CHAPTER 4

Weaving oneself into others

Coordination in conversational systems

Rick Dale and Michael J. Spivey
University of California, Los Angeles / University of California, Merced

We review a range of findings that show how eye movements (and other body movements) exhibit correlated behavior across two or more people during natural interactions. We then synthesize these different results into a more general account of how people's cognitive, sensory and motor systems become coordinated with one another during natural dialogue. We argue that treating conversants as parts of one integrated *system* is a useful explanatory strategy for understanding interaction. We end by describing explicit quantitative conditions for seeking "systemhood" in human interaction. These conditions motivate future research questions on social eye movements and other behaviors.

1. Introduction: The eyes of social systems

When two people interact, they reveal widespread interdependence. This interdependence happens in many ways. The words used by one interlocutor are partly a function of the words that their conversation partner just used. Conversation partners choose words that facilitate understanding given what they know of each other (Clark & Wilkes-Gibbs, 1986). They even combine words into sentences in a manner that adapts to the information flow that might help a comprehender (Jaeger, 2010). This interdependence occurs in other subtler behaviors, too. Their phonological tendencies may become more similar – or less similar – depending on the social goals they have with each other (Manson et al., 2013; Coupland, 1985). The subtle body sway that they reveal may also become correlated, and this correlation may reflect the kind of social interaction they are having (Paxton & Dale, 2013). Under particular conditions, their heart rate may even fluctuate in similar ways (Fusaroli et al., 2016). Sometimes these patterns are powerful indices of a shared understanding, other times they may be a soft but detectable background "hum" of shared multimodal structure (Louwerse et al., 2012).

https://doi.org/10.1075/ais.10.04dal

We do not yet have a good understanding of the processes that govern this interdependence. Some of these similarities between people may be merely correlational, an outcome of other functional correspondences. For example, the precise words chosen by partners may drive shared posture dynamics, suggesting that the bodily signals themselves may not be in a direct causal relationship (Shockley et al., 2007). However, it is unlikely to be so simple a story. For example, eye-movement correspondence can facilitate memory, which in turn may influence eye-movement correspondence – forming a kind of causal circuit or feedback loop (D. Richardson & Dale, 2005). In addition, there may be multiple causal forces driving these correspondences between two people. Shared topics and words in conversation may drive bodily correlation; but so may a desire to be more socially affiliated, such as by occupying similar bodily stances or to nod in understanding in a similar way (e.g., Lakin & Chartrand, 2003).

Because of this complexity, it is important to improve our understanding of how, in natural conversation, the various processes involved work in concert. Though we focus on eye movements in this chapter, we will consider their relationship to the overall system, to other behaviors and cognitive processes. We situate work on visual attention during interaction in this broader landscape of research on language and social interaction.

A first goal of this chapter is a theoretical one. In order to unpack the complexity of interaction, we argue that it is advantageous to consider two people in interaction as constituting their own kind of unitary system. By understanding the dyad as a kind of coupled system, we can investigate relationships between its different parts and processes, whether they are attached to a single interlocutor, or whether they reflect information flowing bi-directionally between two people.

Our second goal is to show that eye movements are an especially fruitful source of data to test this "system" premise. By now, more than a decade of research has been devoted to how two visual attentional systems can become interdependent in the manner described above. We argue that this research has resulted in an important central observation: The dynamics of visual attention between two people reveal that they *can* become a coupled system.

In what follows, we first develop what we mean by "system," in the next section. After this, we provide an extensive empirical review, showing that eye movements and other behavioral signatures indeed reveal patterns of rich social interdependence, even in very basic experimental tasks. In the concluding section of the chapter, we elaborate on the criteria for determining whether two people form a unitary system, and describe how this motivates future research questions.

1.1 Three types of system

As a theoretical commitment, it is insufficient to say simply that two people form a "system." There are many senses of the word "system," and these can be broken down into at least three types: mechanical, computational, and complex. These definitions are not completely independent. Computational systems can also be considered complex dynamic systems (Simon, 1992), and complex systems can certainly, in a way, carry out computation (Crutchfield, 1994). But these notions of "system" offer terminological and conceptual distinctions that are useful for supporting empirical research questions (for an early review and discussion, see: Shenhar, 1990).

The simplest sense of system is a *mechanical system*. A mechanical system is one composed of elementary parts, and these parts have very particular and fixed purposes. These parts carry out such functions amidst other parts, and result in particular behaviors in response to specific states or conditions. If one of these parts is disrupted in some way, it can have an immediate detrimental effect on the functioning of the system. A car's engine, a thermostat, and a typewriter are classic examples of such a system. Even in their modern instantiations, these systems have parts of relatively fixed function, and when these parts are disrupted, the system is not especially adaptive – its operation is disrupted in some fashion. In the parlance of cognitive modeling, these systems do not show "graceful degradation" (Bechtel & Abrahamsen, 1991).

A *computational system* is more flexible. It has components such as variable states and algorithms that can take on more diverse sorts of input/output transformations. In fact, the classical mechanical systems mentioned above can be made more flexible by integrating new computational hardware and software to make these devices more adaptive. For example, a smart thermostat may adapt to the occupants of a home by learning what temperatures they set – even if the thermostat's temperature meter is miscalibrated. Classic examples of computational systems are exemplified by symbolic cognitive models (Newell, 1980; Anderson, 1996). Cognitive models can be highly adaptive. They can deploy context-specific computational rules of operation that can help the system overcome rapid changes in its environment, or to generalize to new environments.

Finally, a *complex system* is one in which its parts may not even have strictly designated functions; these functions may change within the overall system in response to conditions the system finds itself in (M. Richardson et al., 2014). Such systems are highly adaptive (Mitchell, 2009). In fact, adaptive complex systems have been referred to as "anti-fragile," in the sense that they thrive on perturbations (Taleb, 2012). Some have argued that in complex systems, the parts themselves may not be neatly structurally or functionally distinguishable from the whole or its environment (Chemero & Turvey, 2008). Consider a simple, single gesture embedded in linguistic interaction. In form and function, that gesture requires its relationships to other aspects of interaction to bring it significance. The gesture and its context are, in an important sense, mutually reliant (Enfield, 2013).

This system taxonomy has a complicated and sometimes provocative history, which we've synthesized here (for useful review see Adams & Aizawa, 2001; Bechtel & Abrahamsen, 2010; Eliasmith, 1996; Spivey, 2008; Van Orden et al., 2003; Van Orden & Stephen, 2012). Indeed, the three types of system are not unrelated, and not even always mutually exclusive. There is considerable debate about the difference between computational and complex systems (e.g., Eliasmith, 2012; Van Orden & Stephen, 2012). There are deep similarities between these concepts, such that any complex dynamic system can be characterized as computational in some sense or another (e.g., Crutchfield, 1994; Edelman, 2008; Mitchell, 2009; and of course Wiener, 1961). In addition, mechanical systems can be regarded as kinds of very simple computational systems (e.g., Ashby, 1956; Van Gelder, 1998). We therefore do not pretend that these concepts perfectly designate systems that interest cognitive scientists.

However, adopting one of these system concepts can facilitate (or deter) research questions by favoring particular causal narratives. For example, researchers who highlight computational systems often strictly delimit the internal states of the system from the environment in which it functions (Adams & Aizawa, 2001). Those who highlight complex systems, however, see elements of the environment as important parts of the system itself (Chemero, 2011). In cognitive terms, the arrangement of tools or artifacts in our environment, including the position and behavior of other *individuals* in our environment, involve exchanges of information that are an active part of processing (Hutchins, 1995; Tollefsen, 2006).

When one embraces the complex-systems concept, it motivates considering all the exchanges of information – whether within and across individuals and artifacts – as a coherent system that can be investigated. This is the system concept that inspires the current chapter. We hope to convince the reader that this system concept is especially useful in the case of eye movements. Eye movements are *both* a continuous perceptual medium *and* behavioral signal that others can see (Risko et al., 2016). Their semi-continuous dynamics make social eye movements a fruitful signal for investigating "systemhood" between two or more interacting individuals.

Eye-movement research is also an archetype of an emerging general research strategy among those interested in complex and dynamic systems. Eye movements provide semi-continuous dynamics of a system. These dynamics can be investigated for their shared structure across individuals, how processing is perturbed or stabilized in real-time, and more. In general, these behavioral dynamics can reveal the co-variation among parts of a system. The eyes are not alone in this respect. We begin the next section by considering the wealth of data that can be gathered on the dynamics of cognitive systems. This also includes speech, body movement, and more. In order to pursue a complex and dynamic systems approach to cognition, researchers have lately taken on strategies for designing experiments, measuring behavior semi-continuously, and framing their analysis under new techniques. In the next sections, we showcase the wide variety of research that takes on this dynamic approach. We then situate eye-movement research within it.

2. Collecting samples from the temporal dynamics of a cognitive process

In order to measure a complex dynamical system properly, one needs a sequence of samples from the system while it is functioning. Too often, traditional cognitive psychology methods have been designed to collect a single measurement at the end of the process carried out by the cognitive system (e.g., reaction time and/or accuracy). If the dynamics of the process are key to understanding the mechanisms of the system – because it is a *dynamical* system – then that single sample collected after the process is complete does not provide much insight into those mechanisms.

The research paradigms for developing sequences of multiple samples during a cognitive-motor process may be carved up into three basic categories: a) devise a cyclic motor behavior that can be measured continuously, b) treat a series of discrete separated tasks (experimental trials) as though it were one long more-general cognitive performance, and c) develop a methodology that can collect multiple samples within the few seconds that make up each experimental trial: dense-sampling methods.

2.1 Devise a cyclic motor behavior

The first, and perhaps oldest, approach involves designing a behavioral task that is repetitive and cyclic so that the human system can be measured many times during this lengthy process. This could be something as basic as walking on a treadmill while the leg and body movements are recorded with motion capture methods (Hove & Keller, 2015), or something as simple as tapping your finger to the beat of a metronome (Hove, Spivey, & Krumhansl, 2010; Repp, 2005), or it could be as contrived as a bimanual coordination task where one's index fingers are supposed to waggle up and down in anti-phase with one another (one index finger going up while the other index finger is going down; see Kelso, 1984, 1997).

In the case of bimanual coordination, the Haken, Kelso, and Bunz (1985) model provides an elegant mathematical account of the circumstances under which anti-phase synchrony could be maintained in the finger waggling, and those conditions under which the fingers would involuntarily fall into in-phase synchrony (both fingers up at the same time, and then both down at the same time). The fingers can maintain that anti-phase finger waggling at slow and medium speeds. But when the frequency speeds up, these two connected networks naturally slip into in-phase synchrony, where they both produce essentially identical activity patterns at the same time. This pull toward synchrony among systems that are informationally connected with one another is a common phenomenon in an extremely wide variety of contexts, including superconductors, pendulum clocks, fireflies, and neural networks (Strogatz, 2004).

In fact, that natural pull toward synchrony even happens when the two limbs that are trying to coordinate in anti-phase repetition belong to two different brains. Following up on the work of Kelso and colleagues, Schmidt, Carello and Turvey (1990) had *two people* sit on a table and each of them swung one leg in anti-phase synchrony with the other person's leg (e.g., one person's leg swung forward while the other person's leg swung backward). Just as with bimanual coordination, at slow and medium frequencies, this anti-phase pattern could be maintained. But at higher frequencies, the two legs tended to involuntarily slip into in-phase synchrony.

Obviously, this pull toward synchrony was not due a corpus callosum connecting the motor cortices. It was due to the shared perception-action loop connecting the left motor cortex in one person's brain to the right motor cortex in the other person's brain. While each person was producing their own motor output, they were simultaneously perceiving the motoric results of the other person's motor cortex, and this perceptual information continuously influenced their own motor cortex. The continuous information flow between the two people allowed the two of them to function as a single two-legged complex system. Importantly, the same Haken-Kelso-Bunz mathematical model of attractor basins in a relative-phase landscape, which described how a system with two fingers waggling would settle into certain patterns of behavior, also fit the data from two people's legs swinging. That is, the model that treated a person's two fingers like they were *one complex system* also successfully treated two people's different legs like they were *one complex system*.

2.2 A series of experimental trials as dynamics

The second paradigm that allows one to collect many samples from the temporal dynamics of a cognitive process is one where the traditional cognitive psychology experiment remains the same, but the data undergo a very different treatment. In many traditional cognitive experiments, a series of experimental trials are randomized in order, and one data point is collected at the end of each trial. The order of trial types is usually randomized with the intent of ruling out any sequential effects from one trial to the next. The experimenter hopes that, when one trial is over, the participant's brain goes back to a neutral resting state in preparation for the next trail. Any deviations from that expectation will turn into "random noise" in the data once many participants have contributed data to many different sequencings of those experimental trials. But what if there are intrinsic fluctuations in the participant's overall performance of the cognitive task across the duration of the experiment? What if the "noise" in the data wasn't purely random?

When each participant's sequence of reaction times is treated as one long time series of cognitive performance, it turns out that the fluctuations do not exhibit uncorrelated noise, but instead the noise has correlations with itself over time (Gilden, 2001; Kello, Anderson, Holden, & Van Orden, 2008; Van Orden & Holden, 2002). This correlated "pink" noise tends to show an abundance of long-range correlations, as is often observed in the power law distributions of behaviors from self-organized complex systems, such as avalanches, earthquakes, animal swarming behaviors, heart function, and neural networks. Thus, the ubiquitous presence of 1/f power law scaling of the variance (i.e., "pink" noise) in reaction times and other sequential behaviors was taken as evidence that human cognition may not be well described by a traditional *computational system* composed of processing modules linked up with linear signal transmission (i.e., the box-and-arrow computer metaphor for the mind, where variance comes only from white uncorrelated noise). Instead, human cognition may be better described as a self-organized *complex system* in which coherent mental processes emerge, or "soft-assemble", dynamically as a result of many simple sensory and motor processes interacting (Kello et al., 2010; Kloos & Van Orden, 2010; Van Orden, Holden, & Turvey, 2003).

Dotov, Nie and Chemero (2010) offer an example application of this to test whether a person and her environment are acting as a "system." These authors tracked smooth use of an artifact (computer mouse) in an extended task. They found this correlation (pink) noise under conditions when a participant was smoothly utilizing the artifact, and inferred that the artifact had become, in a sense, a participant in those cognitive dynamics.

2.3 Dense-sampling methods

The third paradigm that allows one to collect many samples from the temporal dynamics of a cognitive process is one where laboratory devices are used that can record multiple measurements per second: dense-sampling methods. There are several such methods, and we will mention two here before turning to this chapter's focus: eyetracking methods. For example, electroencephalography (EEG) and magnetoencephalography (MEG) typically collect a thousand or more samples per second of brain activity, and when those data are examined with time series analysis methods, they can reveal the sequence of influences between cortical regions (e.g., Gow & Olson, 2015; see also Spivey, 2016).

Though brain imaging techniques can constrain the experimental task environment, many dense-sampling methods can allow for the experimental task to become more ecologically valid (i.e., more applicable to real-life circumstances). Rather than interrupting a participant's cognitive process and forcing them to provide some explicit meta-cognitive report of their internal processes (e.g., speed-accuracy tradeoff methods), or asking them to perform an unusual observation on what would otherwise be their normal everyday language processing (e.g., phoneme-monitoring tasks, or lexical decision tasks), dense-sampling methods often allow a person to go about their cognitive task in a normal fashion without interruptions. The samples are collected (multiple times per second) as a by-product of the person's natural behaviors while carrying out the cognitive task.

For example, Shockley, Santana, and Fowler (2003) recorded two people's postural sway while they engaged in a joint puzzle-solving task. Each person stood on a pressure plate, and changes in the x,y location of each person's center of mass were recorded 60 times per second. The natural conversation and body language during the 30-minute task were uninterrupted by this dense-sampling measure, and yet millions of data points were collected across the 26 participants. Buried in those millions of data points, Shockley et al. found that when the two people are conversing with each other to solve the puzzle, their postural sway (as a simple index of their body language) gets coordinated in a way that shows recurrent patterns with one another (see also M. Richardson, Marsh, & Schmidt, 2005). By contrast, when each participant is conversing with an unrelated confederate to solve the same puzzle, that coordination in postural sway is not present.

In fact, in those situations where two people become entrained with one another, even with something as simple as tapping together to the beat of a metronome, they can develop an increased affiliation for one another. Hove and Risen (2009) had pairs of participants tap their fingers in synchrony with a visual metronome, and under circumstances where synchrony was greater between the two people, their reported affiliation with one another was also greater.

The direction of causality can go the other way as well. Having that affiliation already, such as with a family member, can in turn cause more synchrony between the two people. For example, Konvalinka et al. (2011) demonstrated that when a person watches their family member do a fire-walking ritual, both the fire-walker's heart rate and the family member's heart rate become coordinated in a way that shows recurrent patterns with one another. By contrast, non-related spectators did not show that heart rate coordination with the fire walker. Thus, similar to Kelso's (1984) waggling fingers and Schmidt et al.'s (1990) swinging legs, a dyad of two people engaged in a conversation or a joint task, or even just sharing a profound moment together, often begin to function in a way that looks a bit more like *one complex system* than two.

3. Eyetracking as a dense-sampling measure of human interaction

Among the various dense-sampling measures that are used to study cognition, the most widespread dense-sampling measure of human linguistic interaction is perhaps eyetracking. Originally called the Visual World Paradigm, this approach started out focusing significantly on adapting traditional experimental frameworks from psycholinguistics into the dense-sampling context. By collecting multiple samples (i.e., visual fixations) during the process of each trial – rather than a solitary data point at the end of a trial – the Visual World Paradigm was also able to use somewhat more ecologically valid tasks that did not involve meta-linguistic queries, the way lexical-decision tasks or phoneme-monitoring tasks do. Instead, the participant could simply follow natural spoken instructions to move objects around in a display, and the eye movements were recorded as a natural by-product of the language comprehension process.

One of the first reported results with this spoken language eyetracking paradigm demonstrated that a visual context could influence the real-time resolution of a syntactic ambiguity (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; see also Farmer, Cargill, Hindy, Dale, & Spivey, 2007). The accepted wisdom of that time suggested that only syntactic principles could influence the syntactic decisions made in the face of a temporary structural ambiguity (Ferreira & Clifton, 1986; Frazier, 1995). For example, while hearing "on the towel" in sentences (1) and (2), there's no way to know whether this prepositional phrase is part of the verb phrase in curly brackets, telling you where to put the apple (as in sentence 1), or whether it is part of the noun phrase in square brackets, telling you where the apple currently is (as in sentence 2). Tanenhaus et al. (1995) showed that when there was one apple in the visual display, the eyetracker recorded people's eyes looking at the irrelevant extra towel, suggesting that people were briefly mis-parsing sentence (2) as initially instructing them to relocate the apple onto that empty towel. By contrast, when the display contained two apples (so that the pragmatics of the situation required further specification for "the apple"), that same syntactically ambiguous sentence no longer elicited eye movements to that other towel.

(1) {Put [the apple] on the towel}.

(2) {Put [the apple (on the towel)] in the box}.

The VWP showed not only that visual context can influence syntactic processing over several seconds of speech, but also that it can influence reference resolution over the one or two seconds it takes to utter a single noun phrase. For example, with sentence (3), when the display has only one starred object, Eberhard, Spivey-Knowlton, Sedivy, and Tanenhaus (1995) found that people would often fixate that correct object before the noun "square" was even uttered. Due to their incremental comprehension of the sentence as it unfolded, and mapping of those adjectives onto the available objects in the display, they were able to determine the object to which that noun phrase referred *without even having to hear its head noun* (see also Reali, Spivey, Tyler & Terranova, 2006).

(3) Touch the starred yellow square.

Altmann and Kamide (1999) showed that a similar kind of linguistic anticipation works for transitive verbs as well. When there was only one edible object in the display (e.g., a cake) and the participant heard sentence (4), participants were often fixating the cake before the word "cake" had even been spoken. Not only does visual context influence these real-time predictions, but so do the semantic relationships between the verb and the relevant nouns (Knoeferle & Crocker, 2006).

(4) The boy ate the cake.

At the finer-grain temporal scale of a few hundred milliseconds, the VWP was even able to show that visual context could influence processing during the recognition of a single spoken word. Partway through hearing the word "candle" in sentence (5), participants frequently fixated a candy when one was in the display (Allopenna, Magnuson, & Tanenhaus, 1998; Spivey-Knowlton, 1996). This highly-replicated finding reveals two key insights about spoken language comprehension: (1) the process of spoken word recognition accrues acoustic-phonetic information continuously over the time scale of dozens of milliseconds, such that objects with similar sounding names can draw one's attention while hearing one of the words, and (2) visual context can influence this speech perception process.

(5) Pick up the candle.

In addition to having inanimate objects in the visual environment that can serve as context for real-time language comprehension, there are frequently people in the environment who can serve as context as well (Brown-Schmidt, Yoon, & Ryskin, 2015; Fitneva & Spivey, 2005). However, in contrast to the inanimate objects as context, people do a lot more than simply help a listener comprehend the linguistic stimuli, they help *co-create* the linguistic stimuli in the first place via natural unscripted conversation. Compared to the unidirectional scripted instructions typically used in the early experiments with the VWP, natural unscripted conversation makes for a significantly different language environment. So different, in fact, that some of the temporary ambiguities that psycholinguists have identified in language processing don't really behave like they are ambiguous.

In an example of this, Brown-Schmidt, Campana and Tanenhaus (2005) used an unscripted joint task that required coordination and conversation, and recorded participants' eye movements. They found that when sentences with complex noun phrases, such as (3), naturally occurred in the conversation (436 of them), the listener at that moment exhibited exactly the incremental comprehension that had been previously observed in the scripted experiments (e.g., Eberhard et al., 1995). However, when Brown-Schmidt et al. identified the 75 times that a spoken word in this unscripted conversation was potentially temporarily ambiguous between two objects, such as a candle and a candy, not a single one of those instances involved the listener making distracted eye movements to the competitor object.

More detailed explorations in this unscripted language scenario revealed that two specific pragmatic factors naturally provide the disambiguation of these potentially ambiguous linguistic events (Brown-Schmidt & Tanenhaus, 2008). Proximity of the potential referents and relevance to the task are two pragmatic factors that tend to constrain the referential domain to a subset of objects in the visual environment. Thus, in natural unscripted conversation, as two people are co-creating their dialog, the dyad routinely uses these pragmatic factors to prevent confusing ambiguities from even happening in the first place.

A dyad develops this ability to behave as one complex system by way of enabling a variety of entrainment processes across the two people; not just postural sway and heart rate, as mentioned above, but even brain activity. Kuhlen, Allefeld, and Haynes (2012) recorded continuous EEG activity from a man and then from a woman while they each retold common fairytales from memory on camera. The video and audio recordings were then replayed, overlaid on one another, for participants who were instructed to either attend to the female speaker or attend to the male speaker. While seeing and hearing the exact same stimulus, but attending to only one of the speakers, these participants' had their brain activity recorded with continuous EEG. The listeners who attended to the female speaker produced EEG time series data that correlated more with the female speaker's EEG time series data, and the listeners who attended to the male speaker produced EEG time series data that correlated more with the male speaker's EEG time series data. Thus, a listener's brain activity becomes similar to a speaker's brain activity (see also Hasson, Nir, Levy, Fuhrmann, & Malach, 2004).

It should not be surprising that when brain activity becomes entrained across two people, eye movements also become entrained across those two people. For example, D. Richardson and Dale (2005) had a participant look at a display of six characters from a television sitcom and tell the unscripted story of their favorite episode into the microphone while their eye movements were tracked. Later, this audio track was played back into the headphones of other participants whose eyes were being tracked. Using time series recurrence analysis in a fashion similar to how Shockley et al. (2003) analyzed postural sway, Richardson and Dale found that listeners produced eye movement sequences that showed substantial similarity to the eye movement sequences that were produced by the speaker – with about a 2-second lag in the time series. That 2-second lag is roughly accounted for by the time it takes for the speaker to look at a face and then utter the name of that face, plus the time it takes for the listener to comprehend that spoken name and then program an eye movement to that face. When that 2-second lag is adjusted for, the eye movements of the speaker and the listener show remarkable synchrony, such that the listener's oculomotor behavior was becoming entrained with the speaker's oculomotor behavior.

Now consider the more common scenario of two people having a two-way live conversation, where the dialog that gets produced is a real-time co-creation of the two people. D. Richardson, Dale, and Kirkham (2007) used two eyetrackers in two separate rooms with the participants each wearing a telecommunication headset. The participants engaged in spontaneous unscripted dialog about the same visual display, and each dyad's eye movements exhibited the same kind of recurrence of sequential patterns as before – but with one important difference. Instead of the peak coordination in eye movements happening at a 2-second lag between the two time series, the peak in eye movement coordination for these live two-way interlocutors happened at zero lag in the time series. Essentially, because the dialog was being co-created, a listener was often able to anticipate the speaker's words and attention enough that they were usually looking at the same thing that the speaker was looking at, at the same time.

Of course, eye movements are not the only behaviors that result from brain activity. And when brain activity gets entrained across two people during conversation, many behaviors will become coordinated. Louwerse, Dale, Bard, and Jeuniaux (2012) recorded a variety of facial and manual gestures, as well as speech acts, while two people participated face-to-face as an Instruction Giver and an Instruction Follower, using a shared cartoon map that had some occasional differences in their landmarks, forcing the two people to discuss and identify the differences between their two maps. At a variety of timescales within the duration of the conversation – be it within a second, or over dozens of seconds, or even minutes – each dyad routinely exhibited temporally coordinated behaviors of all kinds (e.g., smiles, head-nods, pointing gestures, queries and clarifications, etc.). Some of these coordinated behaviors were nearly synchronized, within a second of each other, whereas others exhibited a 20-second lag in their entrainment. Notably, this ubiquitous motor and linguistic coordination tended to increase as the dyad acquired more experience as a team.

By having a shared visual information resource – whether it be a line-up of six faces (D. Richardson et al., 2007), a cartoon map (Louwerse et al., 2012), or an actual nautical map that sailors are looking at together during a boat race (see Hutchins, 1995) – each person that manipulates that co-present visual resource to aid their own cognitive processes winds up also aiding the cognitive processes of the other viewers and interlocutors. As a result, their individual perception-action loops (Neisser, 1976) become entangled with one another via those shared actions and shared language (Spivey, 2012).

3.1 Human interaction as systemic coupling

The research reviewed in the prior section shows that eye movements serve as powerful indices of language comprehension, including when that comprehension takes place in unscripted conversation between two people. Additionally, that tight coupling in eye movements seen in D. Richardson and Dale (2005), D. Richardson et al. (2007) and in other behaviors (e.g., Kuhlen et al., 2012; Louwerse et al., 2012), suggests that the dense-sampling measures are indexing the interdependence of two cognitive systems.

Similar research using this dense-sampling technique aims to show precisely this – that a task analysis of human interaction is best explained as a kind of emergent system. This goal again holds across a variety of measures and behavioral signals. For example, Mønster, Håkonsson, Eskildsen, and Wallot (2016) had three-person teams figure out how to build as many origami boats as possible in assembly-line style, while recording their facial muscles (with electromyography) and sweating (with a galvanic skin response measure). Rather than simply analyzing one time series from one of the team members, Mønster et al. looked at the temporal coordination of those measures across team members. Teams that smiled in synchrony a lot tended to exhibit good team cohesion and positive affect toward the group. Teams that sweated in synchrony a lot tended to exhibit poor team cohesion and negative tensions within the group. Importantly, the key indicators of team success or failure were not merely in the recordings of physiological measures themselves. What provided the index of how coordinated and successful the team was at their joint task was analyzing the degree of synchrony among those physiological measures. Similarly, Fusaroli, Bjorndahl, Roepstorff and Tylén (in press) found that a team of people building with LEGO blocks showed an increase in shared heart rate dynamics as they developed more practice with the task. Improvements in behavioral coordination among the team resulted in increases heart rate synchrony and also produced increases in rapport and group competence.

In addition to this ecologically valid strategy, some have also sought to distill critical ingredients of this two-person complex system in basic experimental designs. One such distillation is from the work of Jordan, Knoblich and colleagues, who designed game-like interfaces for participants to jointly control an icon on a computer screen (Jordan & Knoblich, 2004). A dyad would control an object by sharing the responsibility of its movement (e.g., separate direction vs. velocity keys). In their original findings, Jordan and Knoblich (2003) discovered that smooth dynamic control over the computer icon came from participants coordinating predictive behaviors with each other. The overt actions of one partner served to anticipate or signal to the other what was going to happen next. This meant that not only were participants interdependent in a weak sense (i.e., one affects the other), but that the actions *possible* by another person were being integrated and responded to by a partner. These participants are, one could argue, co-representing the task. Not only do the *actions* of each person tacitly anticipate the competitive response options confronting their partner (Knoblich & Jordan, 2003; Sebanz, Knoblich & Prinz, 2003), but their *neural activity* also reveals this compensation for the partner's challenges at early stages of perceptual processing (Sebanz, Knoblich, Prinz, & Wascher, 2006). In certain circumstances, the pair can become like a two-brained, two-bodied meta-person.

In another example of how eye movements demonstrate this tendency for interaction to "weave" two people into one complex system, Dale et al. (2011) analyzed eye-movement correlation between people playing the so-called tangram game. The tangram game is a psycholinguistic task in which participants learn to refer to unfamiliar objects ("tangrams") to more efficiently pick them out as a team of two. The authors described the coupling of visual attention between two people as reflecting an emerging "tangram recognition system" that comprises two people. They analyzed eye movements of both interaction partners, and also the mouse-movement dynamics of the person tasked with clicking the shape that they were hunting for.

After participants did this over a few rounds, they became highly efficient, picking out tangrams in less than a second. In these expert dyads, eye movements became tightly coupled between the two people. These pairs showed increased maximum correlation between eye movements, in highly compressed periods of time (< 1 s), suggesting that their interaction was organizing their visual attention into bursts of matching. Meanwhile, the computer mouse exhibited a kind of "fixed" lag behind the eyes (approximately 500 ms). These two-person systems had parts with different temporal relationships to each other. Visual attention became coupled, but the manual components of the system (the computer mouse) had a "delay line" that was invariant even in the expert performers.

These results suggest that the task is a critical feature of interacting systems and how they meld. Participants learn how to solve tasks together, and from the simplest anticipation tasks (Knoblich & Jordan, 2003) to complex referential tasks (Dale et al., 2011), they organize their behavior in ways that achieve what the task prescribes. The resulting "two-person system" that emerges is a function of the task parameters.

This can be seen in other interactive experiments in which distinct patterns of joint behavior occur, dependent on task features. For example, when participants are in conflict or argue, the correlation between their body movements may diminish (Paxton & Dale, 2013a,b). In other tasks, participants may even take on complementary patterns of behavior. Fusaroli et al. (2012) find that pairs of participants who perform best on a joint perceptual task do not just copy each other, but become similar in language in task-oriented ways. In fact participants who showed too much copying tended to perform worse. This suggests that participants may sometimes *complement* each other in their dynamics. This can be confirmed by a clever analysis of the time series carried out by Fusaroli and Tylén (2015). They find that dyadic behavior in an interactive task is best accounted for by treating their transcript as a *single time series*, rather than two separate time series.

All this is to say that pairs of participants adapt to a task together. The dense-sampling of their behaviors shows interdependence, and shared dynamics encourage treating them as a kind of system, even if it is only momentary, for the duration of that particular exchange. The parts of this system – whether linguistic, bodily, visual, manual, etc. – come to take on different relative dynamics depending on the task parameters. It is the interaction in these tasks, and the shared goals of the participants, that shape their relative dynamics. This is indicative of the "complex system" concept described in the introduction to his chapter.

But so far, even after this extensive review, we have not fleshed out the explanatory value of describing the two participants as a system. Admittedly, computational systems could account for most behavioral results reviewed above. In the next section, we argue that explicit quantitative conditions can be specified for "systemhood." These conditions are inspired by complex systems, and perhaps more uniquely frame next steps for theoretical and empirical development.

4. Concluding discussion: Criteria for systemhood

We have showcased an array of findings suggesting that two conversants are acting much as a coherent, interdependent unit of two parts. Eye movements and other dynamic behaviors show interdependence during interaction. But this is not quite enough, of course. To a reader who prefers the so-called intracranial "mark of the cognitive" (Adams & Aizawa, 2001), these many indirect or correlational findings are insufficiently convincing.

In this concluding section, we offer further support for this complex systems formulation. Until now we've focused on high-level implications of our three system types, and focused on empirical review. But putting these pieces together, we find some concrete recommendations for cognitive experimentation as it bears on the "systemhood" of two people. The subtle differences suggested by complex nonlinear systems can cash out in empirical studies and theory development. We identify two sets of conditions for systemhood: one weak, and the other strong. Each recommends *specific* conditions in empirical analysis, and hold in eye movement data in particular.

4.1 Implications for research: Weak and strong conditions for systemhood

There are quantitative implications of taking a complex-systems approach to cognitive systems generally, and language use in particular. The first of these weak conditions is that the two-person system shows (WC1) real-time coordination and potential for behavioral reorganization. If we dense-sample behaviors in a natural task, their behaviors become statistically related, demonstrated via some quantitative analysis such as cross-correlation or cross recurrence or lag sequential analysis (see Bakeman & Quera, 2011, for review). The manner in which this correlation unfolds suggests that there is adaptation or short-term learning, in response to cognitive, environmental and social conditions.

When two people are faced with an interactive goal, such as identifying strange tangram shapes (Dale et al., 2011), their visual attentional systems become tightly time locked at 0 ms (see also D. Richardson et al., 2007). If the task changes, such as rendering communication unidirectional (speaker/listener), then the temporal function changes, showing distinct lags (D. Richardson & Dale, 2005). In the example of Brown-Schmidt et al. (2005), real-time two-way conversation leads to a complete absence of the effects of temporary ambiguity of a spoken word that are otherwise routinely observed with decontexualized scripted instructions. This is not just a simple change in the systems, but a reorganization of their patterns of behavior relative to an interlocutor. This reorganization is evident even when no interaction is taking place between people. For example, Crosby et al. (2008) have shown that the expectation that another person is present in a task changes visual attentional strategy (cf. Risko et al., 2011).

The second quantitative implication is how the eyes participate with the rest of the system. A complex systems approach predicts that (WC2) the many potentially varying dimensions of behavior get compressed in their dynamics. Though there are a large number of degrees of freedom in behavior, they are interdependent in a way that their "true" dimensionality is lower (see discussion in Riley et al., 2011). In Louwerse et al. (2012), we find that there is a cascade of constraint over the diverse patterns of behavior. Coordinated gestures become time locked. Coordinated head nods have a distinct temporal organization. Even incidental touching of the face becomes coordinated with its own particular temporal structure. Overall, this suggests there is "correlation structure" across the system. In the case of eye movements, this can be seen again in the tangram task, in which eye movements become tightly constrained alongside the language that participants come to use to identify the unfamiliar shapes (Dale et al., 2011). In fact, one could argue that the eye movements and referential terminology come to form a coupled circuit. As the eyes become better organized together, they better succeed at the task, which amplifies the linguistic strategy that was used, which again further couples the eyes, etc.

The prior two conditions represent weaker inferential conditions because they are easy to establish even in interacting individuals that are obviously not "systems." WC1 would be expected in single channels of behavior that show merely correlated patterns of behavior, and WC2 is simply to say that the correlation structure holds across a variety of behaviors. However weak they may be, they are pervasively observed in natural conversation. The problem, to echo a familiar statistics cliché, is that "correlation does not imply systemhood." Two further conditions strengthen the sense in which two people form a system. The following two strong conditions represent important next steps for research on eye movements in natural interaction and beyond.

First, in order to call two people a system, their behavior does not exhibit only reorganization in observed behavior, but primarily *functional* reorganization. The parts of the two-person system reveal a (SC1) mutual adaptivity that is accounted for by a systematic account of the task constraints. This is perhaps best demonstrated by the goal-oriented alignment that participants show at the linguistic level (Fusaroli et al., 2012). Here we find that participants are not merely reorganizing their behavior but also that it is best accounted for by the constraints from the task. In this sense, the reorganization observed in the two-person systems is in the service of succeeding in their joint task together. This stronger condition requires a solid theoretical understanding of the experimental task. Other examples are audience design tasks, such as those of Brown-Schmidt, Brennan and colleagues, which shows that knowledge of, or experience with, a task partner can sharply influence the pattern of behavior taking place in the task (for review see Brennan et al., 2010).

The second strong condition is that (SC2) the parts that coalesce to support system behavior respond to each other's changes. This is sometimes referred to as *reciprocal compensation* in the motor control literature in which multiple muscles or joints work together as one integrated system (see discussion in M. Richardson et al., 2014 and Riley et al., 2011). In the domain of natural conversation, one person should compensate for the other, or some aspects of each other's behavior should do so, too. Something like this can be seen in the distilled experimental conditions of Knoblich and Jordan (2003), where one participant's actions are in direct response to *anticipated* actions by his or her partner. Similarly, in social attention, Crosby et al. (2008; see also D. Richardson et al., 2012) show that participants anticipate the potential social discomfort of particular conversation topics by looking to potentially offended parties; the eyes are showing a kind of adaptation to the social-cognitive expectations that people have of each other. In general, this is a less attested pattern at the level of natural linguistic interaction, and is an open question: Does behavior, from visual attention to gesture and other linguistic contributions, respond dynamically in reciprocal compensation with one another, including those very behaviors *across* individuals?

These stronger conditions are often offered in other literatures to exemplify system-level operation, from physics (e.g., Haken, 1977) to motor control (e.g., Kelso et al., 1984; Latash et al., 2002). If the latter two conditions could be better attested in ongoing work in social eye movements and other behaviors, it would draw exciting bridges between the performance of two people in interaction with well-developed process theories in other domains.

4.2 Conclusion

The strongest "intracranial" accounts of cognition come perhaps from philosophers of mind. In particular, Rupert (2011) and Adams and Aizawa (2001) argue at length against extending cognitive systems beyond their traditional locus in the head (for counterpoint, see Noë, 2009). We would agree that this traditional boundary is a critical ingredient for cognitive theory. However, we are convinced that multi-person systems offer exciting possibilities for new theory development. By looking upon interaction through a complex-systems lens, it may offer provocative reframing of conversation and other forms of natural interaction as the operation of a single system. Such multi-person systems would be, of course, labile and loosely-coupled – people form ephemeral systems. But as it becomes coupled, what is the system capable of? How does it perform, and how do the parts work together to support that performance?

In the prior section we described some quantitative conditions that support systemhood. Elaborating the relationships within and between interaction partners will help us gain a clearer understanding of social interaction itself. Consider again visual attention, the focus of this chapter. Finding the cognitive processes and behaviors that act in a "causal circuit" with visual attention – such as memory, word choice, social affiliation and so on – are critical for understanding how two people stabilize each other's behavior, and support their interactive goals.

This search for principles could reveal new lines of questioning. For example, once two people form their own complex system, we can then subject it to all sorts of quantitative investigation. This investigation would include the several lines of inquiry that began this chapter, including the computational approach itself (cf. Beer & Williams, 2015). What computations do teams carry out? How can we maximize the computational abilities of a dyad, its transactive memory, or its affect and affiliation? What are the dynamic signatures of the formation and dissolution of dyadic systems, and can we predict their onset? These are questions others have already asked, of course (among many: Fitneva & Spivey, 2005; Fusaroli et al., 2014; Gallagher & Crisafi, 2009; Hutchins, 1995; Spivey, 2008; Tollefsen & Dale, 2012; Wegner, 1987). At minimum, we hope to have convinced the reader that eyetracking research, and related findings in experimental psychology, have much to offer to such theory development.

References

Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical psychology*, 14(1), 43–64. https://doi.org/10.1080/09515080120033571

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of memory and language*, 38(4), 419–439. https://doi.org/10.1006/jmla.1997.2558

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264. https://doi.org/10.1016/S0010-0277(99)00059-1

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *American Psychologist*, 51(4), 355. https://doi.org/10.1037/0003-066X.51.4.355

Ashby, W. R. (1956). *An introduction to cybernetics*. London: Chapman & Hall.

Bakeman, R., & Quera, V. (2011). *Sequential analysis and observational methods for the behavioral sciences*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139017343

Bechtel, W., & Abrahamsen, A. (1991). *Connectionism and the Mind*. Backwell, Cambridge, MA.

Bechtel, W., & Abrahamsen, A. (2010). Dynamic mechanistic explanation: Computational modeling of circadian rhythms as an exemplar for cognitive science. *Studies in History and Philosophy of Science Part A*, 41(3), 321–333. https://doi.org/10.1016/j.shpsa.2010.07.003

Beer, R. D., & Williams, P. L. (2015). Information processing and dynamics in minimally cognitive agents. *Cognitive science*, 39(1), 1–38. https://doi.org/10.1111/cogs.12142

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, 53, 301–344. https://doi.org/10.1016/S0079-7421(10)53008-1

Brown-Schmidt, S., Campana, E., & Tanenhaus, M. K. (2005). Real-time reference resolution by naïve participants during a task-based unscripted conversation. In J. Trueswell & M. Tanenhaus (Eds.), *World Situated Language Use: Psycholinguistic, Linguistic and Computational Perspectives on Bridging the Product and Action Traditions*, (pp. 153–171) Cambridge, MA: MIT Press.

Brown-Schmidt, S., Yoon, S. O., & Ryskin, R. A. (2015). People as contexts in conversation. *Psychology of Learning and Motivation*, 62, 59–99. https://doi.org/10.1016/bs.plm.2014.09.003

Brown-Schmidt, S., & Tanenhaus, M. K. (2008). Real-time investigation of referential domains in unscripted conversation: A targeted language game approach. *Cognitive Science*, 32(4), 643–684. https://doi.org/10.1080/03640210802066816

Chemero, A. (2011). *Radical embodied cognitive science*. MIT press.

Chemero, A., & Turvey, M. T. (2008). Autonomy and hypersets. *Biosystems*,91(2), 320–330. https://doi.org/10.1016/j.biosystems.2007.05.010

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7

Coupland, N. (1985). 'Hark, hark, the lark': social motivations for phonological style-shifting. *Language & Communication*, 5(3), 153–171. https://doi.org/10.1016/0271-5309(85)90007-2

Crosby, J. R., Monin, B. & Richardson, D. C. (2008). Where Do We Look During Potentially Offensive Behavior? *Psychological Science*, 19 (3), 226–228. https://doi.org/10.1111/j.1467-9280.2008.02072.x

Crutchfield, J. P. (1994). The calculi of emergence: computation, dynamics and induction. *Physica D: Nonlinear Phenomena*, 75(1), 11–54. https://doi.org/10.1016/0167-2789(94)90273-9

Dale, R., Kirkham, N. Z. & Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Frontiers in Psychology*, 2. https://doi.org/10.3389/fpsyg.2011.00355

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of psycholinguistic research*, 24(6), 409–436. https://doi.org/10.1007/BF02143160

Edelman, S. (2008). On the nature of minds, or: truth and consequences.*Journal of Experimental & Theoretical Artificial Intelligence*, 20(3), 181–196. https://doi.org/10.1080/09528130802319086

Eliasmith, C. (1996). The third contender: A critical examination of the dynamicist theory of cognition. *Philosophical Psychology*, 9(4), 441–463. https://doi.org/10.1080/09515089608573194

Eliasmith, C. (2012). The complex systems approach: rhetoric or revolution.*Topics in cognitive science*, 4(1), 72–77. https://doi.org/10.1111/j.1756-8765.2011.01169.x

Enfield, N. J. (2013). *Relationship thinking: Agency, enchrony, and human sociality*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199338733.001.0001

Farmer, T. A., Cargill, S. A., Hindy, N. C., Dale, R., & Spivey, M. J. (2007). Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing. *Cognitive Science*, 31(5), 889–909. https://doi.org/10.1080/03640210701530797

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of memory and language*, 25(3), 348–368. https://doi.org/10.1016/0749-596X(86)90006-9
Fitneva, S. A. & Spivey, M. J. (2005). Context and language processing: The effect of authorship. In J. Trueswell & M. Tanenhaus (Eds.), *World Situated Language Use: Psycholinguistic, Linguistic and Computational Perspectives on Bridging the Product and Action Traditions*, (pp. 317–328) Cambridge, MA: MIT Press.
Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. *Journal of Psycholinguistic Research*, 24(6), 437–468. https://doi.org/10.1007/BF02143161
Fusaroli, R., & Tylén, K. (2015). Investigating conversational dynamics: Interactive alignment, Interpersonal synergy, and collective task performance. *Cognitive Science*, 40, 145–171.
Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms quantifying the benefits of linguistic coordination. *Psychological Science*, 23(8), 931–939. https://doi.org/10.1177/0956797612436816
Fusaroli, R., Bjørndahl, J. S., Roepstorff, A., & Tylén, K. (2016). A Heart for Interaction: Shared Physiological Dynamics and Behavioral Coordination in a Collective, Creative Construction Task. *Journal of Experimental Psychology: Human Perception & Performance*, 42(9), 1297–1310.
Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, 32, 147–157. https://doi.org/10.1016/j.newideapsych.2013.03.005
Gallagher, S., & Crisafi, A. (2009). Mental institutions. *Topoi*, 28(1), 45–51. https://doi.org/10.1007/s11245-008-9045-0
Gilden, D. L. (2001). Cognitive emissions of 1/f noise. *Psychological review*, 108(1), 33–56. https://doi.org/10.1037/0033-295X.108.1.33
Gow, D. W. & Olson, B. B. (2015): Sentential influences on acoustic- phonetic processing: a Granger causality analysis of multimodal imaging data, *Language, Cognition and Neuroscience*. https://doi.org/10.1080/23273798.2015.1029498
Haken, H. (1977). *Synergetics. Physics Bulletin*, 28(9), 412. https://doi.org/10.1088/0031-9112/28/9/027
Haken, H., Kelso, J. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological cybernetics*, 51(5), 347–356. https://doi.org/10.1007/BF00336922
Hove, M. J., & Keller, P. E. (2015). Impaired movement timing in neurological disorders: rehabilitation and treatment strategies. *Annals of the New York Academy of Sciences*, 1337(1), 111–117. https://doi.org/10.1111/nyas.12615
Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27(6), 949–960. https://doi.org/10.1521/soco.2009.27.6.949
Hove, M. J., Spivey, M. J., & Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *Journal of Experimental Psychology: Human Perception and Performance*, 36(6), 1525–1534.
Hutchins, E. (1995). *Cognition in the Wild*. MIT press.
Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive science*,19(3), 265–288. https://doi.org/10.1207/s15516709cog1903_1
Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive psychology*, 61(1), 23–62. https://doi.org/10.1016/j.cogpsych.2010.02.002
Jirsa, V. K., Fuchs, A., & Kelso, J. A. S. (1998). Connecting cortical and behavioral dynamics: bimanual coordination. *Neural Computation*, 10(8), 2019–2045. https://doi.org/10.1162/089976698300016954
Jordan, J. S., & Knoblich, G. (2004). Spatial perception and control. *Psychonomic Bulletin & Review*, 11(1), 54–59. https://doi.org/10.3758/BF03206460

Kello, C. T., Anderson, G. G., Holden, J. G., & Van Orden, G. C. (2008). The pervasiveness of 1/f scaling in speech reflects the metastable basis of cognition. *Cognitive Science*, 32(7), 1217–1231. https://doi.org/10.1080/03640210801944898

Kello, C. T., Brown, G. D., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Sciences*, 14(5), 223–232. https://doi.org/10.1016/j.tics.2010.02.005

Kelso, J. A. S. (1984). Phase transitions and critical behavior in human bimanual coordination. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(6), R1000–R1004.

Kelso, J. A. S. (1997). *Dynamic patterns: The self-organization of brain and behavior*. MIT press.

Kelso, J. S., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C. A. (1984). Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. *Journal of Experimental Psychology: Human Perception and Performance*, 10(6), 812.

Kloos, H., & Van Orden, G. (2010). Voluntary behavior in cognitive and motor tasks. *Mind and Matter*, 8(1), 19–43.

Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(5), 1006–1016.

Knoeferle, P., & Crocker, M. W. (2006). The coordinated interplay of scene, utterance, and world knowledge: Evidence from eye tracking. *Cognitive Science*, 30(3), 481–529. https://doi.org/10.1207/s15516709cog0000_65

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E. M., Wallot, S., Van Orden, G., & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences*, 108(20), 8514–8519. https://doi.org/10.1073/pnas.1016955108

Kuhlen, A. K., Allefeld, C., & Haynes, J. D. (2012). Content-specific coordination of listeners' to speakers' EEG during communication. *Frontiers in Human Neuroscience, 6*, *Article* 266.

Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological science*, 14(4), 334–339. https://doi.org/10.1111/1467-9280.14481

Latash, M. L., Scholz, J. P., & Schöner, G. (2002). Motor control strategies revealed in the structure of motor variability. *Exercise and sport sciences reviews*, 30(1), 26–31. https://doi.org/10.1097/00003677-200201000-00006

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*,36(8), 1404–1426. https://doi.org/10.1111/j.1551-6709.2012.01269.x

Manson, J. H., Bryant, G. A., Gervais, M. M., & Kline, M. A. (2013). Convergence of speech rate in conversation predicts cooperation. *Evolution and Human Behavior*, 34(6), 419–426. https://doi.org/10.1016/j.evolhumbehav.2013.08.001

Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.

Mønster, D., Håkonsson, D. D., Eskildsen, J. K., & Wallot, S. (2016). Physiological evidence of interpersonal dynamics in a cooperative production task. *Physiology & Behavior*, 156, 24–34. https://doi.org/10.1016/j.physbeh.2016.01.004

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. San Francisco, California: W. H. Freeman.

Newell, A. (1980). Physical symbol systems*. *Cognitive science*, 4(2), 135–183. https://doi.org/10.1207/s15516709cog0402_2

Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. Macmillan.
Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. *The Quarterly Journal of Experimental Psychology*, 66(11), 2092–2102. https://doi.org/10.1080/17470218.2013.853089
Paxton, A., & Dale, R. (2013a). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods*, 45(2), 329–343. https://doi.org/10.3758/s13428-012-0249-2
Paxton, A., & Dale, R. (2013b). Argument disrupts interpersonal synchrony. *The Quarterly Journal of Experimental Psychology*, 66(11), 2092–2102. https://doi.org/10.1080/17470218.2013.853089
Reali, F., Spivey, M. J., Tyler, M. J., & Terranova, J. (2006). Inefficient conjunction search made efficient by concurrent spoken delivery of target identity. *Perception & psychophysics*, 68(6), 959–974. https://doi.org/10.3758/BF03193358
Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic bulletin & review*, 12(6), 969–992. https://doi.org/10.3758/BF03206433
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. https://doi.org/10.1207/s15516709cog0000_29
Richardson, D. C., Dale, R. & Kirkham, N. (2007). The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. *Psychological Science*, 18, 407–413. https://doi.org/10.1111/j.1467-9280.2007.01914.x
Richardson, D. C., Street, C. N. H., Tan, J. Y. M., Kirkham, N. Z., Hoover, M. A., & Ghane Cavanaugh, A. (2012).Joint perception: gaze and social context. *Frontiers in Human Neuroscience*, 6. https://doi.org/10.3389/fnhum.2012.00194
Richardson, M. J., Dale, R. & Marsh, K. (2014). Complex dynamical systems in social and personality psychology: theory, modeling and analysis. In H. T. Reis & C. M. Judd (Eds.), *Handbook of Research Methods in Social and Personality Psychology* (pp. 253–282). Cambridge University Press.
Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1), 62–79.
Riley, M. A., Richardson, M. J., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in psychology*, 2, 38. https://doi.org/10.3389/fpsyg.2011.00038
Risko, E. F., & Kingstone, A. (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Attention, Perception, & Psychophysics*, 73(2), 291–296. https://doi.org/10.3758/s13414-010-0042-1
Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the Fourth Wall of Cognitive Science Real-World Social Attention and the Dual Function of Gaze. *Current Directions in Psychological Science*, 25(1), 70–74.
Rupert, R. D. (2011). Cognitive systems and the supersized mind. *Philosophical Studies*, 152(3), 427–436. https://doi.org/10.1007/s11098-010-9600-6
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.
Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own?. *Cognition*, 88(3), B11–B21. https://doi.org/10.1016/S0010-0277(03)00043-X

Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in coacting individuals. *Journal of Cognitive Neuroscience*, 18(5), 859–870. https://doi.org/10.1162/jocn.2006.18.5.859

Shenhar, A. (1990). On system properties and systemhood. *International Journal Of General System*, 18(2), 167–174. https://doi.org/10.1080/03081079008935136

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.

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.

Simon, H. A. (1992). What is an "explanation" of behavior?. *Psychological Science*, 3(3), 150–161. https://doi.org/10.1111/j.1467-9280.1992.tb00017.x

Spivey-Knowlton, M. J. (1996). *Integration of visual and linguistic information: Human data and model simulations*. Ph.D. Dissertation, U. Rochester, NY.

Spivey, M. (2008). *The continuity of mind*. Oxford University Press.

Spivey, M. J. (2012). The spatial intersection of minds. *Cognitive Processing*, 13, 343–346. https://doi.org/10.1007/s10339-012-0520-6

Spivey, M. J. (in press). Semantics influences speech perception: Commentary on Gow and Olson (2015). *Language, Cognition, and Neuroscience.*

Strogatz, S. (2003). *Sync: The emerging science of spontaneous order*. Hyperion.

Taleb, N. N. (2012). *Antifragile: Things that gain from disorder* (Vol. 3). Random House Incorporated.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632–1634. https://doi.org/10.1126/science.7777863

Tollefsen, D. P. (2006). From extended mind to collective mind. *Cognitive systems research*, 7(2), 140–150. https://doi.org/10.1016/j.cogsys.2006.01.001

Tollefsen, D. & Dale, R. (2012). Naturalizing joint action: a process-based approach. *Philosophical Psychology*, 25, 385–407. https://doi.org/10.1080/09515089.2011.579418

Van Gelder, T. (1998). The dynamical hypothesis in cognitive science.*Behavioral and brain sciences*, 21(05), 615–628. https://doi.org/10.1017/S0140525X98001733

Van Orden, G. C., & Holden, J. G. (2002). Intentional contents and self-control. *Ecological Psychology*, 14(1–2), 87–109. https://doi.org/10.1080/10407413.2003.9652753

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, 132(3), 331. https://doi.org/10.1037/0096-3445.132.3.331

Van Orden, G., & Stephen, D. G. (2012). Is cognitive science usefully cast as complexity science?. *Topics in cognitive science*, 4(1), 3–6. https://doi.org/10.1111/j.1756-8765.2011.01165.x

Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In *Theories of group behavior* (pp. 185–208). Springer New York. https://doi.org/10.1007/978-1-4612-4634-3_9

Wiener, N. (1961). *Cybernetics or Control and Communication in the Animal and the Machine* (Vol. 25). MIT press.