

Conversation and Coordinative Structures

Kevin Shockley,^a Daniel C. Richardson,^b Rick Dale^c

^a*Center for Cognition, Action, and Perception, Department of Psychology, University of Cincinnati*

^b*Cognitive, Perceptual and Brain Sciences, University College London*

^c*Department of Psychology, University of Memphis*

Received 6 July 2008; received in revised form 14 January 2009; accepted 4 February 2009

Abstract

People coordinate body postures and gaze patterns during conversation. We review literature showing that (1) action embodies cognition, (2) postural coordination emerges spontaneously when two people converse, (3) gaze patterns influence postural coordination, (4) gaze coordination is a function of common ground knowledge and visual information that conversants believe they share, and (5) gaze coordination is causally related to mutual understanding. We then consider how coordination, generally, can be understood as temporarily coupled neuromuscular components that function as a collective unit known as a coordinative structure in the motor control literature. We speculate that the coordination of gaze and body sway found in conversation may be understood as a cross-person coordinative structure that embodies the goals of the joint action system.

Keywords: Coordination; Dynamical systems; Language

1. Introduction

Cognitive phenomena do not always respect the boundaries that are observed by textbooks and conferences in cognitive science. Consider the mirror neuron system—neurons in the premotor cortex that fire both when observing and performing an action. This discovery has challenged the assumed distinction between perception and action in the brain. While the human capacity to perceive opportunities for action (i.e., affordances; Gibson, 1979) has been extensively documented (see Fajen, Riley, & Turvey, in press, for a recent review), it appears that even non-action-related judgments are in the currency of action. For example,

Correspondence should be sent to Kevin Shockley, Department of Psychology, University of Cincinnati, ML 0376, 400 Dyer Hall, Cincinnati, OH 45221-0376. E-mail: kevin.shockley@uc.edu

perceived distances appear to be scaled by the metabolic cost associated with traversing a distance rather than a strict function of visual information (e.g., Proffitt, 2006).

Likewise, action appears to embody cognitive processing. Cognitive performance can interfere with concurrent motor performance, and vice versa (see Barsalou, Simmons, Barbey, & Wilson, 2003, for a recent review). Moreover, the inhibition of motor activity during cognitive processing (Frick-Horbury & Guttentag, 1998) or imposed inconsistency between motor activity and cognitive responses (Glenberg & Kaschak, 2001; de'Sperati & Stucchi, 2000) appears to disrupt cognitive performance. Lastly, the time course of even high-level cognitive processes is reflected in the trajectory of action. When clicking “yes” or “no” to questions ranging in fuzzy truth-value, (e.g., Should you brush your teeth every-day? Is murder sometimes justified?), the degree of motor deviation and complexity, measured in computer–mouse trajectories, reflects the degree of cognitive certainty (McKinstry, Dale, & Spivey, 2008). The implication of all these findings is that a full understanding of perception, action, and cognition will not obtain by studying these domains in isolation.

Cognitive phenomena do not always respect the boundaries that we draw between people either. In theory and in experimental practice, cognitive science has been concerned with the individual in isolation. But there have been striking insights into the brain’s sensitivity to social information (for a review, see Cacioppo, Visser, & Pickett, 2006) and the way in which people think and act cooperatively (Hutchins, 1995). For example, Clark (1996) re-conceptualized language as a process spanning individual and social scales, rather than as exclusively an encapsulated cognitive process within an individual. These observations have led to the recent interest in situated cognition (Robbins & Aydede, 2008) and joint action (Sebanz, Bekkering, & Knoblich, 2006; and elsewhere in this issue). An essential feature of joint action is how each person’s nuanced actions shape the other’s as the joint action unfolds. For example, when two people talk, they spontaneously converge in terms of pausing and speaking duration, speaking rate, turn duration, response latency, vocal intensity, and accent. Movements and postures, likewise, appear to be spontaneously coordinated with speech within and across individuals (see Fowler, Richardson, Marsh, & Shockley, 2008, for a review).

In this paper, we review a variety of findings suggesting that in these phenomena not only is action coordinated across individuals but also is cognition. We consider what this observed coordination may be telling us about cognition and the nature of the coordination itself. One explanation for behavioral coordination in the context of conversation is that it exists to align cognitive representations (e.g., Garrod & Pickering, 2004; Pickering & Garrod, 2004; Sebanz et al., 2006). Here, we explore an alternative possibility: interpersonal coordination is a case of an emergent coordinative structure. This explanation draws on models in physics and biology, and need not involve high-level, cognitive representations. For example, when a thin layer of fluid is heated, small-scale high-dimensional interactions result in a large-scale, lower-dimensional order. The molecules form into rolling columns that efficiently distribute heat across the area of the fluid (e.g., Kelso, 1995). There is no temptation here to attribute such coordination to a “central physics executive” that is directing the fluid molecules. The coordinated behavior of molecules emerges naturally from local interactions and constraints. The possibility we raise at the close of this paper is

whether a similar process can explain the coordinated behavior between people. We first consider two sources of empirical evidence for such coordination during interaction: postural coordination and coordination of gaze.

2. Interpersonal postural coordination

Coordination of postural configurations and body movements in the context of language has been observed for decades. Historically, however, the coordination of movements across people has been evaluated using largely subjective methods, such as hand-scoring video tapes of interactions (e.g., Bernieri, 1988). More recently, the spontaneous coordination of conversants' movements has been objectively quantified using cross-recurrence quantification (CRQ). CRQ is a nonlinear method that quantifies how similarly two observed data series unfold over time (Shockley, Butwill, Zbilut, & Webber, 2002; Zbilut, Giuliani, & Webber, 1998). Like cross-correlation or more recently windowed cross-correlation (Boker, Xu, Rotondo, & King, 2002), CRQ produces measures reflecting coordination, such as how often two data series have similar values or reflect similar system states (called percentage recurrence, or %REC), among other measures. Work in the physical sciences has shown that a benefit of CRQ is its resistance to noise and freedom from assumptions. These features make it well suited to studying noisy behavioral signals generated during naturalistic interaction (see Marwan, Romano, Thiel, & Kurths, 2007, for review of CRQ and related methods). We detail three studies demonstrating postural coordination that use this method.

In a first study, Shockley, Santana, and Fowler (2003) evaluated the shared postural configurations of pairs of people engaged in a conversation by measuring their postural sway trajectories—the subtle, continuous motion of the body during upright standing. Postural sway was taken as a global measure of movement because it reflects the coordination of all body segments required to maintain the center of mass of the body within the base of support (i.e., the extent of the feet on the ground surface) that is required to stand upright. To generate conversation among participants, Shockley et al. (2003) used a puzzle task in which the two participants of a given pair could not see their partner's picture and had to discuss the pictures to discover the differences (Fig. 1, left). Participants were allowed to gesture, and speak freely, which generated a very natural conversation between members of a pair. To control for any influence of the task itself (i.e., looking and speaking while standing) on the coordination, participants were partnered with either the other participant in the pair or each member of the pair was partnered with an experimental confederate (C1 or C2 in Fig. 1, left). To control for any influence of vision (e.g., seeing another participant sway and/or gesture while doing the task), participants were either oriented toward the other member of the participant pair or they faced away from the other member of the pair (confederates were always out of view) (Fig. 1, left). For measurements taken at the waists of participants, as illustrated in Fig. 1 (right), Shockley et al. (2003) found that participant pairs that performed the task together shared more postural configurations (%REC) and maintained similar postural trajectories for longer (indexed by the CRQ measure termed MAXLINE—the length of the longest common trajectory across the two members of a pair)

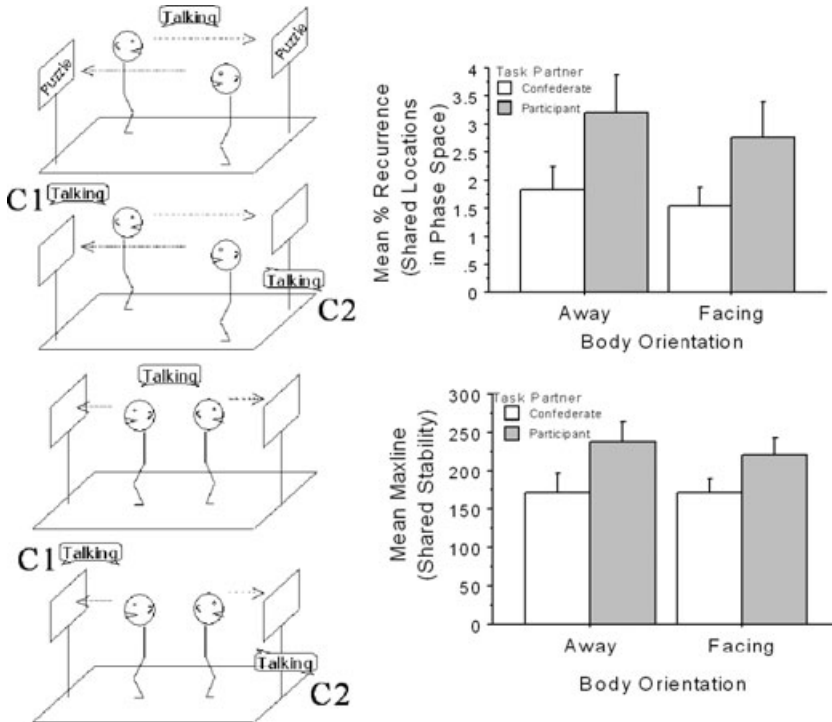


Fig. 1. (Left) Method used by Shockley et al. (2003). (Right) Shared postural activity for the different experimental conditions. (Note: From Shockley et al. (2003, p. 329 [panel a], 330 [panel c]). Copyright 2003 by the American Psychological Association. Adapted with permission.)

than participant pairs that performed the task with an experimental confederate, irrespective of whether they were facing the other participant.

In a subsequent study, Shockley, Baker, Richardson, and Fowler (2007) investigated the nature of the coordination observed by Shockley et al. (2003). They considered the possibility that participants may have converged in their speaking patterns over time and that coordinated speech patterns may have mediated coordinated postural sway patterns. Each member of a pair was presented with a target word that was either the same as his/her partner's (S), a different word for each member of the pair but with the same syllable emphasized (DS) (e.g., *ethnic* and *ancient* have an emphasis on the first syllable), or a different word for each member of the pair with a different syllable emphasized (DD) (e.g., *ethnic* has an emphasis on the first syllable while *deserve* has an emphasis on the second syllable) (Fig. 2, top). As illustrated in Fig. 2 (bottom), they found greater shared postural configurations when words were more similar than when words were less similar, suggesting that the shared postural activity observed by Shockley et al. (2003) may have indeed been mediated by convergent speaking patterns. They also compared postural sway trajectories of participants who uttered the same sequence of words but compared them with a participant who did the same task with a different partner (i.e., virtual pairs; see Fig. 2, bottom) where they found no influence of word similarity. The implication is that the

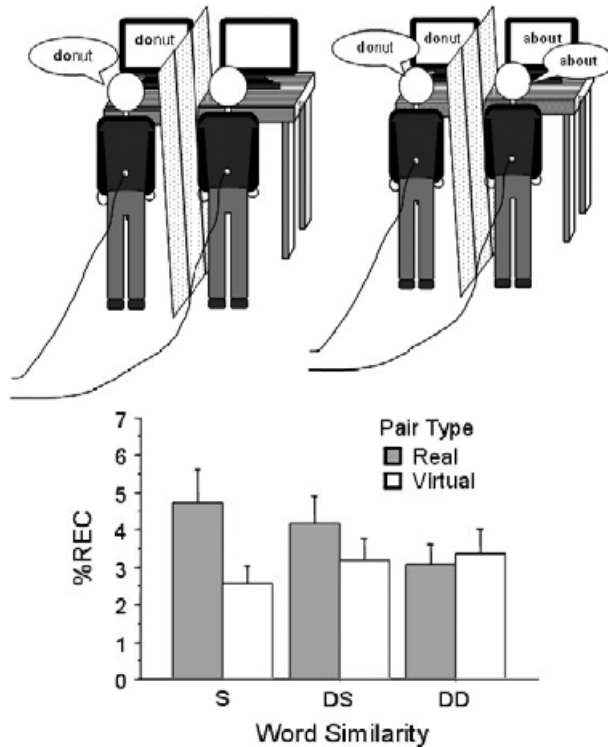


Fig. 2. (Top) Method used by Shockley et al. (2007). (Bottom) Shared postural activity for participant pairs and for “virtual pairs” who uttered the same words (S), different words with the same stress (DS) or different words with a different stress (DD). (From Shockley et al., 2007, p. 203, 204 [panels A and B], and 205. Copyright 2007 by the American Psychological Association. Adapted with permission.)

participants had to be performing the task together, suggesting that the postural coordination observed by Shockley et al. (2003) entailed a social constraint in addition to the constraints of articulation.

Although Shockley et al. (2003) found that it made no difference if participants were facing their task partner, more recently, Giveans, Pelzer, Smith, Shockley, and Stoffregen (2008) did find evidence of visual constraints on shared postural activity. Giveans et al. manipulated the relative distances of the target puzzle pictures for a participant pair and the relative sizes of the target puzzle pictures (both large both small, or one was small while the other was large). These manipulations served to constrain the looking patterns of the conversing participants by having both participants scan a larger or smaller area or have one scan a larger area while the other scanned a smaller area. They found greater shared head configurations when both targets were near participants or large than when both targets were farther away, small, or mismatched in size. The implication of the findings of Giveans and colleagues is that visual constraints influence the observed postural coordination when two persons are conversing. The influence of looking patterns on interpersonal postural

coordination is of particular significance because recent research has demonstrated that gaze patterns embody the joint understanding involved in conversation. The following section showcases these studies, which focus on gaze coordination during interaction.

3. Interpersonal gaze coordination

Consider an argument over a map, a debate over a proof written out on a blackboard, a civilized conversation about a painting at a gallery or an idle chat during a commercial break. In all these cases, a visual scene is shared while people talk to each other. Here, we describe research that quantifies such joint attention by measuring the eye movements of two people discussing a common visual scene. We have found that the coordination between their gaze coordination reveals their shared knowledge, their beliefs about each other, and the success of their communication.

In the first quantification of gaze coordination (Richardson & Dale, 2005), the speech and eye movements of one set of participants were recorded as they looked at pictures of six cast members of a TV sitcom and spoke spontaneously about their favorite episode and characters. One-minute segments were played back unedited to a separate set of participants. The listeners looked at the same visual display of the cast members, and their eye movements were also recorded as they listened to the segments of speech. They then answered a series of comprehension questions. CRQ was used to quantify the degree to which speaker and listener eye positions overlapped at successive time lags (Fig. 3). From the moment a speaker

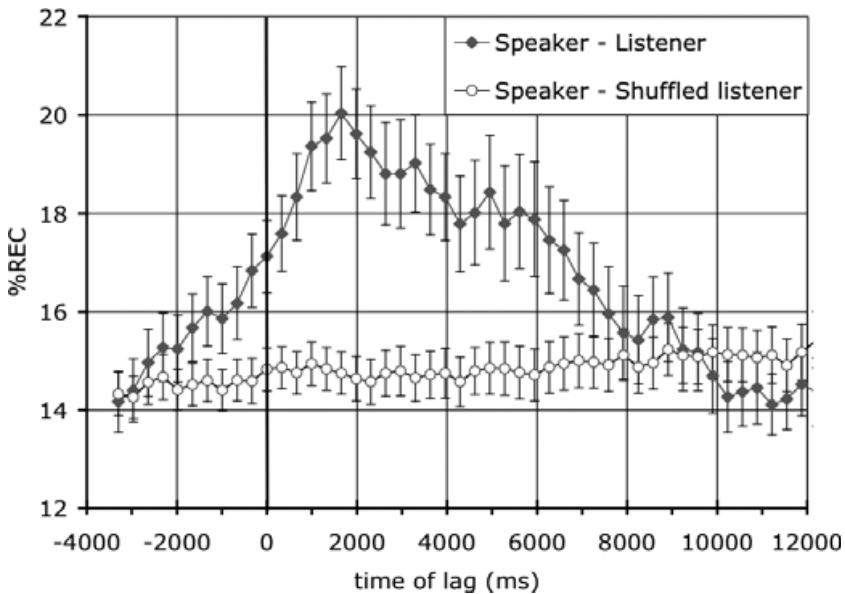


Fig. 3. Average recurrence at different time lags. From Richardson and Dale (2005). Copyright 2005 by the Cognitive Science Society, Inc. Adapted with permission.

looked at a picture, and for the following 6 s, a listener was more likely than chance to be looking at that same picture. The listener was most likely to be looking at the same cast member 2 s after the speaker fixated it. The degree of recurrence between individual speaker–listener pairs correlated with the listeners’ accuracy on the comprehension questions. A second experiment showed that gaze coordination and comprehension were causally connected. Pictures flashing in time with the speakers’ fixations (compared with a randomized version) caused the listeners’ eye movements to look more like the speakers.’ This experimental manipulation improved the speed of listeners’ performance on comprehension questions.

The language use in this study, however, lacked a key element of everyday conversations—interaction. In another study, the gaze of two conversants engaged in a live dialog was tracked simultaneously while they discussed TV shows and surrealist paintings (Richardson, Dale, & Kirkham, 2007a). Conversants’ eye movements were coupled as they looked at a shared display, peaking at a lag of 0 ms. In other words, the conversants were most likely to be looking at the same thing at the same point in time. As in the monolog results, this coupling was above-randomized baseline levels for a period of around 6 s.

Such coordination in conversation can be achieved, Clark (1996) argues, because conversants share not only a stream of words but also a common body of background knowledge that allows them to ground and understand each others’ utterances. This level of “common ground” knowledge was experimentally manipulated, and its effect on gaze coordination assessed (Richardson, Dale, & Kirkham, 2007a). Participants first listened to a 90-s passage describing either the meaning of a specific painting by Dali or facts from his biography. They then saw Dali’s painting and discussed it while their gaze was tracked. Conversational partners who heard the same information had higher gaze coordination (in the region around 0 ms of lag) than those who heard different information. Interestingly, it did not seem to matter which of the passages the conversants heard. What was important was that they had the same information and knew this to be the case.

This raises an interesting question: How is gaze coordination influenced by what conversants believe about each other’s beliefs? We studied the gaze coordination of conversants while varying what they saw and what they believed each other could see (Richardson, Dale, Tomlinson, & Clark, 2008). In previous “reference game” studies of visual common ground (e.g., Hanna, Tanenhaus, & Trueswell, 2003; Keysar, Barr, Balin, & Brauner, 2000), these factors are not separated, and the linguistic interaction is a monolog directed at objects on display. In our study, participants had a spontaneous conversation about a controversial topic. Before their discussion, pairs of participants watched videos of actors giving opposing views on a topic such as the Iraq War. During their discussion, two factors were manipulated (Fig. 4A). While they talked, either both looked at a blank screen or both looked at pictures of the actors. In addition, either both believed that their partner was looking at a blank screen, or both believed that their partner was looking at the actors. These factors were varied at random across eight conversations. On average, both the presence of the visual scene and beliefs about its presence for another influenced participants’ gaze coordination. Recurrence increased overall when the

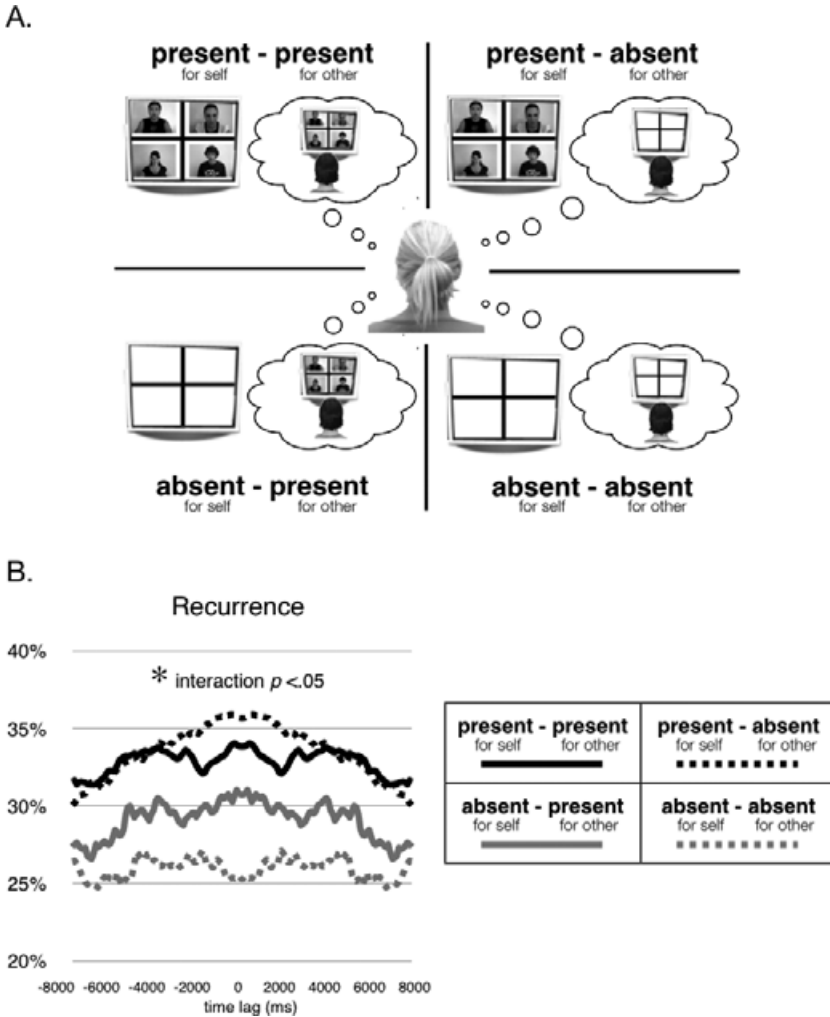


Fig. 4. (A) Experimental conditions and (B) Effects on gaze recurrence in Richardson et al. (2008).

pictures of actors were onscreen, as, unsurprisingly, conversants were less likely to look at a blank screen. When they were looking at pictures of the actors, conversants' eye movements were more tightly coordinated if they believed they were talking to someone who was looking at nothing. Conversely, while looking at an empty screen, their eye movements were more closely coordinated if they believed that each other could still see the actors (Fig. 4B). In other words, when they believed there to be an imbalance between their own view and their partners' view, gaze coordination increased. Analysis of conversants' speech showed that when they believed they were looking at different things, they used a variety of techniques to boost common ground. They tended to refer to actors' viewpoints via a nonvisual route, raise the intonation at the end of their phrases

(as an implicit request for confirmation that their messages had been understood), and signal understanding to each other via back channel responses (such as saying “ok” or “uh-huh”).

Gaze coordination is causally connected to what people know and remember, and what conversants see and what they believe each other can see. In other words, eye movements embody the cognitive processing involved in cooperative conversation and the shared knowledge among interlocutors. Can these phenomena of gaze coordination and shared knowledge be understood in the same terms as postural coordination?

4. Accounts of interpersonal coordination

4.1. Shared representations

Perhaps behavioral coordination—from body posture to gaze direction—is yoked to representational coordination. For example, the activation of the motor system while watching others perform actions suggests a common neural mechanism for motor control and action understanding (cf. Liberman, 1974). Sebanz et al. (2006) suggest that the hypothesized shared representations would serve to facilitate the understanding of the actions of another (e.g., Blakemore & Decety, 2001; Richardson & Dale, 2005), predict the actions of others (e.g., Ramenzoni, Riley, Shockley, & Davis, 2008; Bosbach, Cole, Prinz, & Knoblich, 2005), and to know what joint actions are possible (e.g., Richardson et al., 2007b). Another possibility is that communication can be understood as the alignment of conceptual representations (Garrod & Pickering, 2004; Pickering and Garrod, in press). This alignment can occur in words, syntax, gaze, or posture—but alignment at all levels fuels the alignment of conceptual representations, and furthers the goal of communication. In this characterization, cognitive representations are the motivation and mechanism of coordination. We offer a different interpretation building from the principles of dynamical systems theory. Here, cognition is not assigned a primary status and the phenomena in question are characterized in terms of broader physical principles.

4.2. Coordinative structures

Interpersonal coordination can be conceptualized without an appeal to the cognitive representation of one’s own or another’s actions. It can be thought of as a coordinative structure—a self-organized, softly assembled (i.e., temporary) set of components that behave as a single functional unit (Bernstein, 1967). The power of a coordinative structure is that it is an abstract structure: different components of the body can play different roles under different constraints while preserving the overall collective pattern dictated by the task constraints (e.g., Kelso, Tuller, Vatioksis-Bateson, & Fowler, 1985) (cf. Givens et al., 2008 with Shockley et al., 2003). Coordinative structure is consistent with our understanding of the principles of nonlinear, self-organizing systems that are well known in physics (e.g., Babloyantz, 1986), in that coordinative structures may emerge naturally from the

interaction among the components under certain constraints (e.g., Kelso, 1995; Turvey, 1990).

Proponents of dynamical systems accounts of human performance have offered similar explanations of human motor coordination. For example, self-organized, coordinative structures have been hypothesized to explain interlimb rhythmic coordination (e.g., Fuchs, Jirsa, Haken, & Kelso, 1996). Interlimb rhythmic coordination conforms to a model (generally referred to as the Haken, Kelso, & Bunz [HKB, 1985] model) of coupled oscillators (i.e., oscillating limbs) whose behavior is readily understood in terms of a relational, low-dimensional order parameter (relative phase) that captures the stable relation between the two limbs' oscillation cycles. The model derives from models of physical oscillators and makes no assumptions about executive control systems. Of particular importance is that the dynamics of the coordinative structure emerge from the constraints imposed upon the coordinated units. For example, interlimb coordination dynamics are influenced by physical constraints such as the relative sizes and masses of the limbs and functional constraints such as handedness (e.g., Treffner & Turvey, 1995), perceptual anchors (e.g., Fink, Foo, Jirsa, & Kelso, 2000), and cognition (e.g., Pellicchia, Shockley, & Turvey, 2005).

The power of the HKB model of interlimb coordination becomes evident when considering the fact that 10 nonintuitive predictions of the model have been confirmed experimentally (Schmidt & Turvey, 1995). In fact, this same model can capture interlimb coordination *across* individuals. For example, Schmidt, Carello, and Turvey (1990) investigated interpersonal coordination of rhythmic limb oscillations. They found that when participants viewed the limb oscillations of another person, the coordination of the two limbs exhibited the same stable coordination modes as those observed within an individual (see Marsh, Richardson, & Schmidt, in press). Schmidt et al. further demonstrated that transitions between stable coordination modes across individuals exhibited the same hallmark dynamics found in individual interlimb coordination (e.g., critical fluctuations, critical slowing down, all predicted by HKB). Interpersonal coordination, therefore, need not entail representations of the other person in any classical sense. Rather, it may reflect the spontaneous organization (i.e., coordination) that emerges naturally from the interaction of two (or more) people under certain constraints (e.g., intentions, task constraints, physical constraints).

A comparison of the study of Giveans et al. (2008) with that of Shockley et al. (2003) may illustrate the functional nature of the cross-person organization. Giveans et al. (2008) found that the shared activity of head movements was influenced by visual manipulations, while Shockley et al. (2003) found no influences on head movements but the shared activity at the waist was absent. In other words, different body segments were coordinated in the studies of Shockley et al. (2003) and Giveans et al. (2008) for the same task under different constraints. This may suggest that the body segments involved in postural coordination during conversation changes flexibly with the specific task constraints. The coordination among body segments to both maintain upright stance and to achieve the joint task appears to be a task-dependent organization of the segments supporting stable performance of the task itself (e.g., Oullier, Bardy, Stoffregen, & Bootsma, 2004). This type of functional organization is seen in within-person coordinative structures. For example, vocal tract configuration for a given utterance changes with changes in task constraints. Kelso et al. (1985) demonstrated

that when the jaw is perturbed during the articulation of a particular utterance, an immediate (20–30 ms) compensation by the upper and lower lips occurs to preserve the integrity of the intended utterance. Importantly, this modulation occurs far more quickly than could be controlled by a central executive, suggesting that the modulation is occurring outside of executive control. If joint cognitive tasks are similarly achieved via a functionally defined cross-person organization, then perturbing/constraining the actions of one component of the cross-person coordinative structure of one member (e.g., a relevant body segment, optical information, cognitive constraints) should result in rapid compensatory changes in other components of the cross-person structure (e.g., changes in movement patterns of a body segment, looking patterns, or cognitive kinematics in the other member of the pair). In other words, if the cross-person coordinative structure consists of a certain relation among body segments and cognitive states, then constraints on the (action) effectors of one person should affect the movements and/or cognition of the other member of the pair as readily as cognition can affect one's own effectors.

Interpersonal postural coordination as a functionally defined cross-person coordinative structure was explicitly suggested in a recent study by Ramenzoni, Baker, Riley, and Shockley (2007). Ramenzoni et al. had participant pairs who perform a joint precision pointing task while standing. One participant was required to maintain the position of a pointer (held in the hand) within the boundaries of a target of a particular size (held by the other participant). The motion of the torso, forearm, arm, and hand were recorded. The time series of this collection of body segments of pairs of participants were submitted to principal components analysis to evaluate how many principal components were required to account for the coordination in the various conditions. Fewer components were required to account for 90% of the variance in the behavior of an individual when the task was performed with another participant than when the precision task was performed without a partner. Ramenzoni et al. argued that this intriguing finding of a reduction in the number of modes required to account for the collective behavior of the body segments may reflect an overall reorganization of components (within and across individuals) into a two-person, lower-dimensional (i.e., fewer modes) system that embodies the contributions of both participants.

Given the recent insights into the embodiment of cognitive processes in eye movements (Richardson & Dale, 2005; Richardson et al., 2007a) and the constraints of eye movements on interpersonal postural coordination (Giveans et al., 2008), a speculative hypothesis is that conversational coordination reflects a functional reorganization of body segments and eye movements to support the joint goals/actions of interacting individuals. However, to understand the organization of this joint-action system, the assumption that each participant can be studied in isolation must be relaxed. Moreover, characterizing interpersonal coordination as a cross-person coordinative structure may require a different conceptualization of cognition. Cognition, from a dynamical systems perspective, may be more usefully understood as a set of constraints on action. For example, more knowledge in a particular situation will constrain the looking patterns of an individual, just as more perceptual anchors constrain oscillating limb patterns (Fink et al., 2000). Likewise, more shared knowledge across two individuals will similarly constrain their respective looking patterns to be more similar to one another (Richardson et al., 2007a). Moreover, if one's actions embody cognitive

processes, then two people's actions would be expected to be more similar to one another with similar cognitive constraints when they are interacting in a common environment.

In many ways, this conceptualization of cognition is consistent with the interactive alignment model (Garrod & Pickering, 2004), as both assert that forms of coordination promote communication. We differ mostly in the status or emphasis that is accorded to cognitive representations. Garrod and Pickering (2004) argue that communication, in a broad sense, is the alignment of conceptual representations. When conversing, people will (automatically or intentionally) align their behavior at many levels, from word choice to syntactic structure to body sway. Alignment at each level serves to bring conceptual representations closer together. They offer a plausible account of how "higher" lexical or syntactic representations become aligned through priming and promote linguistic processing. However, in their framework, it is harder to see how alignment of behavior at "lower" levels, such as body sway, can bring about alignment at the loftier level of conceptual representation.

The coordinative structure account, in contrast, does not accord a special status to cognitive representations. They are constraints upon coordination just as any other and can be analyzed in the same language used for physical systems. For example, the type of spontaneous alignment described by Garrod and Pickering is not special to cognitive linguistic systems, and can be observed in many places in physics and biology where there is a tendency for units to coordinate (i.e., pulling coupled units into coordinated modes). This physical principle is known as the *magnet effect* (von Holst, 1973), and the HKB model of interlimb coordination is based on our understanding of this tendency in physically coupled oscillators (see Turvey, 1990, for a detailed discussion). As these principles already exist in nature, and behavioral coordination has already been shown to conform to these principles in many instances, we argue that they lead to a more parsimonious account than proposing a new principle of alignment that is special to cognition.

5. Conclusion

How do we develop a generic and principled understanding of how minds and bodies move together (Sebanz et al., 2006)? We argue, first, that only by abandoning the historic boundaries among the subdomains of psychology will we achieve a principled understanding interpersonal coordination. Secondly, we have provided a speculative account of how behavioral coordination across individuals can be conceptualized as a functionally defined cross-person coordinative structure that embodies the goals of the joint action system. A challenge for this general approach to understanding joint action is to understand how it might scale up to "representationally hungry" (Clark, 1997) domains at higher levels of linguistic description (Garrod & Pickering, 2004). This, of course, is a challenge shared by all embodied accounts of cognition (Wilson, 2002). We argue that meeting this challenge in the case of interpersonal coordination—understanding its cognitive dynamics in terms of coordinative structures—brings two rewards. First, it offers a principled account of how two cognitive systems can become aligned over time. For example, it may offer an understanding of the gradual entrainment of action and cognition using general principles that explain

the entrainment of two physical oscillators or two limb movements. Second, it may allow cognitive coordination to be discussed in the same language as movement coordination. Such a “universal language” of joint action may best capture these itinerant phenomena that wander across boundaries between perception and action, self, and other.

References

- Babloyantz, A. (1986). *Molecules, dynamics, and life: An introduction to self-organization of matter*. New York: Wiley.
- Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Science*, 7(2), 84–91.
- Bernieri, F. J. (1988). Interpersonal sensitivity in teacher-student interactions. *Personality and Social Psychology Bulletin*, 17(1), 98–103.
- Bernstein, N. A. (1967). *Coordination and regulation of movements*. New York: Pergamon Press.
- Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, 2(8), 561–567.
- Boker, S. M., Xu, M., Rotondo, J. L., & King, K. (2002). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods*, 7(3), 338–355.
- Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Understanding another’s expectation from action: The role of peripheral sensation. *Nature Neuroscience*, 8(10), 1295–1297.
- Cacioppo, J. T., Visser, P. S., & Pickett, C. L., & others (2006). *Social neuroscience: People thinking about thinking people*. Cambridge, MA: MIT Press.
- Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.
- Clark, A. (1997). *Being there: Putting brain, body, and the world together again*. Cambridge, England: MIT press.
- Fajen, B. R., Riley, M. A., & Turvey, M. T. (2008). Information, affordances, and the control of action in sport. *International Journal of Sports Psychology*, 40, 79–107.
- Fink, P. W., Foo, P. L., Jirsa, V. K., & Kelso, J. A. S. (2000). Local and global stabilization of coordination by sensory information. *Experimental Brain Research*, 134(1), 9–20.
- Fowler, C. A., Richardson, M. J., Marsh, K., & Shockley, K. D. (2008). Language use, coordination, and the emergence of cooperative action. In A. Fuchs & V. Jirsa (Eds.), *Coordination: Neural, behavioral and social dynamics* (pp. 161–180). Berlin: Springer-Verlag.
- Frick-Horbury, D., & Guttentag, R. E. (1998). The effects of restricting hand gesture production on lexical retrieval and free recall. *American Journal of Psychology*, 111(1), 43–62.
- Fuchs, A., Jirsa, V. K., Haken, H., & Kelso, J. A. S. (1996). Extending the HKB model of coordinated movement to oscillators with different eigenfrequencies. *Biological Cybernetics*, 74(1), 21–30.
- Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, 8(1), 8–11.
- Garrod, S., & Pickering, M. L. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, 1 (2), 292–304.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Giveans, R., Pelzer, C., Smith, A., Shockley, K., & Stoffregen, T. A. (2008, June). Postural support for personal performance and interpersonal coordination. Talk presented at an Annual Meeting of the North American Society for the Psychology of Sport and Physical Activity, Niagara Falls, Canada.
- Glenberg, A. M., & Kaschak, M. P. (2001). Grounding language in action. *Psychonomic Bulletin & Review*, 9 (3), 558–565.
- Haken, H., Kelso, J. A. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics*, 51(5), 347–356.

- Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory & Language*, 49(1), 43–61.
- von Holst, E. (1973). *The behavioral physiology of animal and man*. Coral Gables, FL: University of Miami Press.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT press.
- Kelso, J. A. S. (1995). *Dynamic patterns: The self-organization of brain and behavior*. Cambridge, MA: MIT Press.
- Kelso, J. A. S., Tuller, B., Vatioksis-Bateson, E., & Fowler, C. A. (1985). Functionally specific articulatory cooperation following jaw perturbations during speech: Evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, 10(6), 812–832.
- Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(1), 32–38.
- Lieberman, A. M. (1974). The speech code. In G. A. Miller (Ed.), *Communication, language, and meaning* (pp. 128–140). New York: Basic Books.
- Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, 1(2), 320–339.
- Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438(5-6), 237–329.
- McKinstry, C., Dale, R., & Spivey, M. J. (2008). Action dynamics reveal parallel competition in decision making. *Psychological Science*, 19(1), 22–24.
- Oullier, O., Bardy, B. G., Stoffregen, T. A., & Bootsma, R. J. (2004). Task-specific stabilization of postural coordination during stance on a beam. *Motor Control*, 7, 174–187.
- Pellecchia, G., Shockley, K., & Turvey, M. T. (2005). Concurrent cognitive task modulates coordination dynamics. *Cognitive Science*, 29(4), 531–557.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2), 169–190.
- Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science*, 1(2), 110–122.
- Ramenzoni, V. C., Riley, M. A., Shockley, K., & Davis, T. (2008). Carrying the height of the world on your ankles: Encumbering observers reduces estimates of how high another actor can jump. *Quarterly Journal of Experimental Psychology*, 61(10), 1487–1495.
- Ramenzoni, V. C., Baker, A. A., Riley, M. A., & Shockley, K. (2007, July). *Effects of joint task performance on interpersonal postural coordination*. Paper presented at the 14th International Conference on Perception and Action, Yokohama, Japan.
- Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29(6), 1045–1060.
- Richardson, D. C., Dale, R., & Kirgham, N. Z. (2007a). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science*, 18(5), 407–413.
- Richardson, D. C., Dale, R., Tomlinson, J., & Clark, H. (2008). What eye believe that you can see: Conversation, gaze coordination and visual common ground. Presented at the 2008 Workshop on the Semantics and Pragmatics of Dialogue (LONDIAL2008), London.
- Richardson, M. J., Marsh, K. L., & Baron, R. M. (2007b). Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance*, 33(4), 845–859.
- Robbins, P., & Aydede, M. (2008). *The Cambridge handbook of situated cognition*. Cambridge, England: Cambridge University Press.
- Schmidt, R. C., Carello, C., & Turvey, M. T. (1990). Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance*, 16(2), 227–247.

- Schmidt, R. C., & Turvey, M. T. (1995). Models of interlimb coordination: Equilibria, local analyses, and spectral patterning. *Journal of Experimental Psychology: Human Perception and Performance*, 21(2), 432–443.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *TRENDS in Cognitive Sciences*, 10(2), 70–76.
- Shockley, K., Baker, A. A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 33(1), 201–208.
- Shockley, K., Butwill, M., Zbilut, J., & Webber, C. (2002). Cross recurrence quantification of coupled oscillators. *Physics Letters A*, 305(1–2), 59–69.
- Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 326–332.
- de'Sperati, C., & Stucchi, N. (2000). Motor imagery and visual event recognition. *Experimental Brain Research*, 133(2), 273–278.
- Treffner, P. J., & Turvey, M. T. (1995). Handedness and the asymmetric dynamics of bimanual rhythmic coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 21(2), 318–333.
- Turvey, M. T. (1990). Coordination. *American Psychologist*, 45(8), 938–953.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636.
- Zbilut, J. P., Giuliani, A., & Webber, C. L. Jr (1998). Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification. *Physics Letters A*, 246(1–2), 122–128.