (ISSSM 2007).

II. PRELIMINARY CONSIDERATIONS

The use of robotic applications as tools for investigation of interaction or relevant behavior in general is not established within music research. There are, however, applications, which can be considered relevant for topics currently discussed with respect to music-related behavior, such as the use of gestures and more general bodily movements, their expressivity resp. their relation to musical expressivity, or their temporal coordination / synchronization. In the following, we will offer some preliminary considerations, pointing towards possible restrictions of generality or problems of interpretation, that were inspired by reviewing some of these applications. The remarks represent just a first, rather informal collection without any claim to completeness.

A. Music and AI: Dancing and Drumming Robots

Recently, robots exhibiting regular rhythmic behavior such as dancing or drumming in musical contexts have attracted a substantial amount of interest. Most notorious – e.g. by downloads from YouTube and several awards – is the small "creature-like" (e.g. Michalowski et al. 2007) robot Keepon, explicitly developed to be fitting the expectations of small children. But even the dancing humanoids HRP-2 (e.g. Ikeuchi et al. 2008) and QRIO (see Tanaka et al. 2005) or the anthropomorphic drummer Haile (Weinberg / Driscoll 2006) have gained some fame.

In the context of robot drumming and dancing, music appears to play an analogous role to that taken with regard to auditory scene analysis in the 1990s (e.g. Cooke 1993), providing a clearly temporally structured environment to test more general principles of temporal coordination and alignment of behaviors. Benefits are expected for the understanding of "situated knowledge", "symbol grounding", and "social intelligence" (Aucouturier 2008) or the attribution of observed events such as drum beats to an agent's own actions (Crick et al. 2006) that generalize beyond the musical domain.

Studies of this kind will doubtlessly be valuable for the investigation of movement patterns and their execution, such as beating a drum or performing a dance step and the way they may be recognized. Nevertheless, the stance taken seems to imply domain generality of the underlying cognitive processes, which should at least be based on a conscious decision and checked against relevant research. For pointers to the recent discussion e.g. of the domain specificity of beat-based rhythmic processing and regular rhythmic movement see Patel 2008, Chapter 7.

B. Beat Extraction, Synchronization, and Interactional Synchrony

As indicated in the introduction, one motivation to consider methodological issues in robotics and human-robot interaction concerns the relation of theoretical assumptions thought to be governing a situation to the behaviors actually observed.

A common idea underlying approaches to rhythmic coordination in interaction is the extraction of time points for regularly occurring events, which may then be utilized to estimate appropriate times for the execution of behaviors such as moving the body (e.g. Michalowski et al. 2007) or beating a drum (Crick et al. 2006, Weinberg / Driscoll 2006). The extraction of time points may be based on acoustic data by detecting amplitude maxima or on optic data determining times of maximum / minimum positions (Crick et al. 2006) or the change of movement direction (Michalowski et al. 2007).

It is presupposed in these attempts that a) attunement to an externally presented regular (musical) beat is representative of interactional synchrony in general and that b) this attunement is constituted by the adjustment of frequencies of ongoing regular processes and the temporal matching of characteristic events. Both these suppositions should be considered with care: a) The attunement to an externally presented beat will certainly play an important role in musical practice. Nevertheless, in a typical "listening" situation, the listener will have little or no influence on the beat that is presented and will therefore rather react to than interact with the acoustic input. An alternative (obvious, according to Patel 2008, 100) role of an externally presented beat in an interactive situation may be to facilitate the temporal coordination of several actors by providing a common temporal scaffolding. b) Although the temporal attunement of regularly patterned processes will lead to a common temporal pattern, the observation of such a pattern in an interactive situation does not necessarily mean that independent regular patterns would be observed once the mutual influence of the processes is removed. Rather, the common pattern may be an effect of the interaction of processes that otherwise do not exhibit any regular patterning (cf. Pikovsky et al. 2001, 16).

Leaning too strongly on the idea of beat extraction in the investigation of interactional synchrony (in musical contexts) may entail the danger of letting considerations of feasibility take the lead over the design of systems, at the expense of possible alternative settings, in which different theoretical approaches might be compared.

C. Positive Emotion Bias

A tendency that might be called *positive emotion bias* may result from different reasons within music-related research based on robotic systems.

One of these reasons is related to ethical considerations and the choice of subjects: Apparently children form a particularly attractive group for the investigation of rhythmic behaviors in relation to robot movement and music (e.g. Tanaka et al. 2005, Michalowski et al. 2007). The participation of young children in the tests, however, should preclude the use of systems or behaviors that can provoke strong negative reactions.

Even more strongly, negative emotions should be avoided in therapeutic contexts, which form a growing field of investigation within human-robot interaction (e.g. Kozima et al. 2009, Dautenhahn / Werry 2004).

Bias towards positive emotions may also be a consequence of methodological decisions: Taking up approaches from human-computer or human-robot interaction in general will probably also entail a stronger focus on positive evaluation of the applications under consideration because system development (or interface development, which is at the heart of HCI research, cf. Norman 2008) will typically strive towards removal of aversive aspects (cf. the measures proposed by Bartneck et al. 2008, 2009) to enhance acceptance of a system.

D. Observation in the Field

Besides posing a challenge to integrate theoretical perspectives and to set into relation different cognitive processes thought to underlie music-related behavior, as one motivation to employ robotic applications as modeling tools we pointed out the possibility to arrive at systems with a capacity to exhibit adequate and relevant behavior in realistic musical contexts. In other words, a prospect may be the opportunity to move from purely computer-based modeling in connection with laboratory experimental (e.g. psychological) settings to conditions commonly considered more relevant for the investigation of musical experience and behavior.

In comparison to traditional ethological or social field research, the inclusion of (interactive) technical artifacts provides extended possibilities to manipulate the situation under investigation by changing the design or control parameters of the system employed, even though (see above) the consequences of such interventions may not readily be predictable or attributable to specific parameter changes.

Both in human-computer interaction and human-robot interaction, observational techniques have been applied (see Norman 2008, 99-100). In particular, the approach of *systematic observation* (Bakeman / Gottman 1997) seems to open a way to integrate aspects of experimental work into field research by providing a means to collect data in such a way that it is possible to reason about questions of reliability within an accepted framework (assessing inter-observer agreement or intra-observer consistency, cf. Bakeman / Gottman 1997, Ch. 4; Robson 2002, Ch. 11), and to further subject the data to quantitative (statistical) analysis. Examples can be found in the context of therapeutic applications of robots (Dautenhahn / Werry 2002), the investigation of interaction with robotic toys (e.g. Kahn et al. 2003), and again in the field of music-related rhythmic interaction with robots (e.g. Tanaka et al. 2005, Michalowski et al. 2007). For the sake of illustration, we will give a brief discussion of some aspects of the observational study presented by Michalowski et al. (2007).

To assess the influence of rhythmic coordination on human-robot interaction, children were asked in an open-house situation to dance with the small robot Keepon, whose movements were either in time or out of time with music played in the surrounding space. In the experiment, Keepon was set to perform rhythmic movements to a regular beat extracted from movement patterns captured by a camera that covered the area around the robot. Without any regular movement detected in the environment, Keepon would not move by itself; otherwise, movement patterns were generated that might or might not temporally align with the beat of the music.

Interactions between visitors and Keepon were recorded on video and in an initial analysis coded according to an event-based coding scheme. For every occurring interaction, it was noted whether robot movement was in synchrony with the music played and whether a child approaching the robot would take up regular rhythmic movement / dance (Michalowski et al. 2007, 92). Statistical analysis (Chi-square test) was performed on the obtained counts of the feature combinations, pointing towards a significant increase children's rhythmic movement during encounters when Keepon was moving in time with the beat of the music as compared to periods when Keepon's movements were desynchronized.

The authors make reservations against these results because of "rather subjective impressions on the part of the coder […]" (ibid.), i.e. they doubt the reliability of the analyzed data, and therefore proceed to present a refined coding scheme and behavioral analysis.

But even assuming a high degree of reliability of the code assignments, questions remain as to the conclusions that can be drawn from the data. Granting attunement of the observed children to external regular rhythmic patterns (cf. the remarks above, Section B), the primary source inducing the attunement cannot be determined from the recorded observations: As the beat underlying Keepon's rhythmic behavior is extracted from movement patterns observed in the environment, every time the robot moves in synchrony with the music, a detectable amount of rhythmic movement attuned to the music played must be expected – which may encourage the children's rhythmic engagement in the first place. Regrettably, this aspect is explicitly not taken up in the refined analysis, weakly motivated by a primary interest "in the development of interaction between the children and the robot" (Michalowski et al. 2007, 93).

This example illustrates that observational techniques, and the approach of systematic observation in particular, may be fruitfully applied to the investigation of interactive scenarios involving humans and robots. Nevertheless, the generation of reliable observational data and the correct application of statistical procedures will not ensure validity of the results. Rather, a careful design or adaptation of coding schemes is required, which in turn will benefit from a thorough theoretical and conceptual analysis.

III. PILOT STUDY

The approach of systematic observation was taken up in a pilot study on relations between the appearance of a small mobile robot, characteristics of robot movement patterns, and perceived emotional expressivity of the movements performed. The study was performed with participants of the International Summer School in Systematic Musicology 2007 (ISSSM 2007). Systematic observation comprises observing and analyzing an experimental scene (e.g. behavior of a robot or a human-robot interaction) by external observers. This analysis is performed by using a so-called *coding scheme*, in which certain events and behaviors are listed that are considered important for the experimental task. The procedure consists of first observer training, during which the observers become acquainted with the coding scheme, and second the observation and analysis of the observational task itself after the experimenter is sure that the observers are able to apply the coding scheme correctly. In the observation task, the observers mark every occurrence of the items listed in the coding scheme. As a measure of reliability, the results obtained by at least two observers (e.g. the filled-in coding schemes) are compared by calculating the *user agreement* between both observers. If the user agreement is sufficient, further statistical methods such as significance tests can be applied.

In the study described here, observers were asked to assign category labels to sequences of robot movements. These sequences were designed to display the three different emotional expressions happiness, anger, and sadness, exhibiting features derived from expressive movements of musicians (considering the amount, speed, fluency, and regularity of instrumental gestures) (Dahl / Friberg 2007) and that had already been tested in a previous experiment (Burger 2007). This study addressed two questions: whether 1) the outer appearance of the robot played a significant role for the assignment of labels and whether 2) the assignment of labels was influenced by the kind of labels used.

Due to the technical difficulties typically involved in the design of robots (cf. the discussion of the design process e.g. by Dautenhahn 2007), with advanced robotic systems it will be hard to include specimens of different appearance in a single study. In our case, this was made possible by using the Lego Mindstorms NXT system (see http://mindstorms.lego.com), a modular, programmable robotic kit. Observer training was performed with the same robot that was used in the experiment described in Burger 2007, though its appearance had to be adapted slightly to the experimental conditions. As can be seen in Figure 1, this robot shows some features that were considered anthropomorphic by participants in the experiment reported by Burger (2007), such as “arms” and “eyes” (ultrasonic sensors). The experiment was performed with four observers divided into two pairs, who conducted the experiment separately. During the training phases, both pairs perceived the same movements of the robot, containing each of the three emotional sequences, but they were given different coding schemes. The scheme of group A consisted of the set of descriptive labels *sad*, *angry*, and *happy*, taking up the intended expressive content of the movements, while the other pair, group B, was to assign non-descriptive labels: *1* for the movements that the other pair assigned as *sad*, *2* for the *angry* movements, and *3* for the *happy* movements. After the training was considered successful, i.e. when all the observers were able to assign the labels correctly, the observational task was started. It was performed in two parts. In the first part, the robot used for training was employed; the second part was done with a different robot (see Figure 2) that looked more like a car and less anthropomorphic. Both parts contained 48 movement sequences, each lasting 10 seconds, which were presented in 4 blocks. The three different expressions were presented in randomised order, but equally distributed in the 12 sequences of each block. The order of presentation was changed between parts 1 and 2, but remained the same for both pairs. The observers received a prepared sheet of paper on which they marked the labels they assigned to each presentation.

Figure 1. The anthropomorphic robot that was used in the observer training

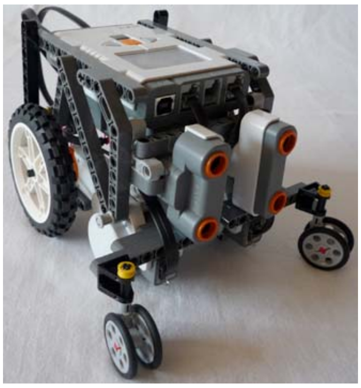

Figure 2. The less anthropomorphic robot that was used in the observational task

In the next step, the completed coding schemes were used to assess the *observer agreement* by calculating the so-called *Kappa coefficient* (see Bakeman / Gottman 1997, Ch. 4). The decisions of the observers were displayed one against the other in a matrix. The matrices for groups A and B can be seen in Figures 3 and 4; data for part 1 are printed in black, data for part 2 in grey. As can be seen in the figures, the results of the first and second parts do not differ much.

The Kappa coefficients for group A are 0.968 (first part) and 0.969 (second part), which can be considered as excellent values according to the rules of thumb quoted by Bakeman / Gottman (1997, 66) resp. Robson (2002, 342). The values 0.125 (part 1) and 0.219 (part 2) calculated for Group B, however, must be considered as insufficient.

The results of group B illustrate the need for careful development of coding schemes: The low observer agreement appears to result from the fact that observer 2 of this group consistently confused the labels *2* and *3* during the observation. Exchanging the assignments of these two labels for observer 2 would result in Kappa values of 0.876 for the first part and 0.781 for the second part, which again could be considered as excellent by the rules of thumb.

<table>
<tr><th rowspan="6">1st observer</th><th colspan="5">2nd observer</th></tr>
<tr><th></th><th>sad</th><th>angry</th><th>happy</th><th>sum</th></tr>
<tr><th>sad</th><td>||||| ||||| ||||| |
||||| ||||| ||||| |</td><td></td><td></td><td>16
16</td></tr>
<tr><th>angry</th><td></td><td>||||| ||||| ||||
||||| ||||| |||||</td><td></td><td>14
15</td></tr>
<tr><th>happy</th><td></td><td>|
|</td><td>||||| ||||| ||||| |
||||| ||||| ||||| |</td><td>17
17</td></tr>
<tr><td>sum</td><td>16
16</td><td>15
16</td><td>16
16</td><td>47
48</td></tr>
</table>

Figure 3. Observer matrix for group A – top row (black): observer agreement in Part 1; bottom row (grey): observer agreement in Part 2

<table>
<tr><th rowspan="6">1st observer</th><th colspan="5">2nd observer</th></tr>
<tr><th></th><th>1</th><th>2</th><th>3</th><th>sum</th></tr>
<tr><th>1</th><td>||||| ||||| ||||| |
||||| ||||| ||||| |</td><td></td><td></td><td>16
16</td></tr>
<tr><th>2</th><td></td><td>||
||||</td><td>||||| ||||| ||||
||||| ||||| |||</td><td>16
17</td></tr>
<tr><th>3</th><td></td><td>||||| ||||| ||||
||||| ||||| ||</td><td>||
|||</td><td>16
15</td></tr>
<tr><td>sum</td><td>16
16</td><td>16
16</td><td>16
16</td><td>48
48</td></tr>
</table>

Figure 4. Observer matrix for group B – top row (black): observer agreement in Part 1; bottom row (grey): observer agreement in Part 2

From the within-group similarity of data obtained in the first and second parts of the observational sessions, the impression is gathered that the differences in outer appearance of the two robots did not influence the (correct) assignments of categories. A formal test of similarity was not performed; a high probability for a Type I error in a test for differences between parts 1 and 2 could be taken as an indirect indication of a low probability for committing a Type II error (cf. Bortz 1999, 160-161). The influence of a robot’s appearance will be highly dependent on tasks contexts of application. Thus, in general a scheme for its investigation will need to be much more elaborate than the ideas sketched here. However, the approach of systematic observation may offer one way of including this aspect in a structured manner in further research.

The apparent confusion of labels by group 2 again illustrates the need for careful development and testing of coding schemes as well as sufficient observer training. Although all observers seemed to be able to differentiate and identify observed behavioral patterns, the more descriptive labels may have facilitated remembering the association of label and pattern from the training phase. The movement patterns and coding schemes described here were rather minimal. If the behavioral sequences to observe and accordingly the coding schemes to be applied become more complex, this kind of problem will be more prominent, especially if coding has to be performed in real time. In addition to a clearly defined coding scheme and sufficient observer training, the use of video recordings as exemplified by the study discussed above will offer the chance to introduce further control by repeated coding of the same sequences.

IV. CONCLUSION

The presentation given here may be symptomatic of a field of research that is still in its initial stage, characterized by diverse studies that are hard to relate to each other and may best be described as exploratory. Accordingly, within robotics and human-robot interaction, concerns about reliability, validity, and generalizability of results are discussed, resulting in a desire to develop standardized methods and measuring tools. As a general trend, attempts to adapt techniques from social and behavioral research to situations including humans and robotic artifacts can be seen. As one interesting approach, we have described an attempt to take up systematic observation in the investigation of human-robot interaction, which may correspond to the extended amount of control in a scenario introduced by the use of robots. The major challenge, however, will remain to be posed by the development of a common conceptual and theoretical framework. An increasing conceptual integration and clarity will enhance the possibilities to develop relevant tasks and to relate experimental findings with each other. On the other hand, the increasing demand for conceptual and theoretical analysis generated by the new opportunities offered by the use of robotic applications will certainly be beneficial for the cognitive science of music.

ACKNOWLEDGMENT

We would like to thank Benjamin Buch, Son-Hwa Chang, Olmo Cornelis,Eva Istók, Jin Hyun Kim, Dirk Moelants, Malte Münster, Jochen Arne Otto, and Uwe Seifert for their invaluable contributions in discussions of experimental setup, building and programming the robots used, participating as observers, and evaluating the data.

REFERENCES

Aucouturier, J.-J. (2008). Cheek to Chip: Dancing Robots and AI's Future. *IEEE Intelligent Systems*, 74-84.

Bakeman, R., & Gottman, J. M. (1997). *Observing Interaction: An Introduction to Sequential Analysis*. Cambridge: Cambridge University Press.

Bartneck, C., Kulic, D., & Croft, E. (2008). Measuring the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. In *Proceedings of the Metrics for Human-Robot Interaction: Workshop in affiliation with the 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI'08)*, 37-44.

Bartneck, C., Kulic, D., Croft, E., & Zoghbi, S. (2009). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics 1(1)*, 71-81.

Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer-Verlag.

Burger, B. (2007). *Communication of Musical Expression from Mobile Robots to Humans: Recognition of Music Emotions by Means of Robot Gestures*. Masters Thesis. University of Cologne / Royal Institute of Technology Stockholm.

Cooke, M. (1993). *Modeling Auditory Processing and Organisation*. Cambridge: Cambridge University Press.

Crick, C., Munz, M., & Scassellati, B. (2006). Synchronization in Social Tasks: Robotic Drumming. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN'06)*, 97-102.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception 24(5)*, 433-454.

Dautenhahn, K. (2007). Methodology & Themes of Human-Robot Interaction: A Growing Research Field. *Journal of Advanced Robotic Systems 4(1)*, 103-108.

Dautenhahn, K., & Werry, I. (2004). Towards interactive robots in autism therapy. Background, motivation and challenges. *Pragmatics & Cognition 12(1)*, 1-35.

Kahn, P. H., Friedman, B., Freier, N. G., & Severson. R. (2003). *Coding Manual for Children's Interaction with Aibo, the Robotic Dog – The Preschool Study*. Technical Report, Department of Computer Science and Engineering, University of Washington. ftp://ftp.cs.washington.edu/tr/2003/04/UW-CSE-03-04-03.pdf

Kozima, H., Michalowski, M. P., & Nakagawa, C. (2009). Keepon. A Playful Robot for Research, Therapy, and Entertainment. *International Journal of Social Robotics 1(1)*, 3-18.

Michalowski, M. P., Sabanovic, S., & Kozima, H. (2007). A Dancing Robot for Rhythmic Social Interaction. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction (HRI'07)*, 89-96.

Nehmzow, U. (2009). *Robot Behaviour. Design, Description, Analysis and Modeling*. London: Springer-Verlag.

Norman, K. L. (2008). *Cyberpsychology. An Introduction to Human-Computer Interaction*. New York: Cambridge University Press.

Patel, A. D. (2008). *Music, Language, and the Brain*. Oxford: Oxford University Press.

Pfeifer, R., & Scheier, C. (1999). *Understanding Intelligence*. Cambridge, MA: MIT Press.

Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). *Synchronization. A universal concept in nonlinear sciences*. New York: Cambridge University Press.

Robson, C. (2002). *Real World Research. A Resource for Social Scientists and Practitioner-Reseachers. Second Edition*. Oxford: Blackwell Publishing.

Schmidt, L. (2005). Towards an "Embodied Cognitive Science of Music": Incorporating Mobile Autonomous Robots into Musical Interaction. *Proceedings of the 2nd International Conference of the Asia Pacific Society for the Cognitive Science of Music (APSCOM-2), Seoul, 2005*, 48-54

Schmidt, L. (2007). Embodied Cognitive Science as a Paradigm for Music Research. In C. Lischka & A. Sick (Eds.), *Machines as Agency. Artistic Perspectives* (pp. 48-62). Bielefed: transcript Verlag

Schmidt, L. (2008). *Embodied Cognitive Science of Music. Modeling Experience and Behavior in Musical Contexts*. Unpublished doctoral dissertation, Cologne University.

Tanaka, F., Fortenberry, B., Aisaka, K., & Movellan, J. R. (2005). Developing Dance Interaction between QRIO and Toddlers in a Classroom Environment: Plans for the First Steps. In *Proceedings of the IEEE International Workshop on Robots and Human Interactive Communication (RO-MAN'05)*, 223-228.

Toda, M. (1982). *Man, Robot, and Society. Models and Speculations.* Boston: Martinus Nijhoff Publishing.

Weinberg, G., & Driscoll, S. (2006). Toward Robotic Musicianship. *Computer Music Journal 30(4)*, 28-45.