The Self-Organization of Human Interaction

Rick Dale[1], Riccardo Fusaroli[1,3,4], Nicholas D. Duran[1,2] and Daniel C. Richardson[2]

[1] Cognitive and Information Sciences, University of California, Merced
[2] Division of Psychology and Language Sciences, University College London
[3] Interacting Minds Center, Aarhus University.
[4] Center for Semiotics, Aarhus University.

Word count: ~13,500 words
Number of figures: 9
Date of this draft: Thursday, December 20, 2012

Address correspondence to:
Rick Dale
Cognitive and Information Sciences
University of California, Merced
5200 N. Lake Rd.
Merced, CA 95343
rdale@ucmerced.edu
cognaction.org/rick

Abstract

We describe a "centipede's dilemma" that faces the sciences of human interaction. Research on human interaction has been involved in extensive theoretical debate, though the vast majority of research tends to focus on a small set of human behaviors, cognitive processes, and interactive contexts. The problem is that naturalistic human interaction must *integrate* all of these factors simultaneously, and grander theoretical mitigation cannot come *only* from focused experimental or computational agendas. We look to dynamical systems theory as a framework for thinking about how these multiple behaviors, processes, and contexts can be integrated into a broader account of human interaction. By introducing and utilizing basic concepts of self-organization and synergy, we review empirical work that shows how human interaction is flexible, adaptive, and structures itself incrementally during unfolding interactive tasks, such as conversation, or more focused goal-based contexts. We end on acknowledging that dynamical systems accounts are very short on concrete models, and we briefly describe ways that theoretical frameworks could be integrated, rather than endlessly disputed, in order to achieve some success on the centipede's dilemma of human interaction.

1. Introduction: the "centipede's dilemma" of interaction research

Next time you have a conversation, pay close attention to what you and your partner are doing. This self-consciousness can be a bit jarring. Like the so-called "centipede's dilemma," attempting awareness of your numerous cognitive processes and behaviors, and those of your conversation partner, can quickly disrupt a natural, flowing performance. Dialogue otherwise seems so easy (Garrod & Pickering, 2004). How do we do it? The famous poem by Katherine Craster has a toad posing to a centipede, "Pray, which leg moves after which?" The centipede goes about some introspection, attempting awareness of this coordination, only to find that she can no longer move.

The same thing seems to happen if we do this in a conversation. The information, and relevant cognitive mechanisms, that can bear on a conversational performance probably outnumber a centipede's legs, especially if you include the array of processes that are not available to conscious report. How do we coordinate everything?

In this review article, we consider a fundamental and still-unsolved puzzle faced by the fields that study human interaction. Much like consciously focusing on ourselves while conversing, the *scientific agenda itself* also suffers from a kind of centipede's dilemma. Ongoing work tends to focus on particular levels of analysis. For example, we know that people use similar vocabulary during interaction, may tend to match in other linguistic styles, or even take on similar bodily postures and movements. Through exploration of these levels, there are many theoretical proposals of candidate cognitive and social processes. Terminologies that populate these theories are diverse, conceptually overlapping, and still don't enjoy consensus definitions: mirroring, simulating, coupling, entrainment, coordination, imitation, mimicry, alignment, synchrony, joint action, theory of mind, perspective-taking, mutuality, accommodation, empathy, contagion, and more.

What is still lacking is a systematic agenda to uncover how these various processes work together to bring about *multimodal coordination* between two interacting people. The current agenda of isolating processes and developing broad theoretical proposals from relatively circumscribed domains is somewhat like the centipede's analysis of its performance leg by leg by leg. There is now a heterogenous assemblage of experimental techniques and observational analyses, and an associated array of diverse theoretical mechanisms that have yet to be integrated. The vibrant debates that ensue focus almost entirely on some subset of experimental paradigms, cognitive processes, or social contexts. The result is an exciting, evolving, but fractured domain.

Admittedly, framing the problem in this way may seem overly grandiose. But if a comprehensive theory of human interaction is our goal, then this is a real puzzle to be solved. We allay any enticement (or skepticism) here: We're not going to solve the puzzle in this paper. We will, however, propose one potential route to a solution. To do this, we look to concepts of adaptive and self-organizing systems, drawing from the tradition of the "dynamical systems framework" as it has come to be known in the cognitive sciences. Importantly, this is a *framework* for thinking about the problem of coordination during linguistic interaction. It does not answer the question of what the precise array of mechanisms is, or the processes of their interaction. It also does not, at least by necessity, replace extant theoretical proposals. Our positive thesis is simple: We will argue that the dynamical systems framework may help to integrate existing theories.

In what follows, we begin with more background on the debate about the mechanisms underlying human interaction. We then introduce some fundamental concepts of dynamics. These concepts bring about some generic expectations about how human interaction should be coordinated and structured. We argue that, in fancy terminology, "self-organization into functional synergies" should be (and is) evident in interactive data (Section 2).

We showcase some evidence for this in a review of empirical literature. This background review focuses on two key aspects of linguistic interaction. The first (Section 3) is that basic social variables *can* sharply modulate the many behaviors involved in interaction, at several levels of analysis. We look to low-level visual attention, and then higher-level spatial perspective-taking, during linguistic interaction. The second empirical review is a glance at complementarity between behaviors of two people interacting, extending current theories of alignment in dialogue and looking to the usefulness of the concept of "synergy" (Section 4).

Following this, in a concluding discussion, we relate dynamical systems to other theories (Section 5). We speculate on some ways that dynamics and adaptive control theory may be a means of becoming more precise about *what exactly* is coordinating during human interaction. As we articulate in this section, a common and important objection to dynamical systems accounts is that they are weak on identifying mechanisms: Dynamical theorists argue "interactions dominate cognition," and focus almost exclusively on relatively indirect measurement outcomes of that interaction; however, where there is interaction, there must be mechanisms making it happen. We address this issue in this final section by arguing that an important way to move forward is to integrate useful concepts from dynamics with some popular computational frameworks already exploring language and complex cognition.

2. An example theoretical debate and the need for integration

There are some prominent theoretical debates in the realms of discourse and psycholinguistics. One of the best known debates revolves around how, and how much, human beings track information about one another as they interact. Some theories posit a two-stage process, with primacy given to egocentric (self-centered) processes, and more social "other-centric" processes coming online only more slowly and strategically (e.g., Barr, 2008; Keysar, Lin, & Barr, 2003; Lin, Keysar, & Epley, 2010). Other theories posit a fundamental sensitivity to a conversation partner, with a rich layering of common ground that emerges while two people talk (e.g., Clark, 1996; Schober & Brennan, 2003; for recent discussion see Brennan, Galati, & Kuhlen, 2010; Brennan & Hanna, 2009; Brown-Schmidt & Hanna, 2011; Shintel & Keysar, 2009).

Other theoretical agendas in this debate have aimed to specify key cognitive processes that permit one person to keep track of, or continually adapt to, their conversation partner. Some of these accounts centralize a process of multi-level "alignment," which can be automatic and often non-conscious, and builds common representational states across individuals while they talk (Pickering & Garrod, 2004; Garrod & Pickering, 2004). Some have taken the suggestion of a human "mirror system" to be central, specifying core social processes that must be in place for us to interact successfully (for review see Gallese, 2008). Recent accounts have articulated the important role of executive function during conversation (Brown-Schmidt, 2009), of memory (Horton, 2005; Horton & Gerrig, 2005), and of the integration of basic contextual parameters of an interaction (Brennan, Galati, & Kuhlen, 2010). Still others have identified *kinds* of coordination, such as the emergent vs. non-emergent linguistic interaction that see different origins in activities done jointly (Knoblich, Butterfill, & Sebanz, 2011).

We see at least three exciting characteristics to this growing literature. First, researchers in these areas are beginning to tap into the cognitive mechanisms underlying interactions like conversation (e.g., Brown-Schmidt, 2009; Gambi & Pickering, 2011; Mehler, Weiß, Menke, & Lücking, 2011; Horton, 2005; Pickering & Garrod, 2009; Reitter, Keller, & Moore, 2011). This advances the valuable work on observational and conversation analysis that has shed great light on the structure of interaction (Sacks, Jefferson, & Schegloff, 1995; Schegloff, 2007), but is not capable of identifying the cognitive processes that drive it.

Secondly, and relatedly, social cognitive neuroscience (e.g., Frith & Frith, 2001; Van Overwalle, 2008) and related areas (e.g., imitation: Wang & Hamilton, 2012) have begun to explore these basic mechanisms at the level of the brain. The growth of this subfield of cognitive neuroscience has been very rapid, with many programmatic proposals for studying the circuits underlying social interaction (e.g., Cooper, Catmur, & Heyes, 2012; Dumas 2011; Dumas, Chavez, Nadel & Martinerie, 2012; Hasson, Ghazanfar, Galantucci, & Garrod, 2012; Konvalinka & Roepstorff 2012; Lieberman, 2007; Wolpert, Doya, & Kawato, 2003).

Thirdly—and this may sound odd—researchers have come to embrace the inherent social nature of language, and to carry out investigations of cognitive processing in more naturalistic circumstances (see Tanenhaus & Brown-Schmidt, 2008 and Fusaroli & Tylén 2012 for a review). The past century has seen some fundamentally different assumptions for a scientific understanding of language. For example, the classic conception of the ideal speaker-hearer, perhaps useful in some circumscribed domains, is an assumption that has outlived any usefulness it may have had in understanding how people actually use language in so wide a circumstance. Conversation analysis and discourse psychology have now been coupled with sophisticated computational and behavioral methods such as natural language processing and computational linguistics (Graesser, Swamer, & Hu, 1997), eye-tracking (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), automated body movement (Paxton & Dale, in press; Schmidt et al., 2012) and acoustic analysis (Oller et al., 2010; Wyatt et al., 2008), dynamical systems methods (Riley & Van Orden, 2005; Shockley et al., 2003), and more. Language is a complex and multidimensional activity, and our understanding of it—how it evolved, is learned and used—must come from integrating such sophisticated methods in naturalistic circumstances, not *only* from abstract assumptions about linguistic structure that rarely manifest themselves except in pre-empirical intuitions.

Although these are exciting developments, we would argue that *theoretical integration* has been less emphasized. We can think of a few reasons why this might be. For one, a researcher's theoretical proposals are usually tied to the specific contexts she or he studies. This is a natural feature of any scientific explanation (see Cartwright, 1999), but it limits the generalizability of the processes proposed. The very fact, for example, that language users can be rendered relatively egocentric, or relatively "other-centric," by experimental design means that something more complex is going on cognitively than simply the deployment of fixed architectures (see also Brennan et al., 2010 for discussion). Another reason, in our opinion, is that the multidimensional and "multimechanism" aspect of human interaction means that traditional conceptions of cognitive explanations are fundamentally challenged. In such a complex circumstance, theories seem unlikely to succeed by anchoring to small set of specific mechanisms, but rather to a context-dependent integration of a wide variety of processes acting together. This is what we mean by the "centipede's dilemma": there tends to be much local and circumscribed analysis, and much less cross-paradigm and inter-theoretical synthesis.

We introduce one way that approaches integration. Specifically, we look to the tools offered by what is often termed the "dynamical systems" approach in cognitive science (M. J. Richardson, Dale, & Marsh, in press; Spivey, 2007; Chemero, 2009; Port & Van Gelder, 1995; Thelen & Smith, 1995; Turvey, 1990; more on this below). A dynamical approach to these phenomena affords a variety of theoretical tools that embrace context-dependent integration, adaptation, and process flexibility. In the study of basic cognitive processing, significant debate has emerged about the usefulness of a so-called "nonlinear interaction-dominant dynamic complex systems" approach (see collection in Van Orden & Stephen, 2012), and whether it really adds much above already-present accounts such as constraint-satisfaction mechanisms (e.g., Eliasmith, 2012). These are all very important concerns, and we will address some of them in discussion below. However, this debate seems to have been unfortunately influenced by overly radical and unrealistic theoretical commitments, and perhaps reactionary tendencies in commentators. In many circumstances, a dynamical systems approach can be integrated in telling ways with existing theories. We argue that there is great value in the approach, with specific benefits to be gained from applying it to conversation. We begin by describing the theoretical framework in a highly introductory manner for those who still haven't read much about it, or have been too skeptical to get into it.

2. Self-organization and human interaction

2.1 The need to integrate accounts of cognition in linguistic interaction

The empirical literature on conversation and relevant interaction can stymie many theoretical proposals. A cursory glance at this empirical literature reveals a more complex story than is typically portrayed in any single theory. The cognitive processes proposed to be centrally involved in social interaction are numerous. In addition, they operate in a highly context-dependent way. Depending on the experimental paradigm chosen by the researcher, one can highlight some capacities over others. As noted above, particular laboratory interaction tasks may produce behavioral patterns indicative of egocentrism (Barr & Keysar, 2002; Keysar et al., 2000; Keysar et al., 2003; cf. Shintel & Keysar, 2009); at the same time, putting two people who are highly acquainted in an interaction may have a similar effect, of highlighting egocentrism, since each person can make assumptions that they are likely to be understood (e.g., Wu & Keysar, 2007b; whether this is an explicit metacognitive assumption is unknown). However, when establishing pointed moments of conversational disruption, a conversation-partner's needs or abilities, or different cultural contexts, these egocentric tendencies can become reversed (see, e.g., Wu & Keysar, 2007a; Brown-Schmidt, 2009; Brown-Schmidt et al., 2008; Galati & Brennan, 2010; Roche et al., submitted; Tanenhaus & Brown-Schmidt, 2008; see for review Brennan et al., 2010; Schober & Brennan, 2003).

These are exciting avenues of investigation, with extremely clever experiments destined to fuel this debate and discussion. Nevertheless, an inference to the best set of unique capacities is not possible from a small set of experiments, or even from a whole literature that highlights specific designs. The range of possible contexts of human interactions are simply too numerous to do so. For these reasons, it is unlikely to be the case that conversational performance and linguistic interaction, in whole, can be accounted for in terms of a small, single subset of mechanisms. Of course not all theories aim to be so comprehensive, as they tend to focus on specific aspects of social interaction. We would argue that, to achieve a more comprehensive account of social interaction, an *integration* of these task contexts, and cognitive capacities, is needed (cf. Brennan et al., 2010). But how can we hook up differing accounts into an overall theoretical framework that can achieve this integration? Here we argue that a dynamical systems framework may serve these questions in valuable ways. In the following section, we describe what we mean by "dynamical systems account," and describe two basic, but important, features: *self-organization* and *synergies*.

2.2 Dynamics, self-organization, and so on

It is widely known that the dynamical systems approach to cognition utilizes some terminology unfamiliar to many cognitive scientists. This concern has been expressed in many critiques. For example, a recent commentary's tongue-in-cheek title uses the comprehensive phrase "[n]onlinearly coupled, dynamical, self-organized critical, synergistic, scale-free, exquisitely context-sensitive, interaction-dominant, multifractal, interdependent brain-body-niche systems" (Wagenmakers, van der Maas, & Farrell, 2012). These are legitimate concerns, because identifying important new theoretical concepts inside an array of unfamiliar terms requires at least some concrete aspects of the agenda (in fact this is the important point expressed in the above-mentioned commentary). To be fair, however, we could say the same thing about classical information-processing accounts as they emerged. One could construct such a title for virtually any theoretical account, as terms that make subtle distinctions or highlight particular nuances are common in all theoretical domains of cognitive science (rich vocabularies appear to be a feature of any domain of human expertise; Tanaka & Taylor, 1991). For example, the classical cognitive approach is the "truth-value preserving, hierarchically organized, discretely symbolic, recursive, satisficing, structure-dependent, information-encapsulated modular" approach. Naturally, as Wagenmakers et al. demand, concrete models help anchor such terms, and it is true that the dynamics approach needs more of them (see, e.g., Kello, in press, for some recent exciting progress; we also discuss this in concluding Section 5).

So before getting lost in a wave of fancy new terms, we expend some energy in this review article discussing why they are used. To do so we describe a simplified and shortened version of a dynamic, self-organized approach to cognition. We do this to present only the most general ideas, and avoid some detailed debate that has emerged even in these areas. Readers may be surprised to discover that even amongst this tribe of cognitive science there is significant dispute, from teeth-gnashing displays that threaten abstract theoretical constructs, to pleas to remain open-minded about such constructs (for reviews see Chemero, 2009; Dale, 2008). Still, the core ideas can be laid out readily in a short section, as we attempt here. Our goal is to showcase the specific aspects of this account of cognition that seem helpful to understand conversation (see M. J. Richardson et al., in press, for a thorough presentation of both theory and methods).

Complex system. We can take "dynamics" for granted here. The dynamical systems approach takes the position that it is important to study the time-evolving properties of systems. A dynamical system is simply one that is changing in time, and can (in some way) be modeled as such (mathematically, computationally, or just conceptually). Instead, let's start with the notion of a "complex system." The phrase itself seems highly relevant to our language abilities. Carruthers (2002) refers to language as an "intersection" system, because, among other functions (e.g., complex thought), effective language use requires a wide variety of mechanisms to successfully intersect. Several mechanisms have already been implicated in theories of perspective-taking during dialogue, and other aspects of conversation. These have included social memory traces (Horton, 2005), memory for shared experiences (Galati & Brennan, 2010; Wu & Keysar, 2007), social-status adaptation (Duran & Dale, 2012), executive control (Brown-Schmidt, 2009), priming processes and alignment (Garrod & Pickering, 2004), the mirror neuron system (Gallese, 2008), forward models of social and linguistic prediction (Pickering & Garrod, in press), rich common-ground representations (Clark, 1996), socially-guided attention (Kingstone, Smilek, Ristic, Friesen & Eastwood, 2003), perceptuomotor linkages (Shockley, Richardson, & Dale, 2009), and even processes at various linguistic levels such as perception of accent (Lev-Ari & Keysar, 2010), and lexical (Brennan & Clark, 1996; Bortfeld & Brennan, 1997, Niederhoffer & Pennebaker, 2002) and syntactic choice (Branigan et al., 2000a, 2000b).

Treated as a system of intersecting mechanisms, our language capacity appears quite complicated. In the parlance of researchers who embrace dynamics and complexity science, this *complex system* is unlikely to be controlled by a central "homunculus." No theory of our language capacity has proposed such a central executive that simultaneously integrates all of these mechanisms. For example, emerging models of sentence processing imply that, even at just this processing level, central processing cannot alone account for our success, and an explanation must derive from exploring the dynamic relationship between memory retrieval, working memory, and focal attention (e.g., Lewis, Vasishth, & Van Dyke, 2006; McElree, 2006, Raczaszek-Leonardi, 2010). In the domain of motor control, where dynamical systems have been and continue to be highly influential, this is sometimes referred to as "Bernstein's problem" or the "degrees of freedom problem" (Turvey, 1990). If the components making up our language system are truly modular, there are simply too many ways in which our overall language system can change, with each mechanism flailing about unto itself unless it is somehow anchored to other processes around it. Put simply, there are too many degrees of freedom in this system for it to be managed by a single control process. These many proposed mechanisms must somehow influence each other, directly and continually, in order for language to function in naturalistic circumstances. In any one experiment we focus on a very deliberately narrowed set of controlling variables, and identify their influence on a very specific set of resultant behaviors. Such is the justifiable nature of experimental science.

Naturalistic language performance seems very unlikely to be based on a single control process. Somehow, our system integrates all of these components simultaneously. There are, at present, a limited number of theories for how this is accomplished (though see, for the closest current approximation of a grand theory, Pickering & Garrod, 2004, 2009, 2012). But this is how we wish to pose the problem: If language, in its naturalistic context, is underlain by such a wide array of processes, then these processes must somehow interact, mutually constrain each other, and act together continually to produce coherent performance. Systems that do this—that have a multitude of parts that mutually interact and constrain each other—are often referred to as "complex systems" (see Gallagher & Appenzeller, 1999, and articles therein, for discussion). The term is only meant to highlight the problem of interdependency that must be present among the system's components for it to function.

Self-organization. So if there is not a control process that "calculates the positions" of all mechanisms (working memory, social judgment, visual attention, etc.), then there must be some other means by which we can understand how they function together. A process that contrasts with the presence of a central controller is *self-organization*. Without a central control process, the mechanisms must mutually constrain each other to behave (in whole) as a stable performance. There are plenty of natural examples that are often raised to exemplify this concept (see Kauffman, 1996, for many examples). For example, the behavior of a beehive, termite or ant's nest is not controlled by a single entity, but is a large self-organizing organism unto itself (see Seeley, 2010 and M. J. Richardson et al., in press, for more discussion). The same may be true for human interaction.

This is often where things get heated between dynamical systems researchers and other cognitive scientists. Isn't working memory that executive controller? Clearly it cannot be, because there is much more being coordinated during conversation than just a handful of manipulated chunks of information (and see note about sentence processing above). What about process threading in working memory extended over time? Recent computational models of complex cognitive control may be relevant here, but even these require articulating the details of interacting components in the system (see, e.g., Salvucci & Taatgen, 2008). Similarly, what we are suggesting is that there may be "chains of influence" between processes of the cognitive system that we tend not to explore. Whatever one's favorite array of theoretical constructs or model formalisms, these processes must be working *simultaneously* to bring about stable cognitive performance, which seems especially true for face-to-face human interaction.

So how does self-organization work? Often when one is trying to convince skeptical colleagues of the *value* of these concepts, already they may point to two issues. First, perhaps this is trivial: "Okay, so you're saying to put it all together, great. That's obvious." Or, the process is far too vague to be even worthy of consideration: "Okay, so you're saying they work together, but that isn't telling us anything new, because we don't know what processes it relies on!" These critiques are entirely legitimate, but the devil is always in the details. The researchers in the dynamics crowd have identified elegant ways of understanding what goes on during the process of self-organization (even in high-level cognition; see, e.g., Dixon et al., 2010). Many of these dynamical concepts are descriptions of *form* rather than *function*. They are characterizations of what is taking place in the system as it self-organizes. Here we consider an important one. When a system of many interacting components self-organizes, it undergoes a *reduction in its degrees of freedom*.

***Synergies and the reduction of degrees of freedom*.** As noted above, a key issue raised in motor control decades ago by Nikolai Bernstein was that in order for a human being to perform any coherent action, a massive array of variables must somehow coalesce in order for it to happen. In the 1940's, Bernstein was trying to understand how motor control harnesses the high number and complexity of the components of the human body (Kelso, 2009; Bernstein, 1967; Turvey, 1990; Latash, Scholz, & Schöner, 2007). There is no way we can micro-manage each and every joint and muscle at the same time. He introduced the idea of *synergy*: a functionally driven reduction of degrees of freedom, where components do not simply align, but also complement and compensate for each other. Instead of a top-down micro-control, he hypothesized that the different components get coupled and constrain each other locally.

A classic example is Bernstein's analysis of chisel and hammer. If we want to strike a chisel with a hammer, this gives direction to and constrains the workings of our body. The exact timing and force of contraction and relaxation of all the individual muscles in our hands, fingers and arm are locally regulated to comply with that overall goal and the unfolding interaction with the environment. This intuition was tested empirically by measuring the precision of movements at all relevant joints in a blacksmith's hammering of a chisel. The variability of the trajectory of the tip of the hammer across a series of strikes turned out to be smaller than the variability of the trajectories of the individual joints on the hammering arm (Bernstein, 1967; Latash, 2008). The joints are not acting independently but correcting each other's errors at the relevant time scale, in order to preserve function, thus supporting the idea that the function itself is the coordinating principle. Importantly, when putting the hammer down and, say, grasping a cup of tea, the very same joints and muscles will flexibly combine in very different ways, thus stressing the functional, that is, task-oriented nature of the synergy.

The same is likely to be true for conversation. The array of mechanisms described above do not *merely* interact. They must have interdependencies operating in a coherent fashion that organizes the system into a lower-dimensional functional unit, and possibly a much smaller number of stable, higher-level behaviors, unexpectedly lower than what would be anticipated from the complexity of the system's composition. For example, perhaps at the coarsest level of description in human interaction, one could see stable modes in the form of *arguing* (Paxton & Dale, submitted), or *flirting* (Grammer et al., 1998), or *joint decision making* (Fusaroli & Tylén in Prep), or *giving-directions* (Cassell et al., 2007; cf. the notion of "oral genres," e.g., Busch, 2007). Beneath these coarse-level quasi-stable characterizations of interaction, we have different levels of coordination taking place. Within one interlocutor, whatever components compose a cognitive system must work together to support coherent individual performance; across two individuals a similar process of systematic reduction of degrees of freedom may organize interactions into stable modes of functioning (Shockley et al., 2009).

2.3 Summary, social modulation, and multimodal coordination

So far we have argued that the study of interaction has faced a kind of "centipede's dilemma," that the field has specialized in specific experimental paradigms, specific behavioral channels and social contexts, but that it has not integrated knowledge in a systematic way. Yet, in an important sense, the cognitive system undoubtedly performs this integration during interaction. One framework for thinking about this, which we have outlined in brief in the previous section, is to import the concepts of self-organization and synergies into this discussion. Our reasoning is that there are far too many degrees of freedom available to a dyad during conversational performance for the cognitive system to compute their activities all at once. In the following sections we offer extensive empirical review, looking to two general features of this issue of conversational performance.

Social modulation of cognitive dynamics. The first is the fundamental role of the social in human conversation. Despite the intrinsic social nature of language and conversation, there has been much debate on whether and to what extent social variables, such as the belief states and presence of another person, modulate the dynamics of performance. The self-organization of some cognitive or conversational performance is shaped by social variables in various ways. For example, the mere presence of a person will greatly change behavior of the visual attentional system (as monitored by eye-tracking). In addition, facts that one knows about a potential social partner may guide perspective-taking strategies. We explore this in review of empirical literature. The dynamical systems approach suggests the following basic insight: Significant reorganization of interactive behaviors should occur under different social contexts. This sees interaction as a process that is *organizing itself* around key variables, such as a social ones; it also means that basic theoretical accounts emphasizing of "egocentric" or "other-centric" processes may be a dialectical veil over the deeper flexibility and self-organization taking place during human interaction.

Coordination, complementarity, synergies. The second is, even if we acknowledge the importance of social variables on the way that cognitive processing unfolds, there must nevertheless be a process of coordination among the channels in conversation. For example, if I learn something new about a conversation partner, it may suddenly shift my focus both in the content of what I am saying, and how I say it. Here is where the concept of synergies becomes most important: Two people interacting in a joint task come to form their behaviors through compensatory, complementary behaviors. These behaviors influence one another locally and incrementally, making the whole conversational performance itself a kind of self-organizing synergy. We review research suggesting this is the case, and offer theoretical discussion that is meant to supplement, not replace, current discussion of coordination and alignment.

3. Cognitive dynamics under social constraints

3.1 Social modulation of the dynamics of low-level visual attention

It can be very lonely, being a participant in a cognitive psychology experiment. If the experimenter wants to study memory, language processing or decision making, a common first step is to exclude as may social factors as possible. The participant sits alone, typically, interacting with a computer, and perhaps a researcher, whom they don't know, follows a rigid script. Like friction in a high school physics problem, social context is discarded by initial simplifying assumptions in cognitive models. Both forces get in the way of more important aspects of phenomena, the argument goes. But both friction and social context are unavoidable in the world outside of the laboratory. In fact, we argue, they are both are essential to understanding many phenomena. The claim is not that what is discovered in a typical cognitive laboratory is invalid because of an absence of social context. Rather, we argue that perhaps there are interesting, dynamic interactions between cognitive processes and social context that occur all the time in the real world (Hutchins, 1995a, 1995b, 2010, 2011), but are left at the door of a cognitive laboratory.

In this section, we review more naturalistic experimental designs that have looked to how social variables constrain the dynamics of low-level visual attention, using eye movements. We survey a range of experiments that have been titrated according to the level of social context that they entail. Each measures visual attention, in the form of eye movements, to see how these varying levels of social context influence cognitive and perceptual processing. In the first, the level of social context was high, as two people interacted with each other while having a conversation or an argument. The researchers measured how their thoughts about each other drove their eye movements around a scene or an empty screen. In the second, a participant alone was asked to listen to the opinions of one person talking on a screen, but the researchers found that their eye movements were influenced by the presence and the identity of others in the background. In the lowest level of social context, the researchers gave participants an explicitly non-social task of looking at a set of pictures while sat alone. The researchers studied the effect of introducing a minimal social context to the task by telling them that another, unseen person was also looking at the same images at the same time. Across all of these levels, the researchers find a pervasive effect of social context on low-level visual processes.

When two people are engaged in a conversation with each other, they display great sensitivity to each others' thoughts and beliefs. But what happens at a lower-level of social context? Here we look at the case where the participant is merely a spectator, watching a prerecorded video of a group of people, and listening to one of them give their opinion. A standard cognitive approach might be to focus on the words of the speaker and how they are processed by the participant. But Crosby et al. (2008) looked at the relevance of the other people in the video, the silent bystanders who provided a social context.

In their experiment, participants watched a video of four people giving their views on Stanford University's admissions policies (Figure 1). All four members of the 'focus group' could be seen onscreen at the same time, in a grid arrangement of cubicles, and they all wore headphones so they could hear what each other said. At one point, the speaker in the top right corner complained that, "certain groups who come from less privileged backgrounds… get an unfair advantage". At this point, participants routinely fixated a man on the bottom row of the screen who was black. It appeared that participants were sensitive to the fact that the speakers' words, criticizing policies of affirmative action which would typically benefit black Americans, might be offensive to members of that group.

A parsimonious explanation, however, is that the participants simply noticed the ethnicity of everyone at the start, and the speaker's words simply activated a memory of one person on the screen. This memory trigger launches an eye movement, as memory representations often do (Richardson & Spivey, 2000; Richardson et al., 2009). But Crosby et al (2008) were able to rule out this 'association hypothesis.' In another condition, before the speaker began talking, an offscreen voice said that the headphones of people on the bottom row were being turned off. Then the participants saw exactly the same video. The association hypothesis predicted that since the black member of the focus group was still onscreen, he would still attract a fixation. In this case, however, participants barely looked at him when the potentially offensive remarks were made. Participants were supposed to be simply listening to the speaker's words. But these results show that they were also keeping track of the social identity of all the other people on screen, monitoring whether or not they could hear the speaker, and, presumably, anticipating how each might respond to the speakers words. In other words, for the participants, the cognitive task of processing speech was embedded in a social context.

Other paradigms too have shown an effect of 'social tuning' (Shteynberg, 2010; Shteynberg et al., 2011). In these paradigms, stimuli are explicitly identified as being relevant by other people, and as a consequence are processed selectively by individuals (He et al., 2011). Recently, researchers have adapted the Simon task and inhibition of return paradigms, splitting these cognitive tasks between pairs of participants. The behavior of the pairs is a remarkably similarity to the individuals', showing the same patterns of response interference (Knoblich et al., 2011). This work suggests that when participant act jointly with each other, in a very simple social context, they immediately represent each others' tasks and goals.

The final set of experiments attempt to reduce social context to it lowest level. Our strategy was to take a simple perceptual task that participants carry out alone or jointly with another person, and to make the difference between those conditions as small as possible (Richardson et al., 2012). Pairs of participants were sat in opposite corners of a lab room, each looking up at a screen while their gaze was tracked (see Figure 2). On each trial of the experiment, they saw four images on screen for eight seconds. Beforehand, they were either told that both they and their partner would be looking at the same images, or that they would be looking at images and their partner would be looking at symbols. Participants could not see each other and could not interact at all. Nevertheless, when they thought that their current perceptual experience was being shared with another, their eye movements were systematically changed. They looked more towards pictures with a negative valance than when looking alone. We believe that people are doing so because they each believe that the other person is looking more at the negative images. At least, when participants are told that this is a memory task and that they will score more points if they recall the same pictures as their partners, they too look more at negative images (unpublished data). When the images are replaced by album covers, people will look more at classical albums when they are looking jointly, and their partner (a confederate, in this case) walked in carrying a violin (ref unpublished data). Across these all experiments we have found that even a minimal social context—the belief that an unseen other was seeing the same stimuli—was enough to manipulate an individuals' visual processing.

From a rich, interactive conversation, to listening to one person against a backdrop of bystanders, to gazing at images alone, believing that someone else is too, social context can have a pervasive influence on visual attention. Across all of these levels, in all of these different settings, there is a consistency to how people respond to social context. It evokes coordination. People take into account each other's knowledge and visual context in order to coordinate their gaze around an empty display. They anticipate each others' responses to potentially offensive remarks. And even with the thinnest slice of social context, when there is no interaction or contact between people, they will still shift their gaze towards where they think each other is looking.

3.2 Social modulation of higher-level processes, like perspective-taking

Social variables can radically alter the dynamics of visual attention. Here we review recent research suggesting that these same dynamic changes take place in relatively higher-level cognitive processing: perspective-taking. In pragmatic models of language processing, an essential component of how people produce and comprehend language depends greatly on communicative function. The social environment thus has a central role in constraining how language is interpreted and used (Brennan, Galati, & Kuhlen, 2010; Brennan & Hanna, 2009; Brown-Schmidt & Hanna, 2011; Brown-Schmidt, & Tanenhaus, 2008; De Jaegher, Di Paolo & Gallagher, 2010; Hanna & Tanenhaus, 2004; Hanna, Tanenhaus, & Trueswell, 2003). A critical source of constraint is in the common ground that may exist between language users. Common ground corresponds to the shared characteristics derived from the local context of being present with another, such as viewing the same scene, to more global, shared histories that arise from being members of the same culture or speaking the same language. This information is brought to bear when interpreting what another says and when choosing words to speak (Clark & Krych, 2004; Clark & Wilkes-Gibbs, 1986; Fussell & Krauss, 1992; Lockridge & Brennan, 2002; Schober & Brennan, 2003).

A central question is *when* common ground information is available in language processing. As described earlier in this review, at one theoretical extreme is a view that people are primarily egocentric, and that even when common ground information is available for influencing a particular interpretation, people initially rely on their own frame of reference, or act to minimize their own difficulty in processing (Epley, Keysar, Van Boven, & Gilovich, 2004, Keysar, Barr, & Horton, 1998). Otherwise, as the argument goes, to integrate common ground early in processing would result in increased cognitive effort and processing times (Barr, 2008; Keysar, Barr, Balin, & Brauner, 2000). But such conclusions stand in contrast to other studies that show people are quite capable of making rapid social judgments based on briefly presented sources of social information, such as dispositional expressions (Ambady, Bernieri, & Richeson, 2000) or gaze direction (Hanna & Brennan, 2007; Tuefel, Fletcher, & Davis, 2010). This social information can also extend to belief attributions about another, such as another's needs, characteristics, or limitations that are initially present in an interaction (Bortfeld & Brennan, 1997), or are emergent factors (Horton & Gerrig, 2002, 2005). Integrating common ground information does not necessarily have to be a cognitively complex process, as simple attributes of another can immediately constrain what and how something is interpreted. Moreover, such integration is commensurate with "incremental models" of language processing. As words are encountered in a sentence, new evidence is provided for the committal, or abandonment, of a particular interpretation. As a sentence unfolds, multiple interpretations are simultaneously activated and competing for expression, with accruing evidence constraining possible interpretations (Seidenberg & MacDonald, 1999). Based on this account, common ground information is tantamount to just another source of potential constraint.

From a "traditional" dynamical systems perspective, what constitutes a relevant constraint is similar, but is more connected to what can be directly perceived from within an interactive, social environment (Marsh, Richardson, Baron, & Schmidt, 2006; Marsh, Richardson, & Schmidt, 2009; Richardson, Marsh, & Schmidt, 2005). Contrary to egocentric accounts of processing, there is no intermediary "representational" stage where another's intentions are first calculated and then acted upon. Rather, the behavior and actions of another hold immediate sway on how people respond to each other. Such direct couplings are possible through the interactive context, by which individuals' processing capabilities are reshaped by the presence and actions of social partners (Ramenzoni, Davis, Riley, Shockley, & Baker, 2011; Riley, Richardson, Shockley, & Ramenzoni, 2011). Such connectivity allows for nimble social coordination that cannot be reduced to individual-level contributions, but instead must be evaluated on the basis of the social unit, where the emergence of meaning is reciprocally caused and maintained by social partners during interaction (Marsh et al., 2009).

On the face of it, this traditional dynamical systems approach should be closely aligned with language theories that allow for common ground constraints to have immediate influence on moment-by-moment linguistic processing. Yet, what constitutes common ground often corresponds to simple inferential states based on what another knows or believes. Such inferences are not easily integrated with a dynamical systems approach, where interpersonal coordination is driven by the actions, or possibilities for action, that are expressed and integrated by physically co-situated agents. Rather, what is found in the social environment are opportunities for merging individual-level perceptuomotor systems into a collective system, with evidence taken from joint action tasks where people rapidly converge on patterns of coordinated, synchronous movements (Knoblich, Butterfill, & Sebanz, 2011; Knoblich & Sebanz, 2006). Put simply, common ground invokes the cognitive, and many dynamical systems theorists avoid this.

So if social interaction is grounded exclusively in coupled perceptuomotor systems, there appears to be little room for the role of informationally-grounded sources in shaping language processes. To push the boundaries of the dynamical approach, explanations need to go beyond motor behaviors alone, to more abstract properties of the social environment (Chemero, 2009). Such attributional properties do not necessarily require elaborate mental operations or representation, but instead can be thought of as spontaneously elicited opportunities for social responding, embodied from past histories of social interaction. As Schmidt (2007) describes, these experiences elicit tendencies to respond in socially-appropriate ways, and are sustained by cultural expectations and reinforced by the immediate social context. This behavior is inextricably defined by the relationship between social partners, and "affords" opportunities for responding, even during language comprehension.

One of the simplest belief attributions in a language task is whether a communicative partner is an intentional agent (Gallagher, Jack, Roepstorff, & Frith, 2002). Although most interactions provide situated cues to determine veridicality, there are scenarios, such as with computer-mediated interactions, where people may be uncertain whether they are interacting with someone real or simulated. When people are told beforehand the true nature of personhood, their response orientations change in systematic ways, even though the actions of the other remain the same (Nass, Fogg, & Moon, 1996). For example, when a partner is thought to be simulated, people are more likely to use language that is less complex, presumably because of perceived communicative limitations of an artificial intelligence (Branigan, Pickering, Pearson, McLean & Brown, 2011). Thus, simple belief attributions have the capacity to guide various modes of responding (Clark & Wilkes-Gibbs, 1992). A challenge for dynamical systems is to explain how such behavior arises through an emergent, self-organized process in which informational couplings within a social environment produce complex but systematic behaviors (also see Di Paolo & De Jaegher, 2012; M. J. Richardson, Marsh, & Schmidt, 2010).

3.3 Perspective-taking as self-organization under social constraint

Duran and Dale (in press) present one such attempt by employing a dynamical simulation of response resolution in a task where participants' beliefs were central in disambiguating utterances spoken by another. In this response data, participants and a simulated agent were "connected" within a virtual environment through an elaborate ruse in which the participant was unaware of whether their partner was actually simulated (Duran, Dale, & Kreuz, 2011). This omission allowed participants to form their own impressions about the reality of their partner. For each trial, the task proceeded with the simulated agent instructing the participants to select one of two objects in the shared environment (see Figure 3). The position of the partner shifted from trial-to-trial, sometimes creating a situation in which the intended referent was ambiguous. When both participants were in the same location, perspectives were ostensibly shared, and an instruction such as, "Grab the object on the right," permitted straightforward identification. However, when participants' positions were across from each other, the same instruction created an ambiguous referent as to whose "right" was the basis for interpretation. That is, instruction-receivers could either select the object on their partner's right, thereby taking into consideration the perspective of the other, or they could select the folder on their own right, in what would be an egocentric interpretation.

Despite an informationally-situated social context, simple belief attributions, in the form of whether a communicative partner was thought to be real or simulated, the researchers were able to shape perspective-taking strategies across participants. Specifically, when people believed that their partner was simulated, other-centric responding was facilitated in three key ways: (a) the likelihood for other-centrism increased, (b) responses were faster within and across trials, and (c) decreased competition from the egocentric response option. Such behavioral change was also seen in a dynamical systems simulation that treated attributional factors as a control parameter within a low-dimensional attractor landscape of partly stable perspective-taking modes. In this simulation, when a person takes the perspective of another, this response stabilizes through the graded accumulation and competition of both other- and ego-centric factors. Thus, contrary to existing models of language processing where these factors are mostly independent components, here they are explicitly allowed to interact from the very start of processing to influence behavior over time (e.g., multipotentiality).

To conduct the simulation, Duran and Dale (in press) borrowed from the Haken-Kelso-Bunz (HKB) model that was originally developed to capture the relative coordination of bimanual motor movements in time and space (Kelso, 1981; Kelso, 1995; see Figure 4). This model has been extended to a variety of domains, revealing widespread commonalities between perceptual, cognitive, and motor systems (e.g., Engstrom, Kelso, & Holroyd, 1996; Frank, Richardson, Lopresti-Goodman, & Turvey, 2009; Tuller, Case, Ding, & Kelso, 1994; van Rooij, Bongers, & Haselager, 2002; see Schmidt & Turvey, 1995, Chemero, 2009 for reviews). It draws from core principles of bistable dynamics to allow complex behaviors to self-organize over time, with responses unfolding within a low-dimensional attractor landscape. Thus, perspective-taking during communication, much like in the previous research, could be described as following coordinative dynamics similar to those observed in perceptuomotor coupling.

In the current instantiation of the model, a control parameter in a potential function is set to initiate bistable attractor basins of other-centric or egocentric interpretations. These basins are a reduction of system complexity to a quantifiable and transparent outcome variable of the two perspective types. In other words, perspective-taking is characterized as in a system of substantially reduced degrees of freedom, a "lower-dimensional" cognitive space in which choices are made. The particular shape of each basin corresponds to the likelihood and speed in which the system can settle into a particular response, with deeper and steeper basins indicating a stronger pull and therefore more rapid stabilization. During the time course of a single trial, "settling" occurs through nonlinear competition between landscape shape, initial conditions (i.e., starting position in the landscape), and a subtle noise impulse. When a response threshold is met, the control parameter is adjusted, and a new trial is allowed to run. This is analogous to updating a belief about one's partner being real or simulated at the end of each trial, with beliefs becoming stronger and more stable over time. In doing so, the global characteristics of response choice stability are captured, as well as the competition effects that influence the moment-by-moment processes involved in response execution. By capturing the response dynamics also exhibited by human participants, simple social constraints, in the form of belief attributions, are essential pieces of information that bias a system's "perspectival" landscape and thus its eventual behavioral strategy (see example in Figure 5).

Perspective-taking in communication is fundamental to how language is used and understood. This was certainly evident in the Duran et al. (2011) study. Somewhat counterintuitively, participants were more likely to consider the other's perspective when they believed the other to be simulated. This result makes sense when communication is viewed as a collaborative process between language users. The goal of communication is to maximize mutual understanding, and when one partner is unable or hard-pressed to do so, the other will compensate by putting in increased effort, such as engaging in other-centric perspective-taking behavior (Clark, 1996; Clark & Krych, 2004; Clark & Wilkes-Gibbs, 1986, Goodwin, 2003). This tendency emerges though past histories of social interaction where people actively attempt to establish mutual understanding. Attributions about others' abilities to cooperate are eventually embodied by language users and brought to bear in responding, even in simple communicative scenarios such as the one described above (also see Schober 1993, 1995). When participants think they are interacting with a simulation, where cooperation by the other is not even possible, the "afforded" response is to assume the other's perspective in interpreting their ambiguous instructions.

Such spontaneous perspective-taking occurs despite enacting increased cognitive demands. However, demand is minimized in communication where assessments about a partner can be reduced to simple alternatives, such as whether a conversational partner is very young or not, or has a language disorder that changes the goals of mutual understanding (Newman-Norlund et al., 2009; Perkins & Milroy, 1997). Of course, reciprocal and emergent constraints occur during the course of an interaction that subsumes any individual-level sources of difficulty. The dyad operates as a unit that collaboratively minimizes processing load, with success depending on the level of coordination shared between the language users (Fusaroli et al., 2012; Louwerse et al., in press).

4. Coordination, complementarity, and interactive performance

Language is increasingly acknowledged as a social coordination device, a way of accomplishing otherwise difficult or impossible coordination of actions and cognitive processes (H.H. Clark, 1996; Fowler, Richardson, Marsh, & Shockley, 2008; Fusaroli, Gangopadhyay, & Tylén, in review; Fusaroli & Tylén, 2012; Galantucci, 2009; Galantucci & Sebanz, 2009; Hasson, Ghazanfar, Galantucci, Garrod, & Keysers, 2012; Hutchins & Johnson, 2009; Louwerse, Dale, Bard, & Jeuniaux, 2012; Pickering & Garrod, 2004; Tylén, Fusaroli, Bundgaard, & Østergaard, accepted; Tylén, Weed, Wallentin, Roepstorff, & Frith, 2010). Through language we can easily entertain a friend while waiting for her bus to arrive, exchange words with a stranger in an elevator, coordinate in carrying a heavy piano down a flight of stairs, negotiate the price of an apartment, share information and make joint decisions. However, as described in the introduction to this review, coordinating via language is a complex business. A great number of studies have been dedicated to unveiling the crucial subtleties in coordinating not only topics, lexical choices and syntax, but also gestures, gaze management, head movements and postural sways, which underlie conversations (Goodwin, 2000, 2011; Louwerse, et al., 2012). Such a multimodal richness seems to imply serious uncertainty (Jaeger, 2010) and cognitive load (Garrod & Pickering, 2009) for the participants: How does a conversant choose between all the possible linguistic and non-linguistic behaviors on so many different channels at once? How does a person focus his or her attention and interact in a meaningful way? In other words, how can interlocutors seemingly effortlessly orchestrate all these dimensions (each level, presumably, with its own numerous degrees of freedom) in tight intra- and interpersonal coordination?

As a reaction to computationally heavy models of conversations, requiring theory of mind and full accommodation of models of the other, there has been a strong focus on low-demanding bottom-up models of linguistic coordination, such as the model of interactive linguistic alignment (Pickering & Garrod, 2004). After a brief presentation of this model, we will argue that it should be integrated and complemented in the larger model of interpersonal synergies, presenting evidence supporting this and a study aimed at comparing alignment and synergies.

4.1 Alignment

One intuitive way of reducing the complexity in interpersonal interactions is to diminish the range of possible behaviors via a progressive adaptation to each other. By becoming increasingly similar the interlocutors greatly simplify the cognitive load needed to interact with the other. Indeed, there is strong evidence for behavioral mimicry (Chartrand & van Baaren, 2009) and interactive linguistic alignment (Pickering & Ferreira, 2008; Pickering & Garrod, 2004). Interacting human beings have been observed to mimic each other's posture, gestures and other behaviors (Chartrand & van Baaren, 2009). A prototypical example of an experimental investigation of this kind of human unconscious mimicry is the 'Chameleon effect' (Chartrand & Bargh, 1999; Dijksterhuis & Bargh, 2001). In this research, participants interacted with an unknown confederate in two consecutive picture-describing sessions. In one session, the confederate either rubbed her face or shook her foot while describing the pictures with the participants, while the second confederate performed the behavior that the first confederate did not. The behavior of the participants, 'secretly' recorded on videotape, showed that participants shook their foot more in the presence of the foot-shaking confederate, and rubbed their faces more in the presence of the face-rubbing confederate. Debriefing indicated that participants were unaware of their mimicry. Analogously, facial expressions, gestures and yawns have been observed to spread across interlocutors and around a room (Louwerse, et al., 2012; Platek, 2010).

Pickering and Garrod have argued that mimicry is commonly co-opted in linguistic interactions through what is called "interactive linguistic alignment": interlocutors tend to imitate each other's choice of linguistic forms. Participants primed with a specific syntactic structure are more likely to produce new sentences employing the same syntactic structure under circumstances in which alternative non-syntactic explanations could be excluded (Bock, 1986; Branigan, Pickering, & Cleland, 2000; Gries, 2005; Hartsuiker & Westenberg, 2000; Levelt & Kelter, 1982; Pickering & Branigan, 1999; Smith & Wheeldon, 2001; Szmrecsanyi, 2005, 2006). Analogously, topics (Angus, Smith, & Wiles, 2012; Angus, Watson, Smith, Gallois, & Wiles, 2012) and lexical choices (Brennan & Clark, 1996; H. H. Clark & Wilkes-Gibbs, 1986; Garrod & Anderson, 1987; Garrod & Clark, 1993; Garrod & Doherty, 1994; Orsucci, Giuliani, & Webber, 2006; Orsucci, Giuliani, & Zbilut, 2004; Orsucci, Walter, Giuliani, Webber, & Zbilut, 1997; Wilkes-Gibbs & Clark, 1992) tend to be imitated across interlocutors. Linguistic alignment can also be found at more subtle levels of linguistic coordination: Interlocutors align accent and speech rate (Giles, Coupland, & Coupland, 1991). More recently, a lot of effort has also been put in showing that the organization of pauses in and between interlocutors' speech and their average pitch, intensity and voice quality tend to become similar over time (C. De Looze & Rauzy, 2012; Kousidis & Dorran, 2009; Lee et al., 2010; Lelong & Bailly, 2011; Levitan & Hirschberg, 2011; Nishimura, Kitaoka, & Nakagawa, 2008; Pardo, Gibbons, Suppes, & Krauss, 2011; Truong & Heylen, 2012; Vaughan, 2011). In a single conversation many of these channels will be aligned, as recently shown in a massive study by Louwerse and colleagues (Louwerse, et al., 2012). These channels have been argued not to be independent. On the contrary, aligning on one channel in many cases seems to facilitate alignment on others. For instance syntactic priming is enhanced when the same lexical items or even just semantically related ones are also repeated (Branigan, Pickering, & Cleland, 2000; Branigan, Pickering, Stewart, & McLean, 2000; Cleland & Pickering, 2003).

Several mechanisms have been proposed to underlie these phenomena: Most researchers seem to agree on an unconscious priming mechanisms, a "perception-action link" (Chartrand & Bargh, 1999; Dijksterhuis & Bargh, 2001), or "structural priming" (Pickering & Ferreira, 2008), in other words the overlapping between mechanisms involved in perceiving a behavior and producing it, which implies that by perceiving a behavior the participant pre-activates the production of the same. Other authors prefer to focus on the conscious aspects of alignment, where interlocutors try to take each other into account, developing conceptual pacts on which words to use (Brennan, Galati, & Kuhlen, 2010; H.H. Clark & Brennan, 1991).

Whatever the mechanisms at work, alignment within and across modalities is a very effective way to reduce degrees of freedom—namely, the possible behaviors from which to choose. Not only the other's behavior can be used as guide in how to behave, but also the repertoire of possible behaviors is reduced over time. The mechanism of alignment might be differently motivated, but it is sensible to argue that once it is established, it plays an important role, making linguistic interactions more manageable. However, a few problems arise when we take it at face value as the fundamental motor of linguistic coordination. A few studies are pointing out that not all conversations contain the same amount of linguistic alignment (Healey, Howes, & Purver, 2010; Reitter, Moore, & Keller, 2006), and that coordination might not rely on alignment across neighboring speech turns, but on the contrary across many speech turns, thus escaping the tight temporal constraints of automatic priming (Reitter & Moore, 2007). At a more intuitive level, a conversation constituted exclusively of reciprocal repetitions does not strike as a very productive one.

4.2 An alternative model: interpersonal synergies

As described in the introduction, the reduction of degrees of freedom is not a new problem. Bernstein (1967) proposed that functional units of motor control are established through mutual constraint among the parts of the body and motor control system, effectively reducing the degrees of freedom of the system into "synergies." Recently, Ramenzoni and colleagues (Ramenzoni, Davis, Riley, Shockley, & Baker, 2011; Ramenzoni, Riley, Shockley, & Baker, 2012; Riley, Richardson, Shockley, & Ramenzoni, 2011) have been exploring interpersonal motor synergies. They showed that in joint actions participants increasingly coordinate hands, forearms, and torsos, forming reciprocally compensating synergies spanning across individuals. While studies of interpersonal motor coordination are not rare (Marsh, Richardson, & Schmidt, 2009; Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007; R. C. Schmidt, Bienvenu, Fitzpatrick, & Amazeen, 1998; R.C. Schmidt & Richardson, 2008), only very recently this approach has been applied to linguistic coordination. The rest of this section will argue that linguistic coordination is achieved through interpersonal synergies, that is, through functionally driven reduction of the degrees of freedom involved in the interaction. This approach does not dispense with alignment, but introduces additional mechanisms—complementarity and interactional patterns—and integrates alignment with this dynamical-inspired perspective.

4.3 Complementarity

Several studies have pointed out that interlocutors strive to complement each other's behavior in order to develop a structured conversation. For example, turn-taking seems the most elementary, and very important, example of complementarity: A remarkable—and seemingly universal (Sidnell & Enfield, 2012; Stivers et al., 2009)—ability of humans to *not* do the same thing at the same time, i.e., stay quiet when the other speaks. Simultaneous starts are reported to be surprisingly rare in dyadic conversations (Jefferson, 1988), even if over 50% of the pauses between interlocutors are below the usual threshold for reactions (300 ms). Wilson & Wilson (2005) have been developing a model of turn-taking that explains this fine-tuned complementarity: The beginning of an interaction sets up an oscillator in each of the interlocutors' cognitive systems establishing a shared frequency of speech rate (see also Buder & Eriksson, 1999). This cyclic pattern governs the potential for initiating speech at any given instant for both interlocutors. The interlocutors, in other words, have to keep the same pace (alignment). However, if the oscillators were simply entrained in phase, simultaneous starts would be frequent. Therefore the oscillators must be entrained in *antiphase*, giving the participants both a common rhythm, constituted by speech rate and length of comfortable pauses, and complementarity—readiness to take the floor must be opposite at any given moment for speaker and hearer. This ability seems to appear at a very early developmental stage (Gratier & Devouche, 2011; Murray & Trevarthen, 1985; Nadel, Carchon, Kervella, Marcelli, & Réserbat-Plantey, 1999; Spurrett & Cowley, 2004; Warlaumont, 2012).

Recent work on conversations involving patients with speech impairment shows the importance of complementarity. Expert interlocutors—e.g. family members—tend to engage compensatory procedures to keep the conversation fluent despite the impairment (Dressler, Buder, & Cannito, 2009; Goodwin, 2003, 2011; Wilkinson, Beeke, & Maxim, 2003). For example, Goodwin reports on Chil, who, after having suffered a severe stroke, can only speak three words: 'yes', 'no' and 'and.' Despite this clear impairment, Chil is able to engage in complex conversations by coordinating other people's utterances. Chil thus relies on different types of reciprocal compensatory moves to restore the dialogue: On the one hand, interlocutors have to actively produce utterances completing and supporting Chil's conversational moves. On the other hand, Chil's three words are relational ones: They do not communicate much on their own, but make sense only in a conversational situation. Together with a host of non-verbal means such as facial expressions and gesture, Chil employs his minimal vocabulary to couple with the other interlocutors' communicative activity. Relying on three words, he is able to coordinate, support, supplement, and sometimes reject his interlocutors' utterances (Goodwin, 2011). Similarly, Dressler and colleagues (Dressler, et al., 2009) have explored prosodic patterns in conversations with aphasic patients. They report that conversation with familiar interlocutors displays overall prosodic rhythms which are much more fluent and regular than conversations with unfamiliar interlocutors.

4.4 Interactional patterns

Beyond this rhythm of interaction, conversation analysis has persuasively shown how speech turns are often organized in functionally structured sequences of turns, such as adjacency pairs: Questions are ordinarily responded to with an answer, not with another question; offers and invitations are ordinarily followed by acceptances or declinations, etc. (Schegloff, 1986). Turns and adjacency pairs are themselves not free-floating entities, but often fulfill a role in larger interactional patterns, locally unfolding routines which scaffold and constrain the possibilities of actions and interpretation in joint activities (H.H. Clark, 1996; Levinson, 1983). Interactional patterns are often conceived of as normative static phenomena already shared—or assumed to be shared—by interlocutors (Sacks, Schegloff,

& Jefferson, 1974; Schank & Abelson, 1977). The synergy approach, however, implies that these elements are part of a dynamic, context-sensitive interaction. Interactional patterns vary in formality and flexibility from free and relatively unconstrained conversation over the morning coffee, to tightly structured and sometimes even explicitly codified task oriented conversations (Hutchins, 1995a, 1995b; Perry, 2010). Interactional patterns work to reduce the overall degrees of freedom of the system in a functionally driven way and enable a smoother flow of the interaction.

A number of recent studies indirectly show that *ad hoc* interactional patterns emerge and are maintained in task-related interactions. In a version of 'the maze game' (Healey & Mills, 2006; Mills & Gregoromichelaki, 2010), it was observed that, over the course of 12 games, participants radically structured and shortened their linguistic exchanges from more than 150 turns to brief and efficient exchanges. Through a shared history of interaction, the structure of their interaction is stabilized. This enabled participants to smoothly produce and interpret highly elliptical, and fragmentary utterances without much negotiation or clarification. Extending this work, Mills (2011) systematically investigated how these interaction patterns emerge and spread in a small speech community. Each participant played a number of games with shifting partners within a 'community'. Then, in a critical test trial, half of the participants were paired with a member from another community. This perturbation seriously disrupted the interaction in the affected groups. Participants were found to edit their utterances to a much higher degree, were observed to explicitly acknowledge each other's utterances more often and overall performed less accurately. The findings suggest that interactional patterns emerge from a shared history of interaction and come to implicitly constrain the degrees of freedom of the interlocutors, diminishing ambiguity and supporting a smoother and more effective flow of the coordination (for a more comprehensive discussion of these issues, cf. Mills (in press)).

4.5 Alignment and interactional patterns: testing the synergy model

The review so far suggests that complementarity, in the form of systematized patterns of interaction between two people, is a crucial component of human interaction. Two recent studies based on the same experimental design have tried to test implications of the model (Bahrami et al., 2010; Fusaroli et al., 2012; Tylén & Fusaroli, in preparation). In the experiment, pairs of participants were instructed to individually indicate in which of two brief visual displays they had just been shown a contrast oddball. If their individual decisions diverged, they were prompted to discuss and reach a joint decision. This paradigm generated a corpus of task-oriented conversations—which emphasizes the development of interactional patterns to quickly solve the repeated tasks—as well as an accurate measure of cooperative performance—to assess the efficacy of linguistic coordination. In order for a pair to achieve a cooperative benefit, that is, to perform better than the best of the individuals, they had to find ways of assessing and comparing their individual levels of confidence so as to choose, on a trial-by-trial basis, the decision of the more confident participant. In other words, they had to develop an interactional pattern for accurately expressing confidence and smoothly taking joint decisions relying on that.

The first study investigated lexical alignment. As described earlier, this notion of alignment predicts that the more people use the same words, the better they will perform ("indiscriminate" lexical alignment). By contrast, a model of coordination as *synergy* would predict that the alignment of confidence expressions only—serving the interaction's goals of sharing confidence in order to make a joint decision—would correlate with performance. The analysis did reveal prominent "indiscriminate" alignment in all pairs: Interlocutors displayed a high probability of picking up and employing words used by the other in the previous interaction. However, the more a dyad indiscriminately repeated each other's words the *lower* the collective benefit they gained from cooperation. Automatic linguistic alignment seemed to be deleterious to coordination on the task. In contrast, the participants' reciprocal, selective adaptation to vocabularies of expressing confidence (task-motivated selective alignment), turned out to correlate positively with the collective benefit gained from linguistic coordination (see Figure 7).

The second study (Fusaroli & Tylén, in preparation) more systematically compared linguistic repetitions at three levels. First repetition of triplets of phonemes, second the repetition of patterns of pitch, and finally repetition of patterns of speech pause sequences. A model of coordination as alignment would predict the structure of repetitions across subjects to correlate with performance, while a model of coordination as synergy would predict the structure of repetitions at the interaction level—that is, not discriminating between interlocutors—to be correlated with performance. In other words, a synergy model would predict that the relevant coordination happens in interactional patterns where it does not matter which interlocutor shares confidence and which makes a decision, as long as somebody fills those roles in each joint decision.

Employing a combination of information theory and recurrence plots (Marwan, Carmen Romano, Thiel, & Kurths, 2007), the authors quantified these repetitions both across interlocutors and in the overall interaction. The results show that the relevant coordination happens at the level of interactional patterns, but not simply across interlocutors: The more the interlocutors develop a regular pattern of lexical choices, pitch and speech pause sequences which repeats across joint decisions, no matter who is producing its different parts, the better they perform. On the contrary, indices of repetitions across interlocutors did not correlate with performance.

5. Conclusion: time for more models

5.1 Summary

We have offered some discussion and review of how interaction can be understood as a process of self-organization. First, we showed that social variables, when perceived, and when taking particular forms, can fundamentally change the cognitive processes and behaviors of a conversation partner. These emerging patterns can be described in the form of "phase transition," where lower-level systems become organized differently in a manner that is shaped by these social variables. But how do these lower-level processes constrain each other, and act together? Akin to the centipede's dilemma, rather than understanding the interaction "leg by leg by leg," we entertained the notion of a synergy between interacting human beings: The behaviors—turn-taking and rhythms, use of particular words, emergence of adjacency pairs, and so on—can be seen as an array of levels that are mutually constraining, and dynamically evolving, as two people come to form, in an important way, a "unit of analysis," and the interaction itself a stable, if temporary, synergy itself.

5.2 Moving forward: models of these processes

The two key features we have articulated mostly describe the *form* of interaction, rather than the underlying mechanisms that give way to it. This is an issue raised often in discussion and critique of dynamical systems approaches to cognition (e.g., Bechtel, 1997; Dale, 2010; Eliasmith, 1996, 2012; Wagenmakers et al., 2012). In fact, we described that one exciting aspect of growing approaches to social interaction is that these approaches factor in mechanism. One critique of the current review could be that we have simply advocated for a wholesale integration of as much as can be gleaned about mechanism—and this doesn't really tell us much about mechanism. We have advocated instead for conceptualizing human interaction as a system that self-organizes and adapts to particular contexts, such as social variables, and organizes itself through evolving local interactions, such as in incremental contributions to a dialogue, including even non-verbal channels, like winks and nods. These are important critiques, and they should be addressed directly. So we end this paper with a brief review of some modeling endeavors that will help to guide integration of many channels, helping to solve the centipede's dilemma.

5.3 Surface network analysis, and mechanistic models

One way to get at the synergies directly is to carry out integrative analysis of "multimodal" (multi-person, multi-behavior, multi-level) corpora. The past decade has seen a growing agenda to build large-scale corpora of human interactions, capturing a variety of interpersonal behaviors, linguistic contributions, contextual variables, and so on (e.g., Carletta et al., 2006; Carletta, 2007; Calhoun et al., 2010; Fusaroli et al., 2012; Louwerse et al., in press; see Galatica-Perez, 2009, for some review). After such collection of data, researchers often go about identifying the relationship among particular variables, such as gestures and group collaboration or prosodic contours in particular discourse situations (e.g., Charfuelen et al., 2010). These agendas are important for understanding interactions at particular levels of analysis—the manner in which gestures are deployed, and in what context, and how prosody may index particular modes of interaction.

The argument we have made is that it is a nontrivial mission, both methodologically and theoretically, to discover the manner in which these multiple behaviors, and cognitive processes, are integrated during ongoing interaction. One way to do this "at the surface" is to translate corpora into a form that allows the analysis of the temporal relationship between behaviors. In other words, different behaviors such as nodding, gestures, use of particular words, and so on, can be rendered into analyzable *time series.* This was done by Louwerse et al. (in press), who, at a rate of 250ms, tracked patterns of synchrony between interlocutors in a direction-giving dialogue (e.g., participants tended to laugh and smile together, nod one after the other at a particular timescale, etc.). Extraction of time series would permit an exploration of the dynamic interaction between different channels, and between people, and exploring how these change over the course of an interaction. One way to do this is to project the channels into a network structure, with nodes representing the behaviors, and edges representing their relationship (e.g., strength of connection).

Consider the following hypothetical research scenario: investigating bouts of human interaction along a set of 4 behaviors (A, B, C, and D), and measuring these behavioral channels at 250ms intervals. Such a hypothetical data set is presented in the Figure 8. Various circumstances may arise during interaction. The channels may exhibit only weak coincidental structure, with each "degree of freedom" of this system being one of these channels. However, if systems exhibited pure "synchrony," then behavioral channels across individuals serve to constrain each other. So instead of 2x4=8 degrees of freedom in the interaction, we have only 4, since each channel serves to constrain that in the other person. If a process of alignment were to cascade across levels, as predicted in Pickering and Garrod (2004) for example, we would have a continued shrinking of the degrees of freedom. As displayed beneath the middle panel in Figure 9, that saturation would result in coupling across behaviors. We would effectively have only a single degree of freedom, as behaviors fluctuate now all together as one unit.

We know through extensive explorations described earlier in this paper that speech, gesture, and other features of interaction will exhibit coordination (e.g., see Louwerse et al., 2009). This can be identified as clusters in the network that become tightly entrained over time. These are portrayed in the rightmost panel in Figure 9. The degrees of freedom relevant to this interaction are now constrained by the number of unconnected subgraphs (in Figure 9, top right right). As an interaction changes across time, the network structure may change, but the degrees of freedom may stay the same (see lower right figure, "Transition"). We could imagine this sort of thing occurring during face-to-face interaction. Imagine two students discussing lecture briefly, which one of them missed. In this bout of interaction, nodding and gesturing and speaking may have a characteristic temporal interaction. However, if this part of the conversation ended, and one asks the other for directions, suddenly their gaze and gesture may take on that "clamped degrees of freedom" property, while others may change.

This network analysis approach may serve as a powerful means of visualizing and quantifying the "surface configurations" of an interaction. The authors are engaged in some early work exploring this possibility (Dale & Louwerse, 2012; Duran and Dale, in preparation; Fusaroli et al., in preparation; Paxton & Dale, 2012; see also related work in Kopp, 2010; Bergmann & Kopp, 2009). There are considerable details in need of investigation if this agenda were to be carried out in the naturalistic context (here we have only sketched this hypothetically using random point processes). For example, what temporal functions best characterize the linking between channels? Gesture and nodding (for example) have a different timescale from, one would suspect, explanation or querying, referred to as "dialog moves." Another issue is what the appropriate measures are to determine that these channels are indeed coupled. Methods such as vector autogression (Dixon & Stephen, 2012), Bayes nets (Bergmann & Kopp, 2009), and related techniques (Shalizi et al., 2007) may do a better job at capturing the cross-covariation among so many channels.

This "surface network" analysis may be a useful way of proceeding, to extract the "hidden" degrees of freedom that are guiding the behavioral structure of a conversation. Still, it is important to note that, in some ways, this research agenda is already unfolding in some prominent projects. For example, consider the case of the Augmented Multi-party Interaction corpus (AMI; Carletta et al., 2007), which tracked several sham meetings at many behavioral levels (not unlike the above simulation). Computer vision and speech automation techniques permitted the extraction of a wide range of behaviors among several people while discussing topics such as designing artifact prototypes. From extraction of multiple channels, researchers have been developing automated techniques for capturing argumentation (Hakkani-Tur, 2009), the structure of the meeting (Murray, Renals, Carletta, & Moore, 2006), the emergence of particular emotions (Reidsma, Heylen, & Ordelman, 2006), and so on. This work is beginning to leverage, in essence, the probabilistic relationship among multiple channels during interaction. Perhaps these probabilistic models will help solve the centipede's dilemma.

Once we have this surface structure, and potentially even an estimate of the number of "freely moving parts" of an interaction, there is still the open question of the specific *cognitive processes* that underlie the control of these degrees of freedom. Lack of space precludes a detailed review, but there are many exciting possibilities that may be pursued. In models of reading and sentence processing, Miyake and colleagues have used latent-variable modeling to relate cognitive processing tendencies with individual differences measures like the Wisconsin card sort and dual-task batteries (Miyake et al., 2000). This individual differences approach, through statistical modeling, may be useful in the interactive context to identify the cognitive constraints on specific forms of interaction. This agenda has begun in the work of Brown-Schmidt (e.g., 2009) who has identified the role of executive control in predicting the extent to which one person is likely to integrate knowledge of another during interaction. These are statistical models, but there are also some computational possibilities. For example, rational models and adaptive control theory have allowed some researchers to tap into the dynamic relationship among hypothesized cognitive processes and constraints to capture, for example, reading and other language comprehension (Lewis et al., 2006; Bicknell & Levy, 2010; Smith & Levy, 2008). It may be possible to induce something akin to a Hidden Markov process, beneath the "surface structure" we articulated above, and specifying in greater detail the cognitive interactions taking place that control that surface behavior. As noted earlier, mathematical development of these models in the motor control literature has reached a very sophisticated level (e.g., Todorov & Jordan, 2002).

5.5 Conclusion

We have not advocated for an approach that supplants existing theoretical accounts of interaction. The role of memory (Horton, 2005), executive control (Brown-Schmidt, 2009), alignment and priming processes (Pickering & Garrod, 2004), coordination and adaptation (Brennan et al., 2010; Schober & Brennan, 2003), perceptuomotor coupling (Richardson et al., 2009; Shockley et al., 2003), accessibility accounts (Barr & Keysar, 2002), and so on, are all crucial for accounting for interacting persons. Though the authors of this article may wrestle with each other on this grander point, it seems instructive to proclaim that *all of these theories* have central contributions to play in accounting for interaction. One major, and very simple, reason for this could be advanced in the following way: These theories have been usefully deployed in specific experimental contexts investigated by the researchers who have advocated for them. This means they have strong empirical backing in some subset of human interactive situations; a corollary of this is that they are *predictive* of human interaction in similar situations.

We have argued that it is time to integrate, to go beyond the centipede's dilemma, and gain an understanding of the manner in which processes coordinate and act together. Motivated by basic concepts of self-organization and synergy (Section 2), we described a series of experiments that show the flexibility of human interaction. Under different social situations, low- and high-level cognitive processes flexibly adapt (Section 3). Under different task conditions, the dyad self-organizes through local exchanges, incrementally emerging, that develop whole new "synergies" (Section 4). By exploring the structure and underlying control mechanisms (Section 5), perhaps an integration of these theories will be possible. We haven't done this here, but we have provided some clues that seem useful to us. We hope some readers feel the same way.

References

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream, 201–271.

Amorim, M. A. (2003). " What is my avatar seeing?": The coordination of" out-of-body" and" embodied" perspectives for scene recognition across views, 157–199.

Angus, D., Smith, A., & Wiles, J. (2012). Human Communication as Coupled Time Series: Quantifying Multi-participant Recurrence. IEEE Transactions on Audio, Speech and Language Processing, 20, 1795–1807.

Angus, D., Watson, B., Smith, A., Gallois, C., & Wiles, J. (2012). Visualising conversation structure across time: insights into effective doctor-patient consultations. PloS one, 7, e38014. doi:10.1371/journal.pone.0038014

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. Science, 329, 1081–5. doi:10.1126/science.1185718

Bailly, G., & Lelong, A. (2010). Speech dominoes and phonetic convergence (pp. 1153–1156). Presented at the Interspeech 2010.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. Journal of Memory and Language, 42(1), 1–22.

Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground, 18–40.

Barr, D. J., & Keysar, B. (2002). Anchoring comprehension in linguistic precedents. Journal of Memory and

Language, 46(2), 391–418.

Barron, J. M., Bishop, J., & Dunkelberg, W. C. (1985). Employer search: The interviewing and hiring of new employees. The Review of Economics and Statistics, 43–52.

Bechtel, W. (1997). Representations and cognitive explanations: Assessing the dynamicist's challenge in cognitive science. Cognitive Science, 22(3), 295–318.

Bergmann, K., & Kopp, S. (2009). GNetIc–Using bayesian decision networks for iconic gesture generation. In Intelligent Virtual Agents (pp. 76–89).

Bernstein, N. A. (1967). Coordination and Regulation of Movement. New York: Pergamon Press.

Bicknell, K., & Levy, R. (2010). A rational model of eye movement control in reading. In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (pp. 1168–1178).

Black, J. B., Turner, T. J., & Bower, G. H. (1979). Point of view in narrative comprehension, memory, and production, 187–198.

Bock, J. K. (1986). Syntactic persistence in language production. Cognitive Psychology, 18, 355–387.

Bortfeld, H., & Brennan, S. E. (1997). Use and acquisition of idiomatic expressions in referring by native and non-native speakers, 119–147.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic co-ordination in dialogue. Cognition, 75, 13–25.

Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F., & Brown, A. (2011). The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers, 41–57.

Branigan, H. P., Pickering, M. J., Stewart, A. J., & McLean, J. F. (2000). Syntactic priming in spoken production: linguistic and temporal interference. Memory and Cognition, 28, 1297–302.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. Journal of experimental psychology. Learning, memory, and cognition, 22, 1482–93.

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.

Brennan, S. E., & Hanna, J. E. (2009). Partner-Specific Adaptation in Dialog, 274–291.

Brown-Schmidt, S. (2009a). Partner-specific interpretation of maintained referential precedents during interactive dialog. Journal of Memory and Language, 61(2), 171–190.

Brown-Schmidt, S. (2009b). The role of executive function in perspective taking during online language comprehension. Psychonomic bulletin & review, 16(5), 893–900.

Brown-Schmidt, S., Gunlogson, C., & Tanenhaus, M. K. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. Cognition, 107(3), 1122–1134.

Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. Dialogue & Discourse, 2(1), 11–33.

Brown-Schmidt, S., & Tanenhaus, M. K. (2008). Real-Time Investigation of Referential Domains in Unscripted Conversation: A Targeted Language Game Approach, 643–684.

Brunyé, T. T., Rapp, D. N., & Taylor, H. A. (2008). Representational flexibility and specificity following spatial descriptions of real-world environments, 418–443.

Buder, E. H., & Eriksson, A. (1999). Time-series analysis of conversational prosody for the identification of rhythmic units. In Proceedings of the 14th International Congress of Phonetic Sciences (Vol. 2, pp. 1071–

1074).

Busch, M. W. (2007). Task-based pedagogical activities as oral genres: A systemic functional linguistic analysis. ProQuest.

Carlson, L.A. (1999). Selecting a reference frame, 365–379.

Carlson, Laura A., & Covey, E. S. (2005). How far is near? Inferring distance from spatial descriptions, 617–631.

Carlson-Radvansky, L. A., & Irwin, D. E. (1994). Reference frame activation during spatial term assignment.

Carlson-Radvansky, L. A., & Logan, G. D. (1997). The influence of reference frame selection on spatial template construction, 411–437.

Carruthers, P. (2002). The cognitive functions of language. Behavioral and Brain Sciences, 25(06), 657–674.

Cartwright, N. (1999). The dappled world. Cambridge, UK: Cambridge University Press.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: the perception-behavior link and social interaction. J Pers Soc Psychol, 76, 893–910.

Chartrand, T. L., & Van Baaren, R. (2009). Human Mimicry. Advances in Experimental Social Psychology, Vol 41, 41, 219–274.

Chemero, A. (2009). Radical embodied cognitive science. MIT Press.

Clark, H. H. (1996). Using language. Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick (Ed.), Perspectives on Socially Shared Cognition (pp. 127–149). Washington: American Psychological Association.

Clark, H. H., & Krych, M. A. (2004). Speaking while monitoring addressees for understanding, 62–81.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process, 1–39.

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun- phrase structure. Journal of Memory and Language, 49, 214–30.

Cooper, R. P., Catmur, C., & Heyes, C. (2012). Are Automatic Imitation and Spatial Compatibility Mediated by Different Processes? Cognitive Science.

Crosby, J. R., Monin, B., & Richardson, D. (2008). Where do we look during potentially offensive behavior? Psychological Science, 19(3), 226–228.

Dale, R. (2008). The possibility of a pluralist cognitive science. Journal of Experimental & Theoretical Artificial Intelligence, 20(3), 155–179.

Dale, R. (n.d.). Critique of Radical Embodied Cognitive Science. Journal of Mind and Behavior, 31, 127–140.

Dale, R., & Louwerse, M. M. (n.d.). Human interaction as a multimodal network structure. Presented at the Conceptual Structures, Discourse, and Language.

Dale, R., Warlaumont, A. S., & Richardson, D. C. (n.d.). Nominal cross recurrence as a generalized lag sequential analysis for behavioral streams. International Journal of Bifurcation and Chaos, 21, 1153–1161.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition?, 441–447.

De Looze, C., Oertel, C., Rauzy, S., & Campbell, N. (2011). Measuring dynamics of mimicry by means of prosodic cues in conversational speech (pp. 1294–1297). Presented at the International Conference on Phonetic Sciences (ICPhS).

De Looze, C., & Rauzy, S. (2012). Measuring speakers' similarity in speech by means of prosodic cues: methods and potential (pp. 1393–1396). Presented at the Interspeech 2011.

Di Paolo, E., & De Jaegher, H. (2012). The interactive brain hypothesis.

Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. Advances in Experimental Social Psychology, Vol 33, 33, 1–40.

Dixon, J. A., & Stephen, D. G. (2012). Multi-scale interactions in dictyostelium discoideum aggregation. Physica A: Statistical Mechanics and its Applications.

Dixon, J. A., Stephen, D. G., Boncoddo, R., & Anastas, J. (2010). The self-organization of cognitive structure. Psychology of Learning and Motivation, 52, 343–384.

Dressler, R. A., Buder, E. H., & Cannito, M. P. (2009). Rhythmic patterns during conversational repairs in speakers with aphasia. Aphasiology, 23, 731–748.

Duran, N. D., Dale, R., & Kreuz, R. J. (2011). Listeners invest in an assumed other's perspective despite cognitive cost. Cognition, 121, 22–40.

Duran, N. D., & Dale, R. (in press). Perspective-taking in dialogue as self-organization under social constraints. New Ideas in Psychology.

Edlund, J., Heldner, M., & Hirschberg, J. (2009). Pause and gap length in face-to-face interaction (pp. 2779–2782). Presented at the Interspeech.

Eliasmith, C. (1996). The third contender: A critical examination of the dynamicist theory of cognition. Philosophical Psychology, 9(4), 441–463.

Eliasmith, Chris. (2012). The Complex Systems Approach: Rhetoric or Revolution. Topics in Cognitive Science, 4 (1), 72–77. doi:10.1111/j.1756-8765.2011.01169.x

Engstrom, D. A., Kelso, J., & Holroyd, T. (1996). Reaction-anticipation transitions in human perception-action patterns, 809–832.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment., 327.

Fowler, C. A., Richardson, M. J., Marsh, K. L., & Shockley, K. (2008). Language use, coordination, and the emergence of cooperative action. In A. & J. Fuchs (Ed.), Coordination: Neural, Behavioral and Social Dynamics. Springer.

Frank, T. D., Richardson, M. J., Lopresti-Goodman, S. M., & Turvey, M. T. (2009). Order parameter dynamics of body-scaled hysteresis and mode transitions in grasping behavior, 127–147.

Franklin, N., & Tversky, B. (1990). Searching imagined environments., 63–76.

Freeman, J. B., Dale, R., & Farmer, T. A. (2011). Hand in Motion Reveals Mind in Motion.

Frith, U., & Frith, C. (2001). The biological basis of social interaction. Current Directions in Psychological Science, 10(5), 151–155.

Fusaroli, R., Bahrami, B., Olsen, K., Rees, G., Frith, C. D., Roepstorff, A., & Tylén, K. (2012). Coming to terms: an experimental quantification of the coordinative benefits of linguistic interaction. Psychological science, 23.

Fusaroli, R., Gangopadhyay, N., & Tylén, K. The dialogically extended mind: making a case for language as skilful intersubjective engagement.

Fusaroli, R., & Tylén, K. (2012). Carving Language for Social Interaction: a dynamic approach. Interaction studies, 13, 103–123.

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know., 378.

Galantucci, B. (2009). Experimental Semiotics: a new approach for studying communication as a form of joint action. Topics in Cognitive Science, 1, 393–410.

Galantucci, B., & Sebanz, N. (2009). Joint Action: Current Perspectives. Topics in Cognitive Science, 1, 255–259.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? Journal of Memory and Language, 62(1), 35–51.

Gallagher, H. L., Jack, A. I., Roepstorff, A., & Frith, C. D. (2002). Imaging the intentional stance in a competitive game, 814–821.

Gallagher, R., & Appenzeller, T. (1999). Beyond Reductionism. Science, 284(5411), 79–79. doi:10.1126/science.284.5411.79

Gallese, V. (2008). Mirror neurons and the social nature of language: The neural exploitation hypothesis. Social neuroscience, 3(3-4), 317–333.

Gambi, C., & Pickering, M. J. (2011). A cognitive architecture for the coordination of utterances. Frontiers in Psychology, 2.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. Cognition, 27, 181–218.

Garrod, S., & Clark, A. (1993). The development of dialogue co-ordination skills in schoolchildren. Language and Cognitive Processes, 8, 101–126.

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. Cognition, 53, 181–215.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? Trends in Cognitive Sciences, 8(1), 8–11. doi:10.1016/j.tics.2003.10.016

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. Topics in Cognitive Science, 1, 292–304.

Giles, H., Coupland, J., & Coupland, N. (1991). Contexts of Accomodation. New York: Cambridge University Press.

Goodwin, C. (2000). Action and Embodiment within Human Situated Interaction. Journal of Pragmatics, 32, 1489–1522.

Goodwin, C. (2003). Conversation and brain damage. Oxford; New York: Oxford University Press.

Goodwin, C. (2011). Building Action in Public Environments with Diverse Semiotic Resources. Versus, 112-113.

Graesser, A. C., Swamer, S. S., & Hu, X. (1997). Quantitative discourse psychology. Discourse Processes, 23(3), 229–263. doi:10.1080/01638539709544993

Grammer, K., Kruck, K. B., & Magnusson, M. S. (1998). The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. Journal of Nonverbal Behavior, 22(1), 3–29.

Gratier, M., & Devouche, E. (2011). Imitation and repetition of prosodic contour in vocal interaction at 3 months. Dev Psychol, 47, 67–76. doi:10.1037/a0020722

Gries, S. T. (2005). Syntactic priming: a corpus-based approach. Journal of psycholinguistic research, 34, 365–99. doi:10.1007/s10936-005-6139-3

Hakkani-Tur, D. (2009). Towards automatic argument diagramming of multiparity meetings. In Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference On (pp. 4753–4756).

Hanna, J. E., & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation, 596–615.

Hanna, J. E., & Tanenhaus, M. K. (2004). Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements, 105–115.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation, 43–61.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. Cognition, 75, 27–39.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. Trends in cognitive sciences. doi:10.1016/j.tics.2011.12.007

Healey, P. G. T., Howes, C., & Purver, M. (2010). Does structural priming occur in ordinary conversation? Presented at the Linguistic Evidence.

Healey, P. G. T., & Mills, G. (2006). Participation, Precedence and Co-ordination in Dialogue. Presented at the 28th Annual Conference of the Science Society.

Horton, W. S, & Keysar, B. (1996). When do speakers take into account common ground?, 91–117.

Horton, W. S. (2005). Conversational common ground and memory processes in language production, 1–35.

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: knowing< i> when</i> and knowing< i> how</i> to adjust utterances to addressees, 589–606.

Horton, William S., & Gerrig, R. J. (2005). The impact of memory demands on audience design during language production. Cognition, 96(2), 127–142. doi:10.1016/j.cognition.2004.07.001

Hutchins, E. (1995a). Cognition in the Wild. Cambridge, Mass.: MIT Press.

Hutchins, E. (1995b). How a cockpit remembers its speeds. Cognitive Science, 19, 265–288.

Hutchins, E., & Johnson, C. M. (2009). Modeling the Emergence of Language as an Embodied Collective Cognitive Activity. Topics in Cognitive Science, 1, 523–546. doi:10.1111/J.1756-8765.2009.01033.X

Jaeger, T. F. (2010). Redundancy and reduction: speakers manage syntactic information density. Cognitive psychology, 61, 23–62. doi:10.1016/j.cogpsych.2010.02.002

Kauffman, S. (1996). At Home in the Universe: The Search for the Laws of Self-Organization and Complexity: The Search for the Laws of Self-Organization and Complexity. Oxford University Press.

Kello, C. T. (n.d.). Critical branching neural networks. Psychological Review.

Kelso, J. (1981). On the oscillatory basis of movement, 9.

Kelso, J. A. . (1995). Dynamic patterns: The self-organization of brain and behavior. The MIT Press.

Kelso, J. A. S. (2009). Synergies: atoms of brain and behavior. Adv Exp Med Biol, 629, 83–91. doi:10.1007/978-0-387-77064-2_5

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension, 32–38.

Keysar, B., Barr, D. J., & Horton, W. S. (1998). The egocentric basis of language use: Insights from a processing approach, 46–50.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. Cognition, 89(1), 25–41.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention, Researchers! It Is Time to Take a Look at the Real World. Current Directions in Psychological Science, 12(5), 176–180. doi: 10.1111/1467-8721.01255

Knoblich, G., Butterfill, S., & Sebanz, N. (2011a). 3 Psychological Research on Joint Action: Theory and Data, 59.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011b). Psychological Research on joint action: theory and data. Psychology of Learning and Motivation-Advances in Research and Theory, 54, 59.

Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action, 99–104.

Kopp, S. (2010). Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. Speech Communication, 52(6), 587–597.

Kousidis, S., & Dorran, D. (2009). Monitoring convergence of temporal features in spontaneous dialogue speech.

Kousidis, S., Dorran, D., Mcdonnell, C., & Coyle, E. (2009). Time Series Analysis of Acoustic Feature Convergence in Human Dialogues.

Kousidis, S., Dorran, D., Wang, Y., Vaughan, B., Cullen, C., Campbell, D., … Coyle, E. (2008). Towards measuring continuous acoustic feature convergence in unconstrained spoken dialogues.

Latash, M. L. (2008). Synergy. Oxford; New York: Oxford University Press.

Latash, M. L., Scholz, J. P., & Schoner, G. (2007). Toward a new theory of motor synergies. Motor control, 11, 276–308.

Lee, C., Black, M., Katsamanis, A., Lammert, A., Baucom, B., Christensen, A., … Narayanan, S. (2010). Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples (Vol. 793–796). Presented at the Interspeech.

Lee, P. U., & Tversky, B. (2005). Interplay between visual and spatial: The effect of landmark descriptions on comprehension of route/survey spatial descriptions, 163–185.

Lelong, A., & Bailly, G. (2011). Study of the phenomenon of phonetic convergence thanks to speech dominoes. Analysis of Verbal and Nonverbal Communication and Enactment. The Processing Issues, 273–286.

Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. Journal of experimental social psychology, 46(6), 1093–1096.

Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. Cognitive Psychology, 14, 78–106.

Levinson, S. C. (1983). Pragmatics. Cambridge Cambridgeshire; New York: Cambridge University Press.

Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions (pp. 3081–3084). Presented at the Interspeech 2011.

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. Trends in Cognitive Sciences, 10(10), 447–454.

Lin, S., Keysar, B., & Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. Journal of Experimental Social Psychology, 46(3), 551–556.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices, 550–557.

Louwerse, M. M., Benesh, N., Watanabe, S., Zhang, B., Jeuniaux, P., & Vargheese, D. (n.d.). The multimodal nature of embodied conversational agents. In Proceedings of the 31st Annual Meeting of the Cognitive Science Society (pp. 1459–1463).

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior Matching in Multimodal Communication Is Synchronized. Cogn Sci. doi:10.1111/j.1551-6709.2012.01269.x

Marsh, K. L., Richardson, M. J., Baron, R. M., & Schmidt, R. C. (2006). Contrasting approaches to perceiving and acting with others, 1–38.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. Topics in Cognitive Science, 1, 320–339.

Marwan, N., Carmen Romano, M., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. Physics Reports, 438, 237–329.

McElree, B. (2006). Accessing recent events. Psychology of learning and motivation, 46, 155–200.

McNamara, T. P., Sluzenski, J., & Rump, B. (2008). Human spatial memory and navigation. In J. Byrne (Ed.), (pp. 157–178). Oxford: Elsevier.

Mehler, A., Weiß, P., Menke, P., & Lücking, A. (2010). Towards a simulation model of dialogical alignment. In Proceedings of the 8th International Conference on the Evolution of Language (Evolang8), 14-17 April 2010, Utrecht (pp. 238–245).

Mills, G. (2011). The emergence of procedural conventions in dialogue. Presented at the 33rd Annual Conference of the Cognitive Science Society.

Mills, G., & Gregoromichelaki, E. (2010). Establishing coherence in dialogue: sequentiality, intentions and negotiation. Presented at the SemDial (PozDial).

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. Cognitive psychology, 41(1), 49–100.

Mou, W., Liu, X., & McNamara, T. P. (2009). Layout geometry in encoding and retrieval of spatial memory, 83.

Murray, G., Renals, S., Carletta, J., & Moore, J. (2006). Incorporating speaker and discourse features into speech summarization. In Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics (pp. 367–374).

Murray, L., & Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. Social perception in infants, 177–197.

Nadel, J., Carchon, I., Kervella, C., Marcelli, D., & Réserbat-Plantey, D. (1999). Expectancies for social contingency in 2 month olds. Developmental Science, 2, 164–173.

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates?, 669–678.

Newell, K. M., Broderick, M. P., Deutsch, K. M., & Slifkin, A. B. (2003). Task goals and change in dynamical degrees of freedom with motor learning. Journal of Experimental Psychology: Human Perception and Performance, 29(2), 379–387.

Newman-Norlund, S. E., Noordzij, M. L., Newman-Norlund, R. D., Volman, I. A. C., Ruiter, J. P., Hagoort, P., & Toni, I. (2009). Recipient design in tacit communication, 46–54.

Nishimura, R., Kitaoka, N., & Nakagawa, S. (2008). Analysis of relationship between impression of human-to-human conversations and prosodic change and its modeling (pp. 534–537). Presented at the Interspeech 2008.

Oller, D., Niyogi, P., Gray, S., Richards, J., Gilkerson, J., Xu, D., … Warren, S. (2010). Automated vocal analysis of naturalistic recordings from children with autism, language delay, and typical development. Proceedings of the National Academy of Sciences, 107(30), 13354.

Orsucci, F., Giuliani, A., & Webber, C. (2006). Combinatorics and synchronization in natural semiotics. Physica A: Statistical Mechanics and its Applications, 361, 665–676.

Orsucci, F., Giuliani, A., & Zbilut, J. (2004). Structure & Coupling of Semiotic Sets (Vol. 742, pp. 83–83).

Orsucci, F., Walter, K., Giuliani, A., Webber, C. L., & Zbilut, J. P. (1997). Orthographic Structuring of Human Speech and Texts: Linguistic Application of Recurrence Quantification Analysis. Arxiv preprint cmp-lg/9712010.

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2011). Phonetic convergence in college roommates. Journal of Phonetics.

Paxton, A., & Dale, R. (submitted). Argument disrupts interpersonal synchrony.

Paxton, A., & Dale, R. (n.d.). Frame-differencing methods for measuring bodily synchrony in conversation. Behavior Research Methods.

Perkins, L., & Milroy, L. (1997). Sharing the communicative burden: A conversation-analytic account of aphasic/non-aphasic interaction, 199–215.

Perry, M. (2010). Socially distributed cognition in loosely coupled systems. AI & Society, 1–14.

Pickering, M. J., & Branigan, H. P. (1999). Syntactic priming in language production. Trends Cogn Sci, 3, 136–141.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: a critical review. Psychological bulletin, 134, 427–59. doi:10.1037/0033-2909.134.3.427

Pickering, M. J., & Garrod, S. (in press). An Integrated Theory of Language Production and Comprehension". Behavioral and Brain Sciences.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences, 27, 169–190.

Pickering, M. J., & Garrod, S. (2009). Prediction and embodiment in dialogue. European Journal of Social Psychology, 39(7), 1162–1168.

Platek, S. M. (2010). Yawn, yawn, yawn, yawn; yawn, yawn, yawn! The social, evolutionary and neuroscientific facets of contagious yawning. Front Neurol Neurosci, 28, 107–12. doi:10.1159/000307086

Ramenzoni, V. C., Davis, T. J., Riley, M. A., Shockley, K., & Baker, A. A. (2011). Joint action in a cooperative precision task: nested processes of intrapersonal and interpersonal coordination. Experimental brain research. Experimentelle Hirnforschung. Experimentation cerebrale, 211, 447–57. doi:10.1007/s00221-011-2653-8

Ramenzoni, V. C., Riley, M. A., Shockley, K., & Baker, A. A. (2012). Interpersonal and intrapersonal coordinative modes for joint and single task performance. Hum Mov Sci. doi:10.1016/j.humov.2011.12.004

Reidsma, D., Heylen, D. K. J., & Ordelman, R. J. F. (2006). Annotating emotions in meetings.

Reis, H. T., Sheldon, K. M., Gable, S. L., Roscoe, J., & Ryan, R. M. (2000). Daily Well-Being: The Role of Autonomy, Competence, and Relatedness. Personality and Social Psychology Bulletin, 26(4), 419–435. doi:10.1177/0146167200266002

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. Cognitive Science, 35(4), 587–637.

Reitter, D., & Moore, J. D. (2007). Predicting success in dialogue (Vol. 45, pp. 808–815). Presented at the Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics.

Reitter, D., Moore, J. D., & Keller, F. (2006). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation (pp. 685–690). Presented at the Proceedings of the 28th annual conference of the cognitive science society.

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: A Multidisciplinary Journal, 29(6), 1045–1060.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). 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., & Spivey, M. J. (2000). Representation, space and Hollywood Squares: looking at things that aren't there anymore. Cognition, 76(3), 269–295.

Richardson, D. C., Street, C. N. H., Tan, J. Y. M., Kirkham, N. Z., Hoover, M. A., & Cavanaugh, A. G. (2012). Joint perception: gaze and social context. Frontiers in Human Neuroscience, 6.

Richardson, Daniel C., Dale, R., & Tomlinson, J. M. (2009). Conversation, Gaze Coordination, and Beliefs About Visual Context. Cognitive Science, 33(8), 1468–1482. doi:10.1111/j.1551-6709.2009.01057.x

Richardson, M. J., Dale, R., & Marsh, K. L. (n.d.). Complex dynamical systems in social and personality psychology: theory, modeling and analysis. In Handbook of Research Methods in Social and Personality Psychology.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C. (2007). Rocking together: dynamics of intentional and unintentional interpersonal coordination. Human movement science, 26, 867–91. doi:10.1016/j.humov.2007.07.002

Richardson, M. J., Marsh, K. L., & Schmidt, R. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination., 62.

Richardson, M. J., Marsh, K. L., Schmidt, R. C., & Richardson, M. J. (2010). Challenging the egocentric view of coordinated perceiving, acting and knowing, 307–333.

Riley, M. A., Richardson, M. J., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. Frontiers in psychology, 2, 38. doi:10.3389/fpsyg.2011.00038

Riley, M. A., & Van Orden, G. C. (2005). Tutorials in contemporary nonlinear methods for the behavioral sciences. National Science Foundation.

Roche, J., Jaeger, T. F., Dale, R., & Kreuz, R. J. (n.d.). Don't rush the navigator: disambiguating strategies are hard to establish but easier to maintain. Submitted manuscript.

Sacks, H., Jefferson, G., & Schegloff, E. A. (1995). Lectures on conversation (Vol. 1). Wiley Online Library.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. Language, 50, 696–735.

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: an integrated theory of concurrent multitasking. Psychological Review, 115(1), 101.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see., 1255.

Schank, R. C., & Abelson, R. P. (1977). Scripts, plans, goals, and understanding: an inquiry into human knowledge structures. Hillsdale, N.J. New York: L. Erlbaum Associates; distributed by the Halsted Press Division of John Wiley and Sons.

Schegloff, E. A. (1986). The Routine as Achievement. Human Studies, 9, 111–151.

Schegloff, E. A. (2007). Sequence organization in interaction: Volume 1: A primer in conversation analysis (Vol.

1). Cambridge University Press.

Schmidt, R. C. (2007). Scaffolds for social meaning, 137–151.

Schmidt, R. C., Bienvenu, M., Fitzpatrick, P. A., & Amazeen, P. G. (1998). A comparison of intra- and interpersonal interlimb coordination: coordination breakdowns and coupling strength. J Exp Psychol Hum Percept Perform, 24, 884–900.

Schmidt, R. C., Morr, S., Fitzpatrick, P., & Richardson, M. (2012). Measuring the Dynamics of Interactional Synchrony. Journal of Nonverbal Behavior, 36(4), 263–279. doi:10.1007/s10919-012-0138-5

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In Coordination: Neural, Behavioral and Social Dynamics (pp. 281–308).

Schmidt, R. C., & Turvey, M. T. (1995). Models for interlimb coordination: Equilibria, local analyses, and spectral patterning: Comment on Fuchs and Kelso (1994).

Schober, M. F. (1993). Spatial perspective-taking in conversation* 1, 1–24.

Schober, M. F. (1995). Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about locations?, 219–247.

Schober, M. F., & Brennan, S. E. (2003a). Processes of interactive spoken discourse: The role of the partner, 123–64.

Schober, M. F., & Brennan, S. E. (2003b). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), Handbook of Discourse Processes (pp. 211–232). Lawrence Erlbaum.

Seidenberg, M. S., & MacDonald, M. C. (1999). A probabilistic constraints approach to language acquisition and processing, 569–588.

Senay, I., & Keysar, B. (2009). Keeping Track of Speaker's Perspective: The Role of Social Identity, 401–425.

Shalizi, C., Camperi, M., & Klinkner, K. (2007). Discovering functional communities in dynamical networks. Statistical Network Analysis: Models, Issues, and New Directions, 140–157.

Shintel, H., & Keysar, B. (2009). Less is more: A minimalist account of joint action in communication. Topics in Cognitive Science, 35(2), 281–322.

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. Topics in Cognitive Science, 1(2), 305–319.

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.

Shteynberg, G. (2010). A silent emergence of culture: The social tuning effect. Journal of personality and social psychology, 99(4), 683.

Shteynberg, G., & Galinsky, A. D. (2011). Implicit coordination: Sharing goals with similar others intensifies goal pursuit. Journal of Experimental Social Psychology, 47(6), 1291–1294.

Smith, M., & Wheeldon, L. (2001). Syntactic priming in spoken sentence production - an online study. Cognition, 78, 123–64.

Smith, N. J., & Levy, R. (2008). Optimal processing times in reading: a formal model and empirical investigation. In Proceedings of the 30th Annual Conference of the Cognitive Science Society (pp. 595–600).

Song, J. H., & Nakayama, K. (2009). Hidden cognitive states revealed in choice reaching tasks, 360–366.

Spivey, M. J. (2007). The continuity of mind. Oxford, UK: Oxford University Press.

Spurrett, D., & Cowley, S. J. (2004). How to do things without words: infants, utterance-activity and distributed cognition. Language Sciences, 26, 443–466.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. Proc Natl Acad Sci U S A, 106, 10587–92. doi: 10.1073/pnas.0903616106

Szmrecsanyi, B. (2005). Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. Corpus Linguistics and Linguistic Theory, 1, 113–149.

Szmrecsanyi, B. (2006). Morphosyntactic persistence in spoken English: A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis. Berlin/New York: Mouton de Gruyter.

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? Cognitive Psychology, 23(3), 457–482. doi:10.1016/0010-0285(91)90016-H

Tanenhaus, M. K., & Brown-Schmidt, S. (2008). Language processing in the natural world. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1493), 1105.

Tanenhaus, M., 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.

Tanya L. Chartrand, & Jessica L. Lakin. (2012). The Antecedents and Consequences of Human Behavioral Mimicry, 1–28.

Taylor, H. A., Naylor, S. J., Faust, R. R., & Holcomb, P. J. (1999). "Could you hand me those keys on the right?" Disentangling spatial reference frames using different methodologies, 381–397.

Taylor, H. A., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions* 1, 261–292.

Taylor, H. A., & Tversky, B. (1996). Perspective in Spatial Descriptions, 371–391.

Teufel, C., Fletcher, P. C., & Davis, G. (2010). Seeing other minds: attributed mental states influence perception, 376–382.

Thelen, E., & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge, MA: The MIT Press.

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. Nature neuroscience, 5(11), 1226–1235.

Truong, K. P., & Heylen, D. K. J. (2012). Measuring prosodic alignment in cooperative task-based conversations. Presented at the Interspeech 2012.

Tuller, B., Case, P., Ding, M., & Kelso, J. A. (1994). The nonlinear dynamics of speech categorization, 3–16.

Turvey, M. T. (1990). Coordination. The American psychologist, 45, 938–53.

Turvey, M. T. (2007). Action and perception at the level of synergies. Human Movement Science, 26(4), 657–697.

Turvey, M. T., Fitch, H. L., & Tuller, B. (1982). The Bernstein perspective: I. The problems of degrees of freedom and context-conditioned variability. In Human Motor Behavior: An Introduction (pp. 239–252).

Tversky, B., & Hard, B. M. (2009). Embodied and disembodied cognition: Spatial perspective-taking, 124–129.

Tversky, B., Lee, P., & Mainwaring, S. (1999). Why do speakers mix perspectives?, 399–412.

Tylén, K., & Fusaroli, R. Testing models of linguistic coordination.

Tylén, K., Fusaroli, R., Bundgaard, P., & Østergaard, S. Making Sense Together: a dynamical account of linguistic meaning making. Semiotica, xxx.

Tylén, K., Weed, E., Wallentin, M., Roepstorff, A., & Frith, C. D. (2010). Language as a tool for interacting minds. Mind & Language, 25, 3–29.

Van Orden, G., & Stephen, D. G. (2012). Is cognitive science usefully cast as complexity science? Topics in cognitive science. Van Overwalle, F. (2008). Social cognition and the brain: A meta-analysis. Human brain mapping, 30(3), 829–858.

Van Rooij, I., Bongers, R. M., & Haselager, W. (2002). A non-representational approach to imagined action, 345–375.

Vaughan, B. (2011). Prosodic synchrony in co-operative task-based dialogues: A measure of agreement and disagreement.

Wagenmakers, E. J., Van der Maas, H. L. J., & Farrell, S. (2012). Abstract Concepts Require Concrete Models: Why Cognitive Scientists Have Not Yet Embraced Nonlinearly Coupled, Dynamical, Self-Organized Critical, Synergistic, Scale-Free, Exquisitely Context-Sensitive, Interaction-Dominant, Multifractal, Interdependent Brain-Body-Niche Systems. Topics in cognitive science.

Wang, Y., & Hamilton, A. F. C. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. Frontiers in Human Neuroscience, 6.

Warlaumont, A. S. (2012). A spiking neural network model of canonical babbling development. Presented at the 2012 IEEE International Conference on Development and Learning (ICDL).

Wilkes-Gibbs, D., & Clark, H. H. (1992). Coordinating beliefs in conversation. Journal of Memory and Language, 31, 183–94.

Wilkinson, R., Beeke, S., & Maxim, J. (2003). Adapting to conversation. In C. Goodwin (Ed.), Conversation and Brain Damage (pp. 59–89). Oxford University Press.

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. Psychonomic bulletin & review, 12, 957–68.

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358(1431), 593–602.

Wu, S., & Keysar, B. (2007). The effect of information overlap on communication effectiveness. Cognitive Science, 31(1), 169–181.

Wyatt, D., Bilmes, J., Choudhury, T., & Kitts, J. A. (2008). Towards the automated social analysis of situated speech data. In Proceedings of the 10th International Conference on Ubiquitous Computing (pp. 168–171).

Zacks, J. M., & Michelon, P. (2005). Transformations of visuospatial images, 96.

[1] Though the brilliant corpus strategies used by researchers like Bard, Aylett, and others can reveal substantial clues about cognitive mechanism through, for example, acoustic properties of what one person says to another (see, e.g., Bard et al., 2000).[2] This is a deliberately mild way of putting it. Others, such as anonymous reviewers of some journal articles, have described the vocabulary as "Star Wars terminology." Readers may have other examples.[3] We will also take "system" for granted. No further words are offered on this. You have to start from somewhere. Readers wishing to have pure operationalized definitions of all things can consult the success of, for example, Rudolf

Carnap's sense-data grammar (Carnap, 1928; Quine, 1960). [4] Complex systems are also figured to involve interactions among components which produce collective, higher-level behaviors not reducible to properties of the components themselves. We wish to avoid this debate here ("emergence" or "emergentism") though it seems likely a natural consequence of the perspective we describe here (see also Knoblich et al., 2011).[5] Note that this has sparked decades of exciting work and debate in motor control. Though Bernstein's solution has, in many respects, become standard in broad strokes, how it is solved can be the subject of some debate (see, among many, Latash et al., 2007; Newell et al., 2003; Todorov & Jordan, 2002; Turvey, 2007).[6] Here we are using "degree of freedom" in a very informal way, simply to specify whether a channel, or set of channels, are "free to vary," or whether they constrain each other *in some fashion*. Of course network analysis can involve gradient aspects of these couplings, but we ignore this for simplicity here.

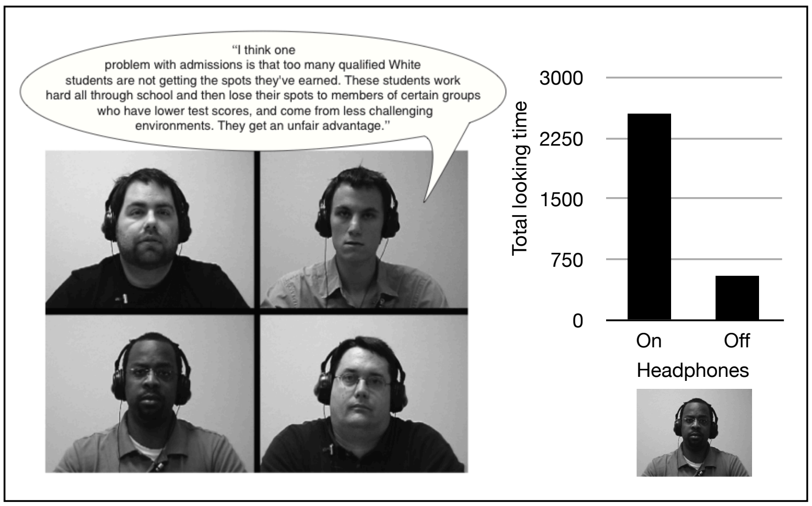


Figure 1: Example 2x2 grid viewed by subjects watching a discussion. How much subjects fixate a visible minority is predicted by whether that person is able to hear an offensive statement.

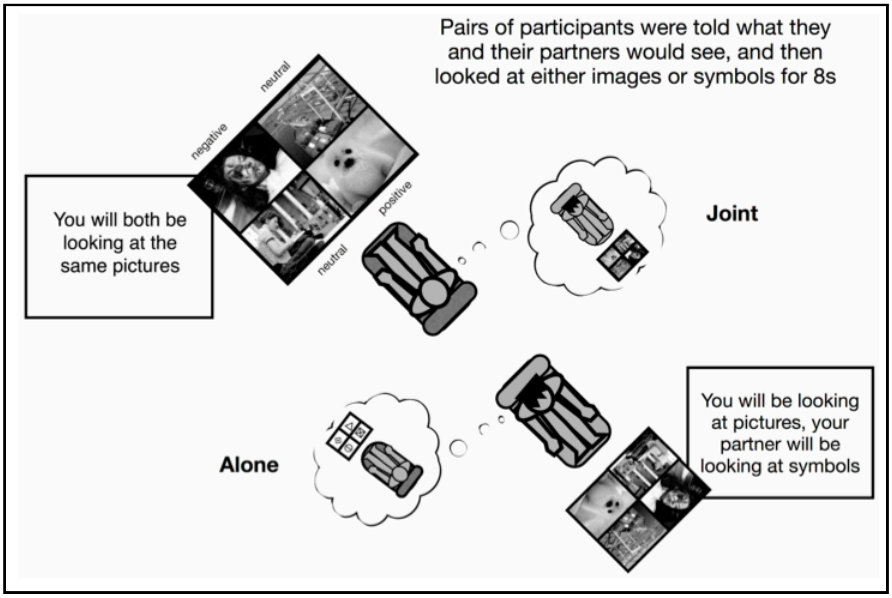


Figure 2: Example context in which 2 people were given different beliefs about what their "socially co-present" partner could see. Imagines of differing emotional valence were presented.

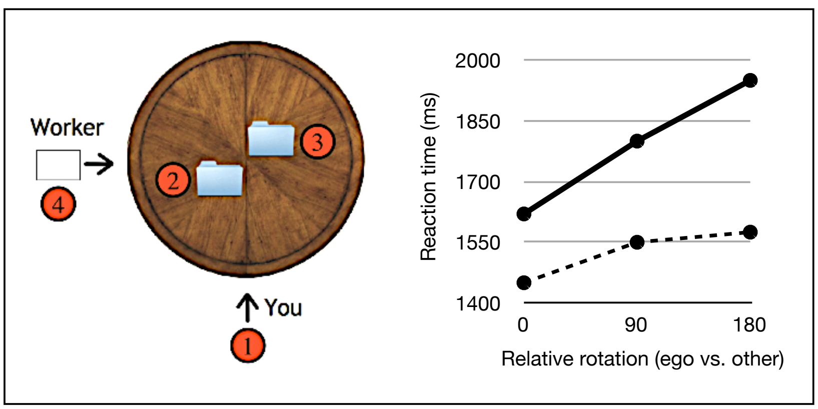


Figure 3: When participants are asked to retrieve a folder in an "ostensibly social" computer task as shown on the left, they tend to exhibit different mental rotation functions, predicted by whether they are taking theirs or another person's perspective.

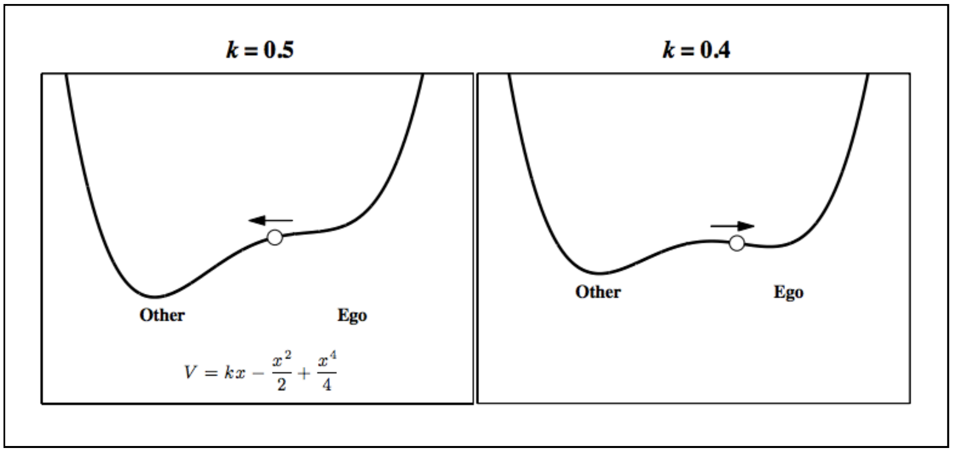


Figure 4: Perspective-taking can be modeled as a low-dimensional dynamic process. Self-organization into "other" or "ego" perspectives can be seen as traversing a low-dimensional landscape. Simulating this model allows qualitative fits to three timescales of human data: (i) decisions, (ii) response times, and (iii) response dynamics.

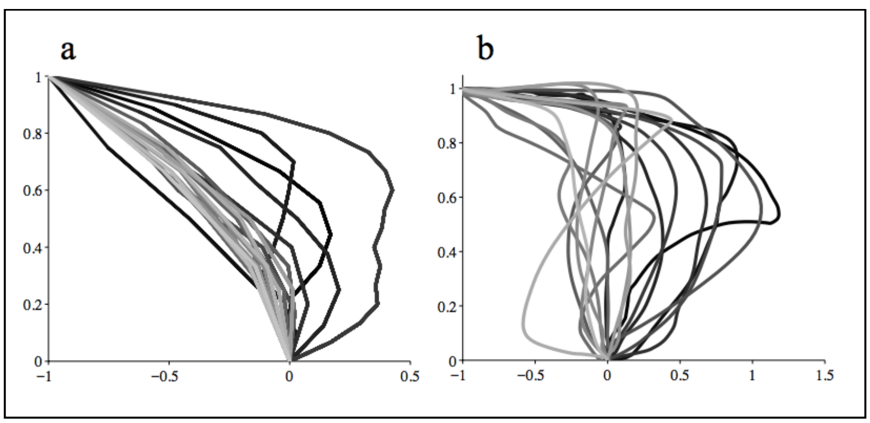


Figure 5: On the left, model responses (from bottom left, up to top right) are shown, with grayscale representing how deeply into the simulated experiment models have proceeded; on the right, the human mouse-movement data for the real experiment (i.e., selecting a folder). Human data appear noisy due to the only very small possible curvature in each trial. Nevertheless, this level and the levels of reaction time and decision distribution show compelling qualitative fits.

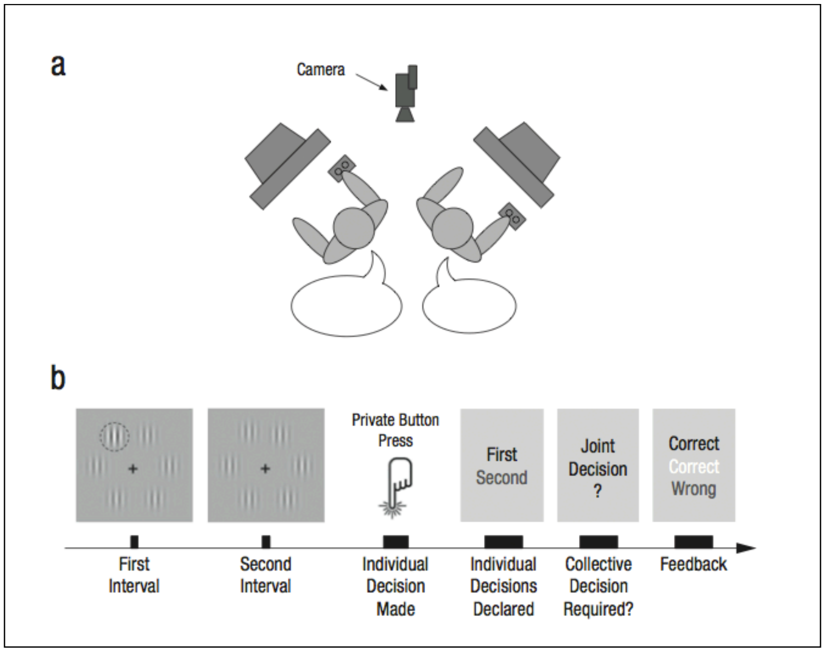
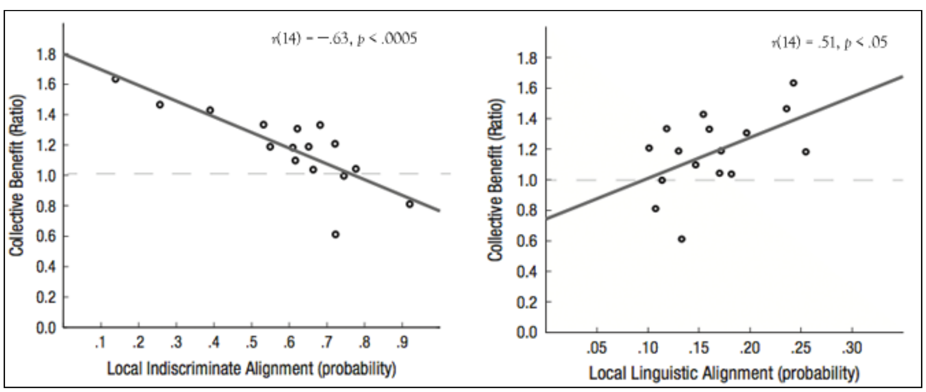


Figure 6: Interactive perceptual detection task. (a) Participants both view noisy stimuli and can communicate regarding the presence of a target. (b) The sequence of events in the task, from stimulus presentation, to the presentation of feedback. Image adapted with permission from Bahrami 2010.Figure 7: Results from the alignment of word usage during the task. The Collective Benefit (y-axis) is a measure of how much dyads benefited from their interaction; local linguistic alignment was a benefit (right plot), however rampant widespread "indiscriminate" alignment predicted a drop in joint performance (left plot).



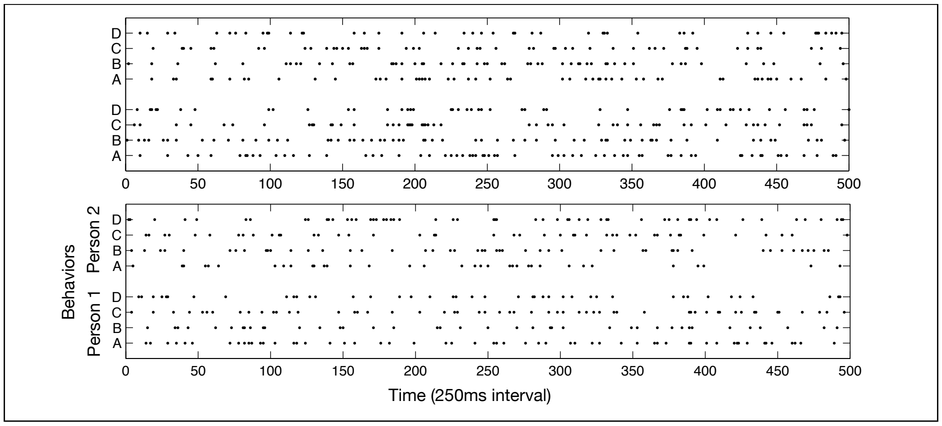


Figure 8: *Simulated* point processes of behavioral events. “Person 1” and “Person 2” have four behavioral dimensions (A-D). In human data these would correspond to delimitable actions such as nods, laughter, or gesture (Louwerse et al., 2009, 2012). Across time, these events occur and may serve to coordinate behavior both *within* and *across* modeled processes. In the top plot, all processes are random; in the bottom, there is a greater probability of “alignment” (e.g., Person 1’s A occurring with Person 2’s A). This may not be evident by mere visualization, but by inducing a network through (for example) temporal correlation, we can extract the interactive structure of the model (see Figure 9).

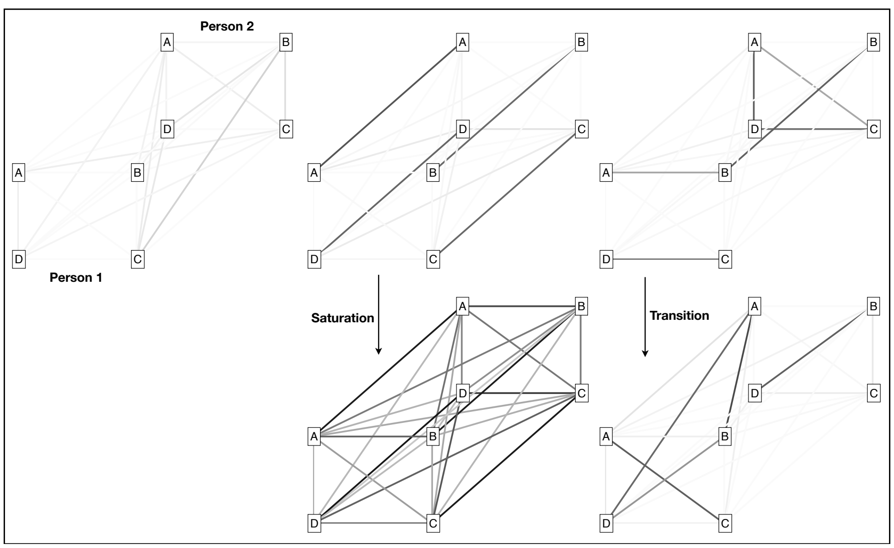


Figure 9: An illustration of different graph (network) structures induced from different interrelationships among *simulated* point processes (Person 1, Person 2). On the top left, only very light gray edges reflect a weak connection across all pairs of nodes (no behaviors are coupled). In the middle column, an illustrate of alignment (A’s, B’s, etc. go correlate), with saturation in the behavior (cascading such that behavioral events all occur together in a kind of synchronous multi-level alignment). In the rightmost column, a synergistic structure, where there is occasional alignment, but amidst a variety of other interconnections that may fluctuate from moment to moment (e.g., top panel: {$A_{Person\ 2}$, $D_{Person\ 2}$, $C_{Person\ 2}$},{$A_{Person\ 1}$, $B_{Person\ 1}$, $B_{Person2}$},{$D_{Person\ 1}$,$C_{Person\ 1}$}, transitions to bottom panel: {$D_{Person\ 1}$,$B_{Person\ 1}$,$A_{Person\ 2}$},{$A_{Person\ 1}$,$C_{Person\ 1}$},{$D_{Person\ 2}$,$B_{Person\ 2}$}).