Journal of Experimental Psychology: General

Complexity Matching in Dyadic Conversation

Drew H. Abney, Alexandra Paxton, Rick Dale, and Christopher T. Kello

Online First Publication, October 6, 2014. http://dx.doi.org/10.1037/xge0000021

Journal of Experimental Psychology: General

0096-3445/14/$12.00 http://dx.doi.org/10.1037/xge0000021

Complexity Matching in Dyadic Conversation

Drew H. Abney, Alexandra Paxton, Rick Dale, and Christopher T. Kello
University of California, Merced

Recent studies of dyadic interaction have examined phenomena of synchronization, entrainment, alignment, and convergence. All these forms of *behavioral matching* have been hypothesized to play a supportive role in establishing coordination and common ground between interlocutors. In the present study, evidence is found for a new kind of coordination termed *complexity matching*. Temporal dynamics in conversational speech signals were analyzed through time series of acoustic onset events. Timing in periods of acoustic energy was found to exhibit behavioral matching that reflects complementary timing in turn-taking. In addition, acoustic onset times were found to exhibit power law clustering across a range of timescales, and these power law functions were found to exhibit complexity matching that is distinct from behavioral matching. Complexity matching is discussed in terms of interactive alignment and other theoretical principles that lead to new hypotheses about information exchange in dyadic conversation and interaction in general.

Keywords: conversation, dyadic interaction, dialog, alignment, complexity matching

Conversation is a complex coordination of human behavior (Shockley, Richardson, & Dale, 2009). Interlocutors need to attend to each other flexibly and continuously over the course of conversation so that they know what to say and when to say it in order to satisfy their conversational goals. One prominent model of dyadic conversation is Pickering and Garrod's (2004) interactive alignment model. The model emphasizes the importance of aligning different linguistic representations and holds that interlocutors match representations at different linguistic levels.

There are numerous schemes for dividing linguistic processing into levels, but Pickering and Garrod (2004) discussed six: phonetic, phonological, lexical, syntactic, semantic, and situational. In support of this model, a range of studies has shown that interlocutors match speech behaviors at various scales of linguistic structure. Interlocutors have been shown to match productions of phonemes (Pardo, 2006), speech pauses (Cappella & Planalp, 1981), syntactic structures (Bock, 1986), and descriptive utterances (Garrod & Anderson, 1987). In these cases, there are direct correspondences between particular instances of behaviors, such as mimicking individual utterances, syntactic phrasings, accented words, and so on. We use the term *behavioral matching* to refer to these phenomena alternately known as alignment, entrainment, convergence, and synchronization (Louwerse, Dale, Bard, & Jeuniaux, 2012).

A growing body of literature supports the existence of behavioral matching, but the specifics and interpretation are matters of debate. Some argue that behavioral matching and related processes are integral to dyadic interactions (Pickering & Garrod, 2004), while others emphasize the role of behavioral matching in facilitating mutual comprehension (Brennan & Clark, 1996; Brennan & Hanna, 2009). Others argue that principles and processes of perception and action give rise to behavioral matching (Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007; Sebanz, Bekkering, & Knoblich, 2006). Still others contend that human communication is a general framework for situated action in which interlocutors maximize detection and sensitivity to others (Suchman, 2007).

These ongoing debates have been useful and informative because they suggest that behavioral matching plays some role in establishing common ground and, more generally, facilitating communication. However, opportunities for behavioral matching in natural conversation are limited because interlocutors do not simply mirror each other's behaviors. Each person makes unique, individual contributions to dyadic interactions, and effective communication necessitates that interlocutors share common ground and coordinate behaviors (cf. Healey, Purver, & Howes, 2014; Mills, 2014). Thus many aspects of conversational behavior may be expressed by more indirect, subtle forms of coordination. Even turn-taking is more complex than synchronization or syncopation. Turns often do not alternate cleanly and evenly (Stivers et al., 2009), and interlocutors often speak and gesture simultaneously during periods of so-called "back channeling" (McClave, 2000).

The irregular, complex nature of dyadic interaction raises the question of whether behavioral matching may be generalized to more indirect forms of matching. That is, the drive to establish common ground and facilitate communication may be addressed through other

Drew H. Abney, Alexandra Paxton, Rick Dale, and Christopher T. Kello, Cognitive and Information Sciences, University of California, Merced.

We would like to thank J. P. Gonzales for his help in data collection and Stephanie Frewen and Jamie Faria for their help in data preparation. We would also like to thank Anne Warlaumont, Ramesh Balasubramaniam, Teenie Matlock, and Riccardo Fusaroli for comments on a previous version of this article. This work was supported by National Science Foundation Grants BCS-1031903 (to Christopher T. Kello), BCS-0826825 (to Rick Dale), and BCS-0926670 (to Rick Dale).

Correspondence concerning this article should be addressed to Drew H. Abney, Cognitive and Information Sciences, University of California, Merced, School of Social Sciences, Humanities and Arts, 5200 North Lake Road, Merced, CA 95343. E-mail: drewabney@gmail.com

means that can be viewed as extensions of behavioral matching. One natural extension is *distributional matching*—the idea that behaviors may match at the level of statistical, ensemble characterizations, rather than the level of particular behavioral acts. For instance, mean speech rates may converge during conversations (Webb, 1969), or two interlocutors may converge in their proportions of slang expressions, without directly matching each other slang for slang.

The concept of distributional matching is consistent with Pickering and Garrod's (2004) interactive alignment model. Perhaps the best example comes from the well-known phenomenon of syntactic priming (Bock, 1986; Pickering & Branigan, 1998, 1999), in which hearing or seeing the usage of a given syntactic form (e.g., active vs. passive) increases the likelihood that speakers will use it themselves. Syntactic priming can arise from behavioral matching or distributional matching. In the latter case, the probability distributions over syntactic forms may converge between interlocutors (see Jaeger & Snider, 2008).

The hypothesis of distributional matching takes on a new dimension when the distributions being matched follow *power law* functions (Clauset, Shalizi, & Newman, 2009). A power law function expresses one variable as a nonlinear function of another variable raised to a power, $f(x) \sim x^{\alpha}$. The heterogeneities and irregularities of language behaviors are reflected in many different power laws. For instance, power law relations are found in frequencies of word usage (Zipf, 1949), frequencies of n-grams in text corpora (Kello & Beltz, 2009), frequencies of syntactic links to words (Ferrer-i-Cancho, Solé, & Köhler, 2004), correlations and burstiness across vowels/consonants, letters, words, and topics (Altmann, Cristadoro, & Degli Esposti, 2012), and spectral density of fluctuations in music and human speech (Voss & Clarke, 1978). We provide an illustrative example in the next section, but suffice it to say here that these power laws reflect the heterogeneity of language in terms of variability across a wide range of measurement scales. They correspond to the irregularity of language in terms of rough stochastic patterns, unlike the highly regular fractals (i.e., power laws) of snowflakes and Mandelbrot sets.

In the present study, we find evidence for a new power law distribution in conversational speech signals. The power law is hypothesized to reflect hierarchical clustering and levels of linguistic information in conversational speech (Grosjean, Grosjean, & Lane, 1979), akin to levels proposed for the interactive alignment model. The speech data come from dyadic conversations designed to be either affiliative or argumentative (Paxton & Dale, 2013), and the speech signals are analyzed in terms of their temporal dynamics, as captured by acoustic onset events and subsequent periods of acoustic energy.

The power law in event clustering is measured by the Allan Factor (AF) function, which computes coefficients of variation across multiple timescales. We find that AF functions measured from interlocutor speech signals *converge* in dyadic conversations, particularly for affiliative conversations and not argumentative conversations. We call this convergence *complexity matching* as a special case of distributional matching when distributions are power laws. The term comes from studies in statistical mechanics (West, Geneston, & Grigolini, 2008) showing maximal information exchange between coupled complex systems that individually produce similar power laws.

We test whether conversational speech signals exhibit the conditions predicted from statistical mechanics on the approach that complexity matching can provide a unique angle into naturalistic conversation. We compare behavioral and complexity matching to test whether they make distinct contributions toward explaining dyadic interaction, and whether complexity matching yields useful evidence beyond behavioral matching.

Power Law Clustering in Conversational Speech

A simple way to approximately describe a power law distribution is to say that variability occurs across a wide range of measurement scales, including timescales. For the latter, imagine that a time series of measurements is windowed and the average measured value is computed for each window of size S. Variability across scales means that measures of variance scale up with window size S, e.g., small variations for millisecond windows, larger variations over seconds, even larger variations over hours, and so on. Variability across scales is unexpected for most types of simple systems. For instance, if one measures the temperature fluctuations in a refrigerator, variations would actually *decrease* with larger time windows, because larger windows would yield averages that converge on or near the temperature setting.

Variability that spans measurement scales is indicative of power laws, and such power laws will emerge from more complex systems, namely, ones that display hierarchically nested structures and processes (Simon, 1973). In the case of language, sentences are collections of syntactic phrases, phrases of words, words of syllables, syllables of phonemes, and so on. Such nested levels of linguistic representation are integrated in the interactive alignment model, as illustrated in Figure 1. We therefore expect the hierarchical nesting of language to be physically manifested as power laws in speech signals.

Hierarchical nesting in speech signals can be illustrated as follows. At the coarsest timescales, when two people converse, each interlocutor produces turns—long, clustered periods of acoustic speech energy interspersed with mostly no acoustic energy while the other person is talking. At finer timescales, there are breaks in the signal due to thinking time, phrase boundaries, rhetorical effects, and the like. At still finer timescales, breaks occur sometimes at word boundaries, and sometimes at phonemes with little or no sonority, such as plosive consonants (e.g., p, t, k, b, d, g), quiet fricatives (e.g., f, h, th), and even voiced fricatives and nasal stops in some cases (e.g., v, m, n, ng). All of these breaks are defined as falling below some threshold of acoustic energy, i.e., we do not assume total silence or even a total lack of perceptible sound during breaks.

The three illustrative scales just listed are visualized in the speech waveform displayed in Figure 2. It is important to note that one could posit additional or different scales as well. Whatever the case, their physical manifestations are likely to overlap and blend such that one simply observes clusters of acoustic energy across a continuous range of scales in the raw speech signal. In fact, a continuous range of scales is expected to emerge when interactions propagate across levels of representation (Holden, Van Orden, & Turvey, 2009; Mitzenmacher, 2004), as posited in the interactive alignment model. Phonetic processes interact with lexical processes, which interact with syntactic processes and feedback to phonetic processes, and so on.

Complexity Matching in Speech Signal Clustering

Our discussion so far leads us to expect power law clustering in speech signals due to the hierarchical nesting of language representations and processes. Thus we need a method for measuring

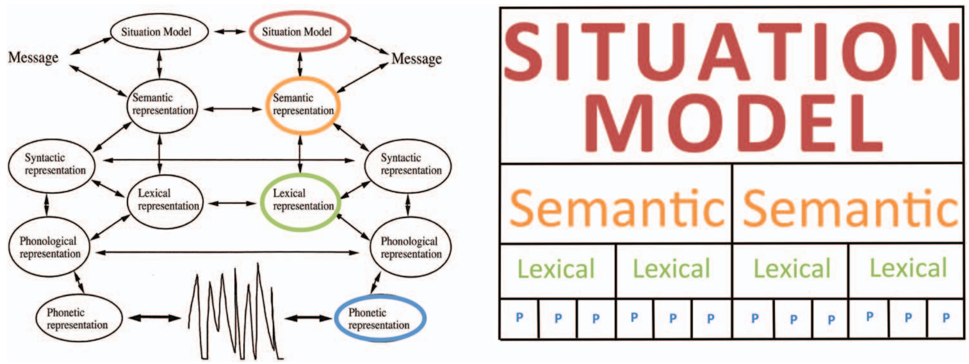

Figure 1. Left: Pickering and Garrod's (2004) schematic representation of the stages of comprehension and production processes according to the interactive alignment model. Right: An illustration of the nesting of different scales of linguistic representations, using four levels of the six from the interactive alignment model: phonetic, lexical, semantic, and situation model. Left panel reprinted from "Toward a Mechanistic Psychology of Dialogue," by M. J. Pickering and S. Garrod, 2004, *Behavioral and Brain Sciences, 27*, p. 177. Copyright 2004 by Cambridge University Press.

and quantifying clustering in speech signals across different timescales. Clustering is expected specifically in the timing of periods of acoustic energy interspersed with breaks as defined by some threshold. Such temporal clustering can be measured in the onset times when acoustic energy crosses from below to above threshold. Acoustic onset times are not only appropriate for measuring temporal clustering, but they also are highly salient and important events in speech perception (Cummins & Port, 1998; Cutting &

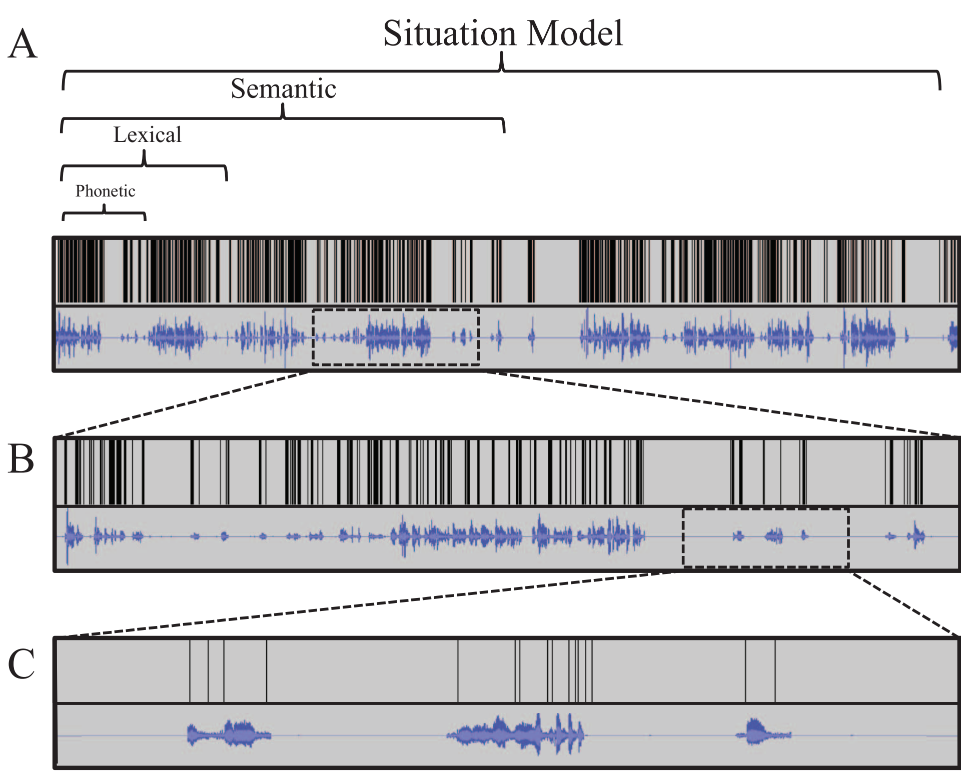

Figure 2. An example segment of a conversational speech signal, shown at three different levels of temporal resolution. A: The longest scale corresponds with conversational turns, and brackets show roughly the timescales at which different levels of linguistic representation reside: phonetic, lexical, semantic, and situation model. B: The middle scale illustrates, e.g., thinking pauses and phrase boundaries. C: The shortest scale illustrates word, syllable, and phoneme boundaries. Vertical lines shown for each speech signal signify acoustic onsets relative to a threshold chosen by visual inspection. See the online article for a color version of this figure.

Rosner, 1974; Liberman, Harris, Hoffman, & Griffith, 1957). Clustering in acoustic onset times is visible in Figure 2.

The interactive alignment model holds that interlocutors "align" representations across levels of linguistic processing. The particular nature of alignment is an ongoing area of research, and as mentioned earlier, behavioral matching is one manifestation of alignment that is well-documented in the literature (Louwerse et al., 2012). But also, as mentioned earlier, behavioral matching is limited because direct correspondences alone cannot explain the rich behavioral diversity in natural conversations (e.g., Healey, 2008; Howes, Healey, & Purver, 2010).

Temporal clustering of acoustic onsets across scales, as a physical expression of linguistic processing across levels of representation, affords the possibility for a kind of distributional matching distinct from behavioral matching. The overall *amount* of temporal clustering can be quantified as a function of timescale, as we explain more formally below. Conversational speech signals may converge in terms of the distribution of temporal clustering across timescales. Such convergence would constitute a *complex coupling* in the dynamics of linguistic processing. This coupling would be complex partly because it would go beyond synchronization and other simple phase relations between time series, and partly because it would constitute the coupling of two power law distributions that reflect nested, interactive scales of processing.

Power law distributions are defining of *complex systems* in general (Sales-Pardo, Guimera, Moreira, & Amaral, 2007; Simon, 1977). Specifically, a complex system is one in which microscopic events may cascade up to alter macroscopic patterns of activity, which in turn may constrain and shape its microscopic events (Stanley, 1987). By this definition, both humans and human languages are demonstrably complex systems (The "Five Graces Group" et al., 2009; Kugler & Turvey, 1987; Mitchell, 2009; Spivey, 2007; Swenson & Turvey, 1991). Molecular and cellular events cascade up to affect behavior via myriad genetic and physiological processes, and behavior helps shape those processes via evolution and learning, for example. Likewise, microscopic changes in phonetic features may alter entire words, sentences, and conversations as macroscopic patterns, and the latter provide higher level constraints on how phonemes are phonetically realized.

Classes of complex systems can be formalized statistically, relative to the dynamics of their interacting components. West et al. (2008) recently analyzed the coupling of complex systems in terms of their event dynamics, which amounts to temporal clustering of point processes analogous to acoustic onsets. Interestingly, analyses have shown information exchange between coupled systems to be maximal when the exponents of their power laws are similar (Aquino, Bologna, Grigolini, & West, 2010; Aquino, Bologna, West, & Grigolini, 2011; Turalska, West, & Grigolini, 2011). For power laws in the temporal clustering of point processes, convergence of exponents corresponds with convergence in the amounts of temporal clustering across timescales. Thus West et al. (2008) provided independent theory and rationale for expecting convergence in the temporal clustering of conversational speech signals—under these conditions, information exchange should be maximized between interlocutors as complex systems (see also Stephen & Dixon, 2011; Stephen, Stepp, Dixon, & Turvey, 2008).

The formal analysis conducted by West et al. (2008) relies on statistical physics and mechanics and its elaboration is outside the scope of the current article. However, we can draw an intuitive analogy with simple oscillators designed to illustrate coupling beyond synchronization. Imagine two metronomes whose kinematics are coupled through a physical medium such as a sliding platform (Figure 3a). Provided that their frequencies are sufficiently similar, and coupling is sufficiently strong, the beats of the metronomes will tend to synchronize over time (Kelso, 1981; Strogatz & Mirollo, 1991). The phase-coupled oscillations that result from these interacting forces can be seen as idealized forms of behavioral matching, and a number of dyadic interaction studies have drawn this parallel (for a review, see Schmidt & Richardson, 2008).

Now imagine two sets of metronomes at each end of the platform (Figure 3b) whose resonant frequencies span a wide range of timescales and do not correspond one-to-one across the two ends of the platform. Coupling may still yield a system for which synchronization is an inherently low energy state, but synchronization and other simple phase relations may no longer be sufficiently strong attractors to create stable dynamical states of the system. This is more likely to be true especially when coupling is relatively weak. In such cases, the system instead is prone to exhibiting intermittent, irregular transitions from one *metastable* state to the next (Kelso, 1995). Such complex dynamics are readily observable in systems as simple as coupled oscillators, and coupled oscillators provide only a simple model of human interlocutors. Thus the metronomes serve to illustrate how complex couplings are not exotic or rare but, rather, are quite expected for interactions between such richly heterogeneous systems like humans.

Expectations of phase couplings and more complex couplings lead us to predict behavioral matching *and* complexity matching in human interactions. To our knowledge, this prediction has not been tested previously for conversational interactions, but we can find support for a similar hypothesis in human perceptual-motor interactions (Coey, Washburn, & Richardson, in press; Marmelat & Delignières, 2012). Marmelat and Delignières (2012) recently

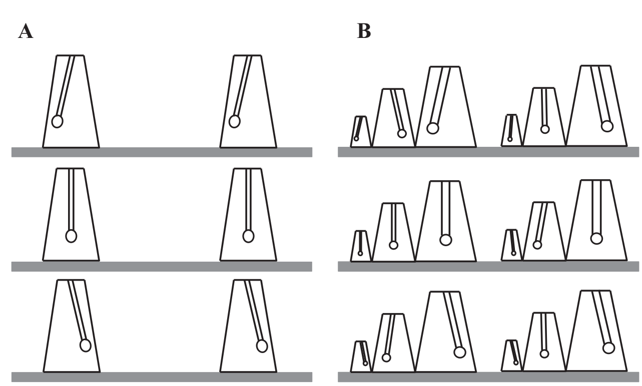

Figure 3. Examples of synchronization and behavioral matching with toy metronome systems. A: Illustration of two metronomes interacting along a sliding platform, as a simple model of synchronization and a form of behavioral matching. B: Illustration of interactions between multiple metronomes with differing frequencies, to aid the intuition of complexity matching.

conducted an experiment in which each participant in a dyad swung a hand-held pendulum, with instructions to swing in synchrony. Synchronization is a form of behavioral matching, but deviations from synchrony were analyzed for power law fluctuations in the form of $1/f^{\alpha}$ noise. Results showed that α estimates for each member of a dyad were correlated to the extent that coupling was facilitated by visual and physical contact. These α correlations served as a direct measure of complexity matching, and they could not be explained in terms of behavioral matching because dyadic time series of deviations from synchrony were uncorrelated at all lags—i.e., there were no cross-correlations.

Current Study

The new contributions of the current study are tests of (a) power law clustering in the temporal patterning of acoustic onsets in conversational speech and (b) complexity matching in the temporal clustering of speech across timescales. Power law clustering is expected to manifest due to the hierarchical nature of language processes. Complexity matching is expected to extend and complement behavioral matching, as part of a broader basis for interactive alignment that enhances communication through increased information exchange.

Our study was designed to investigate complexity matching through a number of different conditions and analyses. First, we analyzed data from a recent study by Paxton and Dale (2013), in which participants who previously did not know each other were asked to have two conversations (order counterbalanced). One was a casual, *affiliative* interaction about popular media. The other was on provocative issues based on participants' closely held beliefs and designed to evoke more *argumentative* conversations. Beforehand, participants were given questionnaires to gauge their opinions on these provocative issues, and specific issues were chosen if participants had strong but differing opinions about them. Partners were instructed to converse for 10 min in each condition, which provided ample time for long stretches of speech to be analyzed. The original aim of the study was to investigate alignment in asymmetric contexts, that is, interactions between interlocutors who have conflicting, differing, or opposing goals and opinions.

These experimental data serve the current goals quite well, because the time series that can be extracted from audio data are long enough to afford measurements of temporal clustering across a wide range of scales. In addition, we can test for a relationship between complexity matching and a high-level discourse constraint: conversation type. Testing for such a relationship is important for providing converging evidence that temporal clustering of acoustic onsets is reflective of levels of linguistic processing rather than just matching of low-level acoustic properties of speech. The experiment also allowed us to compare matches between two speech signals from an originally paired dyad, with mismatches between signals from two different dyads. The latter provides a baseline for measuring complexity matching above chance and is a common baseline among dyadic interaction researchers (e.g., Bernieri, Reznick, & Rosenthal, 1988).

Another important feature of the experiment by Paxton and Dale (2013) is that it allows us to compare our measure of complexity matching with a more traditional measure of behavioral matching, where the latter can be quantified through cross-correlations in speech signals. As elaborated below, greater behavioral matching in our case corresponds to the negative peak of the cross-correlation function, which reflects the complementary turn-taking relationship between the temporal patterns of acoustic speech energy produced by each member of a dyadic conversation. We directly test whether complexity matching can be reduced and attributed to behavioral matching as measured by negative peaks in cross-correlations, or whether the two reflect distinct aspects of coordination in dyadic conversation.

Method

Participants

A total of 28 undergraduate students (mean age = 20.14 years; females = 22) from the University of California, Merced participated in return for extra course credit. Individual participants signed up for time slots anonymously, and participants were not informed of their partner's identity beforehand. Dyads included eight female-female, six mixed-sex pairings, and no male-male pairings (by chance; the uneven breakdown of gender pairings prohibited a post hoc analysis of gender). All participants reported conversational fluency in English and normal or corrected hearing and vision. Participants also reported their native language as English ($n = 10$), Spanish ($n = 10$), or other ($n = 6$; an additional two participants did not disclose their native language).

Procedures

Before conversing with one another, each participant completed a brief series of questionnaires, including an opinion survey on political, social, and personal topics (e.g., abortion, death penalty, gay/lesbian marriage, legalization of marijuana). For each topic, participants were asked to write a brief synopsis of their opinion and mark how strongly they held their opinion from 1 (*feel very weakly*) to 4 (*feel very strongly*) on a Likert-style scale. Experimenters determined the topic of argument by comparing the two participants' survey answers to identify the topic on which participants held strong but opposing views. This topic was chosen as the dyad's argumentative prompt, given along with an instruction to convince one another of their opinion. Two additional prompts were also selected by those criteria but were given only if the participants were unable to continue the conversation on the topic at hand. Of the 14 dyads analyzed here, 10 required additional prompts (secondary = 9; tertiary = 1).

In addition to the argumentative conversation, each dyad also had an affiliative conversation. The affiliative prompt instructed each dyad to identify and discuss popular media that both participants enjoyed. Affiliative prompts were designed to emphasize the common ground between partners, whereas the argumentative prompts were designed to emphasize their differences of opinion.

Following the questionnaires, participants were brought together in a private room and seated facing each other. To provide an opportunity for partners to become acquainted with each other, they were left alone for about 3 min to introduce themselves outside the context of the experiment, without yet knowing the nature of the experimental task. To make introductions as natural as possible, participants were told that experimenter had to step out of the room to complete last-minute paperwork before beginning the experiment.

After the introduction period, the experimenter entered the room and delivered the first conversation prompt. The order of prompts was counterbalanced across dyads, and participants were not informed of upcoming prompts. During each 10-min conversation, the experimenter monitored recording equipment from a seat on the periphery of participants' range of vision. After each conversation, participants were separated and asked to complete postconversation questionnaires. At the end of the experiment, participants were thanked and debriefed.

Apparatus, Data Collection, and Data Preparation

Conversations were recorded on a Canon Vixia HF M31 HD Camcorder, mounted on a Sunpak PlatinumPlus 600PG tripod. Audio for each participant was recorded separately at 44-kHz sample rate, using an Azden CAM-3 mixer and Audio-Technica ATR 3350 lapel microphones affixed to the upper portion of each participant's shirt. Two audio files were recorded per conversation (one for each interlocutor), which yielded four files per dyad and 56 files altogether across the 14 dyads.

After truncating audio files to contain only the conversations, Audacity was used to remove nonspeech signals, as well as any partner cross-talk so that each file contained only one participant's speech signal. The Audacity "sound finder" was then used to locate acoustic onset and offset events in each file. The signal/no-signal threshold of acoustic intensity was set at –30 db for all dyads, which was judged to be the lowest threshold that resulted in less than ~5% spurious onset events. This threshold yielded an average of 764 paired onset and offset events per partner, per conversation. For every audio file, the resulting event time series was highly irregular and clustered, based on visual inspection. Each event series was unique, as expected given that each partner made unique contributions to their conversations. However, we are interested in statistical quantities that abstract away from particular event times and characterize their temporal properties.

Interevent Intervals

The interactive alignment model, along with its hierarchically nested levels of linguistic processing, leads us to predict complexity matching in the temporal clustering of acoustic onset events. However, West et al. (2008) showed that complex systems in general are expected to exhibit complexity matching when their *interevent intervals* (IEIs) are power law distributed with an exponent near two, P(IEI) $\sim 1/\text{IEI}^{\gamma}$, where $\gamma \sim 2$. West et al.'s analysis suggests that we test IEIs for the predicted power law.

A histogram of IEIs was computed for the time series from each participant in each conversation, where the position of the smallest bin was set relative to the shortest IEI value in each given time series. The nine subsequent bins were logarithmically spaced to capture IEIs of all lengths for each time series. Logarithmic spacing accounted for the anticipated power law in IEI distributions—that is, greatest resolution in the histogram is needed for at the small end of the scale because the vast majority of IEIs are relatively short, and resolution can become coarser as IEIs become larger and less frequent.

Figure 4 shows the resulting histograms for each participant, plotted together in a single graph. Plotting individual histograms together provides a picture of the overall trend of the

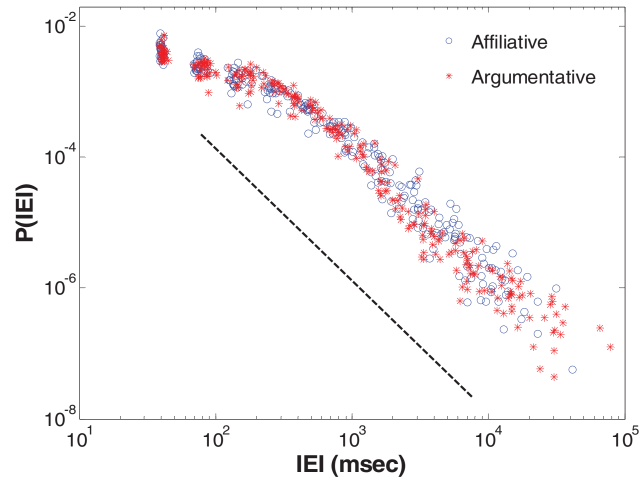

Figure 4. Interevent interval (IEI) probability density functions for individual interlocutors in individual conversations, plotted in logarithmic coordinates using logarithmic binning. Dashed line shows idealized slope of –2 (per West, Geneston, & Grigolini, 2008).

distributions, as well as the individual variability around that trend. The figure shows a clear trend of a negatively sloped line in logarithmic coordinates that flattens out for the shortest IEI values on the left. The slope of the trend is about –2 for both conversation types, as can be seen by comparing with the dashed line that has a slope of exactly –2. Thus the data closely resemble the theoretically derived precondition for complexity matching,[1] i.e., the power law P(IEI) $\sim 1/\text{IEI}^{\gamma}$, where $\gamma \sim 2$.

Temporal Clustering in Acoustic Onsets

To quantify temporal clustering in acoustic onsets and test for a power law across timescales, we adopted *Allan Factor* (AF) analysis that has been used to measure temporal clustering in neural spike trains (Teich, Heneghan, Lowen, Ozaki, & Kaplan, 1997). Spikes and acoustic onsets are both examples of *point processes*, i.e., time series of events treated as occurring at instantaneous points in time. A Poisson process is one whose events occur unpredictably through time, i.e., for which knowledge of any and all event times up to a given point in time t provides no information about when future events may occur. AF is a statistical method that distinguishes between Poisson processes and those whose events occur nonrandomly. In our case, we are interested in non-Poisson processes whose events cluster at different timescales more than would be expected by a Poisson process.

AF analysis is partly illustrated in Figure 2. Time series are tiled with adjacent windows of given size T; in the figure, each bracket represents one window of a given size. Events are simply counted within each window, and a measure of variance—AF variance, $A(T)$—is derived from the differences in counts between adjacent

[1] IEI distributions were tested via multimodel inference using Akaike information criterion (AIC) values and maximum-likelihood estimation and showed that the lognormal function was most likely to generate the distributions. The lognormal distribution is heavy-tailed and known to provide good fits to power law distributions with truncated tails such as the IEI distributions (Edwards, 2008).

windows. A(T) is calculated for a range of window sizes (i.e., timescales T), and Poisson processes are those for which A(T) $\sim$ 1 for all T. Clustering at a given scale results in A(T) $>$ 1, and more specifically, clustering across scales means that A(T) $\sim T^{\alpha}$, where $\alpha > 0$. Finally, complexity matching is measured as the difference between two A(T) functions, where *more* matching corresponds to *smaller* differences.

Formal Description of AF Analysis

A formal description of AF analysis is as follows. A given point process is segmented into M adjacent windows of size T (enough to span the entire series), and the number of events N_j is counted within each window indexed by $j = 1$ to M. The differences in counts between adjacent windows of a given size T is computed as $d(T) = N_{j+1}(T) - N_j(T)$. $d(T)$ values are computed for each of a range of values for T. The AF variance A(T) for a given timescale T is the expected value of the squared differences, normalized by mean counts of events per window (i.e., a type of coefficient of variation),

$$A(T) = \frac{\langle d(T)^2 \rangle}{2\langle N(T) \rangle}.$$

Poisson processes yield A(T) $\sim$ 1 for all T, whereas power law clustering yields A(T) $\sim (T/T_1)^{\alpha}$, where T_1 is the smallest time scale considered and α the exponent of the scaling relation. Point processes with $\alpha \sim 0$ are Poisson-distributed, whereas power law clustering means meaning $\alpha > 0$ over the measureable range of timescales (Thurner et al., 1997).

A(T) was computed for each event time series from each interlocutor in each conversation. Each time series was 10 min long, and time windows varied as a power of 2, $T = 2^t$ where t ranged from 4 to 12. The resulting timescales ranged from 160 ms to 41 s. Smaller timescales were excluded because they are heavily affected by measurement error, and larger timescales could not be reliably estimated given the length of time series. A(T) values were averaged across participants for each conversation, and averages are plotted as a function of T in Figure 5.[2]

Results of AF Analysis

A clear power law is evident in the roughly linear relationship in logarithmic coordinates for both conversation types. This power law is evidence of nested clustering of events over the measured timescales, as expected for nested language processes. The exponent of the AF power law was estimated for each individual time series by taking the slope of a regression line fit to each AF function in logarithmic coordinates. Mean exponent estimates for affiliative conversations (M = .53, SE = .02) were reliably less than those for argumentative conversation (M = .63, SE = .02), $t(27) = 4.57$, $p < .001$. This effect can be seen in Figure 5 as deriving from A(T) differences at the largest timescales. In general, this is evidence that the clustering of acoustic onsets reflects linguistic processing during conversations, rather than purely acoustic structure.

More specifically, results showed *greater* temporal clustering of onsets in argumentative conversations relative to affiliative ones, at longer timescales. Longer timescales mainly reflect turn-taking

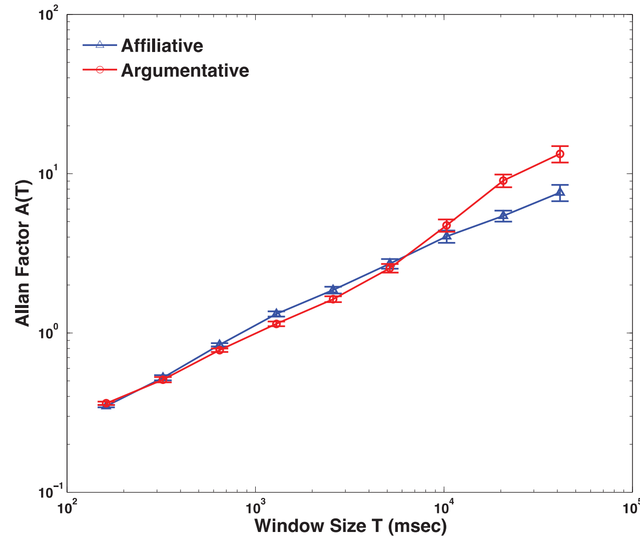

Figure 5. Mean Allan Factor functions for argumentative versus affiliative conversation types, with standard error bars. See the online article for a color version of this figure.

dynamics, which suggests that there were fewer, longer turns in argumentative conversations. To confirm this interpretation of the observed difference in AF functions, we compared the number and mean duration of IEIs greater than 4 s. Four seconds was approximately where the AF functions diverged and was a cutoff that should capture mostly turn intervals, i.e., an utterance without a break in acoustic energy, followed by a pause before the partner begins the next turn. We did not expect this automated method to capture turns perfectly—some turns will be missed or cutoff, and some intervals will reflect utterances within turns—but it is safe to assume that the majority of these few very long intervals (less than 5% of all intervals on average) mostly correspond with turns. As expected, estimated turns for argumentative conversations were found to be fewer (M = 21.7 vs. M = 26.8), $t(27) = 2.9$, $p < .01$, and longer (M = 12.6 vs. M = 8.0), $t(27) = 5.2$, $p < .001$, compared with affiliative conversations.

Complexity Matching

The previous two sections established two preconditions necessary to test for complexity matching, i.e., (a) power law distributions in IEIs that approach an exponent of two and (b) power law clustering of acoustic onsets, as expressed in the AF function, that ostensibly reflects the hierarchical nesting of linguistic processing during conversation. Now, to test for complexity matching, we need a measure of similarity between two AF functions and a baseline for the amount of complexity matching expected by chance.

Our measure of AF similarity is the summed absolute difference between two AF functions a and b, with a negative log transformation:

$$D_{a,b} = -\sum_{T} \log \left| A(T_a) - A(T_b) \right|$$

[2] For example code, please see http://cogmech.ucmerced.edu/downloads.html

The log transformation takes into account the scaling law over T, and the negative simply makes greater values correspond with greater complexity matching, relative to a baseline control.

For baseline controls, we used surrogate comparisons between event series. Specifically, *condition controls* were created by comparing event series of two interlocutors from the same conversation type (either both affiliative or both argumentative) but who did *not* converse with each other. Fifty-two condition controls were created for each dyad in each condition, and the resulting $D_{a,b}$ values were averaged for each dyad. Mean $D_{a,b}$ functions are plotted in Figure 6 for original pairings and condition controls, separated by conversation type.

$D_{a,b}$ values for original pairings in the affiliative conversation ($M = 11.91$, $SE = 1.40$) were greater than their condition controls ($M = 9.80$, $SE = .44$), $t(13) = -1.95$, $p_{one\text{-}tailed} < .05$. However, there was no such effect in the argumentative conversation, $t(13) = -0.06$, $p_{one\text{-}tailed} = .48$. These results provide evidence for complexity matching in the power law clustering of acoustic onsets in affiliative conversations, but not argumentative conversations. A qualitative inspection showed that AF differences generally occurred across timescales between affiliative originals and controls. Thus matching did not vary significantly over the range of timescales in which phonological, lexical, syntactic, and discourse processes unfold.

Finally, we note that the effect of conversation type was so strong that complexity matching for affiliative *controls* was a little *more* than that for argumentative *original* pairings, albeit not reliably so, $t(13) = 1.59$, $p = .134$. The reason for this result needs further investigation, but one possibility is that argumentative conversations create a repelling dynamic that opposes complexity matching, thereby making speech signals no more similar than chance. This possibility is supported by analyses of behavioral matching reported next.

Behavioral Matching

The previous section reported evidence for complexity matching, but it is important to test whether this evidence can be attributed to behavioral matching. Interlocutors' speech signals may exhibit "align-able" patterns in their periods of acoustic energy, possibly with some temporal lag between the signals. Phase-shifted alignment would constitute behavioral matching, and if the patterns are power law clustered, the same signal similarity that yields behavioral matching would also yield complexity matching. Here we test for behavioral matching in conversational speech signals and compare with complexity matching results to determine whether the complexity matching results may be attributed more simply to behavior matching.

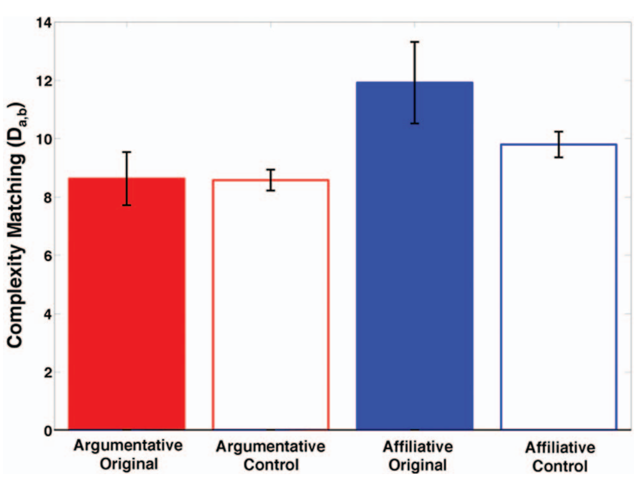

Figure 6. Mean summed Allan Factor difference functions plotted for the two conversation types, separately for original pairings versus randomized controls, with standard error bars.

The most directly related measure of behavioral matching would be to use the same time series of acoustic onsets as used for complexity matching and simply cross-correlate them. However, point processes are theoretically instantaneous, which means their lack of duration complicates direct use of the cross-correlation function. Rather than assign each onset a temporal range, we used the duration of ongoing acoustic energy that followed each acoustic onset, i.e., the periods of acoustic energy from each onset to each subsequent offset. The resulting time series of acoustic energy periods were then cross-correlated to test for evidence of behavioral matching. Surrogate cross-correlation functions were also computed using the same method as that for AF functions.

Cross-correlations yielded no evidence of alignment at any lag: Peak positive correlation coefficients for affiliative pairings ($M = .06$, $SE = .006$) were not reliably different from their surrogate controls ($M = .06$, $SE = .001$), $t(13) = -1.14$, $p = .274$, and the same was true for argumentative pairings ($M = .09$, $SE = .01$) compared with their surrogate controls ($M = .07$, $SE = .002$), $t(13) = -1.67$, $p = .119$. These null results provide an initial suggestion that our complexity matching results cannot simply be attributed to behavioral matching.

Inspection of the cross-correlation functions revealed that, unlike peak positive correlations, peak *negative* correlations are far greater in magnitude for original pairings compared with surrogate pairings. This effect held for both affiliative and argumentative conversations, $t(13) = 4.77$ and 5.92 (respectively), both $p < .001$. These negative peaks reflect complementarity in the time series of acoustic energy periods, which likely derives from turn-taking in conversational speech. Thus maximal misalignment is also a kind behavioral matching, albeit one where each speech act is matched with a lack thereof. While this may be considered as behavioral *mis*-matching, it is demonstrative of a strict temporal coordination between partners that is very much in line with the spirit of behavioral matching research.

This turn-taking measure of behavioral matching was stronger for argumentative conversations ($M = -.32$, $SE = .16$) compared with affiliative conversations ($M = -.23$, $SE = .14$), $t(13) = 4.16$, $p < .001$. Thus the effect of conversation type on behavioral matching was *different and opposite* from its effect on complexity matching: Behavioral matching results point to stricter turn-taking in argumentative conversations, whereas complexity matching highlights stronger coupling across levels of linguistic processing in affiliative conversations.

It is possible that our measure of complexity matching is somehow the converse of our measure of behavioral matching. If so, the two measures should be negatively correlated. Results did not bear out this hypothesis: A correlation of $D_{a,b}$ values with peak minimum cross-correlations yields a coefficient of $r(28) = .23$, which is not reliable ($p = .243$). This null result suggests that the complexity matching we observed cannot be straightforwardly attributed to behavioral matching.

General Discussion

Perhaps the most salient coordination we experience in conversations is behavioral matching. We take turns, echo speech acts of our partners, and strive for mutual understanding by sharing and in some sense matching our states of knowledge. The saliency of behavioral matching in firsthand experience has an analog in the scientific study of interpersonal coordination. Synchronization is a salient form of coordination dynamics, and one that is relatively easy to formalize and investigate mathematically (Schmidt, Morr, Fitzpatrick, Richardson, 2012). Other phase relations—like antiphase (Haken, Kelso, & Bunz, 1985; Keller & Repp, 2004)—are also investigated in this area, but both phase relations can be conceptualized as different types of behavioral matching.

If we introspect further into the nature of conversational interactions, we find other more indirect forms of coordination in speech. The "tone" of a conversation, for instance, is not just carried by particular matches between turns, words, or other speech acts. Tone can be partly expressed as an approximate statistical convergence in, for instance, pitch, loudness, and pace of speech (Manson, Bryant, Gervais, & Kline, 2013; Neumann & Strack, 2000; Webb, 1969). Similarly, regional accents and dialects can be considered as a kind of convergence (Coupland, 1980) in the temporal dynamics of speech over multiple timescales and partly stem from common allophonic variations that are coordinated among populations of speakers over countless conversations.

In this study, we introduced *complexity matching* to the interpersonal interaction literature. Complexity matching was adopted from West et al. (2008) to measure broad, statistical forms of coordination in conversational speech. By analyzing data from naturalistic conversations, we found that complexity matching provides a new window into interpersonal coordination beyond behavioral matching. We measured temporal dynamics in speech as expressed through clustering of acoustic onset events across timescales. We chose this measure in part because it is a purely temporal index of speech dynamics—each acoustic onset varies only in time and nothing else—and in part because it expresses temporal dynamics across timescales, from phonetic to lexical to turn-taking variations in speech timing.

Using AF analysis, we found evidence for multiscale dynamics in the power law clustering of acoustic onsets, as measured by the AF function, and we found greater clustering at longer timescales for argumentative conversations, as measured by greater AF exponent estimates. This effect of conversation type on AF exponents indicates that multiscale clustering reflects more than just low-level acoustic properties of speech. It also indicates that argumentative conversations are more structured at the larger timescales of turn-taking, and this interpretation is supported by cross-correlation analyses indicating stricter turn-taking in argumentative conversations.

While argumentative conversations show stricter turn-taking, only affiliative conversations demonstrated complexity matching, i.e., convergence in multiscale clustering. We interpret this difference as reflective of the more subtle forms of coordination in speech that we mentioned earlier. When people engage in affiliative interactions to converge on some mutual understandings and opinions, this convergence can be reflected in subtle aspects of their speech dynamics that operate similar to constructs like tone, pace, and style. AF analysis of acoustic onsets was able to capture such subtle aspects of convergence.

The present findings also are consistent with previous multimodal analyses of the conversations. As mentioned above, herein we found no evidence of complexity matching in argumentative conversations, yet there was more behavioral matching compared with affiliative conversations, as measured by peak negative cross-correlations. Consistent with this difference, analyses of movement dynamics also found no behavioral matching during argumentative conversations (Paxton & Dale, 2013). In the future we plan to work on complexity matching analyses that may be applied to both movement and speech dynamics, in order to investigate whether multimodal coordination may further illuminate the coupling of interlocutors during affiliative conversations, and lack thereof during argumentative conversations.

Complexity Matching and Theories of Conversation, Coordination, and Development

As discussed in the Introduction, our interpretation of complexity matching is consistent with a multilevel view of influence in interpersonal interaction, like that found in the interactive alignment model (Pickering & Garrod, 2004). Language systems and processes are inherently multilevel, i.e., multiscale, and the interactive alignment model posits coupling across levels. The concept of complexity matching is exactly this—a kind of coupling across the scales of two interactive systems. The concept comes from work in statistical mechanics (West et al., 2008) that connects to the idea of interactive alignment. Multiscale systems with interactive levels of processing generally are expected to exhibit power laws as signatures of the complexity that is concomitant in such cross-scale interactions. West and colleagues report formal analyses to show that two multiscale, complex systems are most responsive to each other when their power laws converge, particularly near a specific exponent in the power law distribution of interevent intervals.

We find it somewhat remarkable that data from dyadic conversations fit the theoretical predictions of a theory from statistical mechanics that was formulated for a broad class of physical systems. In the current study, we found the estimated exponents of interevent intervals during conversations indeed were near the predicted exponent value of two. Thus our study is an example of how work from statistical mechanics can inform and enhance specific theories (such as interactive alignment) in the psychological and cognitive sciences, and how interdisciplinary research can yield new disciplinary insights.

To illustrate this point further, formal analyses of complexity matching yield another theoretical prediction that has been pursued in other behavioral studies. As noted earlier, complexity matching is predicted to correspond with increased information exchange between two complex systems. The experiment analyzed herein did not include a direct measure of information exchange, but Fusaroli, Abney, Bahrami, Kello, and Tylén (2013) reexamined data from a joint perceptual decision-making task (Bahrami et al., 2010) in which dyads collaborated on visual discrimination judgments. Speech signals were analyzed using similar methods to those herein, and measures of complexity matching were found to correlate with increased performance derived from joint decision making. These results suggest that greater increased complexity matching can correspond with enhanced joint decision making, and by extension with enhanced mutual comprehension as well (Brennan & Clark, 1996; Brennan & Hanna, 2009).

Thus far, we have discussed complexity matching primarily in the context of the interactive alignment model, but it is also related to theories of *interpersonal synergy* (Fusaroli, Rączaszek-Leonardi, & Tylén, 2014; Ramenzoni, Riley, Shockley, & Baker, 2012; Riley, Richardson, Shockley, & Ramenzoni, 2011) that have grown from synergies as theorized in motor systems (Bernstein, 1967; Turvey, 1990, 2007). Synergy is the emergence of coordination via reduction in the degrees of freedom in a system of many interacting components. This concept can be extended from physical systems into the interpersonal interaction domain. By extension, Fusaroli et al. (2014) proposed that interpersonal synergies should (a) be highly sensitive to conversational context, (b) adapt flexibly to changing needs of the task, and (c) self-assemble to minimize variance to manage the degrees of freedom within the interaction (Riley et al., 2011). These entailments of synergies may lead to specific, testable hypotheses about functional specificity and reciprocal compensation in interpersonal coordination that could be meaningfully explored with complexity matching.

Complexity matching may also shed light on mechanisms proposed to explain behavioral matching. In particular, some researchers hypothesize that behavioral matching arises from shared internal representations (Sebanz et al., 2006), while others hypothesize coordinative structures and dynamics (Richardson et al., 2007). Complexity matching cannot be readily explained in terms of sharing representations or mimic dynamics. Complexity matching pushes these hypotheses to become more multiscale in how dyadic interactions affect representations and dynamics.

Beyond these insights into moment-to-moment interaction, we believe that complexity matching could provide significant advances in our understanding of communication and social interaction more broadly. For example, developmental researchers hypothesize that rhythmic coordination between infant and caregiver—akin to behavioral matching—supports infant language learning (Feldman, 2007; Jaffe et al., 2001). But like speech in adult conversations, there is evidence that infant vocalizations also are organized into hierarchical clusters during typical development (Lynch, Oller, Steffens, & Buder, 1995; Oller, 2000). Recent results using AF analyses revealed complexity matching between infant prelinguistic vocalizations and caregiver speech (Abney, Warlaumont, Oller, Wallot, & Kello, 2014). Taken together, these studies suggest that complexity matching may be foundational to the learning and development of interpersonal coordination and communication, unveiling insights not captured by other methodological lenses. We imagine that similar important discoveries in other domains of human interaction and communication may be uncovered when explored in this new light.

Conclusion

Interaction research has relied on measures of behavioral matching as a measure of interpersonal coordination for decades. Complexity matching is a new, *complementary* measure of coordination. Behavioral and complexity matching provided unique insights into the different interactions that occur during affiliative versus argumentative conversations—arguments were characterized by stricter turn-taking, whereas friendly conversations yielded distributional similarities that may reflect the establishment of common ground. Together, these analyses provide a richer view of interaction than either alone. These complementary analyses may be generalized and applied to yield similar insights in other areas of language and interaction research, wherever hierarchical nesting may yield power law scaling in the temporal dynamics of behavior.

References

Abney, D. H., Warlaumont, A. S., Oller, D. K., Wallot, S., & Kello, C. T. (2014, June). *The multiscale dynamics of infants' and caregivers' vocalizations*. Paper presented at the conference Finding Common Ground: Social, Ecological, and Cognitive Perspectives on Language Use, Storrs, CT.

Altmann, E. G., Cristadoro, G., & Degli Esposti, M. (2012). On the origin of long-range correlations in texts. *Proceedings of the National Academy of Sciences, 109,* 11582–11587. doi:10.1073/pnas.1117723109

Aquino, G., Bologna, M., Grigolini, P., & West, B. J. (2010). Beyond the death of linear response theory: Criticality of the 1/f-noise condition. *Physical Review Letters, 105,* 040601. doi:10.1103/PhysRevLett.105.040601

Aquino, G., Bologna, M., West, B. J., & Grigolini, P. (2011). Transmission of information between complex systems: 1/f resonance. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, 83,* 051130. doi:10.1103/PhysRevE.83.051130

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010). Optimally interacting minds. *Science, 329,* 1081–1085. doi:10.1126/science.1185718

Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother–infant interactions. *Journal of Personality and Social Psychology, 54,* 243–253. doi:10.1037/0022-3514.54.2.243

Bernstein, N. A. (1967). *The coordination and regulation of movements*. London, England: Pergamon Press.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18,* 355–387. doi:10.1016/0010-0285(86)90004-6

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1482–1493. doi:10.1037/0278-7393.22.6.1482

Brennan, S. E., & Hanna, J. E. (2009). Partner-specific adaptation in dialog. *Topics in Cognitive Science, 1,* 274–291. doi:10.1111/j.1756-8765.2009.01019.x

Cappella, J. N., & Planalp, S. (1981). Talk and silence sequences in informal conversations III: Interspeaker influence. *Human Communication Research, 7*(2), 117–132.

Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM Review, 51,* 661–703. doi:10.1137/070710111

Coey, C. A., Washburn, A., & Richardson, M. J. (in press). Recurrence quantification as an analysis of temporal coordination with complex systems. In N. Marwan, C. L. Webber, Jr., M. A. Riley, & A. Giuliani (Eds.), *Translational recurrences: From mathematical theory to real-world applications* (pp. 173–186). Chicago, IL: Springer International.

Coupland, N. (1980). Style-shifting in a Cardiff work-setting. *Language in Society, 9,* 1–12. doi:10.1017/S0047404500007752

Cummins, F., & Port, R. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26,* 145–171. doi:10.1006/jpho.1998.0070

Cutting, J. E., & Rosner, B. S. (1974). Categories and boundaries in speech and music. *Perception & Psychophysics, 16,* 564–570. doi:10.3758/BF03198588

Edwards, A. M. (2008). Using likelihood to test for Lévy flight search patterns and for general power-law distributions in nature. *Journal of Animal Ecology*, *77,* 1212–1222. doi:10.1111/j.1365-2656.2008.01428.x

Feldman, R. (2007). Parent–infant synchrony and the construction of shared timing; Physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry, 48,* 329–354. doi:10.1111/j.1469-7610.2006.01701.x

Ferrer-i-Cancho, R., Solé, R. V., & Köhler, R. (2004). Patterns in syntactic dependency networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, 69*, 051915. doi:10.1103/PhysRevE.69.051915

The "Five Graces Group". Beckner, C., Blythe, R., Bybee, J., Christiansen, M. H., Croft, W., . . . Schoenemann, T. (2009). Language is a complex adaptive system: Position paper. *Language Learning, 59,* 1–26.

Fusaroli, R., Abney, D. H., Bahrami, B., Kello, C. T., & Tylén, K. (2013). *Conversation, coupling and complexity: Matching scaling laws predict performance in a joint decision task*. Poster presented at the annual Cognitive Science Conference, Berlin, Germany.

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology, 32,* 147–157. doi:10.1016/j.newideapsych.2013.03.005

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition, 27,* 181–218. doi:10.1016/0010-0277(87)90018-7

Grosjean, F., Grosjean, L., & Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology, 11,* 58–81. doi:10.1016/0010-0285(79)90004-5

Haken, H., Kelso, J. S., & Bunz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51,* 347–356. doi:10.1007/BF00336922

Healey, P. G. T. (2008). Interactive misalignment: The role of repair in the development of group sub-languages. In R. Cooper & R. Kempson (Eds.), *Language in flux* (pp. 13–39). London, United Kingdom: College Publications.

Healey, P. G. T., Purver, M., & Howes, C. (2014). Divergence in dialogue. *PLoS ONE, 9*(6): e98598. doi:10.1371/journal.pone.0098598

Holden, J. G., Van Orden, G. C., & Turvey, M. T. (2009). Dispersion of response times reveals cognitive dynamics. *Psychological Review, 116,* 318–342. doi:10.1037/a0014849

Howes, C., Healey, P. G. T., & Purver, M. (2010). Tracking lexical and syntactic alignment in conversation. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd annual conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Jaeger, T. F., & Snider, N. (2008). Implicit learning and syntactic persistence: Surprisal and cumulativity. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the Cognitive Science Society* (pp. 827–812). Washington, DC: Cognitive Science Society.

Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P., & Stern, D. N. (2001). Rhythms of dialogue in infancy: Coordinated timing in development [Monograph]. *Monographs of the Society for Research in Child Development, 66* (No. 2).

Keller, P. E., & Repp, B. H. (2004). When two limbs are weaker than one: Sensorimotor syncopation with alternating hands. *The Quarterly Journal of Experimental Psychology. A. Human Experimental Psychology, 57,* 1085–1101. doi:10.1080/02724980343000693

Kello, C. T., & Beltz, B. C. (2009). Scale-free networks in phonological and orthographic wordform lexicons. In I. Chitoran, C. Coupé, E. Marsico, & F. Pellegrino (Eds.), *Approaches to phonological complexity* (pp. 171–192). Berlin, Germany: Mouton de Gruyter. doi:10.1515/9783110223958.171

Kelso, J. A. S. (1981). On the oscillatory basis of movement. *Bulletin of the Psychonomic Society, 18,* 9–63.

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

Kugler, P. N., & Turvey, M. T. (1987). *Information, natural laws, and self-assembly of rhythmic movement.* Hillsdale, NJ: Erlbaum.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54,* 358. doi:10.1037/h0044417

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

Lynch, M. P., Oller, D. K., Steffens, M. L., & Buder, E. H. (1995). Phrasing in prelinguistic vocalizations. *Developmental Psychobiology, 28,* 3–25. doi:10.1002/dev.420280103

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

Marmelat, V., & Delignières, D. (2012). Strong anticipation: Complexity matching in interpersonal coordination. *Experimental Brain Research, 222,* 137–148. doi:10.1007/s00221-012-3202-9

McClave, E. Z. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics, 32,* 855–878. doi:10.1016/S0378-2166(99)00079-X

Mills, G. J. (2014). Dialogue in joint activity: Complementarity, convergence and conventionalization. *New Ideas in Psychology, 32*, 158–173. doi:10.1016/j.newideapsych.2013.03.006

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

Mitzenmacher, M. (2004). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics, 1,* 226–251. doi:10.1080/15427951.2004.10129088

Neumann, R., & Strack, F. (2000). "Mood contagion": The automatic transfer of mood between persons. *Journal of Personality and Social Psychology, 79,* 211–223.

Oller, D. K. (2000). *The emergence of the speech capacity*. Mahwah, NJ: Erlbaum.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119,* 2382. doi:10.1121/1.2178720

Paxton, A., & Dale, R. (2013). Argument disrupts interpersonal synchrony. *The Quarterly Journal of Experimental Psychology, 66,* 2092–2102. doi:10.1080/17470218.2013.853089

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39,* 633–651. doi:10.1006/jmla.1998.2592

Pickering, M. J., & Branigan, H. P. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences, 3,* 136–141. doi:10.1016/S1364-6613(99)01293-0

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27,* 169–190. doi:10.1017/S0140525X04000056

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, 31,* 1253–1267. doi:10.1016/j.humov.2011.12.004

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science, 26,* 867–891. doi:10.1016/j.humov.2007.07.002

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

Sales-Pardo, M., Guimera, R., Moreira, A. A., & Amaral, L. A. N. (2007). Extracting the hierarchical organization of complex systems. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 104,* 15224–15229. doi:10.1073/pnas.0703740104

Schmidt, R. C., Morr, S., Fitzpatrick, P., & Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *Journal of Nonverbal Behavior, 36,* 263–279. doi:10.1007/s10919-012-0138-5

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In A. Fuchs & V. K. Jirsa (Eds.), *Coordination: Neural, behavioral and social dynamics* (pp. 281–307). Berlin, Germany: Springer-Verlag.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences, 10,* 70–76. doi:10.1016/j.tics.2005.12.009

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science, 1,* 305–319. doi:10.1111/j.1756-8765.2009.01021.x

Simon, H. A. (1973). The organization of complex systems. In R. S. Cohen & M. W. Wartofsky (Eds.), *Models of discovery* (pp. 245–261). Dordrecht, the Netherlands: Springer.

Simon, H. A. (1977). How complex are complex systems? *Proceedings of the 1967 Biennal Meeting of the Philosophy of Science Association, 2,* 507–522.

Spivey, M. (2007). *The continuity of mind*. New York, NY: Oxford University Press.

Stanley, H. E. (1987). *Introduction to phase transitions and critical phenomena*. New York, NY: Oxford University Press.

Stephen, D. G., & Dixon, J. A. (2011). Strong anticipation: Multifractal cascade dynamics modulate scaling in synchronization behaviors. *Chaos, Solitons, & Fractals, 44,* 160–168. doi:10.1016/j.chaos.2011.01.005

Stephen, D. G., Stepp, N., Dixon, J. A., & Turvey, M. T. (2008). Strong anticipation: Sensitivity to long-range correlations in synchronization behavior. *Physica A: Statistical Mechanics and its Applications, 387,* 5271–5278.

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. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 106,* 10587–10592. doi:10.1073/pnas.0903616106

Strogatz, S. H., & Mirollo, R. E. (1991). Stability of incoherence in a population of coupled oscillators. *Journal of Statistical Physics, 63,* 613–635. doi:10.1007/BF01029202

Suchman, L. (2007). *Human-machine reconfigurations: Plans and situated actions.* Cambridge, England: Cambridge University Press.

Swenson, R., & Turvey, M. T. (1991). Thermodynamic reasons for perception-action cycles. *Ecological Psychology, 3,* 317–348. doi:10.1207/s15326969eco0304_2

Teich, M. C., Heneghan, C., Lowen, S. B., Ozaki, T., & Kaplan, E. (1997). Fractal character of the neural spike train in the visual system of the cat. *Journal of the Optical Society of America A, 14,* 529–546. doi:10.1364/JOSAA.14.000529

Thurner, S., Lowen, S. B., Feurstein, M. C., Heneghan, C., Feichtinger, H. G., & Teich, M. C. (1997). Analysis, synthesis, and estimation of fractal-rate stochastic point processes. *Fractals, 5,* 565–595. doi:10.1142/S0218348X97000462

Turalska, M., West, B. J., & Grigolini, P. (2011). Temporal complexity of the order parameter at the phase transition. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics, 83,* 061142. doi:10.1103/PhysRevE.83.061142

Turvey, M. T. (1990). Coordination. *American Psychologist, 45,* 938–953.

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

Van Orden, G., Hollis, G., & Wallot, S. (2012). The blue-collar brain. *Frontiers in Physiology, 3.* doi:10.3389/fphys.2012.00207

Voss, R. F., & Clarke, J. (1978). "1/f noise" in music: Music from 1/f noise. *The Journal of the Acoustical Society of America, 63,* 258. doi:10.1121/1.381721

Webb, J. T. (1969). Subject speech rates as a function of interviewer behaviour. *Language and Speech, 12*(1), 54–67.

West, B. J., Geneston, E. L., & Grigolini, P. (2008). Maximizing information exchange between complex networks. *Physics Reports, 468,* 1–99. doi:10.1016/j.physrep.2008.06.003

Zipf, G. K. (1949). *Human behavior and the principle of least effort: An introduction to human ecology*. Reading MA: Addison-Wesley.

Received January 16, 2014
Revision received August 6, 2014
Accepted August 7, 2014 ■