

How to Support Action Prediction: Evidence from Human Coordination Tasks

C. Vesper

Abstract— When two or more people perform actions together such as shaking hands, playing ensemble music or carrying an object together, they often naturally adjust the spatial and temporal parameters of their movements to facilitate smooth task performance. This paper reviews recent findings from experiments with human participants to demonstrate ways in which individuals strategically modulate their own action performance to support a task partner in predicting their actions and thereby facilitate coordination. Based on this evidence, it is discussed how strategic action modulation (“action signaling”) might be a useful approach also for robotic systems to assist human users, thereby reducing cognitive load and flexibly supporting the acquisition of new skills.

I. INTRODUCTION

Human collaboration is not a passive process where two or more people simply happen to be coordinated. Instead, it often involves active mutual adaptation such that each co-actor modifies her action performance to support reaching the desired collaborative action outcome [1, 2]. What are the cognitive processes underlying this highly developed ability?

Various findings from research in cognitive psychology and social neuroscience suggest the existence of specific mechanisms facilitating physical real-time action coordination in humans. For example, by reducing action variability over repeated interactions, people make their own actions as predictable for others as possible [3, 4]. This is especially useful if only little information about another person is available such as when co-actors cannot see each other’s actions. Most of our daily interactions, however, take place in contexts in which rich perceptual information is available. For instance, when moving furniture together, people have visual and auditory access to each other’s actions and also receive haptic feedback about each other’s actions. In these cases in which perceptual information is shared between interaction partners, the available information is often used to achieve smooth and successful action coordination. Such forms of coordination involve two processes:

On the one hand, co-actors must be able to recognize and use the information that is available about another person. Visual cues from a co-actor’s action allow the formation of internal models of the other’s behavior that can be used to generate predictions about an upcoming action [5]. Internal predictive models have first been described in the field of individual human motor control where they explain how fast

real-time movement planning comes about [6]. First, internal inverse models are responsible for calculating the required motor commands necessary to reach a desired movement outcome. These motor commands are then issued to actually execute the movement. Second, in parallel with motor execution, internal forward models use copies of the motor commands to estimate the sensorimotor consequences of performing the action. Thus, this generates a short-term prediction into the near future. This is relevant also for the field of social interaction as the same principles apply to the perception of others’ actions – with the exception that the formation and updating of forward and inverse models is managed through the use of external perceptual information instead of directly through motor commands [5]. Thus, while the motor commands used by the internal forward model are directly available within one individual these need to be inferred indirectly from perceptual evidence from another person in the case of action perception. Using internal predictive models for collaborative action is useful because it allows very fast action planning and performance as co-actors can predict each other’s actions in advance and do not need to rely on slower reactive processes [7, 8].

On the other hand, besides using available information to predict another person’s actions, co-actors can also facilitate each other’s action prediction processes by modifying their action performance in a way that makes it easier for an interaction partner to recognize the action. This is referred to as signaling [9, 10] or strategic action adaptation [2]. Signaling is an active form of supporting collaborative action because people perform their actions differently compared to how they would perform them individually.

The following sections will briefly discuss why action prediction can be challenging and then provide a brief overview about research on signaling in humans. Based on this, it will be discussed how human-robot interaction can benefit from this research. Specifically, using signaling might decrease cognitive load in human users interacting with a machine and thereby facilitate acquisition of relevant interaction skills.

II. CHALLENGES FOR ACTION PREDICTION

Action prediction can be easy and often happens automatically when two human adults collaborate. One reason why this works so well is that humans are physically and functionally very similar. Thus, if one person observes another person perform a reach-and-grasp movement, the observer’s motor system easily “resonates” with the observed movement, thereby providing the internal forward model with accessible data to model the actor’s behavior and generate predictions about the unfolding of the action. Much

C. Vesper is with the Central European University, Budapest, Hungary (phone: +36-1-887-5146; e-mail: vesperc@ceu.hu). The work presented here is supported by a research grant from the Central European University Foundation Budapest (Budapesti Közép-Európai Egyetem Alapítvány).

empirical evidence indeed suggests, first, that the similarity of motor acts is important for forming predictions such that people are better in predicting an action the closer it matches their own performance [11, 12] and, second, that action prediction is improved by own motor experience such that actions can be more accurately predicted if one knows how to and is trained to perform them [13, 14].

In contrast, if a human adult interacts with someone who is physiologically and functionally quite different action prediction becomes more challenging. Already when interacting with a young infant who has limited motor capacities we do not always succeed in estimating what the infant will do, how fast and accurately she can perform an action and what the final outcome of the action will be. The challenge even increases when interacting with artificial agents whose action repertoire and movement kinematics might differ substantially from those of a human person. For instance, an industrial manufacturing robot might not be constrained in its physiology in the same way that human movements are constrained, allowing it to perform complete 360° arm rotations or very fast and highly precise movements. When a human observer sees an action that is impossible for her to perform this creates a processing conflict [15, 16], suggesting that the default prediction mechanisms are impaired or fail.

Given the fast developments in the field of interactive robotics, more and more tasks will be performed by artificial agents and humans together in the near future, for example, with household robots or in professional domains such as industrial or medical applications [17, 18]. As a consequence, the challenge of how to design robots in a way that they display autonomous and flexible behavior while still being predictable for humans becomes increasingly relevant.

III. SUPPORTING ACTION PREDICTION THROUGH SIGNALING

Signaling is a way of helping an observer to predict one's immediate behavior. In contrast to, for instance, using co-speech gestures during a conversation [19] or explicitly communicative actions such as pointing [20], a signaling action is special in that it involves modifying a standard, non-communicative action in a way that makes it informative for another person. Thus, a signaling action involves both a pragmatic goal (for example, moving a cup towards a coffee pot held by someone else) and a communicative goal (facilitating prediction of the endpoint of the movement so that the other can predict the final goal position) [10]. There are two ways in which this can be achieved: First, one can exaggerate one's own action to allow another person to easily detect and track it. Second, one can provide additional information by using one's own action to distinguish a set of possible action alternatives.

A. Exaggerating Action Performance

Action prediction can be facilitated for an interaction partner by overdoing one's own action performance. For example, moving the cup towards the coffee pot could be done with a direct, straight trajectory (which would be the standard way of performing such an action individually) or it

could be performed with higher amplitude, thereby making the movement more salient for another person.

Fig. 1 provides an example for this type of signaling movement. It displays artificially created (upper panel) and a real human participant's (lower panel) movement trajectories based on a data set from a recent study [21] in which pairs of participants performed a sequence of arm movements towards four different target locations. Their goal was to arrive at each target at the same time. However, only one person in each pair knew which of the four possible targets was to be reached for the upcoming movement. Thus, there was an asymmetry between co-actors with respect to the amount of task information – one person knew which target to move to and the other one did not. The aim of the study was to identify ways in which people would solve this coordination task in real time (participants could see each other, but were not allowed to talk to each other as the study intended to investigate non-verbal signaling). A first finding was that the person who had information about the location of the upcoming target exaggerated her movement trajectory towards that target. This allowed the task partner to more easily track, predict and coordinate with her movement because it became salient and easier to recognize (Fig. 1).

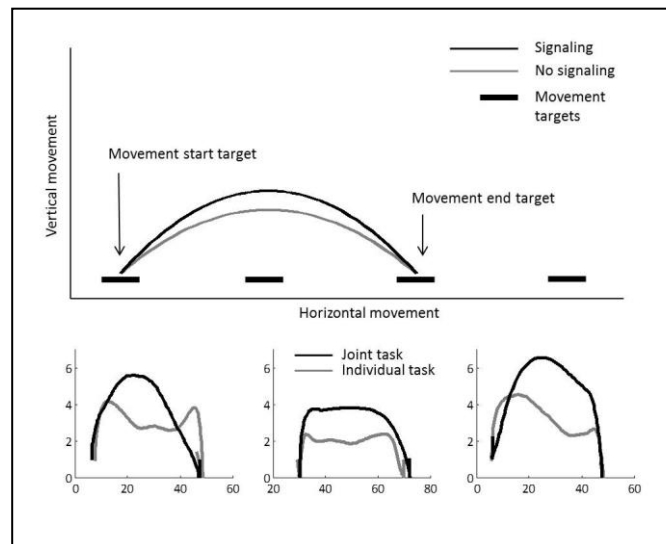


Figure 1. Signaling by exaggeration. The black line depicts a hand movement trajectory, demonstrating one form of signaling that involves exaggerating one's movement path towards a target location (thick black lines) in comparison to a "standard" task performance without signaling (grey line). The task required pairs of participants to make arm movements towards different targets and coordinate the arrival times at the targets while only one of them knew which target was the correct next one. The data displayed in the upper panel are artificially created for demonstration purposes, based on human movement data from [21]. The lower panels show example movements from one real participant of that study.

Similar findings have been reported by other studies. For instance, in one case [22], two expert musicians played a piano duet together under different auditory feedback conditions – either mutually hearing each other, hearing only their own playing or hearing only the other's playing. An analysis of the musicians' movement kinematics indicated that they lifted their fingers higher specifically in the conditions with reduced auditory feedback. This suggests

that they switched from an auditory to a visual information channel to achieve the close temporal coordination that is required for expert music performance. By increasing the amplitude of their finger movements, they made it easier for each other to predict the timing of the next key press and thereby facilitated coordination.

A slightly different way of exaggerating information was found by a study [23] in which two people had to move a pendulum together such that each of them was responsible for moving the pendulum to one side by pulling a rope. Thus, on each swing of the pendulum one person had to pull it over to achieve smooth and regular pendulum motion. When comparing task performance in this interactive condition to bimanual action (one person manipulated both sides of the pendulum), it was found that participants maintained more force overlap of their movements. Thus, while one co-actor pulled the pendulum towards her side, the other maintained some force on the other side of the pendulum. This could be interpreted as a way to communicate through haptic coupling, thereby creating an additional perceptual channel for coordination (or exaggerating the information flow through the existing haptic channel).

B. Disambiguating Action Alternatives

Another way to facilitate action prediction through signaling is by providing additional task information that another person does not have or cannot easily access. People are very accurate in extracting information from the kinematics of observed actions [24, 25, 26]. This ability can be used in an interaction situation if another person modulates her actions to provide information that can be read and understood by others. For instance, when it is unclear whether a person will move a cup towards the person with the coffee pot or towards the dishwasher, changing her movement in a way that differentiates the two movement paths already early on will help the person with the coffee pot to decide whether she will be asked to pour coffee or not.

Similarly, in the case of the earlier described study [21], i.e. in a situation in which only one person knows the target of an interactive coordination task, this person could use her movements to provide information about the location of the upcoming target. Fig. 2 shows three artificially created example trajectories of this type of signaling as well as two examples from real human data. The central observation is that signaling is used here as a form of disambiguating action alternatives by making the possible movement paths as dissimilar as possible [10]. Therefore, movement amplitude “encodes” target location and allows another person to easily detect the correct target location already at an early point in time. (Note that a relation between movement amplitude and target distance also exists in individually performed action; however, in a signaling action this relation is strongly enhanced.)

Another study [27] showed this form of signaling in a task in which two people grasped a bottle-shaped object. One of them knew where to grasp (the bottle could be grasped at the top part with a precision grip or at the lower part with a power grip) and the other had to either perform a grasp to the same or the opposite position on the bottle, as specified by

the task instructions. The main finding revealed that also in this situation the person who had task information which the other one lacked made an effort in disambiguating the action alternatives. In particular, when she grasped the top part of the bottle, she moved relatively higher than at baseline and when she grasped the bottom part of the bottle, she moved relatively lower. This made it easier for the task partner to choose and perform her own complementary action.

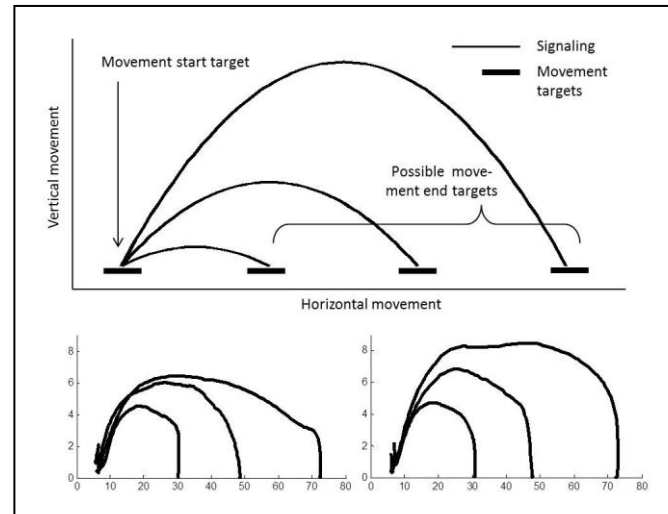


Figure 2. Signaling by disambiguation. The black lines depict different hand movement trajectories towards each of the three target alternatives (thick black lines). This demonstrates a second form of signaling that differentiates movement targets based on increasing amplitude differences. The data displayed in the upper panel are artificially created for demonstration purposes, based on human movement data from [21]. The lower panels show example movements from one real participant of that study.

Pezzulo and colleagues [10] recently suggested a computational model based on a Bayesian approach to describe this form of action disambiguation. The central idea is that signaling can be modelled as an optimization problem with two conflicting goals: On the one hand, an action should be executed efficiently which, in most cases, would mean choosing a direct movement path between a movement start and an end position. On the other hand, facilitating action recognition for another person requires a deviation from the straight movement path, where larger deviations mean easier recognition. Under the assumption that humans really strive for optimal behavior, this optimization problem can then be solved and thereby generates the best movement path that optimally satisfies both constraints. A similar formal approach has been suggested that differentiates ‘legible’ motion from everyday, minimal effort motion as is evident when comparing how a handwritten note to oneself looks very different from a note that is written in a way that another person can read it [28].

C. Effects on Temporal Coordination

In most of the reviewed studies, action signaling significantly improves coordination. For example, in [21], coordination between co-actors was better in a condition in which signaling was possible compared to one in which the informed person’s movements were hidden so that signaling was not an available strategy. Coordination was measured in

terms of absolute asynchronies between co-actors' landing on the targets and also in terms of continuous synchronization (instantaneous relative phase [29]).

How does signaling improve coordination? It is intuitive that signaling supports action prediction in terms of the spatial layout – signaling disambiguates alternative action targets so that more information about another's action goal is available. In addition, however, signaling also supports temporal coordination such that co-actors' actions are more aligned in time when they choose to signal. How does that come about? If signaling allows co-actors to make fast predictions about others' actions, it creates the opportunity for co-actors to quickly and efficiently adjust their behavior to each other [30, 31]. Thus, providing more and earlier information supports action prediction that allows fast adjustments and consequently leads to better coordination.

IV. SIGNALING IN COLLABORATIVE LEARNING?

What implications does research on signaling in human-human interaction have for human-robot interaction? Specifically, can a robot assist a human user to learn a new interactive skill by employing signaling, for instance how to operate a new work piece together? Signaling might help human users interacting with artificial agents by 1) reducing cognitive load, 2) supporting skill acquisition, and 3) increasing safety.

A. Reducing Cognitive Load

As outlined before, human-robot interaction is especially challenging because of differences in physiological and functional constraints. Thus, robots perform movements that humans cannot perform or would perform in very different ways. This makes predicting the robot's actions difficult. One way to solve this challenge is to design human-like robots [32, 33] that perform actions that humans can do in a way that humans would do them, e.g. with biologically plausible action kinematics. However, it is often not possible or not desirable to design robots that have the same movement constraints as humans. Especially in professional domains such as in industrial applications, robots need to do work that a human cannot or should not have to do, for example because the tasks pose health risks or require exceptional strength and precision [17].

In such cases, signaling could be a way in which the robot makes its actions more easily readable by a human user. This would facilitate recognizing and predicting the robot's action and would therefore leave more cognitive capacities to the human user for other, more primary task aspects related to the manufacturing goal. It can be expected that this would increase efficiency of task performance – both for the human user individually and for the interaction with the robot.

B. Supporting Skill Acquisition

When a human user needs to acquire a new skill related to operating and interacting with a machine, this can be done explicitly, e.g. verbally or in written form as direct teaching, or it can be taught implicitly as learning by observation and

by acting together. This latter form of learning is often beneficial in real-time physical interaction because it does not interrupt the work flow and is often faster than explicit verbal instruction [1].

Signaling might be helpful for implicit skill acquisition in human users because it would increase the accuracy and speed of recognizing the to-be-learned action. As an example, if a robot needs the human user to perform a task on a specific part of a large work piece, moving this work piece in an exaggerated way towards the user might make the location where the interaction should take place more salient. Explicit verbal instruction or additional markings like flashing lights become unnecessary. Similarly, also in cases of direct haptic interaction where a human and a robot move an object together, signaling can be a form of teaching the human user about the exact way of performing the task.

C. Increasing Safety

Last but not least, signaling might be useful to increase safety for the human user during human-robot interaction. Especially industrial robots are often very strong and very fast and therefore pose risks to human users. Although considerable progress has been made to make industrial robots safer, e.g. by including sophisticated context-dependent detection and stopping mechanisms [34], it is desirable to prevent accidents rather than to reduce the possible damage. Signaling, as a way of increasing the predictability of the robot's movements, could reduce the occurrence of situations in which the human user incorrectly predicts a robot movement and could therefore make interaction safer.

V. DISCUSSION

The aim of this paper was to provide a brief introduction to recent psychological research on action signaling and to discuss possible ways in which signaling might be useful for human-robot interaction. Signaling is a way in which human interaction partners facilitate real-time coordination by exaggerating their movement paths and by disambiguating different action alternatives. It therefore allows an interaction partner to more easily recognize important aspects of the action and to use this information to make predictions about how the action will unfold in the near future.

Given that robots, especially in professional domains like industrial manipulation, are often very different from humans in terms of their appearance and movement parameters, facilitating predictions is an important research goal. If signaling could be implemented in robots that directly interact with human users, this might be a fruitful way to achieve better prediction performance. Moreover, signaling could increase interaction efficiency, support the learning of new skills and increase safety during physical, real-time interaction.

To better understand the principles underlying signaling for both human-human and human-robot interaction, future research should focus especially on the limitations and constraints of this approach. Some relevant questions involve the practicality and acceptance of robots that use signaling. First,

how different can a robot be from a human user to still make signaling and action prediction feasible? Given the close link to the theory of internal models [6], it would be most useful to employ a combined strategy in which the robot performs biologically plausible movements that allow the built-up of an internal model and, in addition, use signaling to enhance the recognition of the robot action by the human user. Second, what expectations do human users have about a robot? Would they actually expect it to adapt its movements as a form of signaling? On a more theoretical note, this raises the questions whether some form of awareness that an action is meant as a signaling action is required to use its information content.

Finally, whereas this paper only addressed the usefulness of implementing signaling production into a robotic system, it might be similarly fruitful to also include signaling recognition. Thus, if human users perform signaling actions, it might be beneficial if these could be recognized and used by a robot to achieve smooth, efficient and safe real-time collaborative human-robot interaction.

REFERENCES

- [1] G. Knoblich, S. Butterfill, and N. Sebanz, "Psychological research on joint action: theory and data," in *The Psychology of Learning and Motivation*, vol. 54, B. Ross, Ed. Burlington: Academic Press, 2011, pp. 59-101.
- [2] C. Vesper, S. Butterfill, G. Knoblich, and N. Sebanz, "A minimal architecture for joint action," *Neural Networks*, vol. 23, no. 9, pp. 998-1003, 2010.
- [3] C. Vesper, R. P. R. D. van der Wel, G. Knoblich, and N. Sebanz, "Making oneself predictable: reduced temporal variability facilitates joint action coordination," *Experimental Brain Research*, vol. 211, pp. 517-530, 2011.
- [4] C. Vesper, L. Schmitz, N. Sebanz, and G. Knoblich, "Joint action coordination through strategic reduction in Variability," in *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, 2013, pp. 1522-1527.
- [5] D. M. Wolpert, K. Doya, and M. Kawato, "A unifying computational framework for motor control and interaction," *Philosophical Transactions of the Royal Society of London*, vol. 358, pp. 593-602, 2003.
- [6] D. M. Wolpert, and Z. Ghahramani, "Computational principles of movement neuroscience," *Nature Neuroscience*, vol. 3, pp. 1212-1217, 2000.
- [7] G. Knoblich, and J. S. Jordan, "Action coordination in groups and individuals: learning anticipatory control," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 29, no. 5, pp. 1006-1016, 2003.
- [8] C. Vesper, R. P. R. D. van der Wel, G. Knoblich, and N. Sebanz, "Are you ready to jump? Predictive mechanisms in interpersonal coordination," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 39, pp. 48-61, 2013.
- [9] G. Pezzulo, and H. Dindo, "What should I do next? Using shared representations to solve interaction problems," *Experimental Brain Research*, vol. 211, pp. 613-630, 2011.
- [10] G. Pezzulo, F. Donnarumma, and H. Dindo, "Human sensorimotor communication: A Theory of signaling in online social interactions," *PLoS ONE*, vol. 8, no. 11, pp. e79876, 2013.
- [11] P. E. Keller, G. Knoblich, and B. H. Repp, "Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization," *Consciousness and Cognition*, vol. 16, pp. 102-111, 2007.
- [12] G. Knoblich, and R. Flach, "Predicting the effects of actions: interactions of perception and action," *Psychological Science*, vol. 12, no. 6, pp. 467-472, 2001.
- [13] S. M. Aglioti, P. Cesari, M. Romani, and C. Urgesi, "Action anticipation and motor resonance in elite basketball players," *Nature Neuroscience*, vol. 11, no. 9, pp. 1109-1116, 2008.
- [14] E. S. Cross, A. F. de C. Hamilton, and S. T. Grafton, "Building a motor simulation de novo: Observation of dance by dancers," *NeuroImage*, vol. 31, pp. 1257-1267, 2006.
- [15] R. Blake, and M. Shiffrar, "Perception of human motion," *Annual Review of Psychology*, vol. 58, no. 1, pp. 47-73, 2007.
- [16] M. R. Longo, A. Kosobud, and B. I. Bertenthal, "Automatic imitation of biomechanically possible and impossible actions: Effects of priming movements versus goals," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 34, no. 2, pp. 489-501, 2008.
- [17] K. Kosuge, and Y. Hirata, "Human-robot interaction", *Proceedings of the 2004 IEEE International Conference on Robotics and Biomimetics*, pp. 8-11, 2004.
- [18] S. Thrun, "Toward a Framework for Human-Robot Interaction," *Human Computer Interaction*, vol. 19, pp. 9-24, 2004.
- [19] A. B. Hostetter, "When do gestures communicate? A meta-analysis," *Psychological Bulletin*, vol. 137, no. 2, pp. 297-315, 2011.
- [20] D. Peeters, M. Chu, J. Holler, A. Özyürek, and P. Hagoort, "Getting to the point: The influence of communicative intent on the kinematics of pointing gestures," in *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, 2013, pp. 1127-1132.
- [21] C. Vesper, and M. J. Richardson, "Strategic communication and behavioral coupling in asymmetric joint action," *Experimental Brain Research*, in press.
- [22] W. Goebel, and C. Palmer, "Synchronization of timing and motion among performing musicians," *Music Perception*, vol. 26, no. 5, pp. 427-438, 2009.
- [23] R. P. R. D. van der Wel, G. Knoblich, and N. Sebanz, "Let the force be with us: Dyads exploit haptic coupling for coordination," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 37, no. 5, pp. 1420-1431, 2011.
- [24] C. Becchio, L. Sartori, and U. Castiello, "Toward you: The social side of actions," *Current Directions in Psychological Science*, vol. 19, pp. 183-188, 2010.
- [25] M. Graf, B. Reitzner, C. Corves, A. Casile, M. Giese, and W. Prinz, "Predicting point-light actions in real-time," *NeuroImage*, vol. 36, pp. T22-T32, 2007.
- [26] V. Manera, C. Becchio, A. Cavallo, L. Sartori, and U. Castiello, "Cooperation or competition? Discriminating between social intentions by observing prehensile movements," *Experimental Brain Research*, vol. 211, pp. 547-556, 2011.
- [27] L. Sacheli, E. Tidoni, E. Pavone, S. Aglioti, and M. Candidi, "Kinematics fingerprints of leader and follower role-taking during cooperative joint actions," *Experimental Brain Research*, vol. 226, no. 4, pp. 473-486, 2013.
- [28] A. Dragan, K. Lee, and S. Srinivasa, "Legibility and predictability of robot motion", *8th ACM/IEEE International Conference on Human-Robot Interaction*, 2013.
- [29] A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Science*. Cambridge University Press: New York, 2001.
- [30] I. Konvalinka, P. Vuust, A. Roepstorff, and C. D. Frith, "Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping", *Quarterly Journal of Experimental Psychology*, vol. 63, pp. 2220-2230, 2010.
- [31] A. M. Wing, S. Endo, A. Bradbury, and D. Vorberg, "Optimal feedback correction in string quartet synchronization", *Journal of the Royal Society Interface*, vol. 11, pp. 20131125, 2014.
- [32] S. Schaal, "Is imitation learning the route to humanoid robots?," *Trends in Cognitive Sciences*, vol. 3, pp. 233-242, 1999.
- [33] A. Schubö, C. Vesper, M. Wiesbeck, and S. Stork, "Movement coordination in applied human-human and human-robot interaction", *Lecture Notes in Computer Science*, vol. 4799, pp. 143-154, 2007.
- [34] M. Giuliani, C. Lenz, T. Müller, M. Rickert, and A. Knoll, "Design principles for safety in human-robot interaction," *International Journal of Social Robotics*, vol. 2, pp. 253-274, 2010.