



Cognitive Science (2012) 1–23

Copyright © 2012 Cognitive Science Society, Inc. All rights reserved.

ISSN: 0364-0213 print / 1551-6709 online

DOI: 10.1111/j.1551-6709.2012.01269.x

Behavior Matching in Multimodal Communication Is Synchronized

Max M. Louwerse,^a Rick Dale,^b Ellen G. Bard,^c Patrick Jeuniaux^d

^a*Department of Psychology, University of Memphis*

^b*Cognitive and Information Sciences, University of California, Merced*

^c*Department of Psychology, University of Edinburgh*

^d*University of Memphis*

Received 13 November 2011; received in revised form 21 February 2012; accepted 29 February 2012

Abstract

A variety of theoretical frameworks predict the resemblance of behaviors between two people engaged in communication, in the form of coordination, mimicry, or alignment. However, little is known about the time course of the behavior matching, even though there is evidence that dyads synchronize oscillatory motions (e.g., postural sway). This study examined the temporal structure of nonoscillatory actions—language, facial, and gestural behaviors—produced during a route communication task. The focus was the temporal relationship between matching behaviors in the interlocutors (e.g., facial behavior in one interlocutor vs. the same facial behavior in the other interlocutor). Cross-recurrence analysis revealed that within each category tested (language, facial, gestural), interlocutors synchronized matching behaviors, at temporal lags short enough to provide imitation of one interlocutor by the other, from one conversational turn to the next. Both social and cognitive variables predicted the degree of temporal organization. These findings suggest that the temporal structure of matching behaviors provides low-level and low-cost resources for human interaction.

Keywords: Behavior matching; Synchronization; Mimicry; Alignment; Coordination; Multimodal communication; Face-to-face conversation; Entrainment

1. Introduction

The spatial and temporal structure of behavior is integral to the functioning of living communities. Such structure is striking in non-human animals. Fish in schools synchronize

Correspondence should be sent to Max M. Louwerse, Department of Psychology/Institute for Intelligent Systems, University of Memphis, 202 Psychology Building, Memphis, TN 38152. E-mail: maxlouwerse@gmail.com

Patrick Jeuniaux is now at Institut National de Criminalistique et de Criminologie, Brussels, Belgium.

direction and speed with their neighbors (Partridge, 1981). Birds in flocks synchronize take-off and landing (Ward, Axford, & Krause, 2002). Male and female mosquitoes synchronize wing beats (Cator, Arthur, Harrington, & Hoy, 2009). In humans, there is evidence that individuals match their behaviors in spatial organization (i.e., they can imitate each other's behaviors) and in temporal organization (i.e., they can coordinate their behaviors) (Bernieri & Rosenthal, 1991; Grammer, Kruck, & Magnusson, 1998; Richardson, Marsh, & Schmidt, 2005). We refer to combined spatial and temporal behavior matching as *synchronization* in the broad sense of the term.

The manner and extent to which people synchronize behavior matching during natural behavior and the mechanisms responsible for that synchronization are of growing theoretical significance. In particular, face-to-face interaction involves multiple channels—language, gesture, and other behaviors—within which individuals may synchronize matching behavior. Despite research showing synchronization of matching behavior in some individual channels (Wang, Newport, & Hamilton, 2011), we do not yet know whether such synchronization characterizes all channels during face-to-face interaction, what precise temporal structure synchronization follows, and what functions it may serve. The goal of this article is to embark on preliminary answers to these questions by focusing on the synchronization of matching behaviors in face-to-face interactions. Whether *different*, nonmatching behaviors (e.g., nodding and gesturing) are also synchronized is outside the scope of the current article and is the subject of another study.

There is already evidence that people *imitate* others' behavior, but it is unclear whether they necessarily align these matched behaviors in real time. Imitation is perhaps most obvious in intentional mockery, purposeful replication of selected characteristics of another's behavior at lags of up to years from the original execution of behavior and often not in the presence of the individual being imitated. More interesting is the *unintentional* and fairly *immediate* "mimicry" (Hurley & Chater, 2005) that might make a yawn spread around a room (Platek, et al., 2003). Subtler still is the correlation in frequency between interlocutors in incidental mannerisms like shaking a foot or rubbing the nose, even when neither is aware of imitating the other (Chartrand & Bargh, 1999). Such unconscious imitation can be found within different modalities (Chartrand & Bargh, 1999; Chartrand, Maddux, & Lakin, 2005; Dijksterhuis & Bargh, 2001). People imitate lexical choice (Garrod & Anderson, 1987), accents (Giles & Powesland, 1975), pauses (Cappella & Planalp, 1981), speech rate (Webb, 1969), tone of voice (Neumann & Strack, 2000), syntax (Branigan, Pickering, & Cleland, 2000), emotions (Hatfield, Cacioppo, & Rapson, 1994), and moods (Neumann & Strack, 2000). Even newborns copy adults' facial gestures (Meltzoff & Moore, 1977). With the possible exception of mockery, imitation has social benefits (Van Baaren, Holland, Steenaerts, & van Knippenberg, 2003). For instance, Chartrand and Bargh (1999) showed that participants whose foot-shaking or nose-rubbing was mimicked perceived the interaction as running more smoothly than did participants who were not mimicked. Bailenson and Yee (2005) demonstrated that participants who were presented with an argument from a virtual embodied agent who mimicked them were more persuaded and liked the agent more than participants who were interacting with an agent that used prerecorded nonmimicked behavior.

There is also evidence that people *coordinate* their behavior, without necessarily imitating one another, when they collaborate to solve problems with mutually understood structure (Sebanz, Bekkering, & Knoblich, 2006). People may perform delicately aligned but complementary actions. When two people lift a box, have a telephone conversation, or conduct a financial transaction (Clark, 1996), they must perform slightly different actions in a properly time-aligned way or the box falls, the conversation fails, and the business transaction ends in confusion. Similarly, people coordinate their different dialog turns in the correct sequence, for example, when one asks a question and the other answers it (Schegloff & Sacks, 1973).

Synchronization, imitation, and coordination are widely recognized in the literature on co-action. *Synchronization* is also called entrainment (Shockley, Baker, Richardson, & Fowler, 2007). *Imitation* is also called mimicry (Chartrand & Bargh, 1999) or contagion (Gambetta, 1988; Hatfield et al., 1994). *Coordination* is also called joint action (Clark, 1996) or cooperative action (Fowler, Richardson, Marsh, & Shockley, 2008). The distinction between synchronization, imitation, and coordination is a subtle one. Our distinction comes closest to Semin and Cacioppo's (2008) proposal. Coordination appears to be a largely conscious mechanism, an intentional attempt to participate in a joint activity. When people are working toward the same goal in the same timescale and on the same physical objects, their teleologically coordinated joint activity is time aligned. Because coordinated activities are often complementary, coordination need not be imitative. Imitation, on the other hand, is often unconscious and automatic, and because it can serve as social glue over an extended timescale, it does not have to be time aligned. Synchronization of matching behavior (synchronization henceforth) is both time aligned like coordination and form aligned like imitation. It is therefore possible that synchronization shares features of both coordination (problem solving) and imitation (empathy, affiliation, etc.) while offering any individual the benefits of a system that is sensitive to the behavior of any other, even if no affiliative or coordinating action is required.

There is evidence people synchronize several behaviors. Unintentional synchronization of movements has been found for the swinging of hand-held pendulums (Richardson et al., 2005), postural sway (Shockley, Santana, & Fowler, 2003), and the motion of rocking chairs (Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007). For instance, when participants were involved in a joint puzzle task, they synchronized their postural sway, sharing more postural configurations, and maintaining similar postural trajectories longer when they could see each other and worked together on the puzzle task, than when they did not (Shockley et al., 2003). While these findings show synchronizing behavior in dyads, they intentionally address behaviors that are not integral to human interaction. Such results do not tell us whether synchronized behavior is naturally pervasive or functional.

Pickering and Garrod's (2004) interactive alignment account suggests that synchronization should be common and functional. They argue that linguistic representations (situation model, semantic, syntactic, phonological and phonetic representations) employed by dialog partners become aligned at different stages of comprehension and production processes as a result of a largely automatic process. As interlocutors align their linguistic representations, overt linguistic behaviors also align in form. For instance, a dyad of interlocutors converging on a spatial description scheme in a route-navigation task has probably aligned their

mental representations of the route display (Garrod & Anderson, 1987; Garrod & Doherty, 1994). Branigan et al. (2000) demonstrated that individuals align more strongly at one level (syntactic representations) when they align at another (semantic representations). They also demonstrated that alignment of syntactic form can come about quickly enough to be genuine synchronization: When speakers alternate in describing a sequence of objects, one will use the structure just employed by the other.

In genuine dialog, Garrod and Pickering (2009) suggest a dyad's behaviors dovetail even more closely: They propose that as the first speaker produces a structure, the second is analyzing it by generating a similar one. In fact, to be of use in the comprehension process, one person's uttered structure must immediately prime its counterpart within the other person. Thus, the evidence for interactive alignment reviewed in Pickering and Garrod (2004) is indicative of synchronization of linguistic channels. This extensive work has generated considerable evidence for alignment but has focused on one or two linguistic channels at a time. There has yet to be a demonstration that synchrony of matching behavior pervades many channels in the simultaneous manner that the interactive-alignment theory predicts. Moreover, this work has not yet explored alignment across individuals either for paralinguistic gestures that are interpretable in a linguistic representation, such as pointing, iconic hand shape, or body posture, or for behaviors of uncertain semantic importance that people may align with their own linguistic output, such as brow raising (Flecha-García, 2010). Nor has existing work examined behaviors that indicate internal states, such as frowning, or aperiodic movements with no obvious symbolic representation, such as touching one's cheek. If dialog mechanisms are part of a larger system of imitation or coordination, interspeaker synchronization could be pervasive.

The purpose of this study was two-fold: first, to discover the extent to which synchronization occurs within each of the many different channels unfolding during naturalistic interaction; and second, to determine whether any such synchronization lends itself to social and communicative goals. Although there have been long-standing attempts at identifying this kind of time alignment (e.g., Condon & Ogston, 1967), the channels explored in each case have been few in number and the results have been, at best, mixed (e.g., McDowall, 1978). If there is a pervasive phenomenon at work, then we should find dyads synchronizing in linguistic (e.g., dialog acts, lexical items), expressive (e.g., facial expressions, manual gestures), and nonexpressive actions (e.g., use of mannerisms). If synchrony serves a dialog function, it should be characterized by latencies short enough to allow the perceptions involved to contribute to next turns or coordinating actions. If cross-speaker within-behavior synchrony unites imitation and coordination, we should find that all kinds of synchronization respond to many of the same factors, both social and communicative. To test these predictions, we investigated a multimodal spectrum of channels in a corpus of face-to-face route-communication dialogs.

1. Experiment

This study used a version of the Map Task (Anderson et al., 1991), an unscripted route-communication task, in which pertinent knowledge is distributed between a person whose

map includes a route (the Instruction Giver) and another person (the Instruction Follower) who has to reproduce it on a similar but not identical map. The Map Task makes it possible to control base conditions, genre, topic, and goals of dialogs while allowing interlocutors full freedom of expression at all times.

Three features of the present design made it possible to test for common effects on synchronization. First, assignment of roles produced social interactive asymmetry: the Instruction Giver always knew what the next subgoal of the dialog was and characteristically initiated subtasks and determined strategy. As imitation of socially dominant individuals is likely (cf. Van Baaren, Janssen, Chartrand, & Dijksterhuis, 2009) and synchronization shares features with imitation, any asymmetry in roles should bring an asymmetry in synchrony: The Instruction Follower is more likely to do what the Instruction Giver has just done than vice versa. Second, dyads participated in multiple dialogs. There is evidence that imitation tends to increase over time, with accents, for instance, converging as people speak to one another more (Giles, Coupland, & Coupland, 1991). If synchronization shares features with imitation, it should also increase over time. Finally, the maps were designed to vary the difficulty of communication. If synchronous behavior has the potential for use in social interaction, it should be robust. When communication becomes more difficult, the phenomenon should not disappear. If all modalities respond to the same variables in the same way, we will have good reason to attribute them to a common mechanism.

1.1. Method

1.1.1. Participants

A total of 48 students (24 dyads; 30 females and 18 males; 19 African American, 1 Asian, and 28 Caucasian) from the University of Memphis participated for payment. All were native speakers of English.

1.1.2. Materials

The maps were designed so that description was necessary: occasional landmarks differed between the Instruction Giver's and the Instruction Follower's maps. Half the basic maps showed cartoon "landmarks" of homogeneous type (e.g., all birds or all bugs), while the other half showed a mixed array of landmarks (e.g., birds and bugs). To make clear distinctions where all landmarks shared a type, interlocutors needed to use landmark features (color, number, location) in addition to the basic name for the type (e.g., *bird*, *bug*). Each map pair had a different route shape and pattern of differences (Fig. 1 shows examples of Instruction Giver and an Instruction Follower maps).

Access to landmark color differed between interlocutors. While the Instruction Giver maps were always fully colored, irregularly shaped "inkblots" grayed out some landmarks on the Instruction Follower maps, while leaving number and shape visible. To vary the difficulty interlocutors would have in establishing a common view of common objects, the location of inkblots was varied. In half the basic maps, the grayed items were covered by a single irregular inkblot, which the Instruction Follower might explain to the Instruction

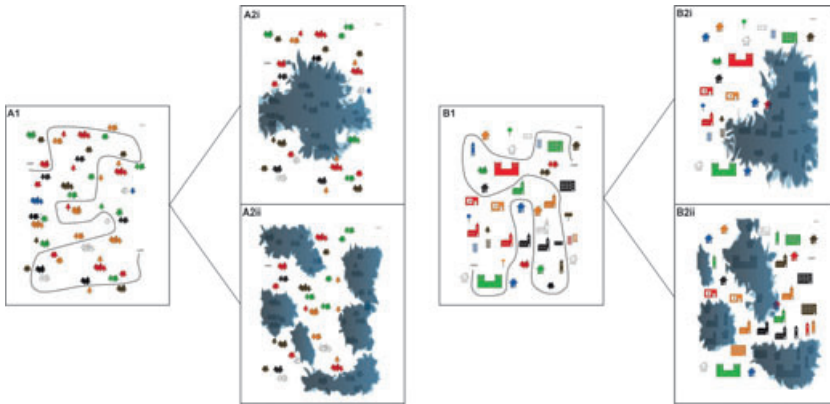


Fig. 1. A sample of six maps used in the experiment. The maps for the Instruction Giver (1) are on the left, and the maps for the Instruction Follower (2) are on the right. The goal of the task is to reproduce the Instruction Giver's route on Instruction Follower's map. Figure A shows the homogeneous objects condition with orderly inkblots (i) and disorderly inkblots (ii), Figure B shows the mixed objects condition with orderly ink blots (i) and disorderly inkblots (ii). In total 2 (orderly and disorderly) \times 2 (homogeneous and mixed objects) \times 8 (tree, bird, alien, bug, car, fish, house, traffic sign) = 32 different maps were used.

Giver fairly simply, while in the other half, an equal number of grayed items lay under scattered inkblots whose locations had to be established piecemeal. Each dyad worked on eight maps, half homogeneous, half with grouped grayed items (inkblots), and sampling all the types of landmarks. Interlocutors exchanged roles after four maps.

1.1.3. Apparatus

Each interlocutor's actions were captured by two individual camcorders, one for the face Osaka, Japan (Panasonic PV-GS31), another for the upper torso Osaka, Japan (Panasonic PV-GS150). A fifth camcorder captured both participants from overhead (Osaka, Japan Panasonic PV-GS150). Each interlocutor's speech was recorded via a headset microphone (Vienna, Austria AKG C420) on a separate audio channel (Kanagawa, Japan Marantz PMD670 recorder). Two high-resolution webcams provided each interlocutor with a view of the other's upper body. The Instruction Follower's drawings of the routes on the screen were recorded both spatially and temporally.

1.1.4. Procedure

Participants were seated face to face but separated by a divider to ensure that they could not see each other directly. The left half of the computer monitor in front of them displayed the upper torso of the dialog partner. The right half of the monitor showed the map. This design allowed us to monitor eye gaze, facial expression, and gestures in relation to the dialog partner and the map.

The Instruction Giver and the Instruction Follower were told to work together to enable the Instruction Follower to reproduce on his or her on-screen map the route available on the Instruction Giver's version. To maximize interaction, participants were promised extra

payment if the Instruction Follower reproduced the Instruction Giver's route perfectly. Participants were told that they and their partner had maps of the same location drawn by different explorers and so potentially different in detail. They were not told where or how the maps differed. Participants could not view each other's maps.

Equipment was calibrated before the start of each conversation. The five camcorders were positioned and focused to best capture the facial and the upper torso movements of each participant. Each conversation started with a flash of light and the sounding of a brief tone, to permit precise temporal alignment of all recorded channels.

1.1.5. Coding

All video and audio recordings were coded for all behaviors listed in Table 1 by modality group (e.g., Language, Manual Gesture, Face and Head) and channel (e.g., eyes, brows, dialog acts). In each case an explicit system was applied by all coders to a subset of dialogs, and intercoder agreement established before coders worked individually on the remaining materials. All coded events were marked for onset and offset. To prevent cross-contamination within dyads, dyad identifiers were removed before coders analyzed each participant's record separately.

1.1.5.1. Face: Coders classified video recordings of interlocutors' faces for 14 units inspired by the Ekman, Friesen, and Hager (2002) facial action coding scheme. Facial expressions can be classed according to mouth movements (such as pushing the lips forward and pulling medially to make them pucker), eye movements (such as squinting or rolling the eyes), and eyebrow movements (such as lowering the eyebrows as in frowning or raising the outer eyebrows as if in surprise). In addition, three facial movements were identified that could not be directly linked to the Ekman et al. (2002) scheme (head nodding, head shaking, and asymmetrical eyebrows).

A total of 16 facial movements from four main categories (head [2], eyes [4], eyebrows [3], and mouth [7]) were time stamped and coded. Cohen's κ (.78) showed high agreement among three coders working on the 32 dialogs produced by four dyads (one-sixth of the corpus). Because coding is extremely time intensive (approximately an 8:1 ratio for coding time to elapsed dialog time), the remaining dialogs were coded by individual coders whose high interrater reliability had been established.

1.1.5.2. Manual gesture: Gesture categories followed McNeill's (1992) coding system at the level of gesture types, rather than specific movements. Five types were distinguished: beat, deictic, iconic, metaphoric, and symbolic (Louwerse & Bangerter, 2010). Beat gestures mark speech rhythm with beating of a finger, hand, or arm; deictic gestures point at a referent (e.g., pointing out a location on the map with one's finger); iconic gestures illustrate what is being said (e.g., gesturing the overall path of the route or the shape of the landmark); metaphoric gestures concretely convey a concept is being explained (e.g., moving hands toward one another conveying smallness); symbolic gestures are conventional markers (e.g. thumbs up). Cohen's κ for these five gestures was .82 for three trained judges coding four dyads.

Table 1
 Overview of behavioral coding schemes by modality group and channel (Ekman et al., 2002 codes in parentheses)

Modality Group	Channels		
	Mouth	Eyes	Eyebrows
Face and Head			Head
	Laughing	Blink (AU45)	Nodding
	Lip tightener (AU23)	Rolling eyes (M68)	Shaking
	Mouth in ‘‘o’’-shape (AU27)	Squinting (AU44)	Asymmetrical
	Mouth open (AU25/26)	Widening eyes (AU5)	Down-frowning (AU4)
	Pout (AU17)		Outer brow raiser (AU2)
Manual gesture	Pucker (AU23)		
	Smile (AU12)		
	Beat	Deictic	Iconic (route/landmark)
Touch face	Touching cheek	Chin rest	
Language	Dialog Acts		Discourse Connectives
	Acknowledgment	Query-W	Color
	Align	Query-YN	Compass direction
	Check	Ready	Digit
	Clarify	Reply-N	Relative direction
	Explain	Reply-W	Spatial prepositions
	Instruct	Reply-Y	
			<i>Alright</i>
			<i>No</i>
			<i>Ok</i>
		<i>Um</i>	
		<i>Well</i>	
		<i>Yes</i>	

1.1.5.3. Touching face: In addition to clearly communicative or potentially expressive actions, coders noted two frequent behaviors found in face-to-face communications: touching the cheek with the fingers and resting the chin on the palm or fist. These were coded according to the standard process.

1.1.5.4. Language: All speech was orthographically transcribed. It was coded systematically for three channels: (a) dialog acts, (b) discourse connectives, and (c) landmark descriptions.

Dialog acts classify the meaning of an utterance at the pragmatic level (Austin, 1962). Twelve dialog acts typically used for Map Task scenarios (Carletta et al., 1997) were coded (see Table 1). These dialog acts can be classed as initiation of a new discourse purpose, response to a previous turn, or preparation for a new dialog. Initiation acts are dialog acts that instruct the interlocutor to carry out an action (e.g., *Go between x and y*), that explain that query, or that check with the interlocutor to confirm information. Responses, on the other hand, do not initiate a new dialog purpose but respond to a previous dialog act. Examples are acknowledgments showing understanding, replies to a question, and clarifications. Finally, the preparation for a new dialog includes the dialog act “ready” (*Okay, let’s move on to the next one*). Automatic classification of the dialog acts (Louwerse & Crossley, 2006) was followed by a manual cross-check by three trained judges, as described above (Cohen’s κ for 32 dialogs was .67).

Discourse connectives are linguistic cues that display the relationship between chunks of language (Louwerse & Mitchell, 2003; Schiffrin, 1987). The common connectives *alright*, *okay*, *um*, *yes*, and *well* were coded because they occurred throughout the corpus.

1.1.5.5. Landmark descriptions: The transcripts were examined for color terms for the landmark (the two *red* trees), digits (*three* birds), and various expressions of direction—relative direction (e.g., *left*, *right*), compass direction (e.g., *north*, *south*), and spatial prepositions (e.g., *above*, *below*).

1.2. Results and discussion

Conventional statistical techniques like cross-correlation and classical regression are unsuited to examining the temporal alignment of the coded channels because their use requires continuous-scale variables (or simple bivariate codes) and because experimental observations are not independent in time. The data considered here, however, represent high-dimensional nominal event codes. Cross-recurrence analyses are useful for such data, because they can reveal their temporal dynamics and can quantify temporally non-independent observations (Marwan, Romano, Thiel, & Kurths, 2007). Cross-recurrence plots quantify the recurrences of states (e.g., nodding head) between two time series. This nonlinear data analysis allows for comparisons between two streams of events, for example, the Instruction Giver’s nodding and the Instruction Follower’s nodding, as they unfold over time, revealing how often and at what lags the matching behavior occurs. The technique has been used successfully in a number of studies, for instance, illustrating the coupling of eye movements (Richardson, Dale, & Kirkham, 2007) and syntactic patterns (Dale & Spivey,

2006) in dialog, and is akin to a generalized lag sequential analysis (cf. Dale, Warlaumont, & Richardson, 2011).

All coded actions were polled at 250-ms intervals. This gives two time series per dyad (per dialog) for each channel. For example, nodding would be coded 1 for any 250 ms interval when nodding occurred, and null for any where it was absent. These two time series were then subjected to cross-recurrence analysis to determine whether the Instruction Giver and the Instruction Follower synchronized their nods.

Fig. 2 provides an example. It portrays the average cross-recurrence values over all dialogs for within-dyad synchronization in nodding. Along the y-axis is a measure of the *proportion of behavioral matches* between two people while the x-axis shows the offsets of these matching behaviors in time. Scales on the y-axis may change, because they will be proportional to the relative frequency of the behavior. Two unrelated channels will simply have a flat cross-recurrence profile; two synchronous channels will show curvilinear patterns revealing differences over time in rate of matching behaviors. The middle of the horizontal axis marks time $t = 0$ and represents a lag of 0, where participants are both nodding *at precisely the same time*. Both to the left and to the right of this 0, the x-axis marks lags $t \pm n$. These lags reflect the different relative times at which one participant's nod lags behind the other's. To the right of lag 0 the Instruction Follower's nod follows the Instruction Giver's, while to the left the Instruction Giver's nod follows the Instruction Follower's. In Fig. 2, the cross-recurrence curve for nodding has an asymmetrical shape, with one peak higher than the other: the Instruction Follower's nod follows the Instruction Giver's more often than the Instruction Giver's follows the Instruction Follower's. In Fig. 3, the pattern for touching one's check is more symmetrical: The Instruction Follower and the Instruction Giver seem to lead and follow to the same degree. Figs. 2 and 3 show

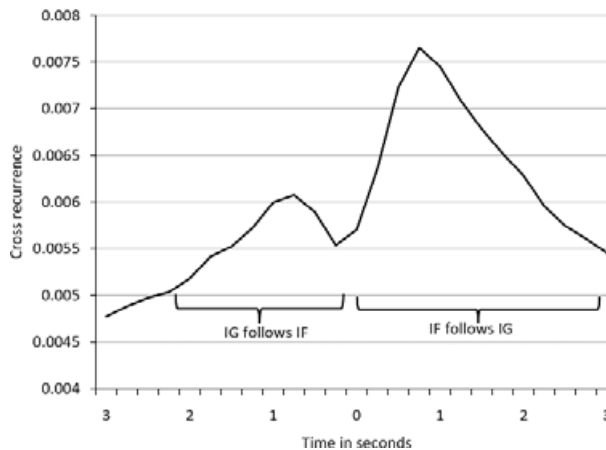


Fig. 2. Cross-recurrence of nodding by the Instruction Giver and the Instruction Follower. Y-axis is proportional to overall frequency of a given behavior. The shape of the cross-recurrence profile reflects the pattern of synchrony between people. The peak of this profile reflects the relative point in time at which behaviors are being matched. The relative infrequency of the behaviors (and behavior matching) explains the low values on the y-axis.

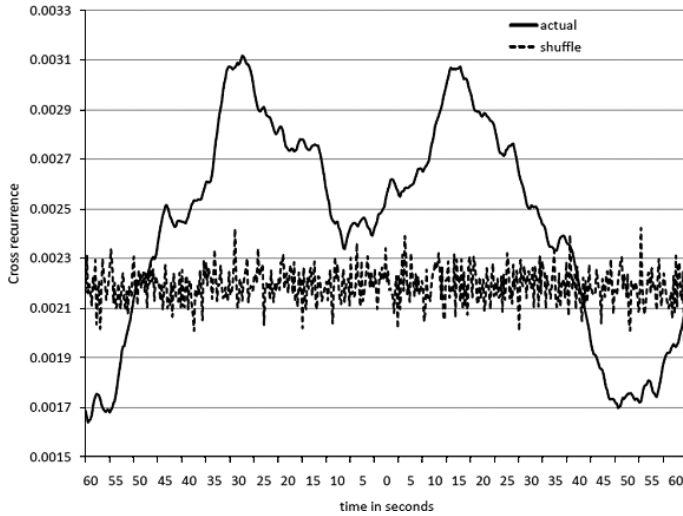


Fig. 3. Cross-recurrence of touching the cheek by the Instruction Giver and the Instruction Follower. The solid line is the cross-recurrence curve. The dashed line is formed by randomizing the order of each interlocutor's temporal series of events and performing a cross-recurrence analysis on the resulting series. Importantly, whether this is significant synchrony is assessed by comparison to the random baseline. The synchrony profile drops beneath random baseline at the edges because the probability of matching behaviors across more disparate time range is often less probable than raw baseline occurrence itself.

different peak lags. While Fig. 2 shows that the Instruction Follower is most likely to repeat the Instruction Giver's nodding behavior 750 ms later, and the Instruction Giver is likely to repeat the Instruction Follower's behavior at about the same latency, Fig. 3 shows that the Instruction Follower and the Instruction Giver are most likely to touch their respective cheeks at a latency of 15–30 s.

To obtain a statistical test for synchrony, we created baseline cross-recurrence data for each channel by randomizing the order of its data points across time. The process turns events into a random string of categories, removing all temporal dependencies in the data.¹ The output appears as an aperiodic dashed “shuffle” line in Fig. 3. To determine whether the dyads synchronize, the statistical analysis tests for a significant difference between the synchronization pattern and the randomized baseline pattern.

Before this can be done, an analysis window must be fixed. The dangers of analyzing a fixed temporal sample of such materials are two-fold: We may mistake a random spike within the sample for organized behavior (i.e., commit a Type I error), and we may miss important events at lags outside the fixed sample (i.e., commit a Type II error). We therefore compared the randomized baseline to the genuine cross-recurrence curve only within the regions where a domed excursion from baseline was observed. The crossing of the actual (synchronization) and shuffled (baseline) cross-recurrence lines at to the left and right ends of the domed excursion bounded the time window for the comparison (e.g., 3 s for nodding behavior in Fig. 2 and 45 s for touching the cheek in Fig. 3).

1.2.1. Synchronization of multimodal channels

A mixed-effects regression analysis was conducted on the recurrence data with synchronization (cross-recurrence vs. baseline) as the fixed factor and both dyads and dialogs as random factors (Baayen, Davidson, & Bates, 2008). The model was fitted using the restricted maximum likelihood estimation (REML). *F*-test denominator degrees of freedom were estimated using the Kenward–Rogers adjustment to further reduce the chance of Type I error (Littell, Stroup, & Freund, 2002). Similar analyses were run independently for each behavior listed in Table 1. Reported results use alpha at .05.

Table 2 reports results for all actions whose recurrence and baseline differed significantly. Evidence for synchronizing behavior was found across all modality groups studied. Within approximately half of the behaviors, speakers synchronized matching behavior. Although this number suggests that synchronization of modalities goes well beyond one or two specific modalities, it is worth considering why some behaviors give significant results while others in the same channels and modality groups do not. A preliminary answer involves the frequency of the coded behaviors. If a behavior is infrequent, synchronized events may be far too rare to produce a significant outcome, or the rare instances might not provide enough exposure for a conversational partner to synchronize to. To test this hypothesis, we ran an independent-samples *t*-test with the frequency of behavior as the dependent variable and the presence versus absence of an effect for synchronization as the independent variable. To do justice to the hypothesis, we excluded the dialog act *Instruction* from this analysis, because the *Instruction Followers* are not in the position to *give* instructions. As predicted, behaviors that yielded synchronization were more frequent than those that did not, $M = .04$ ($SD = .04$) versus $M = .02$ ($SD = .04$), respectively, $t(96) = 2.11$, $p = .04$. This outcome suggests that the absence of a synchronization effect in some behaviors (e.g. mouth in ‘‘o’’ shape, pucker, pout) could be attributed to a lack of opportunity to synchronize rather than to an inherent quality of the behavior.

1.2.2. Latencies

Cross-recurrence allows for nonzero latencies, and a lag can be identified at which maximal cross-recurrence is achieved. As Table 2 shows, the observed peak lags ranged from 0 to just under 28 s, with 10 of the 19 channels under 5 s and 13 (68%) under 10 s. Facial expressions and head movements tend to be matched quickly, typically within 1.5 s. The same is true for those dialog acts and discourse connectives that the interlocutor can act upon immediately, such as the acknowledgment of the dialog turn, a ‘no’ reply, and an agreeing to move on.

In addition to the behaviors whose matching occurs within approximately 1.5 s, there are behaviors whose matching takes longer although typically not more than 25 s after the execution of the behavior by the interlocutor. These seem to be behaviors that are difficult to act upon immediately, such as deictic gestures or explanations, or are awkward to act upon immediately, such as touching the cheek or resting the chin on the palm or fist.

Overall, latencies are generally within a few seconds, the span of at most two conversational turns. The pairing of identical behaviors over this time span indicates that the first

Table 2
Significant cross-recurrence between interlocutors for all actions

Channels	Modality	Action	Observed from Base Line		Excursion (fSec)	Peak lag	Cross-Recurrence versus Baseline ($v_1 = 1$)		Order		Peak Height		Dialog No.	Difficulty
			Start - End	End			v_2	F	IG-IF	IF-IG	IG>IF			
Face & Head		Laughing	0-4.75		0		13,792	122.61	++	++				++
		Smile	0-7.75		0		23,008	1,333.88	++	++		-		++
		Eyebrow down	0-2.75		1.25		5,344	3.85	+	+			++	++
		Eye squint	0-1.75		1		4,576	7.54	++	++				++
		Nodding head	0-3.75		0.75		9,952	62.21	++	++			++	++
Gesture		Shaking head	0-3.00		1		37.82		++	++			++	++
		Deictic	10.00-37.50		25.00		83,680	107.99	++	++			++	++
Touch face		Chinrest	12.50-50		27.50		172,000	11.42	++	++			++	++
		Touch cheek	0-40.00		18.75		114,400	181.25	++	++			++	-
Language:		Acknowledgment	0.25-1.75		0.75		3,808	33.73	++	++			++	++
		Clarify	2.25-8.00		6.75		17,632	12.54	++	++			++	++
Dialog acts		Explain	2.75-27.75		15.00		49,888	111.75	++	++			++	++
		Query-YN	10.50-22.75		16.25		37,600	35.19	++	++			++	++
Discourse		Reply-N	0-1.75		1.00		4,576	60.18	++	++			++	++
		Alright	0.75-4.75		1.50		6,112	3.86	+	+			++	-
Connectives		No	0-2.50		0.75		7,648	56.25	++	++			++	++
		Compass	1.00-16.75		8.75		45,280	59.25	++	++			++	-
Descriptions		Color	1.75-17.00		9.75		32,992	366.30	++	++			++	-
		Digit	2.75-27.75		17.50		68,320	226.25	++	++			++	++

Note. Pluses and minuses mark positive and negative regression coefficients. IG, Instruction Giver; IF, Instruction Follower. The number of symbols indicates p -level: $++p < .01$, $+p < .05$, $-p < .01$, $-p < .05$.

interlocutor's example was processed by the second interlocutor either while the latter was processing what the first said and formulating a response (e.g., face and head movements), or very soon thereafter (e.g., manual gestures and touching face).

1.2.3. Social asymmetry

Table 2 also shows a social effect. While in 15 cases synchrony was bidirectional, that is, each interlocutor imitated the other, in all the asymmetrical cases (4) the Instruction Follower imitated the Instruction Giver—the interlocutor leading the conversation.

To further determine the role of social asymmetry on synchronization, we compared the degree of temporal organization in the cases where there were two cross-recurrence peaks. If social asymmetry explains synchronization, we expect that the Instruction Follower will imitate the Instruction Giver more often than the Instruction Giver imitates the Instruction Follower. Mixed-effects models on genuine recurrence used order of actions as a fixed effect and dyads and dialogs as random effects. As Table 2, Column 9 indicates, for the majority of the behaviors, 12 cases or 63%, the Instruction Follower followed the Instruction Giver to a significantly greater degree than the Instruction Giver followed the Instruction Follower. Two behaviors yielded opposite patterns for social asymmetry, namely smiling (which had very short lags) and color descriptions (less available to the Instruction Follower, whose landmarks were often grayed out). For the remaining five behaviors, where no significant difference was obtained, the pattern was in the predicted direction. Overall, we can conclude that social asymmetry played a role in synchronization. Although synchronization is bidirectional in communication, the interlocutor leading the conversation is more likely to be synchronized to than the interlocutor following directions.

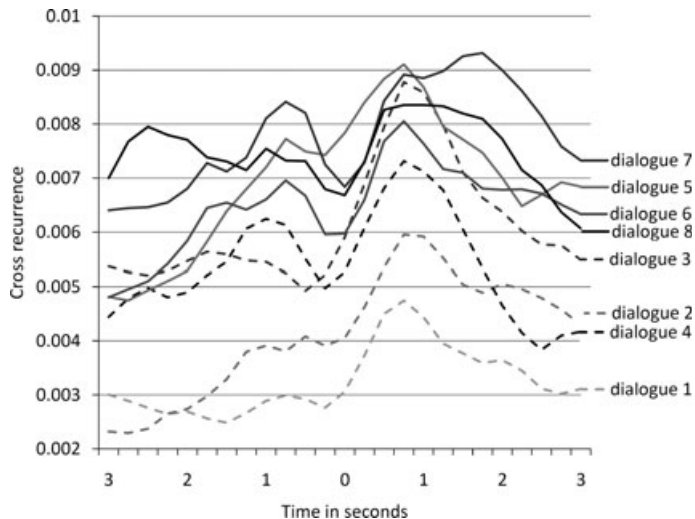


Fig. 4. Cross-recurrence for nodding by dialog.

1.2.4. Cross-recurrence enhances over time

Fig. 4 displays a case of convergence in real cross-recurrence curves for nodding behavior during successive dialogs in a session. A mixed-effects regression model was created for analogous cross-recurrence data for each behavior using dialog number as the fixed factor, and dyads as random factor. In Table 2, the penultimate column shows that in 12 of the 19 behaviors examined, or 63%, synchrony rises over successive dialogs. The effect reverses in two cases, compass directions and colors. For the 12 positive cases, the more interlocutors interacted with each other, the more they synchronized matching behaviors with one another.

This pattern is consistent with a relationship between synchronization and social affiliation (Hove & Risen, 2009), with each enhancing the other. Whether the relationship is causal or not, convergence appears in behaviors from all modality groups, face and head, manual gesture, and language.

1.2.5. Difficulty in communicating

In Table 2, the rightmost column shows the significant changes in cross-recurrence peaks when route navigation became more difficult as a result of irregularly shaped inkblots. A mixed-effects regression model for cross-recurrence with orderly (i.e., easy) versus disorderly (i.e., difficult) inkblots as the fixed factor, and dyads and dialogs as random factors, showed that synchrony not only functioned but strengthened when communication became more difficult. For 11 of the coded actions, or 58%, synchrony increased from the orderly to the disorderly inkblot condition. The three that lost organization included two which the manipulation directly challenged. After all, the distribution of inkblots varied how difficult it was for the Instruction Follower to help the Instruction Giver understand where inkblots obscured color. Without this knowledge, the Instruction Giver, whose landmarks were all colored, could often use a color term where the Instruction Follower, with scattered gray inkblots, could not. Hence, it is no surprise that interlocutors did not synchronize their use of color terms in the more challenging condition. As this same condition was designed to increase confusion, it gave a dyad less opportunity to join in saying *alright*, a discourse connective for agreeing and moving on.

It is worth noting here that the response to difficulty reflects on the convergence effect described above. If convergence over time were due to practice making the task easier, then we would expect that the easier inkblot condition would also show more synchronization within a dyad. But the opposite is true. In eight cases, both practice and difficulty strongly enhance synchronization.

2. Discussion

At the outset of this article, we posed two broad questions: (a) To what extent are matching behaviors synchronized across people who are communicating in a face-to-face setting, and (b) could any such synchrony be a functional part of communication?

In response to the first question, to what extent are matching behaviors synchronized across people who are communicating in a face-to-face setting, we found that during unscripted collaborative face-to-face communication, people synchronize within multiple behaviors: About half of the measured behaviors exhibited synchrony. Those that did not were of lower overall frequency of occurrence, so that lack of temporal pattern might well be due to lack of opportunity rather than to some inherent quality of the behavior. The results show that a dyad's behavior is entrained within each of several modalities, including linguistic expressions, facial expressions, manual gestures, and noncommunicative postures. Although prominent theoretical frameworks predict findings like these (Chartrand & Bargh, 1999; Clark, 1996; Pickering & Garrod, 2004), there has been, to our knowledge, no previous empirical evidence for synchronization in most of these behaviors, and certainly none showing that multiple channels are affected in the same individuals within the same dialogs.

In posing the second question, whether any such synchrony could be a functional part of communication, we proposed that if synchrony serves communicative functions, then it should change in different communicative contexts and occur on a time scale that could implicate the behavior in turn-to-turn communication. Our results indeed show that dyadic synchrony may be involved in the process of communication. First, two effects known for imitation, convergence over time and influence of social role, also characterize the detected synchrony. Second, latencies are generally within a few seconds, the span of at most two conversational turns, putting the behaviors under immediate analysis and potentially accessible to interlocutors as they formulate their utterances. Finally, synchrony tends to become more common when a coordinated view becomes harder to achieve. As task constraints get more difficult, synchrony in fact *increases* in most cases.

The present results show at the very least that people react quickly and similarly to multiple behavioral streams while communicating. If nothing else, this work means that many potential signals are processed at some level during the same interactive task. The fact that the response is imitated means that at least the form of the interlocutor's action is swiftly grasped. However disorganized a person's use of such signals may seem to be, our results support the notion that the multidimensional behavior of one member of a dyad is quickly, robustly, and increasingly available to the other member. This study therefore establishes a core theoretical prediction which merits further testing on laboratory and naturalistic data: For multiple channels present during communication, behavior matching is synchronized within channel, and this synchronization is sensitive to social and task variables.

As in any study, there are some limitations to this work that future research may resolve. We discuss these below in the context of further theoretical consideration. Certainly, no single study can unveil all the potential mechanistic and contextual factors that play into the naturalistic analyses we present. Whatever those factors are, the pervasiveness of synchrony raises theoretical questions. We consider a number of these below, discuss how the results in this article motivate each, and offer potential insights into their solution.

As synchronization is pervasive and surprisingly uniform whatever modality type we investigate, we need to examine the underlying architecture of the systems that drive it. Some theorists, like Pickering and Garrod (2004) and Shockley et al. (2009), would suggest

that synchronization both within and across behaviors is an emergent phenomenon of deeply interconnected processes. Pickering and Garrod (2004) explicitly argue that levels of linguistic representation cascade across each other during interaction (e.g., lexical to syntactic, and back), fueling gradual multilevel alignment in time. Shockley et al. (2009) argue that an individual cognitive system is a self-organizing entity composed of many interacting parts, each constraining the degrees of freedom of the others. This reduction of degrees of freedom means that one behavior, such as deictic gesture, could serve to constrain the space of possible other behaviors, such as dialog moves. During interaction, when two cognitive systems come together, these multicomponent constraints, feeding into synchrony, would produce a cascading multimodal synergy across interlocutors in a task. Minimally, any behavioral channel for one person constrains the same channel for another. Reduction of total degrees of freedom occurs as this process operates over many channels during interaction. Another possibility is that a cognitive or motivational “central executive” may implicitly or explicitly “turn up” or “turn down” synchronization in various interactive contexts while monitoring multiple channels.

Although these are possibilities, we have so far focused on alignment within channels. As the rest of communicative behavior operates on a principle of redundancy control, attenuating one signal when another carries similar information (Aylett & Turk, 2004; Bard & Anderson, 1994; Ferreira, 2003; Levy & Jaeger, 2007; Lieberman, 1963; Lindblom, 1990), it may be that few channels will synchronize at any one instant, so that each requires some separate connection to task variables. One could argue that the channels are controlled by functionally *separate* mechanisms that selectively raise or lower a particular channel’s synchrony during the task. Yet the current results show that the synchronization of channels is in most cases modulated in the same way by the same factors. If synchrony were driven by multiple functionally separate systems, it is unlikely that cross-recurrence instances would covary so reliably. Some common mechanism would appear to influence many behaviors. Whether the mechanism is synergistic or merely widely connected, it may induce channels to act together toward particular ends (whether social affiliation or problem-solving benefits). This leads to a second theoretical question.

2.1. *Functional benefit of multimodal synchrony?*

Whatever the underlying architecture, the presence of synchrony within so many modalities raises the question *why* such behavioral patterns should emerge in the first place. As we have shown here, synchrony correlates significantly with task difficulty and number of conversations, but our results are consistent with a variety of possible functions. Given the assumption that we have evolved systems capable of perception–action coupling across interlocutors, a natural prediction is that synchronization strengthens social affiliation, whatever modality is involved (e.g., Hove & Risen, 2009). Yet we propose that another, and perhaps complementary, function may be relevant: Synchrony may be a recovery device. When participants are actively coordinating in a task, there is a nontrivial possibility of communication breakdown, at which point participants must recover to succeed at the task. Synchronous states may build “at-the-ready” bookmarks for use when higher level,

coordinative processes cannot succeed. This perspective on the function of synchrony sees it as an active and adaptive background process supporting an interactive task. There could be a functional trade-off between low-level, automatic synchrony across levels, and higher level coordinative processes that participants employ.

2.2. Background “hum” of naturalistic interaction?

The data presented here suggest that synchronization is immediate and unintentional, rather than strictly intentional. That is, even though mockery is typically intentional mimicry, it is difficult to explain how the mimicry of so many features in so many multimodal channels—from eyebrow movements to mannerisms—can be intentional and can be under the control of the participant. In fact, as we have seen, previous theoretical and empirical work predicts a pervasive tendency to synchrony. The mechanisms proposed are often themselves of a pervasive nature. For example, the mirroring hypothesis sees perception–action coupling, from low-level action to higher level goals, as the basis for a wide variety of cognitive capacities and their breakdown (e.g., Bekkering et al., 2009; Rizzolatti & Craighero, 2004). The chameleon effect and other results have been explained as part of a “perception-behavior expressway,” predicting that our behaviors are being constantly influenced by socially relevant factors around us (Dijksterhuis & Bargh, 2001). The mere presence of another individual during a task can influence how we represent the task cognitively (e.g., Sebanz et al., 2006). Whatever the functional reason, these cognitive mechanisms may create a widespread synchrony between individuals that is represented at each channel of behavior. So, when we are interacting with another person, any perceptible behavioral channels produce a “background hum” of alignment continually and automatically sustained during that interaction. That background hum could slightly but significantly enhance the probability that participants will choose to use the same behavioral task moves at about the same time during cooperative interaction. This hum may be multifunctional, amplified during affiliation building, and trading off during problem solving. From this perspective, pervasive synchrony is cognitively cheap but potentially useful across contexts and functions.

2.3. Coupling between dyad and task environment?

The results also suggest that there could be coupling beyond the dyad, in a functional relationship between the task environment and the participants themselves. For example, in the case of smiling and laughter, the synchronization occurs almost simultaneously (at lag = 0). As 0 lag gives no time for one to react to the other, the result suggests that both are reacting to a common external stimulus which occurred some time earlier. If participants grow in their awareness of these events, they will laugh together more as the task proceeds. As described by Hutchins (1995) and other approaches to distributed cognition (Rogers & Ellis, 1994), cognitive systems become coupled not just to each other but also to the events and artifacts in their environment. In effect, synchronization need not be *primarily* representational: it may indicate increasingly aligned perception of the external situation.

2.4. Reduction of degrees of freedom

If there is a coherent synergy of synchronization via within- and across-person constraints (Shockley et al., 2009), as well as rapid automatic processes producing a “background hum” of synchronization during interaction, our findings may be pertinent to what has been termed “Bernstein’s problem” (Bernstein, 1967). The problem historically relates to how muscles act together to produce coherent actions when there are so many degrees of freedom that each, individually, may take on (and when, consequently, the overall system of muscles has even more such degrees of freedom, in principle). Nonetheless, through constraints that connect muscle groups and joints, the action system is capable of producing coherent functioning without a “controller” to carry out the work of calculating each muscle’s activity and position relative to all others (Kugler, Kelso, & Turvey, 1980).

Recent perspectives on the reduction of uncertainty (Jaeger, 2006, 2010) and cognitive load (Garrod & Pickering, 2009) are consistent with the idea that synchronization solves this problem for interaction. In the same sense, verbal interaction across people may profit from an active constraining of the space of possible behaviors by cognitive mechanisms such as priming, mirroring, imitation, and so on. Emergent synchronization within any number of modalities is the general description of a solution to the dangerous degrees of freedom of interaction. When two people meet face to face, perplexity is high: A very large selection of possible linguistic and non-linguistic behaviors *could* take place. Multimodal synchronization can reduce degrees of freedom markedly, when one person serves as the constraint for another, and they become in an important (but approximate) sense a functional, coordinative unit (sometimes termed “coordinative structure,” Kugler et al., 1980). Synchronizing within many behaviors may relieve the cognitive system of the burden of constantly computing the next behavior in each of classes 1 to some large n during a task. “Joint cognitive offloading” from one person onto another may assist the cognitive system by reducing detailed planning for each behavioral channel during interaction (Garrod & Pickering, 2009).

Although a radical suggestion, this one is not inconsistent with some other mechanistic accounts (e.g., reduction of uncertainty or perplexity: Jaeger, 2006, 2010; Levy & Jaeger, 2007), grouping linguistic communication with the many natural systems for which the reduction of degrees of freedom is a central problem. Entertaining this description of our synchronization results may fruitfully connect disparate domains, from motor control to linguistic representation. In other words, the “why” of synchrony may be a multifunctional emergent phenomenon from coordinative dynamics. Once established, it could be employed for at-the-ready purposes during communication breakdown, to build affiliation, and so on. This suggests that there is not one stable “mode” in which the system is functioning; instead, the different levels of organization of interaction, from motor behavior to linguistic descriptions, may actively constrain but also adapt to each other in different contexts. The task-based modulation of synchrony here suggests this.

3. Conclusion

The current data find in several modality groups the synchronization of matching behaviors that many theoretical accounts predict for social interaction. The mechanisms underlying this widespread synchronization seem to have a unitary character, given the simultaneous modulation of the synchrony in our results. This modulation may serve many functions. The observed correlations with continued collaborative dialog implicate social affiliation as a driver, but the added correlations with task difficulty suggest that synchronization provides adaptive capacity for the complex process of problem solving in groups. These are not mutually exclusive possibilities, and it would not be surprising if interactive mechanisms were relevant to multiple interactive functions. This article suggests that exploring synchronous behavior matching in the many channels available in naturalistic interaction is a frontier issue in our understanding of how people interact successfully.

Note

1. The shuffled baseline simply reflects the raw probability that separate individual events registered in two behavioral records (e.g., one nod time slice) will overlap if we randomize their locations in time. In a series of simulations, we compared this ‘shuffled’ baseline to a baseline created by a ‘surrogate’ method in which ‘pseudo-dyads’ were created by randomly pairing members of different dyads, but each temporal record was left intact. The surrogate baseline has the benefit of preserving natural sequences of events in each record. In these simulations, the surrogate method estimates a lower average baseline than the shuffled method does, thus increasing the difference between the cross-recurrence pattern in real dyads and the baseline measure, and with it the impression that real dyads’ behaviors were temporally coordinated. Conservatively, we used the smaller differences between real cross-recurrence and the shuffled baseline in testing for the significance of observed effects. See Dale, Richardson, & Kirkham (2011, Appendix) and Richardson & Dale (2005, Figure 3) for other examples.

Acknowledgments

This research was supported by grant NSF-IIS-0416128 and NSF-BCS-0826825. We thank Nick Benesh, Markus Guhe, and Mark Steedman for their help in the design of the experiment and analysis of the data, and many graduate and undergraduate students for their help with conducting the experiment and coding the multimodal channels, specifically Mohammed Ehsan Hoque, Gwyneth Lewis, Divya Vargheese, Jie Wu, Bin Zhang, and Megan Zirnstein. The usual exculpations apply.

References

- Anderson, A., Bader, M., Bard, E., Boyle, E., Doherty, G. M., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The HCRC Map Task Corpus. *Language and Speech, 34*, 351–366.
- Austin, J. L. (1962). *How to do things with words*. Oxford, England: Oxford University Press.
- Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Language and Speech, 47*, 31–56.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.
- Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16*, 814–819.
- Bard, E. G., & Anderson, A. (1994). The unintelligibility of speech to children: Effects of referent availability. *Journal of Child Language, 21*, 623–648.
- Bekkering, H., Bruijn, E. R. A. de, Cuijpers, R. H., Newman-Norlund, R., Schie, H. T. van, & Meulenbroek, R. (2009). Joint Action: Neurocognitive Mechanisms Supporting Human Interaction. *Topics in Cognitive Science, 1*, 340–352.
- Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rime (Eds.), *Fundamentals of nonverbal behavior: Studies in emotion & social interaction* (pp. 401–432). New York: Cambridge University Press.
- Bernstein, N. (1967). *The coordination and regulation of movements*. New York: Pergamon.
- Branigan, H. P., Pickering, M. J., & Cleland, A. A. (2000). Syntactic coordination in dialogue. *Cognition, 75*, B13–25.
- Cappella, J., & Planalp, S. (1981). Talk and silence sequences in informal conversations III: Interspeaker influence. *Human Communications Research, 7*, 117–132.
- Carletta, J., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., & Anderson, A. (1997). The reliability of a dialogue structure coding scheme. *Computational Linguistics, 23*, 13–31.
- Cator, L. J., Arthur, B. J., Harrington, L. C., & Hoy, R. R. (2009). Harmonic convergence in the love songs of the dengue vector mosquito. *Science, 323*, 1077–1079.
- 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.
- Chartrand, T. L., Maddux, W., & Lakin, J. (2005). Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In R. Hassin, J. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 334–361). New York: Oxford University Press.
- Clark, H. H. (1996). *Using language*. Cambridge, England: Cambridge University Press.
- Condon, W.S., & Ogston, W.D. (1967). A segmentation of behavior. *Journal of Psychiatric Research, 5*, 221–235.
- Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Frontiers in Psychology, 2*, 355.
- 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). Sequential analysis for behavioral streams. *International Journal Bifurcation and Chaos, 21*, 1153–1161.
- Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, Vol. 33 (pp. 1–40). San Diego, CA: Academic Press.
- Ekman, P., Friesen, W. V., & Hager, J. C. (Eds.). (2002). *Facial Action Coding System* [E-book]. Salt Lake city, UT: Research Nexus.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164–203.

- Flecha-García, M. L. (2010). Eyebrow raises in dialogue and their relation to discourse structure, utterance function and pitch accents in English. *Speech Communication*, 52, 542–554.
- Fowler, C. A., Richardson, M. J., Marsh, K. L., & Shockley, K. (2008). Language use, coordination, and the emergence of cooperative action. In A. Fuchs & V. Jirsa (Eds.), *Coordination: Neural, behavioral and social dynamics* (pp. 281–307). Berlin: Springer.
- Gambetta, D. G. (Ed.). (1988). *Trust: Making and breaking cooperative relations*. New York: Blackwell.
- Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic coordination. *Cognition*, 27, 181–218.
- Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. *Cognition*, 53, 181–215.
- Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, 1, 292–304.
- Giles, H., Coupland, N., & Coupland, J. (1991). Communication, context, and consequence. In H. Giles, N. Coupland, & J. Coupland (Eds.), *Contexts of accommodation*. (pp. 49–68). New York: Cambridge University Press.
- Giles, H., & Powesland, P. F. (1975). *Speech styles and social evaluation*. New York: Academic Press.
- Grammer, K., Kruck, K. B., & Magnusson, M. S. (1998). The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. *Journal of Nonverbal Behavior*, 22(1), 3–29.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1994). *Emotional contagion*. Cambridge, England: Cambridge University Press.
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27, 949–960.
- Hurley, S., & Chater, N. (Eds.). (2005). *Perspectives on imitation: From neuroscience to social science*. Cambridge, MA: MIT Press.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Jaeger, T. F. (2006). Redundancy and syntactic reduction in spontaneous speech. Ph.D. thesis, Stanford University, Palo Alto.
- Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61, 23–62.
- Kugler, P.N., Kelso, J. A. S., & Turvey, M. T. (1980). On the concept of coordinative structures as dissipative structures: I. Theoretical lines of convergence. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor behavior* (pp. 1–49). New York: North-Holland.
- Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In B. Schlökopf, J. Platt, & T. Hoffman (Eds.), *Advances in neural information processing systems (NIPS) 19* (pp. 849–856). Cambridge, MA: MIT Press.
- Lieberman, P. (1963). Some effects of the semantic and grammatical context on the production and perception of speech. *Language and Speech*, 6, 172–175.
- Lindblom, B. (1990). Explaining variation: A sketch of the H and H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 403–439). Dordrecht, Netherlands: Kluwer.
- Littell, R. C., Stroup, W. W., & Freund, R. J. (2002). *SAS for linear models*. Cary, NC: SAS Publishing.
- Louwerse, M. M., & Bangerter, A. (2010). Effects of ambiguous gestures and language on the time course of reference resolution. *Cognitive Science*, 34, 1517–1529.
- Louwerse, M. M., & Crossley, S. A. (2006). Dialog act classification using N-Gram algorithms. In G. Sutcliffe & R. Goebel (Eds.), *Proceedings of the International Florida Artificial Intelligence Research Society* (pp. 758–763). Menlo Park, CA: AAAI Press.
- Louwerse, M.M., & Mitchell, H.H. (2003). Towards a taxonomy of a set of discourse markers in dialog: A theoretical and computational linguistic account. *Discourse Processes*, 35, 199–239.
- Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, 438, 237–329.
- McDowall, J. J. (1978). Interactional synchrony: A reappraisal. *Journal of Personality and Social Psychology*, 36, 963–975.

- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75–78.
- Neumann, R., & Strack, F. (2000). “Mood contagion”: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology*, *79*, 211–223.
- Partridge, B. L. (1981). Internal dynamics and the interrelations of fish in schools. *Journal of Comparative Physiology*, *144*, 313–325.
- Pickering, M. J., & Garrod, S. (2004). Towards a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–226.
- Platek, S. M., Critton, S. R., Myers, T. E., Gallup, G. G., & others. (2003). Contagious yawning: The role of self-awareness and mental state attribution. *Cognitive Brain Research*, *17*, 223–227.
- Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers’ and listeners’ eye movements and its relationship to discourse comprehension. *Cognitive Science*, *29*, 39–54.
- Richardson, D. C., Dale, R., & Kirkham, N. (2007a). The art of conversation is coordination: Common ground and the coupling of eye movements during dialogue. *Psychological Science*, *18*, 407–413.
- Richardson, M. J., Marsh, K. L., Isenhower, R., Goodman, J., & Schmidt, R. C. (2007b). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, *26*, 867–891.
- Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception & Performance*, *31*, 62–79.
- Rizzolatti, G., & Craighero, L. (2004). The mirror neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
- Rogers, Y., & Ellis, J. (1994). Distributed cognition: An alternative framework for analysing and explaining collaborative working. *Journal of Information Technology*, *9*, 119–128.
- Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*, 289–327.
- Schiffrin, D. (1987). *Discourse markers*. Cambridge, England: Cambridge University Press.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*, 70–76.
- Semin, G. R., & Cacioppo, J. T. (2008). Grounding social cognition: Synchronization, entrainment, and coordination. in G.R. Semin & E.R. Smith (Eds.), *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches* (pp. 119–147). New York: Cambridge University Press.
- Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*, 305–319.
- Shockley, K., Baker, A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 201–208.
- Shockley, K., Santana, M., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 326–332.
- Van Baaren, R., Janssen, L., Chartrand, T., & Dijksterhuis, A. (2009). Where is the love? Social aspects of mimicry. *Philosophical Transactions of the Royal Society B, Biological sciences*, *364*, 2381–2389.
- Van Baaren, R. B., Holland, R. W., Steenaerts, B., & van Knippenberg, A. (2003). Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology*, *39*, 393–398.
- Wang, Y., Newport, R., & Hamilton, A. (2011). Eye contact enhances mimicry of intransitive hand movements. *Biology Letters*, *7*, 7–1.
- Ward, A. J. W., Axford, S., & Krause, J. (2002). Mixed-species shoaling in fish: the sensory mechanisms and costs of shoal choice. *Behavioral Ecology & Sociobiology*, *52*, 182–187.
- Webb, J. T. (1969). Subject speech rates as a function of interviewer behavior. *Language and Speech*, *12*, 54–67.