Investigating Conversational Dynamics: Interactive Alignment, Interpersonal Synergy, and Collective Task Performance

Riccardo Fusaroli,[a,b] Kristian Tylén[a,b]

[a]*Center for Semiotics, Aarhus University*
[b]*The Interacting Minds Centre, Aarhus University*

Abstract

This study investigates interpersonal processes underlying dialog by comparing two approaches, *interactive alignment* and *interpersonal synergy*, and assesses how they predict collective performance in a joint task. While the interactive alignment approach highlights imitative patterns between interlocutors, the synergy approach points to structural organization at the level of the interaction—such as complementary patterns straddling speech turns and interlocutors. We develop a general, quantitative method to assess lexical, prosodic, and speech/pause patterns related to the two approaches and their impact on collective performance in a corpus of task-oriented conversations. The results show statistical presence of patterns relevant for both approaches. However, synergetic aspects of dialog provide the best statistical predictors of collective performance and adding aspects of the alignment approach does not improve the model. This suggests that structural organization at the level of the interaction plays a crucial role in task-oriented conversations, possibly constraining and integrating processes related to alignment.

Keywords: Interactive alignment; Interpersonal synergies; Communication; Decision-making; Distributed cognition; Social cognition; Pragmatics; Dynamical systems

1. Introduction

As the most fundamental and widespread form of language use, dialog has a pervasive and all-encompassing impact on everyday life. Through conversations we develop and maintain social relations and cultural practices, we plan and coordinate with each other, we share experiences, memories, and attitudes, we educate and learn (Clark, 1996; Tylén,

---

Correspondence should be sent to Riccardo Fusaroli, Center for Semiotics, Aarhus University, Jens Chr. Skous Vej 2, 8000 Aarhus, Denmark. E-mail fusaroli@gmail.com

Fusaroli, Bundgaard, & Østergaard, 2013; Tylén, Weed, Wallentin, Roepstorff, & Frith, 2010). Yet dialog has been a challenge to psycholinguistics, which has traditionally focused on linguistic processing of individuals. As an inherently coordinative social activity, dialog does not fit classical assumptions of individual linguistic processing: In dialog, it is often impossible to unambiguously assign roles of speaker and listener—who produces and who receives—as they tend to overlap and mix (Clark & Schaefer, 1989). Well-formed sentences are rather the exception than the rule. In fact, when transcribed, dialogical utterances often appear elliptic to the extent of becoming ungrammatical (Clark, 1996; Linell, 1998, 2005). In dialog the structural organization straddles not only speech turns but also interlocutors, and it is not uncommon that one interlocutor completes sentences initiated by the other.

To tackle these challenges, diverse interpersonal perspectives on dialog have recently emerged. Although a general framework encompassing individual behaviors and interpersonal dynamics as complementary has been advocated (Levinson, 2006), there is still disagreement and uncertainty about which processes and mechanisms are at play and how they interact, possibly due to the diversity of conceptual frameworks and methods employed (Brennan, Galati, & Kuhlen, 2010; Dale, Fusaroli, Duran, & Richardson, 2013; Krauss & Fussell, 1996).

In order to start tackling these issues, here we (a) identify two main approaches to interpersonal coordination in dialog represented in the literature—interactive alignment and interpersonal synergies; (b) develop a method to quantitatively and comparably investigate the two approaches against a control baseline of individual self-consistency; and (c) assess their ability to capture dialogical dynamics and coordinative efficacy in an experimentally elicited corpus of task-oriented conversations.

2. Interpersonal approaches to dialog

2.1. Interactive alignment

A growing literature characterizes human interaction in terms of reciprocal behavioral and physiological mimicry (Chartrand & Bargh, 1999; Dijksterhuis & Bargh, 2001): Many behaviors such as laughing, smiling, and shaking the head or nodding are more likely to occur if one's interlocutor has just employed them (Louwerse, Dale, Bard, & Jeuniaux, 2012). Such behavior matching has been related to common ground, improved rapport, and better collaborative performance (Fusaroli & Tylén, 2012; Marsh, Richardson, & Schmidt, 2009). Analogously, the *interactive alignment* approach investigates dialog as imitation-like coordination of linguistic behaviors. Through an automatic structural priming mechanism (Pickering & Ferreira, 2008), interlocutors reciprocally align multiple properties of their linguistic behaviors, from prosody to syntax, whereby they come to share linguistic structure and ultimately conceptual representations (Branigan, Pickering, & Cleland, 2000; Brennan & Hanna, 2009; Pickering & Garrod, 2004, 2013). In the interactive alignment approach, the analytical focus is on the network of reciprocal

adaptations and repetitions between individuals: the degree to which interlocutors come to share lexical choices, prosodic and syntactic patterns, and so on. For instance, in the context of joint decision-making, a speaker might say, "I am almost sure we should go for option 2" while her interlocutor reply with "I think it could be option 1." The heterogeneity of the expressions employed to express confidence ("to be sure," "to think") makes it difficult to determine who is more certain, potentially engendering confusion and suboptimal joint decisions (Bang et al., 2014). An effective strategy is thus to align on a shared way of expressing confidence (Fusaroli et al., 2012; cf. also Reitter & Moore, 2014). Crucially, taking the approach at face value implies that higher degrees of alignment are correlated with better coordination, reciprocal understanding, and therefore higher collaborative performance.

2.2. Interpersonal synergies

Other approaches to the psycholinguistics of dialog put special emphasis on more complex interpersonal dimensions of linguistic processing. For instance, sociological traditions—such as conversation analysis (CA) and ethno-linguistics—stress how the interdependence between speakers' behaviors in co-constructing conversation relies on complementarity (Goodwin, 1979, 2011; Levinson, 2006; Sidnell & Enfield, 2012). CA has highlighted the widespread presence of structures that go beyond the single speech turn and across interlocutors. A notable example is that of *adjacency pairs*, sequences of complementary speech turns which together create a coherent pattern: A question is appropriately followed by an answer (and not by another question), an offer by an acceptance or a rejection, and so on (Enfield, 2013; Goodwin & Heritage, 1990). Similarly, Clark and Schaefer note that conversations are often characterized by the tight coordination of many different types of complementary contributions, including clarifications, acceptance, and backchannels (Clark & Schaefer, 1989). In other words, language is a pre-eminent case of joint action (Clark, 1996). Ideas of interdependence and complementarity have also found expression in dynamical systems approaches to social interactions (Dale et al., 2013). Recent studies show how agents' interactions gradually come to constitute an interpersonal action system, a structure that is best observable by assessing patterns of interaction and not just of individual behavior. For instance, if two agents are carrying a big piece of furniture and one of them raises one side too high, the other, to keep the piece balanced, can compensate by raising the opposite side in proportional ways. In this sense, their actions come to complement each other, for example, adjusting in opposite directions, maintaining the furniture balanced. Ramenzoni and colleagues experimentally quantified such motor coupling and demonstrated that it is sensitive to task constraints (Ramenzoni, Riley, Shockley, & Baker, 2012; Riley, Richardson, Shockley, & Ramenzoni, 2011). This *interpersonal synergy approach* has recently been extended to linguistic interactions. Relying on pre-existing and locally negotiated procedural scripts or routines, interlocutors become interdependent in their linguistic behavior. For instance, the dialog between customer and waiter at a restaurant is a familiar example of such a routine: There are fixed, complementary roles, including that of the waiter

taking orders and that customer providing them. But routines can also be less fixed and formal and are often flexibly negotiated for the course of a local joint activity: When painting a fence together, agents might implicitly distribute labor to effectively coordinate the joint activity. Similarly, making joint decisions might require one interlocutor to share information and the other—based on that information—to make a decision (cf. also Table 2 and the discussion section for more examples). Routines can thus be understood as patterns of behavior organized at the level of the interaction itself, rather than being merely a quality of individual behaviors (Dale et al., 2013; Fusaroli, Raczaszek-Leonardi, & Tylén, 2014; Mills, 2014). The overall pattern—whether it is a question followed by an answer or the communication of information followed by a decision—can only be fully observed in the regularities straddling the interlocutors' contributions. In other words, if we only observe the individual interlocutors' behaviors in isolation, we would miss the composite sequences of conversational routines constituted by the tight behavioral couplings.

In order to capture these systemic dynamics, the interpersonal synergy approach suggests focusing analytically on structures of the conversation that straddle interlocutors and speech turns. Crucially, this approach also suggests that better coordination and collaborative performance might be related to tighter routines, that is, more recurrent interactional patterns.

2.3. Baseline: Individual self-consistency

Interactive alignment and interpersonal synergies are interpersonal patterns, which are not observable looking at individual behaviors alone. However, in order to assure that a given interpersonal coordination phenomenon is not reducible to individual behaviors, we here introduce a control baseline: *individual self-consistency*. The idea that dialog presents patterns and regularities at the level of individual behavior is grounded in the rich corpus of psycholinguistic literature, which demonstrates that individuals also rely on strategies of self-consistency in linguistic forms. For instance, consistent pause timing and lexical choices have been shown to lessen cognitive costs for recipients and facilitate understanding and coordination overall (Brennan & Clark, 1996; Scarborough, Cortese, & Scarborough, 1977). Analogously in motor coordination, one agent's work in anticipating and adapting the other is greatly facilitated by the other being behaviorally consistent (Vesper, van der Wel, Knoblich, & Sebanz, 2011). Such observations could motivate the prediction that better collaborative performance would simply be associated with higher degrees of individual consistency. Crucially, an approach which emphasizes the centrality of interactive alignment and interpersonal synergies will have to show effects above and beyond such individual patterns of self-consistency.

2.4. Comparing approaches

While the two approaches (plus baseline) introduced above might capture essential and complementary aspects of linguistic processing in conversation (Konvalinka &

Roepstorff, 2012; Levinson, 2006), there are important pragmatic and conceptual reasons to compare them. While the existing literature does present contrastive discussions of approaches (for an extensive repertoire, cf. the commentaries to Pickering & Garrod, 2004, 2013), no systematic comparative analysis on the same corpus has been performed. In order to answer key questions about dialog, a more analytically rigorous set of analyses, derived from the same corpus, is necessary. With such analyses, we may be better prepared to answer a number of pressing research questions: Are all these structures of organization simultaneously present in dialogs? Do higher degrees of interactive alignment, synergetic structure, and/or self-consistency characterize more effective conversations, thus leading to better performance in collaborative tasks? Do these structures provide largely overlapping or complementary information on the efficacy of the conversations? In other words, do we explain more statistical variance in the interlocutors' coordinative performance by taking into account more than one approach at the same time?

In order to tackle these questions, we need comparable ways to assess and quantify the presence of interactive alignment and interpersonal synergies in conversations. In the following, we define the notion of "structural organization": a general index of structure and dynamics in linguistic dialog measured through recurrence quantification analysis (RQA) (Marwan, Carmen Romano, Thiel, & Kurths, 2007), which can be applied to individual and interactional regularities. We will then address the structural organization (amount and flexibility of regularities) of three central properties of linguistic behavior: lexical, prosodic and speech/pause sequences. Finally, we compare which of the approaches leads to more reliable predictions of collective performance and discuss implications for the study of dialog and joint action.

3. Structural organization in dialog

The different approaches to interpersonal processes in dialog address different aspects of organization: reciprocal alignment, interactional patterns, and individual consistency. The challenge is thus to individuate comparable indices of organization applicable to these different aspects of interaction. Here, we propose the concept of "structural organization" encompassing (a) the extension of stable patterns of linguistic behavior in the conversation and (b) the complexity (or flexibility) of their structure. By "extension" we intend the average length of a repeated pattern, which can vary from very short linguistic forms to full utterances. Given that interlocutors rarely (if ever) parrot each other slavishly, we also measure the "complexity" of structures defined as variability in the range of linguistic forms repeated: If repetitions generally involve the same kind of linguistic forms (e.g., a word), complexity is low. However, if they involve a varying range of forms (e.g., sometimes morphemes, sometimes full utterances)—in other words, if the patterns are flexibly repeated with high local variation—the complexity is high. To exemplify, consider situations where interlocutors have to make joint decisions repeatedly. In response to interlocutor A's utterance "I am really not sure," interlocutor B might answer,

"I am not sure either," repeating four words. The same words might come up in the subsequent decision; however, effective coordination requires interlocutors to flexibly adapt their lexical choices to the changing context. In an easier trial they should modify their expressions of confidence, therefore generating shorter or longer sequences of recurrence. If their individual confidence seems closely matched, they might further elaborate on the expressions used, to more accurately individuate who is more confident. Low complexity would imply a slavish repetition of the same, while high complexity would imply a more flexible reuse and modification of expressions.

These indices of structural organization are both sufficiently general to apply to the different approaches and linguistic properties investigated and sufficiently concrete to be objectively quantified. The quantitative assessment of structural organization is achieved through recurrence quantification analysis (RQA), a non-linear method that is employed to assess the structure and complexity of dynamical systems (Marwan et al., 2007).

Dynamical systems approaches have been successfully applied to human cognition and behavior in a wide range of contexts (Chemero, 2009; Riley & Van Orden, 2005; Van Orden & Stephen, 2012; Ward, 2002). It has been shown that both physiological signals and simple sensorimotor behaviors (Konvalinka et al., 2011; Kuznetsov & Wallot, 2011) as well as conceptual thought and problem solving (Angus, Smith, & Wiles, 2012; Dixon, Stephen, Boncoddo, & Anastas, 2010; Van Orden, Hollis, & Wallot, 2012) display complex patterns of regularities expressed, for instance, by scaling law distributions (Kello et al., 2010). This is thought to reflect the self-organization of cognitive processes: At the same time, an organism has to adhere to contextual constraints while maintaining flexible control (Van Orden, Holden, & Turvey, 2003, 2005). Language is no exception: Dynamical patterns of regularities and irregularities have been shown in several simple linguistic phenomena (Kello, Anderson, Holden, & Van Orden, 2008; Kello et al., 2010) and during conversational exchanges (Abney, Paxton, Dale, & Kello, 2014; Fusaroli, Abney, Bahrami, Kello, & Tylén, 2013). Interlocutors display behavioral regularities orienting to routines and local contextual constraints, but in a flexible and non-rigid way open to adaptation, transition, and change (Fusaroli, Konvalinka, & Wallot, 2014; Fusaroli, Raczaszek-Leonardi, et al., 2014). Non-linear statistical methods sensitive to continuously unfolding patterns of interaction have proven to be very informative about the underlying systems (Riley & Van Orden, 2005). RQA (Marwan et al., 2007) is particularly fit to for our comparative purposes: It can be measured for individuals' behaviors, across individuals, and for interactional routines, thus providing us comparable indices of recurrence for each of the approaches to dialog specified above. Moreover, RQA provides information about the regularity and flexibility of the patterns by quantifying, among other qualities, their temporal extension and complexity. RQA allows us to operationalize structural organization as the extent and complexity with which a given linguistic behavior recurs in a patterned way. We hypothesize that structural organization can diagnose the functional efficiency—in this context, making accurate joint decisions—of the system and that different properties of conversation will provide complementary information about it.

3.1. Three properties of conversation: Lexical choices, prosody, and speech/pause rhythm

In order to gain a reliable picture of the dynamics of linguistic behaviors and how they reflect the underlying system, we focus on three different properties of conversations. Previous analyses have highlighted several properties of linguistic processes in conversation, including lexical choices, prosody, and the distribution of pauses in speech. These three properties often work in parallel and serve overlapping and complementary functions in dialog. Lexical patterns roughly pertain to the semantic/pragmatic content of the conversations (Fusaroli et al., 2012; Garrod & Anderson, 1987). Prosodic patterns are often thought to carry pragmatic/discursive meaning; for instance, marking an utterance as intended imperatively or declaratively (Mushin, Stirling, Fletcher, & Wales, 2003), ironically (Bryant, 2010), or with affective emphasis (Banse & Scherer, 1996). Speech/pause rhythm is indicative of the fluency, disruption, or ease of the discourse, and it relates to syntactic structures (McFarland, 2001). Furthermore, it has been persuasively shown that interlocutors tend to adapt to each other's lexical choices (Fusaroli et al., 2012; Garrod & Anderson, 1987; Krauss & Weinheimer, 1966), prosodic patterns (Kousidis & Dorran, 2009; Levitan & Hirschberg, 2011; Pardo, 2006; Pardo, Gibbons, Suppes, & Krauss, 2012), and rhythms of speech and pauses (McFarland, 2001; Wilson & Wilson, 2005). However, systematic investigations of alignment and complementarity across multiple aspects are sparse and little is known about the impact of such adaptations on coordinative efficacy, such as performance in joint tasks.

3.2. Assessing dialogical mechanisms: Summing up the hypotheses

In order to empirically and contrastively investigate interactive alignment and interpersonal synergy against a baseline of individual self-consistency, we proceed to analyze a video corpus of task-oriented conversations where dyads make joint decisions across repeated trials (Bahrami et al., 2010; Fusaroli et al., 2012). This corpus grants us an objective measure of the efficacy of the interaction, which is used to comparatively assess the predictors motivated by the different approaches. In the following, we quantify the structural organization of prosody, speech/pause rhythm, and lexical choices for interactive alignment, interpersonal synergy, and individual self-consistency. We then assess their corresponding statistical models by testing how each predicts joint performance. We also ask whether mechanisms of alignment and synergy combine in a complementary fashion to provide non-redundant contributions to collective performance. We employ statistical model comparison techniques to (a) compare the predictive power of each approach by assessing which indices of structural organization best explain variance in the joint performance of the collaborative task, and (b) test if additional predictive power is gained by combining the two approaches and the control baseline. By employing performance as a dependent variable, we also make the approaches more comparable; rather than comparing the relative presence of alignment, synergy, and self-consistency directly, we compare the explanatory power of the approaches with regard to joint performance. While we find it likely that all approaches point to important aspects of dialog, a statisti-

cal model comparison analysis lets us compare the approaches systematically and discover which properties are most informative about the functional structure of task-oriented conversations, and how much information can be gained by analyzing more than one kind of organization.

The two approaches and the baseline can be summarized as complementary hypotheses:

Hypothesis I: (*The interactive alignment approach*). Conversation is mainly about mutual linguistic adaptation and synchronization. That is, interlocutors consistently and flexibly reuse each other's words, speech/pause rhythms, and prosodic patterns (see Fig. 1a). Accordingly, structural organization across interlocutors (how they adaptively repeat each other) will be present and predict joint performance above and beyond self-consistency.

Hypothesis II: (*The interpersonal synergy approach*). Interlocutors complement each other in jointly evolved interactional routines, for instance patterns of questions and answers. In these cases it is not crucial who is the questioner as long as the question is complemented by an answer. In other words, we do not distinguish the individual contributions (speech turns) of the interlocutors, but treat the conversation as one coherent, co-constructed text to assess structures of recurrence of prosody, speech/pause, and lexicon (see Fig. 1b). We expect structural organization in the overall *jointly constituted* conversation to be present and predictive of joint performance in the task above and beyond self-consistency (and possibly interactive alignment).

Hypothesis III: (*Self-consistency control baseline*). Conversation is a matter of individual linguistic production and processing (see Fig. 1c). Accordingly, we would expect self-consistency—that is, the structural organization of prosody, speech/pause, and lexicalization of individual interlocutors—to be present and predictive of joint performance. In other words, the extent and complexity of stable patterns in the behavior of individual interlocutors (an individual consistently using the same words, regular patterns of prosody or pauses) are informative indices of the performance of the pair.
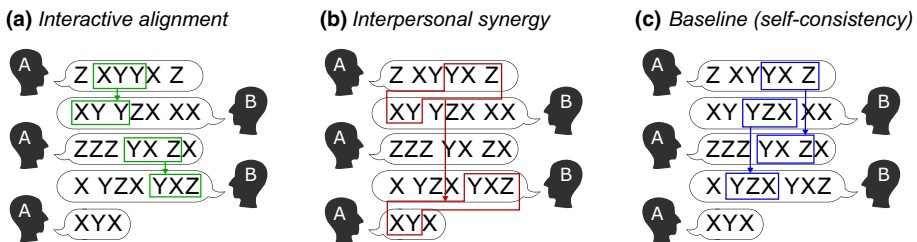

Fig. 1. Schematic depiction of structural organization in interactive alignment, interpersonal synergy, and self-consistency baseline.

4. Materials and methods

4.1. Corpus

The corpus consists of approximately 20 h of video recording of 16 dyads (14 m/18 f, mean age 25.2, *SD* = 6.9, all native speakers of Danish who gave informed, written consent), each performing on average 92 (*SD* = 15.5) joint decision trials for a total of 1,472 joint decision trials.[1] The participants were recorded while sitting at right angles to each other in a darkened room, each in front of a computer screen which displayed exactly the same video output (see Fig. 2a). In all trials participants were shown a sequence of two 85 ms long visual displays containing six Gabor patches (see Fig. 2 and Bahrami et al., 2010 for more details). One of the displays would contain a contrast odd-ball; that is, one of the six Gabor patches would have a stronger contrast and therefore look slightly darker. The degree of visual contrast varied randomly to assess performance at different degrees of difficulty (higher contrast corresponding to an easier task). The participants were instructed to individually and separately indicate, by pressing a button, which of the displays contained the contrast oddball. As long as both participants gave
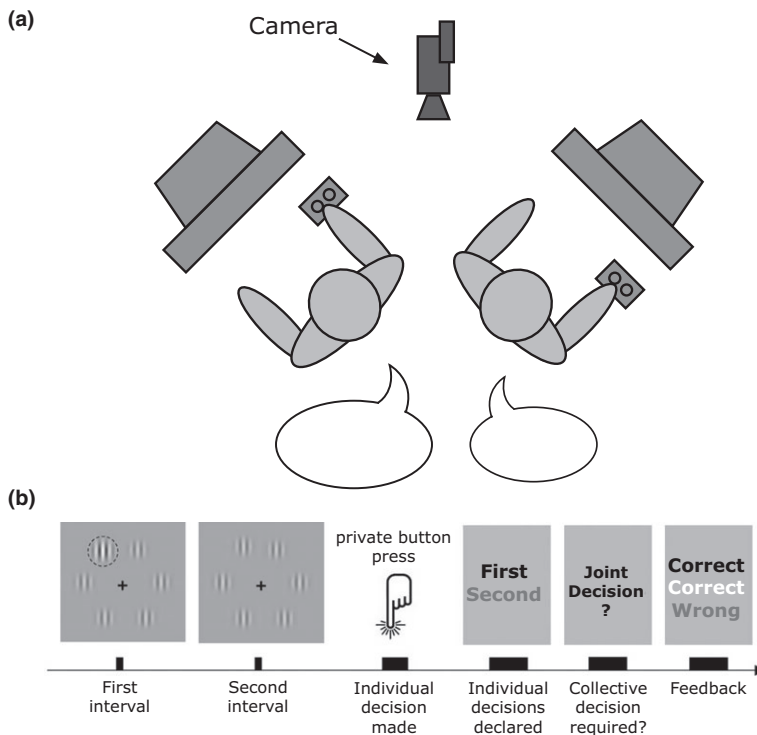

Fig. 2. Experimental paradigm: (a) Experimental setup. (b) Schematic illustration of a typical trial (Adapted with permission from Fusaroli et al., 2012).

the same answer, they would automatically proceed to the next trial. However, if their individual choices disagreed, they were prompted to negotiate, by freely discussing with each other, a joint decision. There was no time or other constraints on the joint decision dialogs.

The videos were coded and transcribed by five research assistants naive to the purpose of the study. Three different data sets were extracted from the videos (cf. Fig. 3): (a) The lexical data set consisted in the transcription at the word level of all the joint decision dialogs. (b) The prosodic data set was produced by sampling the fundamental frequency of the speech data at 100 times per second (every 10 ms), filtering it to only select human conversational voice range (75–600 Hz) employing Praat (Boersma, 2001) and correcting for octave jumps and other artifacts. (c) The speech/pause data set was produced by extracting speech/pause sequences from the speech data sampling at 100 times per second (every 10 ms). Pauses were defined as absence of pitch longer than 200 ms (Kousidis et al., 2008). In order to test self-consistency and interactive alignment, we separated each interlocutor's speech behaviors. The interlocutors' speech turns start and end times were manually tagged through careful listening supplemented by an in-depth examination of the waveform. The manual tags were then adjusted to a 10 ms precision scale through an automated analysis of pitch presence/absence and intensity changes using Matlab (Mathworks Inc., Natick, MA).

4.2. Data analysis

4.2.1. Task performance

We used psychometric functions to assess accuracy across the different degrees of visual contrast following standard procedures of psychophysical research (Bahrami et al., 2010). Psychometric functions were estimated for each interlocutor and for each dyad by plotting per each contrast level the proportion of trials in which the oddball was reported in the second stimulus display. Oddballs in the first stimulus display are attributed negative contrast, and oddballs in the second display positive contrast. Individual interlocutors' functions were calculated relying on the individual decisions that constituted the first part of each trial. Collective psychometric functions were calculated using all joint decisions (collapsed across trials on which interlocutors agreed and disagreed). The psychometric data were fitted to a cumulative Gaussian function whose parameters were bias,
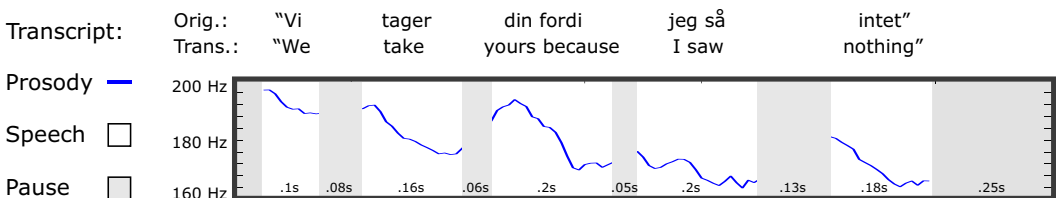

Fig. 3. Illustration of the three properties of linguistic behavior analyzed: (a) word-level transcription, (b) fundamental frequency/prosody, and (c) speech/pause sequence.

b, and variance, σ^2, using a probit regression model created with the glmfit function in MATLAB (Mathworks, Inc.). The cumulative Gaussian function, $P(\Delta c)$, was defined as follows:

$$P(\Delta c) = H\left(\frac{\Delta c + b}{\sigma}\right), \tag{1}$$

where $H(z)$ is the cumulative Normal function,

$$H(z) = \int_{-\infty}^{z} \frac{dt}{(2\pi)^{1/2}} \exp\left[-\frac{t^2}{2}\right] \tag{2}$$

A psychometric function with bias *b* and variance σ^2 would be denoted $P(\Delta c)$, where Δc is the contrast difference between the second and first presentations. Following standard practice, the psychometric curve, $P(\Delta c)$, corresponds to the probability of choosing the second stimulus. Thus, a positive bias indicates an increased probability of saying that the second stimulus had higher contrast. The estimated variance is related to the maximum slope of the psychometric curve, denoted *s*, via

$$S = \frac{1}{(2\pi\sigma^2)^{1/2}} \tag{3}$$

A large slope corresponds to highly sensitive performance. Using the slope measure, we quantified the sensitivity of the individuals and of the dyad. We defined "collective benefit" as the ratio of the dyad's slope to that of the more sensitive dyad member. A collective benefit value above 1 indicated successful cooperation, and values below 1 indicated that collaboration was counterproductive and that the dyad did worse than its more sensitive member.

4.2.2. Quantifying conversational dynamics

Like many behavioral time series, our data sets are characterized by non-stationarity; that is, the mean and the standard deviation of the time series change over time (Dale & Spivey, 2006; Kello et al., 2008; Spivey, 2007). We therefore employed a non-linear method—RQA—developed to quantify dynamical systems and their trajectories, capturing aspects of the conversational dynamics that would otherwise be lost due to averaging in a more traditional correlation analysis. To establish an index of structural organization of the different aspects highlighted by each of the approaches (interactive alignment, interpersonal synergy, and individual self-consistency), we performed three sets of analyses on each of the data sets—lexical transcript, fundamental frequency, and speech/pause sequence: (a) cross-RQA on cross recurrence plots of dyads' speech production, (b) RQA on the recurrence plot of the speech production of the dyad as one coherent time series, not discriminating between interlocutors, and (c) RQA of the speech production of

individual dyad members. Transcripts and speech/pause sequences were treated as nominal sequences, fundamental frequency as a numeric sequence. In order to control for recurrence due to the incidental distribution of the values in the time series, we performed the same analyses on control data sets generated by randomly shuffling the time series. Analyses were performed using the CRP MATLAB toolbox (Marwan et al., 2007).

4.2.2.1. Cross recurrence quantification analysis (CRQA): CRQA is a non-linear analog of cross-correlation used to explore the shared dynamics of two systems, in our case the amount and structure of alignment between interlocutors (Fusaroli, Konvalinka, et al., 2014; Shockley, Butwill, Zbilut, & Webber, 2002). In order to apply CRQA, two steps are necessary: reconstructing of the phase space underlying the time series and production of a cross recurrence plot.

The phase space underlying two time series is an *n*-dimensional space in which all the possible combined states of the two systems are represented, which makes it possible to portray the shared trajectories of the systems' behaviors whether they are periodic (repeatedly crossing the same regions at regular intervals), random, or chaotic. In order to respect the temporal structure of turn taking without introducing regularity artifacts, in each individual's time series, we "blanked out" sequences corresponding to the interlocutor's utterances so that they would not generate recurrence. We then reconstructed the phase space using the time-delay method (Takens, Rand, & Young, 1981). For the analysis of the transcript we employed an embedding dimension—that is the number of dimensions characterizing the phase space—of 3, which minimizes noise from letter pairs recurrence and roughly captures basic Danish morphological structure (Orsucci, Giuliani, & Webber, 2006). For fundamental frequency and speech/pause series, the embedding dimension was estimated using the *false nearest neighbors* algorithm, where the parameter was increased in integer steps until the recruitment of nearest neighbors did not significantly decrease (Abarbanel, 1996). This yielded a value of 4 for fundamental frequency and 2 for the speech/pause. When defining the time-delay, we employed 1 for the transcripts (Dale, Warlaumont, & Richardson, 2011; Orsucci et al., 2006). For fundamental frequency and pause/speech, we estimated time-delay as the first local minimum of the *mutual average information* function of the time series (Abarbanel, 1996). This yielded a value of 4 for fundamental frequency and 1 for speech/pause. Last, a radius (or threshold) was chosen, which represents the radius of the neighborhood in which recurrent states are identified, that is, a criterion of similarity: how close in the phase space two points have to be in order to be conceived as a recurrence of the same state. We chose 0 for transcripts and speech/pause: Given the nominal nature of the time series, only the exact same values should register as recurrence. For fundamental frequency, we chose the threshold separately for each data set, such that it corresponded to a fixed recurrence rate of 4% (Marwan et al., 2007). The recurrence rate represents the percentage of data that are recurrent (i.e., the percentage of dark dots in the recurrence plot). Keeping it fixed allows for the comparison between data sets and assessment of the quality (e.g., extent and complexity) of recurrence.
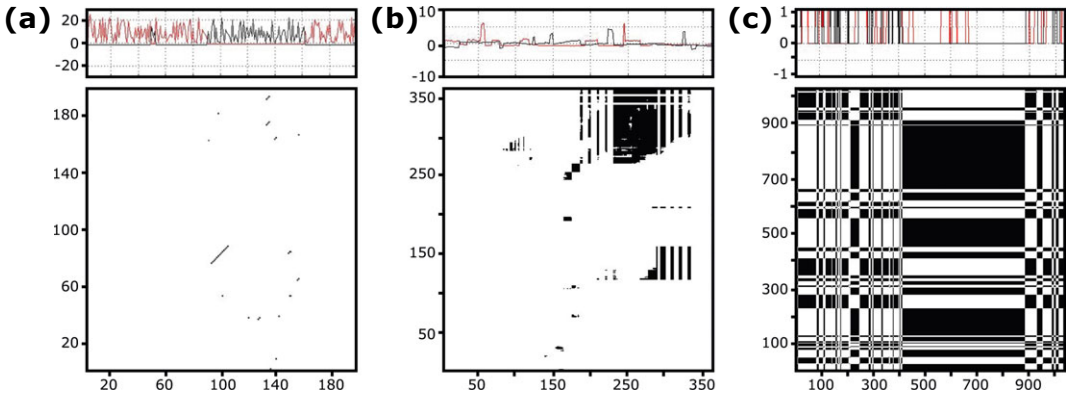
Fig. 4. Example of three cross recurrence plots representing structural organization across the speech production of interlocutors in two joint decisions analyzed at three levels: (a) word-level transcription, (b) fundamental frequency, and (c) speech/pause sequence.

After reconstructing the phase space, we constructed cross recurrence plots (cf. Fig. 4). Black dots on the plots represent every occasion at which a phase space trajectory goes through approximately the same region in the phase space. In mathematical terms, if we represent the trajectory of a system as

$$\{\overrightarrow{x_i y_i}\}_{i=1}^{N} \tag{4}$$

the corresponding recurrence plot is based on the following recurrence matrix:

$$R_{i,j} = \begin{cases} 1 : \overrightarrow{x_i} \approx \overrightarrow{y_j}, \\ 0 : \overrightarrow{x_i} \not\approx \overrightarrow{y_j}, \end{cases} \quad i,j = 1,\ldots,N \tag{5}$$

where N is the number of considered states of the system and $\overrightarrow{x}_i \approx \overrightarrow{y}_j$ indicates that the two systems' states are equal up to an error (or distance) ε. Note that this ε is essential in case of continuous variables (as in fundamental frequency) since systems mostly revisit the same regions of possible states, but rarely repeat the exact same trajectories twice.

To statistically analyze differences in conversational dynamics across dyads, we performed CRQA on the cross recurrence plots. RQA provides several indices quantifying the structure and complexity of coupling between dynamical systems from cross recurrence plots (Marwan et al., 2007). This makes it possible to statistically compare different coupled systems (e.g., different dyads) in terms of their dynamics: for instance, the stability, structure, and complexity in the shared behavior of the coupled elements. Given our interest in the extent of regularities (that is, the length of repeated patterns) and in the complexity of those regularities, we chose to employ two indices: the average diagonal line length (L) and the entropy of the diagonal line lengths distribution (ENTR)

of each data series. L calculates the average length of the diagonal lines parallel to the main diagonal:

$$L = \frac{\Sigma_{l=l_{min}}^{N} lP(l)}{\Sigma_{l=l_{min}}^{N} P(l)} \tag{6}$$

In other words, L calculates the average length of a recurrent trajectory of the system and is a measure of the extent to which a system is structured. Entropy is a measure of the complexity of structure of the system and is calculated as the Shannon information entropy of the length of diagonal lines (Marwan et al., 2007):

$$ENTR = -\sum_{l=l_{min}}^{N} p(l) \ln p(l) \tag{7}$$

In summary, L is a measure of how strongly patterned and regular a behavior is: High L indicates long repeated patterns, that is, stable coordination, while low L indicates short repeated patterns. For instance, in the analysis of the transcripts, high L corresponds to repetitions of long stretches (more than one word, or even full sentences) and low L to repetitions of only short stretches, such as words or morphemes. ENTR is a measure of the complexity of these patterns: High entropy indicates the presence of a plurality of patterns; low entropy indicates the predominance of one kind of pattern. In the analysis of the transcripts, high ENTR corresponds to repetitions of a wide range of lexical patterns: Sometimes a full sentence is repeated, sometimes only a word. Low ENTR corresponds to a more rigid structure of repetitions with patterns all of the same length.

4.2.2.2. Recurrence quantification analysis: RQA is the exact analog of CRQA except that it applies to recurrence plots where the same time series is represented on both *x*- and *y*-axis. In other words, if CRQA could be seen as a non-linear generalization of cross-correlation, RQA generalizes auto-correlation. As interpersonal synergies are expressed as regularities at the level of the interaction, the time series on the plot correspond to the speech production of the dyad as one coherent time series, not discriminating between interlocutors (cf. Fig. 5).

Analogously, self-consistency can be assessed by constructing recurrence plots of the speech production of individual interlocutors (cf. Fig. 6).

We used RQA to evaluate recurrence in interactions (interpersonal synergies) and individuals (self-consistency) with respect to the two indices previously used: L and ENTR.

4.2.3. Structural organization and shuffled control

In order to ensure that the recurrence observed is not due to the incidental distribution of values in the time series but on their actual temporal sequence, we employed indepen-
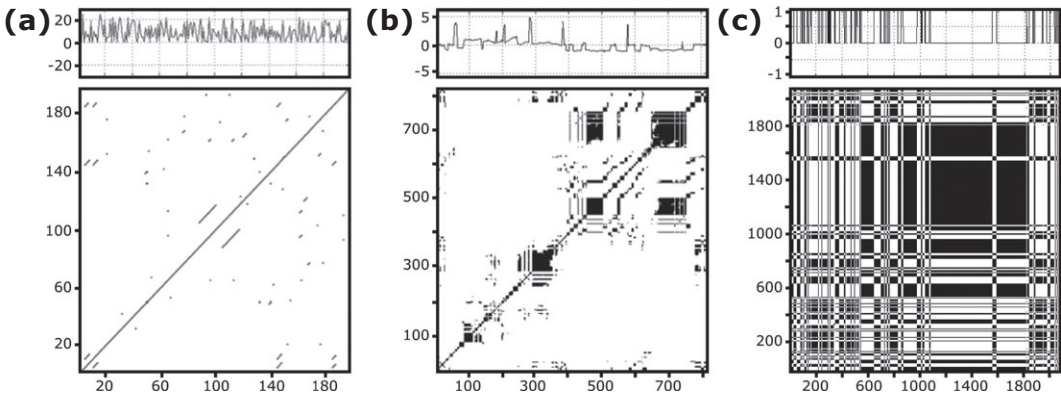
Fig. 5. Example of three recurrence plots representing structural organization in the overall conversation (not discriminating speech turns and interlocutors) in two consecutive joint decision trials analyzed at three levels: (a) word-level transcription, (b) prosodic, and (c) speech/pause sequence.
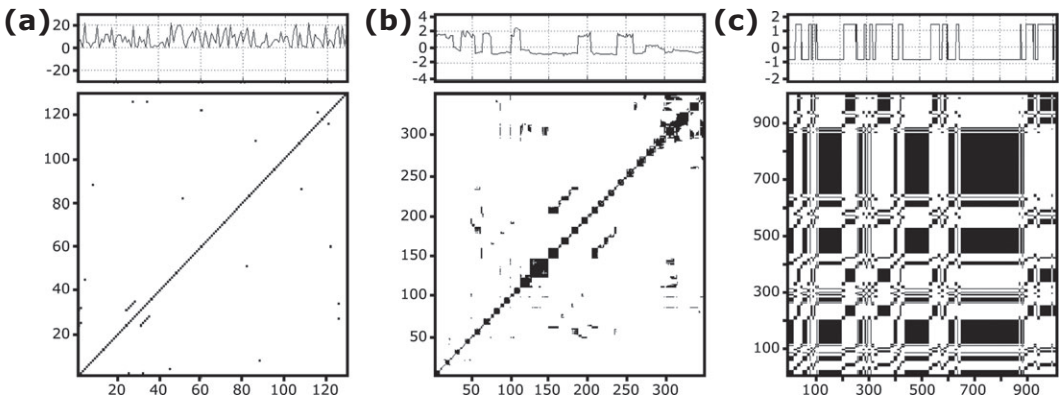

Fig. 6. Example of three recurrence plots representing the structural organization of individuals' speech production in two consecutive joint decision trials analyzed with respect to three properties: (a) word-level transcription, (b) fundamental frequency, and (c) speech/pause sequence.

dent samples t tests to compare the results from the analysis of the actual data sets with those achieved from shuffled controls (Dale & Spivey, 2006).[2]

4.2.4. Structural organization and performance

In order to comparatively assess the predictive power of the aspects highlighted by the two approaches to dialog and the baseline, we employed two complementary analyses. First, we ran three multiple linear regressions per each of the measures (lexical choices, prosody, speech/pause): the first—the interactive alignment approach—included the two indices of structural organization of interactive alignment (L and ENTR of cross-RQA) as predictors; the second—the interpersonal synergy

approach—included the two indices of structural organization of the overall conversation (L and ENTR of overall RQA) as predictors; and the third—the individual self-consistency control baseline—included the two indices of individual structural organization (L and ENTR of individual RQA) as predictors for the collective benefit achieved by the dyads. In order to make individual recurrence comparable to cross and overall recurrence—to create models with comparable degrees of freedom—we selected in each dyad the data sets from the best performing individual interlocutor.[3] We then employed Adjusted R-squared ($AdjR^2$) and Bayesian Information Criterion (BIC) to assess the statistical models developed in each analysis as well as their fit to the data. $AdjR^2$ is a measure of the variance explained by the predictors, and a higher value indicates a better statistical model (Hastie, Tibshirani, & Friedman, 2009). BIC is a measure of the relative goodness of fit of a statistical model—its likelihood according to the data—and a lower value indicates a better model (Schwarz, 1978). $AdjR^2$ and BIC provide complementary information on the likelihood and adequacy of the model in predicting the dependent variable. In order to assess the complementarity of the information of the different aspects—that is, the degree to which we gain more information by combining indexes from the different approaches—we then ran the second analysis: a forward selection stepwise linear regression for each of the measures (lexical, prosodic, speech/pause) including the six indices of structural organization as predictors. The stepwise regression selects a minimal set of indices explaining the maximal amount of variance in the dependent variable, in this case collective performance.

Finally, to investigate the respective role of lexical, prosodic, and speech/pause behaviors, we first ran a statistical model comparison of the multiple linear regressions of each approach (relying on $AdjR^2$ and BIC) including all six indices of structural organization, that is, two per measure. We then ran a forward selection stepwise linear regression including all 18 indices of structural organization (six per measure) at once to reveal the most informative combination of independent variables.

All analyses were run in Matlab (Mathworks, Inc.) relying on the Statistics, Bioinformatics, and CRP toolboxes.

5. Results

5.1. Structural organization and shuffled controls

The conversations displayed a significantly higher degree of structural organization with respect to the relevant measures than the shuffled time series, indicating that recurrent patterns in the conversation data are actually due to their temporal structure and not the incidental distribution of their values (cf. Table 1).

Table 1
Significant comparisons ($p < .0001$) between dialog materials and shuffled controls

RQA Input	Predictors	Lexicon		Prosody		Speech/Pause	
		t	*d*	*t*	*d*	*t*	*d*
Individual	L	10.71	0.88	16.26	0.90	18.77	0.92
	ENTR	12.47	0.91	18.97	0.92	23.39	0.95
Cross	L	14.59	0.93	8.71	0.84	14.99	0.94
	ENTR	18.39	0.96	11.15	0.89	41.45	0.99
Overall	L	19.33	0.96	10.96	0.89	17.01	0.95
	ENTR	52.08	0.99	14.35	0.93	20.46	0.97

5.2. *Structural organization* within *properties of linguistic behavior*

5.2.1. *Lexical choices*

Individually tested models revealed significant fit for interactive alignment ($r = .50$, $AdjR^2 = 0.25$, $p = .01$, $BIC = 1.98$, both L and ENTR with a positive coefficient) and interpersonal synergies ($r = .73$, $AdjR^2 = 0.45$, $p = .002$, $BIC = -2.73$, both L and ENTR with a positive coefficient), but not for self-consistency ($r = .11$, $AdjR^2 = 0.01$, $p = .9$, $BIC = 6.34$). In the forward stepwise model, only ENTR of the structural organization of interpersonal synergies was selected as a significant predictor ($r = .73$, $AdjR^2 = 0.47$, $p = .01$): the higher the ENTR—that is, the more complex and flexible the recurrence in the transcript—the better the performance.

5.2.2. *Prosody*

Individually tested models revealed significant fit for interactive alignment ($r = .40$, $AdjR^2 = 0.31$, $p = .03$, $BIC = 0.64$, both L and ENTR with a positive coefficient) and interpersonal synergies ($r = .40$, $AdjR^2 = 0.31$, $p = .04$, $BIC = 0.63$, both L and ENTR with a positive coefficient), but not for self-consistency ($r = .10$, $AdjR^2 = -0.04$, $p = .6$, $BIC = 7.19$). In the forward stepwise model, only ENTR of the structural organization of interpersonal synergies was selected as a significant predictor ($r = .67$, $AdjR^2 = 0.39$, $p = .02$): the higher the ENTR, that is, the more complex and flexible the recurrence in prosody, the better the performance.

5.2.3. *Speech/pause*

Individually tested models revealed significant fit for interpersonal synergies ($r = .63$ $AdjR^2 = 0.31$ $p = .03$, $BIC = 0.69$, both L and ENTR with a positive coefficient), but not for interactive alignment ($r = .15$, $AdjR^2 = -0.12$, $p = .7$, $BIC = 8.52$) and self-consistency ($r = .19$, $AdjR^2 = -0.11$, $p = .22$, $BIC = 8.31$). In the forward stepwise regression, only L of the structural organization of interpersonal synergies was selected as a significant predictor ($r = .62$, $AdjR^2 = 0.34$, $p = .01$): the higher the L—that is, the longer the average recurring patterns in the speech/pauses sequences—the better the performance. (see Fig. 7a)
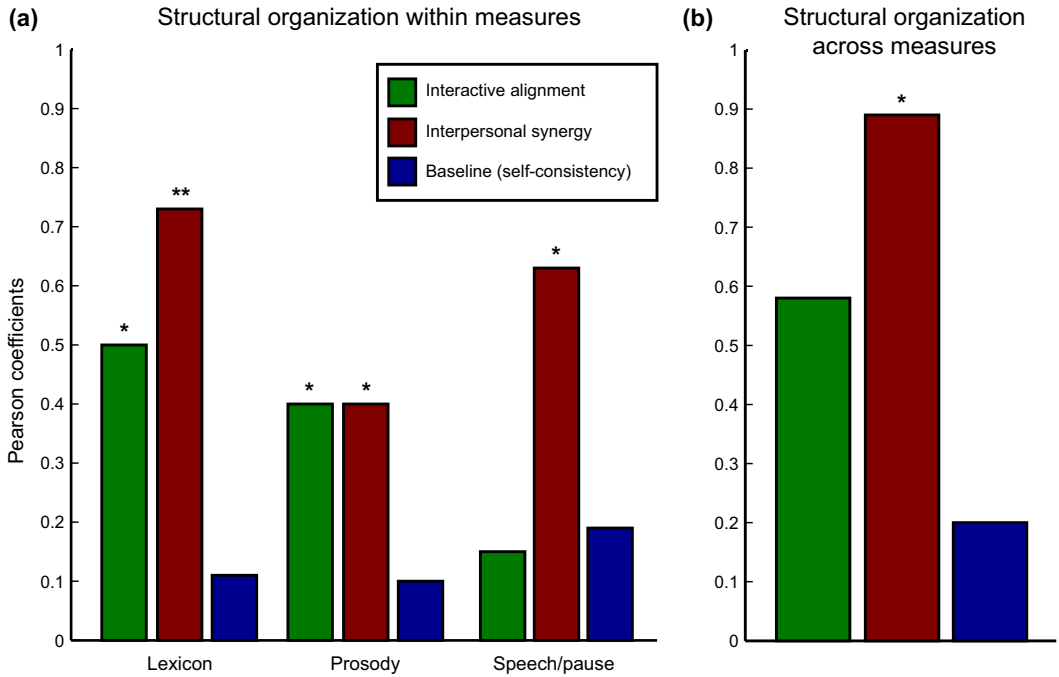
Fig. 7. Predicting collective performance: A comparison of the Pearson coefficients of interactive alignment, interpersonal synergy, and the self-consistency base line to (a) lexical transcript, prosody, and speech/pause sequences; (b) all measures. "*" indicates a p-value below 0.05, "**" a p-value below 0.01.

5.3. Structural organization across properties of linguistic behavior

Including all measures in the same analysis revealed significant fit for interpersonal synergies ($r = .89$, $AdjR^2 = 0.65$, $p = .01$, $BIC = -2.86$), but not for interactive alignment ($r = .58$, $AdjR^2 = 0.303$, $p = .17$, $BIC = 7.91$) and self-consistency ($r = .20$, $AdjR^2 = -0.34$, $p = .3$, $BIC = 18.37$) (see Fig. 7b).

In the forward stepwise regression, only predictors related to interpersonal synergies were selected as significant: ENTR of the lexical time series, L of speech/pause, and ENTR of prosody ($r = .97$, $AdjR^2 = 0.92$, $p < .000001$). For all variables, a higher value corresponds to a better performance.

6. Discussion

6.1. Summary

This study was designed to assess two approaches to interpersonal processes in conversation—interactive alignment, interpersonal synergy, and a baseline of individual

self-consistency—in relation to a corpus of empirical data. Compared to shuffled controls, the conversations analyzed display significant structural organization for all linguistic properties in all approaches—including the control baseline. Both the alignment and synergy approach are more informative than the baseline as to the efficacy of linguistic coordination. However, only structural organization pertaining to interpersonal synergies consistently predicts the coordinative performance of the dyads, and adding predictors from interactive alignment does not significantly improve predictions. Lexical choices, prosody, and speech/pause all contribute significantly to predicting collective benefit. Overall the findings support the idea that interpersonal dynamics across different linguistic aspects are crucial for conversations. They also suggest that, at least in task-oriented conversations, interpersonal synergies, such as procedural scripts and routines, play a crucial role. We speculate that processes at this interactional level of organization might work to guide and constrain other central linguistic processes such as alignment and self-consistency to meet the affordances of the task at hand.

6.2. Interpersonal synergies and task constraints

Based on qualitative approaches, conversation analysis and usage-based linguistics have long advocated for the importance of routines and complementary patterns in conversation. Using a set of novel methods, we provide quantitative support for these positions. Beyond patterns in individuals' linguistic behaviors and reciprocal imitation, effective organization of the interaction is found in the way interlocutors enact interactional routines, which involve complementary contributions and roles. In a recent study, Mills (2011) points out how procedural conventions and routines are often implicitly established as sequences of conversational moves in the context of repeated collective tasks. Analogously, we analyzed the decision-making routines our participants developed to effectively solve the joint decision task. As previously described, pairs of interlocutors have to make joint decisions by sharing confidence in their previous individual decisions, so to be able trial-by-trial to follow the more confident individual. This presents two problems: (a) expressing and comparing confidence; and (b) agreeing on the decision. Closely reviewing the transcripts, we observe that during early trials participants express and iteratively adjust their levels of confidence before they arrive at a consensus regarding the decision. However, through repeated trials most dyads jointly establish an increasingly schematic routine enabling them to enact highly elliptical conversations as they proceed through the later trials. Table 2 portrays the transition from intermediate to late conversations (we did not report very early ones due to their excessive length).

The example shows how, in a representative well-performing pair, the initial redundancy of moves (repeated expressions of confidence) and lexical alignment ("I saw… I saw", "I don't know… I don't know", etc.) is noteworthy. This redundancy and repetition is instrumental in developing a shared lexical scale for the comparison of individual levels of confidence (Fusaroli et al., 2012). However, over time, participants evolve a routine: a stable conversational pattern in which an initial expression of confidence leads directly to the decision. While the overall structure of the routine is relatively stable, the

Table 2
Excerpts of intermediate and late interactions in a well-performing pair

Intermediate Interactions	
Interaction 43	Interaction 43
A: vi tager din, for jeg så intet	A: we take yours, because I saw nothing
B: jeg så heller noget, jeg så…	B: I didn't see anything either, I saw …
A: jeg satsede!	A: I took a bet!
B: sådan!	B: way to go!
A: sådan!	A: way to go!
B: det var også et sats det jeg lavede der	B: mine was also just a bet there
Interaction 44	Interaction 44
B: jeg ved det ikke	B: I don't know
A: jeg ved det heller ikke, jeg så—jeg så både i venstre hjørne og midt for til højre på dem begge	A: I don't know either, I saw—I saw something both in the left corner and in the center on the right in both of them
B: okay jeg s-, jeg synes det var ovre i venstre side, men—øhm—pas	B: okay I s-, I think it was over in the left side, but—uhm—I'll pass
A: nej!	A: no!
B: vi ødelægger scorerne—nu må vi lige…	B: we ruin the scores—now we must…
A: ja nu må vi lige tage os sammen	A: yeah, now we must pull ourselves together
Late Interactions	
Interaction 71	Interaction 71
A: Okay, jeg så kun et blink—vi tager din, så	A: Okay, I only saw a blink—we take yours, then
B: ja	B: yes
Interaction 72	Interaction 72
A: nede i venstre hjørne	A: down the left corner
B: ja [laughs]	B: yes [laughs]
A: Tag den! Tag den! Tag den!	A: Take it! Take it! Take it!
B: ja da! vi saa gode!	B: yes! We are so good!
Interaction 73	Interaction 73
B: oppe i venstre hjoerne	B: up in the left corner
A: ja [laughs]	A: yes [laughs]

interlocutors contribute the parts in interchangeable and complementary ways. The interlocutor expressing confidence in one trial might be the one taking the decision in the next (see Table 2, interaction 72 and 73), or one interlocutor might occasionally do both (interaction 71). These observations can help us explain why interactive alignment and self-consistency seem less informative about coordinative efficacy. The flexible enactment of the routine creates recurrent patterns, but at the level of the interaction as a whole and thus is not fully reflected by the alignment or self-consistency approach. Similarly, the examples can inform intuitions on why the synergic flexibility (ENTR) of lexical behavior is strongly associated with performance: It would seem that effective interlocutors do not just repeat fixed sets of expressions, but modulate and adapt their patterns with sensitivity to, for instance, the changing degrees of difficulty of the task and their degree of confidence (Bang et al., 2014; Fusaroli et al., 2012). In other words, it is not enough to

repeat a particular set of routinized expressions (which would relate to the index L), as that would be too rigid a strategy to adapt to the varying task affordances (task difficulty and degree of confidence). The participants have to flexibly engage in repetitions, varying and developing the expressions over time in response to the ever-changing task environment (ENTR).

6.3. Interactive alignment, interpersonal synergies, and individual self-consistency

It is important to specify that our findings mainly support the interpersonal synergy approach to dialog from an epistemic perspective; in other words, the analyses allow us to answer questions about which indices and levels of description are more informative in the assessment of linguistic coordination. The findings do not exclude a role for structural organization due to interactive alignment and self-consistency. Rather, they suggest that such a role has to be found in the ways alignment and individual behaviors combine to be shaped by the emergent structures and dynamics operating at the level of the interaction itself. This is, indeed, one of the key points of the interpersonal synergy approach to conversation: Inspired by dynamical models of motor coordination, it portrays the way interacting components give rise to functional assemblies not observable or predictable by investigating the individual components themselves (Fusaroli, Gangopadhyay, & Tylén, 2014; Fusaroli, Raczaszek-Leonardi, et al., 2014; Kelso, 2009; Latash, 2008; Riley et al., 2011). The synergy approach aims at explaining why conversations, despite complex timing and the need to coordinate a multitude of possibilities for action and reaction, seem so easy for individuals to enact: How do we manage, at the same time, to observe and produce gestures, linguistic forms, intonations, nods, facial expressions, and bodily postures, in such a smooth and seamless fashion? While the interactive alignment approach suggests that we do so by automatic reciprocal alignment (Garrod & Pickering, 2004), the synergy approach argues that alignment is but one of a variety of mechanisms we employ to manage the complexity of dialogical interaction: We also entrain *across* modalities, assume complementary roles, and enact interactional routines. Indeed, while ample evidence of interactive alignment has been presented (for reviews, cf. Fusaroli & Tylén, 2012; Louwerse et al., 2012; Pickering & Garrod, 2004), there have also been growing concerns about relying on a simple model of interactive alignment as the primary engine for coordination (Dale et al., 2013; Fusaroli, Raczaszek-Leonardi, et al., 2014; Mills, 2014; Reich, Berman, Dale, & Levitt, 2014). For instance, it has been shown that indiscriminate alignment can be detrimental for social coordination (Fusaroli et al., 2012; Healey, Purver, & Howes, 2014; Wang & Hamilton, 2012). Such findings suggest that for alignment to constructively contribute to coordination, it has to be modulated according to contents and contexts. We speculate that rather than serving as the primary engine of linguistic coordination, alignment might be better conceived as a preliminary tool, providing the means to quickly establish common ground and interactional routines (Fusaroli, Raczaszek-Leonardi, et al., 2014; Louwerse et al., 2012; Mills, 2011), while its impact decreases once a synergy is established (Fernández & Grimm, 2014; Fusaroli et al., 2012, 2013; Mills, 2014; Mills & Healey, 2008). The synergy approach thus predicts several

interacting mechanisms through which interlocutors' behaviors become increasingly interdependent. In effect, a superordinate structure of interaction emerges, which constrains complexity and makes individual participation more manageable. In dynamical systems terms, this phenomenon has been designated *soft-assembly* (Kello & Van Orden, 2009). Importantly, the notion of soft-assembly emphasizes that while a superordinate structural organization of the conversation does constrain and re-organize individuals' behaviors, it does not determine them nor deprive them of their autonomy: Individuals can flexibly engage and disengage from the interaction, as well as become part of other soft-assemblies (De Jaegher, Di Paolo, & Gallagher, 2010; Di Paolo & De Jaegher, 2012).

6.4. Quantifying interpersonal dynamics in dialog

This is the first study to systematically quantify and assess different interpersonal conversational mechanisms in different properties of linguistic behavior. It paves the way for the more systematic exploration and understanding of the different components and mechanisms constituting conversational coordination. We argue that synergies and alignment capture different aspects of the dynamics of dialog selectively employed to meet different task constraints. Our findings thus strongly suggest that there is no single mechanism that makes conversations easy and effective. On the contrary, a multiplicity of social mechanisms and environmental (task) constraints work in concert to organize the complex coordinative behaviors of conversation, be it a casual chat, deception, a heated argument, or joint decision-making (Duran, Dale, Kello, Street, & Richardson, 2013; Fusaroli, Raczaszek-Leonardi, et al., 2014; Paxton & Dale, 2012, 2013). The methods applied in this study provide the tools to further disentangle different conversational aspects, their relative presence and impact. Importantly, additional experiments will be required to investigate if synergic structural organization is merely a by-product of effective coordination or if it plays a causal role in dialog and coordination. This study has demonstrated the utility of dynamical systems approaches that can be applied across different aspects of the interaction. As such, it paves the way for systematic explorations and experimental manipulations of conversational coordination.

Notes

1. The same corpus is analyzed with different methods and purpose in Fusaroli et al. (2012).
2. In the study of interpersonal coordination it is sometimes suggested to employ surrogate pairs, that is, pairing of individuals from different dyads as controls. In the case of conversations, intercalating utterances from different interactions would not respect the temporal structure of the conversation (i.e., the length of the replaced speech turns would be different), therefore breaking the mathematical assumptions of RQA (Fusaroli, Konvalinka, et al., 2014; Marwan, 2011; Marwan et al., 2007). Additionally, simulations reported in Louwerse et al. (2012) show that, at least for nominal time series, shuffled data sets provide a more conservative control, that is, higher scores in the recurrence parameters, than surrogate pairs.
3. Analogous control analyses were run on (a) data constituted by repeatedly picking a random interlocutor from each dyad; (b) the average values between the two interlocutors. The results were the same (models with a non-significant fit to the data) as in the reported analysis.

References

Abarbanel, H. D. I. (1996). *Analysis of observed chaotic data*. New York: Springer.

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2014). Complexity matching in dyadic conversation. *Journal of Experimental Psychology. General*, *143*(6), 2304–2315. Advance online publication.

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.

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science*, *329*, 1081–1085.

Bang, D., Fusaroli, R., Tylén, K., Olsen, K., Latham, P., Lau, J. Y. F., Roepstorff, A., Rees, G., Frith, C. D., & Bahrami, B. (2014). Does interaction matter? Testing whether a confidence heuristic can replace interaction in collective decision-making. *Consciousness and Cognition*, *26*, 13–23.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*, 614–636.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glottology International*, *5*, 341–345.

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

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology. Learning, Memory and Cognition*, *22*(6), 1482–1493.

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. *Topics in Cognitive Science*, *1*, 274–291.

Bryant, G. A. (2010). Prosodic contrasts in ironic speech. *Discourse Processes*, *47*, 545–566.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*, 893–910.

Chemero, A. (2009). *Radical embodied cognitive science*. Cambridge, MA: MIT Press.

Clark, H. H. (1996). *Using language*. New York: Cambridge University Press.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, *13*, 259–294.

Dale, R., Fusaroli, R., Duran, N., & Richardson, D. C. (2013). The self-organization of human interaction. *Psychology of Learning and Motivation*, *59*, 43–95.

Dale, R., & Spivey, M. J. (2006). Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning*, *56*, 391–430.

Dale, R., Warlaumont, A. S., & Richardson, D. C. (2011). 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? *Trends in Cognitive Sciences*, *14*, 441–447.

Di Paolo, E., & De Jaegher, H. (2012). The interactive brain hypothesis. *Frontiers in Human Neuroscience*, *6*, 163.

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

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.

Duran, N., Dale, R., Kello, C., Street, C. N. H., & Richardson, D. (2013). Exploring the movement dynamics of deception. *Frontiers in Psychology*, *4*, 140.

Enfield, N. J. (2013). *Relationship thinking: Agency, enchrony, and human sociality*. New York: Oxford University Press.

Fernández, R., & Grimm, R. (2014). Quantifying categorical and conceptual convergence in child-adult dialogue. In P. Bello, M. Guarini, M. McShane, & B. Scassellati (Eds.), *Proceedings of the 36th Annual Conference of the Cognitive Science Society (CogSci 2014)*. Quebec City, Canada: Cognitive Science Society.

Fusaroli, R., Abney, D. H., Bahrami, B., Kello, C., & Tylén, K. (2013). Conversation, coupling and complexity: Matching scaling laws predict performance in a joint decision task. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of 35th Annual Conference of the Cognitive Science Society*. Berlin: Cognitive Science Society.

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*, 931–939.

Fusaroli, R., Gangopadhyay, N., & Tylén, K. (2014). The dialogically extended mind: Making a case for language as skilful intersubjective engagement. *Cognitive Systems Research*, *29–30*, 31–39.

Fusaroli, R., Konvalinka, I., & Wallot, S. (2014). Analyzing social interactions: Promises and challenges of cross recurrence quantification analysis. *Springer Proceedings in Mathematics & Statistics*, *103*, 137–155.

Fusaroli, R., Raczaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157.

Fusaroli, R., & Tylén, K. (2012). Carving language for social coordination: A dynamic approach. *Interaction Studies*, *13*, 103–123.

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., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*, 8–11.

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology*. New York: Irvington.

Goodwin, C. (2011). Building action in public environments with diverse semiotic resources. *Versus*, *112-113*, 169–182.

Goodwin, C., & Heritage, J. (1990). Conversation analysis. *Annual Review of Anthropology*, *19*, 283–307.

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer.

Healey, P., Purver, M., & Howes, C. (2014). Divergence in dialogue. *PLoS ONE*, *9*(6).

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

Kello, C. T., Brown, G. D., Ferrer, I. C. R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., & Van Orden, G. C. (2010). Scaling laws in cognitive sciences. *Trends in Cognitive Sciences*, *14*, 223–232.

Kello, C. T., & Van Orden, G. C. (2009). Soft-assembly of sensorimotor function. *Nonlinear Dynamics in Psychology and the Life Sciences*, *13*, 57–78.

Kelso, J. A. S. (2009). Synergies: Atoms of brain and behavior. In D. Sternad (Ed.), *Progress in motor control: A multidisciplinary perspective*. New York: Springer.

Konvalinka, I., & Roepstorff, A. (2012). The two-brain approach: How can mutually interacting brains teach us something about social interaction? *Frontiers in Human Neuroscience*, *6*, 215.

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjodt, U., Jegindo, E. M., Wallot, S.,..., Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *PNAS*, *108*, 8514–8519.

Kousidis, S., & Dorran, D. (2009). *Monitoring convergence of temporal features in spontaneous dialogue speech*. Dublin: UCD Working Papers.

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. In D. Burnhan (Ed.), *Proceedings of Interspeech 2008*. Brisbane, Australia: International Speech Communications Association.

Krauss, R. M., & Fussell, S. R. (1996). Social psychological models of interpersonal communication. In E. T. Higgins, & A. Kruglanski (Eds.), *Social psychology: A handbook of basic principles* (pp. 655–701). New York: Guilford.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology*, *4*, 343–346.

Kuznetsov, N. A., & Wallot, S. (2011). Effects of accuracy feedback on fractal characteristics of time estimation. *Frontiers in Integrative Neuroscience*, *5*, 62.

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

Levinson, S. (2006). Cognition at the heart of human interaction. *Discourse Studies*, *8*, 85–85.

Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. In R. Pieraccini & A. Colombo (Eds.), *Proceedings of Interspeech 2011*. Brisbane: International Speech Communications Association

Linell, P. (1998). *Approaching dialogue: Talk, interaction and context in dialogical perspectives*. Amsterdam: John Benjamins Publishing Company.

Linell, P. (2005). *The Written language bias in linguistics: Its nature, origins and transformations*. New York: Routledge.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, *36*, 1404–1426.

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. (2011). How to avoid potential pitfalls in recurrence plot based data analysis. *International Journal of Bifurcation and Chaos*, *21*, 1003–1017.

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

McFarland, D. H. (2001). Respiratory markers of conversational interaction. *Journal of Speech, Language and Hearing Research*, *44*, 128–143.

Mills, G. (2011). The emergence of procedural conventions in dialogue. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Boston, MA: Cognitive Science Society.

Mills, G. (2014). Dialogue in joint activity: Complementarity, convergence and conventionalization. *New Ideas in Psychology*, *32*, 158–173.

Mills, G., & Healey, P. G. T. (2008). Semantic negotiation in dialogue: The mechanisms of alignment. In D. Schlangen & B. A. Hockey (Eds.), *Proceedings of the 9th SIGdial workshop on discourse and dialogue*. Columbus, OH: Association for Computational Linguistics.

Mushin, I., Stirling, L., Fletcher, J., & Wales, R. (2003). Discourse structure, grounding, and prosody in task-oriented dialogue. *Discourse Processes*, *35*, 1–31.

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

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustic Society of America*, *119*, 2382–2393.

Pardo, J. S., Gibbons, R., Suppes, A., & Krauss, R. M. (2012). Phonetic convergence in college roommates. *Journal of Phonetics*, *40*, 190–197.

Paxton, A., & Dale, R. (2012). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods*, *45*, 329–343.

Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal alignment. *Quarterly Journal of Experimental Psychology*, *66*, 2092–2102.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological bulletin*, *134*, 427–459.

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

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*, 329–347.

Ramenzoni, V. C., Riley, M. A., Shockley, K., & Baker, A. A. (2012). Interpersonal and intrapersonal coordinative modes for joint and single task performance. *Human Movement Science*, *5*, 1253–1267.

Reich, C. M., Berman, J. S., Dale, R., & Levitt, H. M. (2014). Vocal synchrony in psychotherapy. *Journal of Social and Clinical Psychology*, *33*, 481–494.

Reitter, D., & Moore, J. D. (2014). Alignment and task success in spoken dialogue. *Journal of Memory and Language*, *76*, 29–46.

Riley, M. A., Richardson, M. J., Shockley, K., & Ramenzoni, V. C. (2011). Interpersonal synergies. *Frontiers in psychology*, *2*, 38.

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

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 1.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

Shockley, K., Butwill, M., Zbilut, J. P., & Webber, C. L. (2002). Cross recurrence quantification of coupled oscillators. *Physics Letters A*, *305*, 59–69.

Sidnell, J., & Enfield, N. J. (2012). Language diversity and social action. *Current Anthropology*, *53*, 302–333.

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

Takens, F., Rand, D. A., & Young, L. S. (1981). Detecting strange attractors in turbulence. In D. A. Rand & L.-S. Young (Eds.), *Dynamical Systems and Turbulence*. Berlin: Springer.

Tylén, K., Fusaroli, R., Bundgaard, P., & Østergaard, S. (2013). Making sense together: A dynamical account of linguistic meaning making. *Semiotica*, *194*, 39–62.

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. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive performance. *Journal of Experimental Psychology: General*, *132*, 331.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2005). Human cognition and 1/f scaling. *Journal of Experimental Psychology General*, *134*, 117–123.

Van Orden, G. C., Hollis, G., & Wallot, S. (2012). The blue-collar brain. *Frontiers in Physiology*, *3*, 207.

Van Orden, G. C., & Stephen, D. G. (2012). Is cognitive science usefully cast as complexity science? *Topics in Cognitive Science*, *4*, 3–6.

Vesper, C., van der Wel, R. P., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, *211*(3-4), 517–530.

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*, 153.

Ward, L. M. (2002). *Dynamical cognitive science*. Cambridge, MA: MIT Press.

Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review*, *12*, 957–968.