On Stage: Robots as Performers

Guy Hoffman
School of Communication
IDC Herzliya
Herzliya, Israel 46150
Email: hoffman@idc.ac.il

***Abstract*—This paper suggests to turn to the performative arts for insights that may help the fluent coordination and joint-action timing of human-robot interaction (HRI). We argue that theater acting and musical performance robotics could serve as useful testbeds for the development and evaluation of action coordination in robotics. We also offer two insights from theater acting literature for HRI: the maintenance of continuous sub-surface processes that manifest in motor action, and an emphasis on fast, inaccurate responsiveness using partial information and priming in action selection.**

I. Introduction

Human acceptance of personal, social, and service robots depends on a number of factors in terms of the robot's behavioral and surface traits. Several studies have been dedicated to the effects of robot *appearance* on humans [e.g. 8, 10, 15], and much of the human-robot interaction literature deals with the effects of the verbal and non-verbal *content* of robot behavior, such as the selection of utterances, gestures, and gaze behavior on the part of the robot [e.g. 17, 22, and many more].

However, humans are sensitive not only to the content, symbols, and categories of interaction tokens, but also to their timing. In human-human joint activities, subjects care about when verbal and non-verbal events occur [6]. Timing effects have also been shown in human-robot interaction, most notably with respect to delays of robot activity relative to human actions [24, 26]. In our own work, we have shown that not only discrete post-action delays, but also anticipatory action relative to human activity at sub-action resolution causes subjects to evaluate virtual characters and robots as more fluent, more committed, and more contributing to the team, when compared to robots that were purely reactive, and thus trailing the subjects' behavior [11, 12].

Still, most current HRI models are structured in discrete-action turn-taking frameworks, based on dialog, planning, and state-action models. This results in a rigid stop-and-go interaction, which is often neither efficient nor fluent, but instead imposes unnatural structure on human participants. Research in action fluency, sub-action coordination, and timing could therefore prove valuable for the acceptance of robots by untrained humans.

II. Robots on Stage

For several years we have been using stage performance robots as a framework to develop and evaluate methods for temporally coordinated human-robot activities.

Stage performance, whether theatrical or musical, is valuable both as an implementation platform and as a testing ground for HRI research. On the one hand, it is relatively constrained, be it by script, by score, or by set design, thus limiting the perception and actuation expectations of a robotic system. On the other hand, it can provide for a rich environment in which a robotic agent meshes its actions with a human partner, incorporating dialog, sensory processing, action selection, and action timing. Importantly, stage performance can allow researchers to isolate elements of human-robot interaction, while fixing other aspects as necessary.

A theater robot could, for example, have pre-programmed verbal cues and positions, but would still need to respond in real time to the particular actions and timing parameters of human actors or audiences, which will vary from performance to performance. In music, a robot could be limited to a certain scale, while still be required to improvise in real time based on other musicians' phrases, styles, tempi, and rhythms.

Furthermore, performances are rich in that they may incorporate verbal and non-verbal behavior; and audiences of theatrical and musical performances are highly sensitive to notions of action coordination, timing, and fluency, emphasizing these aspects of human-robot interaction.

Robotic theater may thus prove to be a good candidate for a "grand challenge" of fluent human-robot joint action, dialog, collaboration, and joint activities.

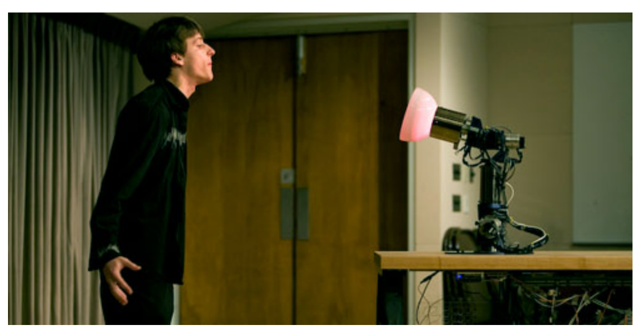

Fig. 1. Scene from a stage production using the described robotic theater system.

That said, historically robotic stage performers have been few and far between. Most work has dealt with fully scripted or extremely simple behavior on one end of the spectrum (for a good review, see Dixon [9]) or with fully teleoperated robots on the other [18, 23]. In several performances, robots have been partnered with each other on stage without the inclusion of a human scene partner [3, 19].

In our own work, we have developed an authoring system for human-robot theatrical performances, using a standard analysis of a play according to modern acting methodology: in this system, a scene is a sequence of short beats, each of which describes a gesture on the robotic character's part. Beats are animated separately, and annotated in terms of their timing with respect to the human actor's action. In the same project, we also developed a hybrid control system, which allows a puppeteer to time the responses of the robot at sub-action resolution, while allowing the robot to autonomously control other aspects of the performance, such as eye-gaze inverse kinematics, animacy layers, and motion layering and interpolation [14].

As for robots used in musical performances, much of the research focuses on sound production, and does not address interactive aspects of musicianship. Most musical robots can be classified into two groups: robotic musical instruments, which are mechanical constructions that can be played by live musicians or triggered by pre-recorded sequences [27, 7]; or anthropomorphic musical robots that attempt to imitate the action of human musicians [28, 32]. Only a few attempts have been made to develop perceptual, interactive robots that are controlled by autonomous methods [1, 33].

Recently, we demonstrated a real-time adaptive and autonomous robotic improvisation system, using anticipatory timing and gesture-based opportunistic improvisation to continuously adapt the robot's improvisation and choreography, while playing simultaneously with a human musician [13]. The system uses anticipatory beat-matched action to enable real-time synchronization with the human player, while still allowing for the dynamic adaptation required in an improvisational music ensemble.

Both our systems, the theatrical and the musical, were used repeatedly in live performances, and demonstrated simultaneous adaptation to the human performance and real-time responsiveness for human-robot fluent coordination.

Fig. 2. Scene from a live performance using *Shimon*, a gesture-based musical improvisation robot.

III. Acting Lessons for HRI

But the utility of theater goes beyond using the stage as a human-robot interaction laboratory. Human acting method and theory holds valuable insights into some of the questions researchers in HRI are also tackling. An actor's preparation of a role includes a systematic investigation of what gesture, body pose or physical action best describes the internal drive and objective of their character in different contexts. Good actors pay attention to conventions of nonverbal communication and often need to take on the difficult task of portraying complex trajectories without words. Much of the repetitive practice actors engage in is aimed to perfect the timing and coordination between a number of agents acting together.

These challenges bear similarity to those one faces when designing behaviors for human-robot interaction.

In a way, personal and social robots are required to "play a part" in their human environments, displaying artificially constructed social signals and accommodating human expectations. Viewing personal robots thus as "actors" on the human stage, with an aim of achieving higher acceptance with their audience through improved timing coordination, it could make sense to look towards acting method and theory for design guidelines and cues.

Indeed, a reading of several canonical texts on modern theater acting holds valuable insights for designers of fluently meshed human-robot coordination systems. Two of these, discussed below, are continuity and responsiveness:

A. Continuity

Nineteenth-century acting was much akin to human-robot interaction today. According to the prevalent method—known as the DelSarte system—behavior was analyzed and separated into discrete action tokens, which were stringed together into a performance [31]. Modern acting—usually called the Stanislavski Method, or simply "the method"—takes a decidedly different approach, in which surface actions are, in fact, representations of continuously evolving sub-surface developments in the arc of a performer.

Sonia Moore, author of the definitive book on the Stanislavski method, quotes Eugene Vakhtangov saying "[a] unit in a role or a scene is a step in moving the through-line of actions toward the goal," [21]. This as opposed to the unit being a single line of dialog or a stage direction by itself.

The method also encourages actors to not memorize the lines, but instead to focus on analyzing a scene in terms of moving powers, objectives, obstacles, and intentions, leading to choosing actions.

Other teachers share this view. Michael Chekhov speaks of "Psychological Gestures" that draw on the character's "definite desire" in a scene [4], and Augusto Boal, too, stresses that any particular action results from the character's desires, will, and needs [2].

This then leads to the concept of the *inner monologue*, which is heavily emphasized by Moore's method.

> "[Your inner monologue] is more important than memorizing your lines [...]. The right inner mono-

logue will bring you to your lines, and you may have entirely different intonations." [21]

The actor's inner monologue must carry on the whole time the actor is on stage, whether they say something or not. This inner monologue should usually be laid out in detail while preparing for a role, and lends the actor credibility of an internal process while they're on stage, leading up to the lines and thus preventing the lines to be uttered in a void.

1) Application to robotics: The notion of maintaining continuity through inner processes has applications for personal and interactive robots, and could be prescriptive when aiming for fluently behaving robotic agents.

Just as an actor is advised to constantly conduct an inner monologue to achieve continuity and realism in their behavior, and to make their externally evident actions internally based, action selection in interactive and personal robotics could also benefit from stemming not only from the most recent input and decision-node, but instead growing out of a continuous and multi-layered stream of constantly changing internal parameters.

Endowing an HRI robot with such an "inner monologue" might have a similar effect on robots as it has on human actors, i.e. result in a more natural and continuous interaction. It could thus help avoid the command-and-response behavior robots usually display, and lead to more fluency and acceptance in human-robot joint activities.

Moreover, the physical manifestations of this internal processing, in the form of semi-pronounced gestures, lexical lookup gestures, and emotional motor activity, could serve both the appeal and the readability of the robot's internal processes.

A possible implementation is an opportunistic action selection mechanism, in which robot activity is continuous, based on the above-mentioned physical manifestations. The robot can, at any point in time, choose to fully produce one of several implied actions. The actual action selection occurs opportunistically based on current perceptual processing, but in concordance with the continuous internal processing and its motor manifestations.

In our work on musical improvisation robots, we have used a similar approach, in which the robot is continuously moving during its improvisation process. In that system, the robot does not respond discretely to a human's phrase, but instead plays continuously. At times, the robot opportunistically responds to a human cue when it fits the current musical movement. At other times, it merely integrates the human's motifs and styles into the parameters specifying the current improvisation algorithm [13].

B. Responsiveness

A second principle appearing in much of the acting theory and method literature is that of quick and intuitive responsiveness. New York City acting guru Sanford Meisner is generally credited with the emphasis on responsiveness in acting. His rule, embodied in a now famous "repetition exercise", states "[d]on't do anything unless something happens to make you do it. What you do doesn't depend on you, it depends on the other fellow." [20]

In another place, he commands that "acting isn't chatter, it's responding truthfully to the other person". This rule, in Meisner's method, is the key to responsiveness, and to meaningful behavior in the on-stage collaboration between two actors.

Meisner is not the only acting theorist to find that the focus of a scene is not happening within any of the actor's minds, or even behaviors, but in the space between the two actors. Viola Spolin calls this notion "communication": "the techniques of the theater are the techniques of communicating. The actuality of the communication is far more important than the method used." [30]

Similarly, Ruth Maleczech speaks of "repercussion": "[t]he other actors are, for me, like the bumpers in a pinball machine. I shoot my pinball, my image, and it goes tch, tch, tch, bouncing off those bumpers, each hit having its repercussion. Often the next image will come directly from the response of the other actor." [in 29]

Moore adds that "[e]nsemble work means continuous inner and external reaction to each other."

These principles are often manifested in actors' exercises, in which pairs of actors need to respond quickly and directly to other actors' behaviors, words, or actions. In the escalation of quick responses, a flow emerges that is then translated into the scripted segments of the final performance.

1) Application to robotics: At first glance, this acting maxim seems trivial for interactive robots. After all, robots are inherently reactive machines. Both the cybernetic tradition formulates robot behavior in terms of responses to the environment, and also early interactive performance robots "moved and emitted different sound patterns in response to movements and light changes going on around them" [9].

However, looking at the particular exercises that embody this principle, emphasis is usually put on the speed, immediacy, and timeliness of the response, rather than merely on the mutual structure of the exchange. In these interactions, synchronizing one actor's actions with the other's, while maintaining mutual responsiveness, is key.

There has been some work on human-robot synchronization, both in the musical and non-musical domain. These include real-time auditory beat detection with gradual step changes in a humanoid [34] multi-modal perception for robotic drummer synchronization [5], interactive human-robot improvisation sessions [33], and most recently, a robot that tracks the beat of both music and a conductor's baton to time its movements [16].

Our own approach to carry over the notion of Meisner "responsiveness" to interactive robotics is a two-tiered framework in human-robot dialog, be it verbal or non-verbal. One part is the immediate, quick response, potentially based on incomplete information; the other is be a more calculated, processed response following later in the interaction.

In previous work on anticipatory action selection, we developed a system that—in some cases—acts quickly and based on partial information. The robot in those experiments could then correct its initial movement when more information accumulated. In human-subject studies we found this to lead to evaluations of higher fluency by human members in human-robot teams [12].

Another method to achieve something akin to "intuitive responsiveness" is to prime the robotic system towards a certain input, be it perceptually or physically. Perceptual priming can lead to reduced action selection response time, as we have shown in our work on top-down anticipatory action selection for human-robot teamwork. Similarly, physically positioning the robot towards an action, a method we have used in our robotic musicianship work, can result in faster motor response, leading to a more concurrent joint action system [13].

IV. Conclusion

In summary, we argue that humans are sensitive to timing and fluent action coordination, both among themselves, and in their interaction with artificial agents and robots. With the aim of improving the timing of coordinated human-robot interaction, we suggest to turn to the performative arts for inspiration, as these have a long history of action-timing methodology.

Specifically, we argue that theater acting and musical performance robotics could serve as useful testbeds for the development and evaluation of action coordination in robotics.

Moreover, we survey acting literature as it applies to developing robots that interact fluently with human counterparts. A reading of acting method texts suggests two insights for HRI: maintaining continuous sub-surface processes that manifest in opportunistic external behavior; and focusing on fast, immediate responsiveness in action selection, even in situations of incomplete information.

References

[1] N A Baginsky. The Three Sirens: a Self-learning Robotic Rock Band. http://www.the-three-sirens.info/, 2004.

[2] Augusto Boal. *Games for Actors and Non-Actors*. Routledge, 2nd edition, July 2002.

[3] Allison Bruce, Jonathan Knight, Sam Listopad, Brian Magerko, and Illah Nourbakhsh. Robot Improv: Using drama to create believable agents. In *Proceedings of ICRA 2000*, volume 4, pages 4002–4008, April 2000.

[4] Toby Cole and Helen K Chinoy. *Actors on Acting : The Theories, Techniques, and Practices of the World's Great Actors, Told in Their Own Words*. Crown, 3th edition, 1970.

[5] Christopher Crick, Matthew Munz, Tomislav Nad, and Brian Scassellati. Robotic drumming : synchronization in social tasks. In *15th IEEE International Symposium on Robot and Human Interactive Communication, (RO-MAN '06)*, 2006.

[6] C. L. Crown. Coordinated Interpersonal Timing of Vision and Voice as a Function of interpersonal Attraction. *Journal of Language and Social Psychology*, 10(1):29–46, March 1991.

[7] R B Dannenberg, B Brown, G Zeglin, and R Lupish. McBlare: a robotic bagpipe player. In *NIME '05: Proceedings of the 2005 conference on New interfaces for musical expression*, pages 80–84, Singapore, Singapore, 2005. National University of Singapore.

[8] Carl F. DiSalvo, Francine Gemperle, Jodi Forlizzi, and Sara Kiesler. All robots are not created equal. In *Proceedings of the conference on Designing interactive systems processes, practices, methods, and techniques - DIS '02*, page 321, New York, New York, USA, June 2002. ACM Press.

[9] Steve Dixon. Metal performance: humanizing robots, returning to nature, and camping about. *The Drama Review*, 48(4), 2004.

[10] J. Goetz, S. Kiesler, and A. Powers. Matching robot appearance and behavior to tasks to improve human-robot cooperation. In *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003.*, pages 55–60. IEEE, 2003.

[11] Guy Hoffman and Cynthia Breazeal. Cost-Based Anticipatory Action-Selection for Human-Robot Fluency. *IEEE Transactions on Robotics and Automation*, 23(5):952–961, October 2007.

[12] Guy Hoffman and Cynthia Breazeal. Effects of anticipatory perceptual simulation on practiced human-robot tasks. *Autonomous Robots*, 28(4):403–423, December 2009.

[13] Guy Hoffman and Gil Weinberg. Gesture-based human-robot Jazz improvisation. *Proceedings of the 2010 IEEE International Conference on Robotics and Automation (ICRA)*, 2010.

[14] Guy Hoffman, Roni Kubat, and Cynthia Breazeal. A hybrid control system for puppeterring a live robotic stage actor. In *Proceedings of the17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, 2008.

[15] T. Kanda, T. Miyashita, T. Osada, Y. Haikawa, and H. Ishiguro. Analysis of Humanoid Appearances in HumanRobot Interaction. *IEEE Transactions on Robotics*, 24(3):725–735, June 2008.

[16] Youngmoo E Kim, Alyssa M Batula, David Grunberg, Daniel M Lofaro, Junho Oh, and Paul Y Oh. Developing Humanoids For Musical Interaction. In *10th IEEE-RAS International Conference on Humanoid Robots.*, 2010.

[17] Sonya S. Kwak. Personality design of sociable robots by control of gesture design factors. In *RO-MAN 2008 - The 17th IEEE International Symposium on Robot and Human Interactive Communication*, pages 494–499. IEEE, August 2008.

[18] Les Freres Corbusier. Heddatron. http://www.lesfreres.org/heddatron/, 2006.

[19] CY Lin, CK Tseng, WC Teng, WC Lee, CH Kuo, HY Gu, KL Chung, and CS Fahn. The realization of robot theater: Humanoid robots and theatric performance. *International Conference on Advanced Robotics, ICAR 2009*, 2009.

[20] Sanford Meisner and Dennis Longwell. *Sanford Meisner on Acting*. Vintage, 1st edition, August 1987.

[21] Sonia Moore. *Training an Actor: The Stanislavski System in Class*. Viking Press, New York, NY, 1968.

[22] Jonathan Mumm and Bilge Mutlu. Human-robot proxemics. In *Proceedings of the 6th international conference on Human-robot interaction - HRI '11*, page 331, New York, New York, USA, 2011. ACM Press.

[23] Robin Murphy, Dylan Shell, Amy Guerin, Brittany Duncan, Benjamin Fine, Kevin Pratt, and Takis Zourntos. A Midsummer Nights Dream (with flying robots). *Autonomous Robots*, 30(2):143–156, October 2010.

[24] Yusuke Okuno. Providing Route Directions : Design of Robot s Utterance , Gesture , and Timing. In *Proceeding of the ACM/IEEE international conference on Human-robot interaction: HRI'09*, pages 53–60, 2009.

[25] K Petersen, J Solis, and A Takanishi. Toward enabling a natural interaction between human musicians and musical performance robots: Implementation of a real-time gestural interface. In *Proceedings of the17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, 2008.

[26] Toshiyuki Shiwa, Takayuki Kanda, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. *How quickly should communication robots respond?* HRI '08. ACM Press, New York, New York, USA, 2008.

[27] E Singer, K Larke, and D Bianciardi. LEMUR GuitarBot: MIDI robotic string instrument. In *NIME '03: Proceedings of the 2003 conference on New interfaces for musical expression*, pages 188–191, Singapore, Singapore, 2003. National University of Singapore.

[28] J Solis, K Taniguchi, T Ninomiya, T Yamamoto, and A Takanishi. The Waseda Flutist Robot No 4 refined {IV}. In *Proceedings of theIEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2007)*, pages 2041–2046. IEEE, 2007.

[29] Janet Sonenberg. *The Actor Speaks : Twenty-four Actors Talk About Process and Technique*. Three Rivers Press, 1st edition, April 1996.

[30] Viola Spolin. *Improvisation for the Theater*. Northwestern University Press, Evanston, IL, USA, 3rd edition, 1999.

[31] Genevieve Stebbins. *Delsarte System of Expression*. Edgar S. Werner, New York, NY, 2nd edition, 1887.

[32] Toyota. Trumpet Robot. \url{'http://www.toyota.co.jp/en/special/robot'}, 2004.

[33] G Weinberg and S Driscoll. Robot-Human Interaction with an Anthropomorphic Percussionist. In *Proceedings of International ACM Computer Human Interaction Conference (CHI 2006)*, pages 1229–1232, Montréal, Canada, 2006.

[34] Kazuyoshi Yoshii, Kazuhiro Nakadai, Toyotaka Torii, Yuji Hasegawa, Hiroshi Tsujino, Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno. A biped robot that keeps steps in time with musical beats while listening to music with its own ears. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1743–1750. IEEE, October 2007.