Title
Convergence of speech rate in conversation predicts cooperation

Journal
Evolution and Human Behavior, 34(6)

ISSN
1090-5138

Authors
Manson, Joseph H
Bryant, Gregory A
Gervais, Matthew M
et al.

Publication Date
2013-11-01

DOI
10.1016/j.evolhumbehav.2013.08.001

Peer reviewed

Abstract

During conversation, interlocutors coordinate their behavior on many levels. Two distinct forms of behavioral coordination have been empirically linked with affiliation and cooperation during or following face-to-face interaction: *behavior matching* and *interpersonal synchrony*. Only the latter form constitutes behavioral entrainment involving a coupling between independent oscillators. We present the first study of the association between spontaneously occurring behavioral coordination and post-interaction economic game-play. Triads of same-sexed strangers conversed for 10 min, after which each participant played an unannounced one-shot prisoner's dilemma (PD) toward each co-participant. When dyads had higher language style matching scores (LSM: Gonzales et al., 2010), the individuals evaluated each other more positively, but they were no more likely to cooperate in the PD. However, when dyads' speech rates (mean syllable duration) converged more strongly from the beginning to the end of the conversation, they were more likely to cooperate in the PD, despite no effect on interpersonal evaluations. Speech rate convergence, a form of rhythmic entrainment, could benefit interlocutors by mutually reducing cognitive processing during interaction. We suggest that spontaneous, temporally-based behavioral coordination might facilitate prosocial behavior when the joint cooperative effort is itself perceived as a form of coordination.

1. Introduction

Conversational interaction is fundamental to human communication, and involves the dynamic interplay of many complex phenomena. While engaged in conversation, interlocutors communicate with their bodies, voices, and language. Research across many disciplines has documented a variety of ways that conversationalists coordinate their actions in the service of mutually beneficial interaction. How people talk together in real time is closely tied to broader interactive goals, which themselves are products of adaptations for navigating the social world.

Two distinct forms of behavioral coordination have been empirically linked with affiliation and cooperation during or following face-to-face interaction (Bernieri & Rosenthal, 1991; Hove & Risen, 2009). The first, *behavior matching*, involves individual *B* copying a behavior of individual *A*, but with neither a particular temporal relation to *A*'s action, nor any implication that *A* responds in any specific fashion to *B*'s copying action. A substantial body of research has established that people unconsciously mimic their interaction partners' postures, gestures, and mannerisms (Lakin et al., 2003), and language use patterns (Niederhoffer & Pennebaker, 2002), and that such mimicry is related to subsequent affiliative behavior. Among a large number of similar findings, people spontaneously mimic an experimental confederate's gestures and report greater liking for a confederate who mimics them (Chartrand & Bargh, 1999), and leave larger tips for a waitress who mimics them (van Baaren et al., 2003). Researchers using the automated Linguistic Inquiry and Word Count algorithm (Pennebaker et al., 2001, 2007) have found that similarity in relative usage frequency of common function word categories (e.g. prepositions, conjunctions) predicts successful hostage negotiations

(Taylor & Thomas, 2008), task group cohesiveness (Gonzales et al., 2010), and the formation and persistence of romantic relationships (Ireland et al., 2011). Coordinated language use and behavior may facilitate mutual understanding (Pickering & Garrod, 2004). Ireland and Pennebaker (2010; see also Meyer & Bock, 1999) argued that function words such as pronouns and articles (unlike content words) are "inherently social," because their comprehension typically depends, not just on the conventions of a speech community, but on shared frames of reference actively established among interlocutors. For example, every English-speaker knows the meaning of *garden*, but the particular garden referred to by *the garden* will be apparent to a listener only when she shares the same immediate frame of reference as the speaker. For this reason, according to Ireland and Pennebaker (2010), pairwise similarity in frequency of function word use is associated with greater affiliation or cooperation.

A second form of behavioral coordination is *interpersonal synchrony*, which typically involves entrainment—a temporal coupling between independent oscillators that enter into some type of phase relationship. Prime examples of this are turn-taking in conversation (Wilson & Wilson, 2005) and playing music with an isochronous beat (Bispham, 2006). In Wilson and Wilson's (2005) model of conversational turn-taking, speech rate entrainment occurs via speakers' syllabic production, which operates interpersonally as a medium for entraining neural oscillators among interlocutors. This facilitates conversational coordination and allows for inter-turn transitions marked by minimal gap and minimal overlap (Stivers et al., 2009). Perceptions of timing in music and speech can affect subsequent productions in these respective domains (Jungers et al.,

2002), and speech rate convergence has been linked to interpersonal judgments (e.g. ratings of competence: Street, 1984).

Talk is just one form of social interaction in which people are sensitive to entrainment. Studies have shown that singing together can increase cooperation in a prisoner's dilemma game (Anshel & Kipper, 1988) and a public goods game (Wiltermuth & Heath, 2009), though the effect can be sensitive to experimental conditions (e.g. Kurzban, 2001). Children who sang and danced together were more likely to assist one another in a later playground incident (Kirschner & Tomasello, 2010). Synchronous tapping, but not asynchronous tapping, generated higher affiliation ratings, but only when the synchrony was with another person, and not just experienced (i.e., tapping to a metronome) (Hove & Risen, 2009). Synchronized training in competitive rowers resulted in increased endorphin release (Cohen et al., 2010), suggesting a proximate mechanism motivating this kind of behavioral coordination. Behavioral entrainment is highly detectable, and can impact people's perceptions of the affiliation between the synchronizers. Hagen and Bryant (2003) showed that better temporal coordination in a music performance positively affected third party judgments of coalition quality between the musicians. While social entrainment may have evolved in many species from the simpler adaptive ability to entrain one's behavior to rhythmic information in the physical environment (Phillips-Silver et al., 2010), human interpersonal synchrony is moderated by many social factors and interacts in complex ways with group membership and the dynamics of alliance formation (Miles, Griffiths et al., 2009; Miles, Lumsden, et al., 2011).

Laughter is another interactive phenomenon that can involve behavioral coordination and may be associated with cooperative behavior. Research has shown that people who have known each other longer tend to laugh together more (Bryant, 2012; Smoski & Bachorowski, 2003a) and familiarity between conversationalists is perceptible in the co-laughter itself (Bryant, 2012). Lynch (2011) found that people with greater similarity in implicit preferences laugh together more, suggesting an association with social cohesion. Gervais and Wilson (2005) argued that laughter functions as a medium for mirthful emotional contagion that recruits partners into resource-building social play. Accordingly, comparative work has demonstrated that chimpanzees use laugh-like vocalizations to manage playful social interactions, and that antiphonal laugh sequences lengthen play time (Davila-Ross et al., 2011). Other scholars have suggested a variety of communicative functions for coordinated laughter that relate to cooperation (Owren & Bachorowski, 2003; Mehu & Dunbar, 2008), bonding (Dezecache & Dunbar, 2012; Platow et al., 2005) and social assortment (Flamson, et al., 2011).

The adaptive significance of these various phenomena remains a matter of debate. Simple mimicry in nonhuman social animals has obvious adaptive advantages (e.g. treating conspecifics' fear responses as reliable cues of imminent danger), and is presumably the phylogenetic source of more elaborate forms of behavioral coordination (Lakin et al., 2003). However, why these should serve as "social glue" is unclear. A number of non-human animal species exhibit inter-individual temporal coordination (Hall & Magrath, 2007), but the functions of these displays often remain unknown. Phillips-

5

Silver et al. (2010) argue that even in cognitively simple species, collective social entrainment can amplify social signals in adaptive ways (e.g. courtship choruses; Greenfield, 1994). In human collective action, social entrainment may be necessary to accomplish work activities that require behavioral coordination. Recent work has shown that engaging in synchronized action facilitates success in later joint activity. For example, people who rocked synchronously in chairs, compared to controls that rocked asynchronously, were better able to subsequently coordinate their action on a collaborative task (Valdesolo et al., 2010). This suggests that synchronizing action may calibrate expectations about others' behavior, and help coordinate action in other domains.

In this study, we examined whether distinct kinds of vocal and verbal convergence in naturalistic social interactions predicted cooperation in a one-shot prisoner's dilemma (PD). In a PD, an actor chooses whether to cooperate or defect toward a recipient. The actor gains the largest payoff when he defects while the recipient cooperates; the second largest when both cooperate; the third largest when both defect; and the lowest when the actor cooperates while the recipient defects. From a strictly monetary perspective, defection is always the best decision in a one-shot PD. However, a sizeable proportion of educated American, European, and Japanese participants treat one-shot PDs as assurance games, gaining the most psychological utility from mutual cooperation (Fehr & Camerer, 2007; Hayashi et al., 1999; Kiyonari et al., 2000), and therefore cooperating if, and only if, they expect their partner to cooperate. This suggests that social preferences transform the PD into a coordination game (specifically, a Stag Hunt – Van Huyck et al., 1990) in

6

which one coordinated outcome (mutual cooperation) yields higher payoffs to both players than the other coordinated outcome (mutual defection).

To assess whether different types of naturally occurring behavioral coordination facilitate cooperation-as-coordination, we measured behavior among strangers in open-ended conversation prior to their playing an unannounced one-shot simultaneous PD. We examined dyadic convergence in three vocal characteristics: (1) fundamental frequency (F_0); (2) variation in F_0, and (3) speech rate (mean syllable duration). We also calculated several measures of coordinated laughter and laughter/speech coordination. Finally, we calculated each dyad's language style matching score (LSM: Gonzales et al., 2010). We also examined the relationships between convergence and coordination in these diverse channels. Based on the empirical literature reviewed above, we expected that greater behavioral convergence would raise expectations of cooperative coordination, and that therefore *dyads showing greater (1) vocal convergence, (2) coordinated laughter and (3) verbal convergence (higher LSM score) would be more likely to cooperate in the PD*. We also elicited ratings of co-participants' *warmth* and *competence*, and predicted that these person perception variables would mediate the relationship between the convergence/coordination variables and PD decisions. This is the first study to examine whether spontaneous (as distinct from experimentally induced) behavioral coordination is associated with post-interaction behavior in an incentivized social dilemma.

7

The analyses presented here build on our previous report of findings regarding the determinants of our conversation participants' PD decisions (Gervais et al., 2013). In a multivariate model, we found two main effects: people were more likely to cooperate (1) if they grew up in a wealthier zip code and (2) towards more facially attractive co-participants. We also found two interaction effects with subclinical primary psychopathy (callous affect, interpersonal manipulation) as measured by a self-report instrument: people who scored higher on primary psychopathy were less likely to cooperate toward co-participants (1) who interrupted them more frequently during the pregame conversation, and (2) with whom they discovered no "common ground" (i.e. reliable cues to future interaction). This model explained 15.6% of the variance in probability of cooperating. Our goal in the present research is to assess which, if any, measures of verbal or vocal convergence improve the predictive power of this model.

2. Methods

More detailed descriptions of (1) the participant pools, (2) the conversation and post-conversation game-play and questionnaire procedures, (3) the attractiveness rating procedure, and (4) the conversation transcription procedures can be found in Gervais et al. (2013).

2.1 Participants

Conversation participants (n = 105) were undergraduates at UCLA. All participants were given a $10 show-up fee; 90% of them were also fulfilling a course requirement. Participants were all native speakers of English, their median age was 19

187 years, and their ethnic composition corresponded closely to that of the multi-ethnic
188 campus population. Conversation groups were same-sex triads (20 female, 15 male).

189 *2.2. Procedures*

190 Conversation participants were grouped equidistantly around a small circular
191 table. After determining that the conversation participants were strangers to each other, an
192 experimenter recited a prepared script asking the participants to converse for 10 minutes
193 on any topic(s) they wished. Participants were informed that their conversation would be
194 videotaped, but were given no details about the post-conversation procedure.
195 Conversations were recorded using a Canon Vixia HV30 camcorder (Audio: MP2
196 compression, 384 kbps) connected to an Audio-Technica U841a omnidirectional
197 condenser boundary microphone (30 Hz – 20 kHz frequency response).

198 Following the conversation, participants sat at visually isolated laptops running z-
199 Tree version 2.1 (Fischbacher, 2007). First, participants played a one-shot PD toward
200 each of their two co-participants. Choices were labeled “Keep” $3 provided by the
201 experimenter (=defect) or “Transfer” the $3 to the co-participant, whereupon it would be
202 doubled to $6 (=cooperate). To ensure the confidentiality of participants’ PD choices,
203 they were instructed, truthfully, that one of the three of them would not receive their
204 earned payoff, but instead a randomly generated but realistic set payoff. Participants then
205 rated each of their co-participants on “warmth” and “competence” using separate sliders,
206 completed a well-validated self-report psychopathy instrument (the LSRP: Levenson et
207 al., 1995), and answered a set of basic demographic questions (age, ethnicity, childhood
208 zip code).

2.3 Data analysis

Perceived Warmth and Competence. Participants' ratings of their co-participants' *warmth* and *competence* were moderately to strongly congruent (Cronbach's $\alpha = 0.69$). Therefore, we averaged the standardized *warmth* and *competence* ratings of each participant towards each co-participant. We refer to this measure as positive person perception (PPP). For all significant results incorporating PPP, separate analyses using *warmth* and *competence* produced qualitatively very similar results.

Language Style Matching (LSM). Because of the large time and training investment required for transcriptions and data analyses reported elsewhere (Gervais et al., 2013), a portion of each 10-minute conversation was pre-selected for transcription and further analysis. This portion included the first 60 seconds of the conversation and two other sections of ≥60 seconds duration from minutes 2-5 and 6-10 of the conversation. Start times of the second and third transcription periods were moved back toward the beginning of the video, if necessary, so that all transcription periods began with the start of a new conversational topic. Total time transcribed per group ranged from 3.02-5.57 minutes ($M \pm SD = 4.08 \pm 0.68$ min). For the LSM analyses, we constrained the transcriptions to yield only English words spelled as indicated in the Linguistic Inquiry and Word Count (LIWC) 2007 program dictionary (Pennebaker et al., 2007). The LIWC algorithm calculates, for a sample of speech or text, the proportion of words in a text that fall into each of 67 categories, not all of which are mutually exclusive. For theoretical reasons reviewed above, Pennebaker and colleagues (Gonzales et al., 2010; Ireland &

10

Pennebaker, 2010; Ireland et al., 2011; Pennebaker et al., 2003) have placed particular emphasis on interpersonal similarity in the usage frequency of nine types of function words: auxiliary verbs (e.g. *am*, *will*, *have*), articles, common adverbs (e.g. *hardly*, *often*), personal pronouns, indefinite pronouns, prepositions, negations, conjunctions and quantifiers.

We ran the LIWC algorithm on each participant's speech output during the transcribed portions of the conversation. To determine overall language style convergence within dyads, we first calculated the correlations between co-participants' usage (i.e. the percentage of each individual's total words uttered) for each of the nine function word categories.

Following Gonzales et al. (2010), we next calculated each co-participant dyad's LSM score based on inter-individual similarities in the proportions of the nine types of function words. The LSM score of a dyad, Person 1 and Person 2, with regard to a particular function word type, e.g. quantifiers, is calculated as:

$$quanLSM = 1 - (|quan1 - quan2|/(quan1 + quan2))$$

where quan1 is the percentage of Person 1's words that are quantifiers, and quan2 is the percentage of Person 2's words that are quantifiers. An LSM score can range from 0.00-1.00. Each dyad's total LSM is calculated as the mean of its LSM scores across the nine categories of function words.

To determine whether dyads with higher LSM scores were more likely to cooperate in the PD, we ran a logistic regression with PD decision (cooperate or defect) as the dependent variable. To account for the non-independence of each individual's two

PD decisions, we calculated robust standard errors of the regression coefficients, clustering by individual, before calculating confidence intervals and *p*-values.

Vocal Characteristics. For the vocal analyses, we also divided the 10-minute conversations into three sections (not the same sections used for LSM analyses). Section 1 was 0:00-3:20, Section 2 was 3:21-6:40, and Section 3 was 6:41-10:00. Audio files were exported from the video recordings using Apple iMovie software and saved as 44.1 kHz, 16-bit wav files. Using Adobe Audition 3, we then extracted the first continuous five-second portion of continuous speech for each speaker and for each section (i.e., three clips per participant) that did not contain overlapping speech, other than cases of one-word backchannels (e.g., uhhuh) or other vocal noises. Of the 315 clips (105 participants × 3 time periods), 10 contained no speech excerpts that met these criteria; these were deleted case-wise in subsequent statistical analyses.

The extracted clips were analyzed acoustically using Praat, version 5.3.01 (Boersma & Weenink, 2011). We measured mean fundamental frequency (F_0) (the acoustic correlate of perceived pitch), fundamental frequency standard deviation (F_0 SD) (acoustic correlate of perceived pitch variability) and mean syllable duration (MSD) (speech rate) for each clip. F_0 was measured using the autocorrelation method in Praat with default pitch settings suggested by Praat for men (100-500 Hz) and women (120-600 Hz). Octave jump errors and other analytical errors, such as F_0 estimates during voiceless segments, were fixed through pitch setting adjustment (never exceeding +/- 20 Hz adjustment in the lower limit, and +/- 60 Hz in the higher limit), or removed manually. In cases where small overlapping vocalizations occurred in the extracted clips, the overlapped portions were removed prior to analysis. On average, >90% of the original

12

clips were analyzed, with most requiring no editing. MSD was calculated by dividing the total time of speech energy determined through visual analysis of the spectrogram in the clip by the number of spoken syllables (i.e., not written word syllables). (Spectrogram: FFT method, Gaussian window shape, dynamic range - 50 dB).

To determine whether co-participants generally converged with respect to F_0, F_0 SD, and MSD, we treated each dyad as a data point. For each of these variables in each conversation section, we regressed the value of one member of the dyad on the value of the other member of the dyad. Significantly positive slopes indicate greater than chance similarity within dyads. To determine whether co-participants became more similar in these variables over time, we used Wald tests to compare slopes across conversation sections.

To test whether convergence in vocal characteristics affected PD play, we first calculated, for every co-participant dyad in each of the three conversation sections, the absolute value of the difference between their values for each variable. For each dyad, we then estimated the slope (β) of the linear regression line formed by the three points (section 1, section 2 and section 3). Negative slopes indicate decreasing differences (i.e. increasing similarity) over time between the two co-participants with regard to that variable. We then ran logistic regressions in which each dyad was a data point, PD decision (cooperate or defect) was the dependent variable and the relevant β value (standardized) was the independent variable. Significantly negative relationships indicate that vocal convergence increases the probability of cooperation.

13

Laughter Analyses. We coded laughter throughout the 10-min duration of every conversation. Two coders counted laugh instances in all conversations using video playback. A bout of laughter was defined as a series of nonverbal, vocalized calls often with successive expiratory elements, though sometimes containing only a single call. We included voiced (i.e., tonal) and unvoiced bouts. Laughter is typically produced with an initial burst amplitude and frequency that decays over time (Titze et al., 2008). Laugh calls judged as a single bout had to originate from the same initial burst. Bouts that were back-to-back without a noticeable pause, as evidenced by perceptible re-initialized energy, were counted as one laugh. Laugh counts across speaker conditions were highly correlated between the two coders (Cronbach's alpha = 0.96) so data from one coder were used in our analyses. We calculated the values of an individual-level variable (raw number of laughs produced) and four dyad-level variables: (1) raw number of simultaneous laughs (co-laughter count), (2) simultaneous laughs divided by the sum of the two dyad members' total laughs (co-laughter proportion); (3) laughs by the first dyad member while the second dyad member was speaking (laughs during other's speech) and (4) summed laughs while the other dyad member was speaking, divided by the sum of the two dyad members' total laughs (bi-directional laughter during other's speech).

We used Akaike's Information Criterion (AIC: Akaike, 1974) to assess the effects of adding independent variables to models that successfully predicted our dependent variables. AIC takes into account the tradeoff between a predictive model's accuracy (which should be maximized) and its complexity, or number of independent variables (which should be minimized). In a comparison between two models, the one with the

14

lower AIC value is better, as it more closely approximates the causal processes that generated the data.

3. Results

3.1 Language style matching

Co-participants generally matched their language styles with respect to function words. Table 1 shows correlation coefficients (Pearson *r*) of co-participants' usage percentages of the nine function word categories. For only two of these categories (conjunctions and quantifiers), co-participant pairs failed to attain highly significant similarity. Because we found, consistent with other research (e.g. Newman et al., 2008), some sex differences in function word use (e.g. compared to men, women used more auxiliary verbs [13.6% vs. 11.9%, Cohen's $d = 0.58$, $p = 0.004$]), we also ran the correlation analyses separately for the two sexes. Among women, co-participant pairs failed to attain significant similarity only for prepositions, conjunctions and quantifiers; among men, co-participant pairs failed to attain significant similarity only for articles, conjunctions and quantifiers. All other within-sex co-participant correlations were significant at $p < 0.01$. Among the 105 dyads, the mean LSM score was 0.82 (SD = 0.08).

Co-participant dyads that were sampled for longer periods of time had higher LSM scores ($\beta = 0.025 \pm 0.011$, $n = 105$, $p = 0.032$), as would be expected if longer sampling periods reduce error variance, i.e. the impact on LSM of random intra-

15

individual variation in function word use. We therefore calculated the residuals of this regression (i.e. LSM score relative to amount of time sampled) and used these values as an independent variable to confirm results obtained using raw LSM scores as the independent variable.

LSM scores were not significantly associated with Prisoner's Dilemma decisions. Bivariate logistic regressions revealed non-significant trends towards more likely defection given higher LSM scores, which is opposite to that predicted (using raw LSM scores: logistic regression with standard errors based on clustering by decision-maker's identity, odds ratio ± SE = 0.062 ± 0.132, $n = 206$, $p > 0.1$; using residuals on time sampled: o.r. ± SE = 0.159 ± 0.333, $n = 206$, $p > 0.3$). When we added either raw LSM scores or residual scores to the multivariate predictive model described in Gervais et al. (2013), (1) neither variable had an independent significant relationship with PD decisions, (2) the resulting models did not increase the proportion of variance explained, and (3) they increased the AIC (Akaike, 1974). Thus, even after taking into account all known effects of independent variables on PD decisions by our participants, LSM scores had no explanatory value with respect to predicting PD decisions.

3.2. Vocal characteristics

Table 2 shows the results (β coefficients with standard errors) of regressing, for each of the three acoustic variables (F_0, F_0 SD, and mean syllable duration) in each

16

conversation section, each participant's value on one of his or her co-participants' value. That is, each data point is a dyad of co-participants. For male F_0, these coefficients were significantly negative, i.e. if one male of a dyad had a high F_0, his co-participants tended to have low F_0 values, at both the beginnings and the ends of conversations. For female F_0, and for F_0 SD in both sexes, there was no relationship between co-participants' values. There were no significant changes between β values from section 1 to section 3.

However, mean syllable duration did show inter-individual convergence over the course of the conversations. In section 1, the regression coefficient was non-significantly negative, whereas in section 3, it was significantly positive and, therefore, significantly different from the β of section 1. Furthermore, during section 2, the β value was intermediate between sections 1 and 3 (0.043 ± 0.092) and not significantly different from either.

For F_0 and F_0 SD, we found no relationship between inter-individual convergence (i.e. the slope over time of the absolute value of the difference between co-participants' values) and probability of cooperating in the Prisoner's Dilemma. Indeed, for F_0, there was a marginally significant trend for higher slopes (i.e. greater inter-individual differentiation over time) to be associated with cooperation (logistic regression with standard errors based on clustering by decision-maker's identity, o.r. ± SE = 1.46 ± 0.30, $n = 178$, $p = 0.061$). For F_0 SD, there was no relationship between convergence and PD decision (o.r. ± SE = 1.33 ± 0.30, $n = 172$, $p > 0.20$). However, in dyads that converged more strongly in mean syllable duration, participants were more likely to cooperate in the

PD (o.r. ± SE = 0.57 ± 1.40, n = 176, p = 0.02). Of the three vocal variables, only MSD convergence improved the predictive power of the multivariate model described in Gervais et al. (2013). When added to this model, MSD convergence had a significant (p = 0.04) independent effect on the probability of cooperating, and adding MSD convergence to the model increased r^2 from 0.156 to 0.196, and decreased AIC from 228.7 to 190.6, indicating a closer approximation of the causal processes that generated the data.

3.3. Laughter

Across all 35 conversations, over 1000 laughs were counted in total (M ± SD = 29.1 ± 13.0), and a substantial percentage of these were produced in coordination (i.e., co-laughter) with at least one other speaker (M ± SD = 41.4% ± 16.8%). Women produced more laughs than men (n_{female} = 60, n_{male} = 45, M_{female} ± SD = 17.2 ± 7.7, M_{male} = 12.3 ± 8.8, d = 0.59, p = 0.003). Among dyads (n = 206 PD decisions in all analyses), neither co-laughter count (o.r. ± SE = 1.03 ± 0.05, p > 0.50), co-laughter proportion (o.r. ± SE = 1.41 ± 1.71, p > 0.50), laughs during other's speech (o.r. ± SE = 1.00 ± 0.06, p > 0.50), nor bi-directional laughter during other's speech (o.r. ± SE = 1.56 ± 1.86, p > 0.50) was associated with the probability of cooperating in the prisoner's dilemma. When added to the multivariate model predicting PD play described by Gervais et al. (2013), none of these independent variables had a significant independent effect on PD play, and all of them increased AIC. Although we made no predictions about sex differences in the relationships between behavioral convergence and PD play, a post-hoc analysis showed that only among men (n = 90 PD decisions), dyads with higher co-laughter counts (o.r. ± SE = 1.14 ± 0.07, p = 0.03) and co-laughter proportions (o.r. ± SE = 26.77 ± 42.53, p =

18

411 0.04) were significantly more likely to cooperate. Running the Gervais et al. (2013)
412 multivariate model separately for men and women revealed that in men only, AIC was
413 reduced by adding co-laugher count or co-laughter proportion as an independent variable.
414 Neither variable had a significant independent effect on PD play in men, but the effect of
415 co-laughter proportion approached significance at $p = 0.08$.
416
417 *3.4. Associations among independent variables*
418
419 LSM score was not associated with any of the vocal or laughter variables, nor
420 were F_0 or F_0 SD associated with any of the laughter variables. However, dyads that
421 converged more with respect to MSD (i.e. had more strongly negative slopes) had higher
422 co-laughter counts ($r = -0.22$, $p = 0.04$).
423
424 *3.5 Positive Person Perception*
425
426 Higher PPP ratings were marginally associated with an increased probability of
427 cooperating in the PD (odds ratio ± SE = 1.43 ± 0.29, $n = 204$, $p = 0.07$).
428 LSM scores were significantly associated in the predicted (positive) direction with
429 participants' positive person perception (PPP) ratings. This held whether the independent
430 variable was raw LSM score (linear regression with robust standard errors clustered by
431 participant: $\beta = 1.93 \pm 0.76$, $N = 208$, $p = 0.013$) or residual of LSM on time sampled ($\beta =$
432 2.33 ± 0.84, $N = 208$, $p = 0.007$). However, LSM scores did not predict behavior in the
433 PD (see above).

None of the vocal characteristics, including MSD, significantly predicted positive person perception scores (F_0 : β = -0.55 ± 0.60, N = 180, p > 0.30; F_0 SD: β = 0.22 ± 0.53, N = 174, p > 0.50; MSD: β = 0.44 ± 4.53, N = 178, p > 0.50).

Only one laughter variable, laughs during other's speech, was associated with positive person perception. PPP ratings were higher in dyads in which the two participants laughed more while the other person was speaking (β = 0.97 ± 0.43, N = 208, p = 0.027). However, laughs during other's speech did not predict behavior in the PDG (see above). Neither co-laughter count (β = 0.019 ± 0.014, N = 208, p = 0.20) nor co-laughter proportion (β = 0.44 ± 0.42, N = 208, p > 0.20) was associated with PPP ratings.

4. Discussion

We examined the relationships between vocal and verbal convergence in a spontaneous conversation and the participants' subsequent decisions in a Prisoner's Dilemma game. Existing empirical work in communication led us to predict that behavior matching in language use and vocal convergence in prosodic features of speech would be associated with cooperation in a PD game. One form of vocal convergence (speech rate) was positively related to the probability of PD cooperation. However, although we found strong evidence for language style matching in zero-acquaintance small groups (consistent with others' findings; Niederhoffer & Pennebaker, 2002; Newman et al., 2008), LSM was unrelated to post-conversation PD decisions. The LSM results are surprising in view of other work (Gonzales et al., 2010; Ireland & Pennebaker, 2010; Ireland et al., 2011) suggesting that cooperation in several contexts (e.g. group task

performance, romantic relationships, even long-term scholarly collaborations) can be predicted using the same language style matching (LSM) metric (Gonzales et al., 2010) that we applied to our data.

One important difference between earlier LSM research and the current study was that we did not cue the importance of substantive cooperation among our participants until after the conversation. We told them only that we were studying "small talk among strangers," and that they would be answering some questions at the end of the conversation. In contrast, the experimental situation of a task group (Gonzales et al., 2010), and the real-life situations of a speed-date (Ireland et al., 2011) or a hostage negotiation (Taylor & Thomas, 2008), presumably foreground the detection of cooperative potential in one's interlocutor(s) before the face-to-face interaction. One possibility is that, with respect specifically to function word use, close style matching (i.e. LSM scores >0.75) is the typical outcome of collaborative conversation, and pairwise style *divergence* follows from declines in affiliation or trust in potentially agonistic situations. In contrast, our study's experimental conversation context was friendly and collaborative, with very little (apparently) at stake. Therefore, style matching occurred (Table 1), but LSM was unrelated to subsequent, and unanticipated, PD decisions. This is consistent with the relationship of LSM to PPP even in the absence of an effect of LSM on cooperation.

Co-participants did tend to cooperate more as a function of how much their speech rates converged. Behavioral convergence that involves entrainment (i.e.,

21

temporally based) might be distinct from other forms of convergence (e.g., behavior matching) because it provides immediate mutual benefits. For example, becoming entrained can introduce synergy that could potentially reduce mutual metabolic costs of interacting (Marsh et al., 2006). This would make entrainment a form of coordination, in which profitable cheating is impossible, but players' uncertainty about each other's choices may still yield suboptimal outcomes (Van Huyck et al., 1990). Successful behavioral entrainment may reduce uncertainty in future coordination by indexing how well co-participants can coordinate their action. For example, speech rate entrainment may be a reliable indicator that an interlocutor can coordinate his or her actions with one's own actions in a rapid and fine-tuned manner, mutually reducing the cognitive processing costs of interaction. Given that our participants may typically view a one-shot PD as a coordination game (Fehr & Camerer, 2007; Hayashi et al., 1999; Kiyonari et al., 2000), coordination in speech rate may increase perceived ability to coordinate on cooperation in a PD, raising rates of cooperation. This is consistent with the effect of speech rate convergence on cooperation even in the absence of an effect on positive person perception – the perceptions of coordination that lead to cooperation do not necessarily require positive interpersonal evaluations. DeSteno et al. (2012) likewise found that disengagement gestures performed by a humanoid robot reduced participants' donations and expectations of donations in a social dilemma, yet did not affect participants' liking of the robot. Future research should use multi-dimensional person and relationship perception measures to tap the relevant attributions and evaluations that underlay perceived coordination capacity.

22

We found that language matching had no impact on cooperative decisions in the PD game, whereas speech rate entrainment did increase the probability of cooperation. It may be that behavior matching is more subject to vigilance against cheating than is synchrony because matching is more used in manipulation (Dawkins & Krebs, 1981). Coordination is mutually beneficial and offers no incentive for defection, while mimicry and other unilateral forms of behavior matching are potentially intentional and manipulative (Bailenson et al., 2008; Bourhis et al., 1975). Pardo et al. (2010) found that when conversationalists were instructed to imitate one another covertly, they often converged phonetically (a form of behavior matching), but simultaneously diverged in articulation rates (a form of entrainment). This suggests greater success at manipulative matching than entrainment. Bailenson et al. (2008) found that mimicry had negative impacts on trustworthiness and warmth judgments when it was explicitly noticed—suggesting a sensitivity to manipulation—whereas even instructed, consciously mediated synchrony (e.g., intentionally walking in time, clapping together, or swinging a cup while singing) can enhance cooperation despite explicit awareness of the behavioral convergence (Valdesolo & DeSteno, 2011; Valdesolo et al., 2010; Wiltermuth & Heath, 2009). Our results fit this pattern, even though language style matching may be less likely than gestural or postural mimicry to be consciously detected, and conscious attempts to match others' language styles are generally unsuccessful (Ireland & Pennebaker, 2010).

We did not find a relationship between coordinated dyadic laughter and game play across all participants, either in the absolute amount of laughing in response to another person, or in the proportion of all laughter in a triad shared by a dyad within it. We did,

23

however, discover an unexpected sex difference. The more a male dyad laughed together, the more likely they were to cooperate in the PD game. Women laughed significantly more than men, a finding consistent with other studies of laughter in small groups of strangers (Bryant 2012; Smoski & Bachorowski, 2003a), but women's laughter was not related to game play. This suggests the intriguing possibility that male co-laughter in zero-acquaintance contexts has relatively higher cue validity for cooperative intentions and/or the ability to coordinate in the future. Kurzban (2001) found that low level social signals such as mutual eye gaze, gentle touching, as well as instant virtual messages, increased cooperation relative to a control condition between men but not between women in a public goods game. The tendency of men, but not women, to cooperate more in response to simple social cues might reflect a difference in the forms and functions of intra-sexual coalitions (e.g., Hess & Hagen, 2006; Rucas et al., 2010; Vigil 2007).

Laughter between established friends, however, does not quite follow the apparent pattern for strangers. Research on laughter in developing friendships revealed that antiphonal laughter (i.e., sequential call and response laughter) occurred earlier in women's friendships than in men's, and was established at least three weeks into the relationship, as opposed to males who took up to six weeks to increase antiphonal laugh frequency (Smoski & Bachorowski, 2003b). Laughter between conversationalists not only increases in frequency as people become friends, but also in form. Bryant (2012) found several acoustic differences in laughter between friends and strangers, and that third parties could detect friendship from very brief (< 2 s) instances of co-laughter. Laughter signals clearly play an important role in social interaction, and the functions of

24

interlocutors laughing together vary depending on relationship context, social strategies, and group composition (Bryant & Aktipis, in review).

The current research illustrates how studies of conversational behavior can inform work on the evolution of cooperation. A limitation of our study is that we traded off experimental control for ecological validity—we therefore cannot document a causal relationship, but we found that some forms of conversational coordination were associated with cooperative behavior in a naturalistic interaction. Future research should vary the protocol described here by cueing the importance of both cooperation and competition before the conversation, without revealing the post-conversation social dilemma. In addition, researchers should explore the perception of affiliation between those engaged in conversation and investigate the possibility that some of these coordinated behaviors are designed to transmit coalition information. Finally, these results are based on the behavior and social interactions of American undergraduates, a subpopulation where many are living away from family and established social networks, and therefore possibly more interested in establishing new friendships with strangers. Further research should explore the cross-cultural validity of these findings, especially in relatively closed societies where social ties are longer in duration, and social mobility is lower. The dynamics of conversation can reveal a great deal about how people interact on many levels, and much work remains.

Acknowledgments

25

We thank Hasti Ahangi, Solange Bohling, Erin Campbell, Natalie Chiengkul, Joe Creason, Kimberly Gon, Erika Hoyle, Brandon Jackson, Kotrina Kajokaite, Alexander Kandel, Mara Ludmer, Sabrina Morsello, Kayla Noonen, Sunaina Rao, Sepehr Shakib, Julia Stone, Beniah Suddath, and Andrew West for research assistance. Funding was provided by the International Society for Human Ethology and the UCLA Academic Senate. Xiao Chen provided statistical advice. James Pennebaker and Athena Aktipis commented on the manuscript.

26

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716-723.

Anshel, A., & Kipper, D.A. (1988). The influence of group singing on trust and cooperation. *Journal of Music Therapy*, *25*, 145-155.

Bailenson, J.N., Yee, N., Patel, K., & Beall, A.C. (2008). Detecting digital chameleons. *Computers in Human Behavior, 24*, 66-87.

Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rime (Eds.), Fundamentals of nonverbal behavior (pp. 401-432). Cambridge, England: Cambridge University Press.

Bispham, J. (2006). Rhythm in music: What is it? Who has it? And why? *Music Perception*, 24, 125-134.

Boersma, P, & Weenink, D. (2011). Praat: doing phonetics by computer. [Computer program]. Version 5.3.01. http://www.praat.org/

Bourhis, R.Y., Giles, H., & Lambert, W.E. (1975). Social consequences of accommodating one's style of speech: a cross-national investigation. *International Journal of the Sociology of Langauge*, *6*, 5-71.

Bryant, G. A. (2012). *Shared laughter in conversation as coalition signaling*. Paper presented at the XXI Biennal International Conference on Human Ethology, Vienna, Austria.

27

Bryant, G. A., & Aktipis, A. (2013). *The animal nature of spontaneous human laughter*. Manuscript under review.

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.

Cohen, E.E.A., Ejsmond-Frey, R., Knight, N., & Dunbar, R.I.M. (2010). Rowers' high: behavioural synchrony is correlated with elevated pain thresholds. *Biology Letters*, *6*, 106-108.

Davila-Ross, M., Allcock, B., Thomas, C., & Bard, K. A. (2011). Aping expressions? Chimpanzees produce distinct laugh types when responding to laughter of others. *Emotion, 11*(5), 1013-1020.

Dawkins, R., & Krebs, J.R. (1981). Animal signals: information or manipulation? In J.R. Krebs & N.B. Davies, Eds. *An introduction to behavioural ecology*. Oxford: Blackwell.

DeSteno, D., Breazeal, C., Frank, R.H., Pizarro, D., Baumann, J., Dickens, L., & Lee, J.J. (2012). Detecting the trustworthiness of novel partners in economic exchange. *Psychological Science, 23*, 1549-1556.

Fehr, E., & Camerer, C.F. (2007). Social neuroeconomics: the neural circuitry of social preferences. Trends in *Cognitive Sciences, 11*, 419-427.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics, 10*, 171-178.

Flamson, T. J., Bryant, G. A., & Barrett, H. C. (2011). Prosody in spontaneous humor: Evidence for encryption. *Pragmatics and Cognition, 19*(2), 248-267.

28

Gervais, M.M., Kline, M., Ludmer, M., George, R. & Manson, J.H. 2013. The strategy of psychopathy: primary psychopathic traits predict defection on low-value relationships. *Proceedings of the Royal Society of London B.*

Gonzales, A.L., Hancock, J.T., & Pennebaker, J.W. (2010). Language style matching as a predictor of social dynamics in small groups. *Communication Research, 31*, 3-19.

Greenfield, M.D. (1994). Synchronous and alternating choruses in insects and anurans: common mechanisms and diverse functions. *American Zoologist*, *34*, 605-615.

Hagen, E.H., & Bryant, G.A. (2003). Music and dance as a coalition signaling system. *Human Nature*, *14*, 21-51.

Hall, M.L., & Magrath, R.D. (2007). Temporal coordination signals coalition quality. *Current Biology*, *17*, R406-R407.

Hayashi, N., Ostrom, E., Walker, J., & Yamagishi, T. (1999). Reciprocity, trust, and the sense of control: a cross-societal study. *Rationality and Society, 11*, 27-46.

Hess, N. H., & Hagen, E. H. (2006). Sex differences in indirect aggression: Psychological evidence from young adults. *Evolution and Human Behavior, 27*, 231-245.

Hove, M.J., & Risen, J.L. (2009). It's all in the timing: interpersonal synchrony increases affiliation. *Social Cognition*, *27*, 949-960.

Ireland, M.E., & Pennebaker, J.W. (2010). Language style matching in writing: synchrony in essays, correspondence, and poetry. *Journal of Personality and Social Psychology*, *99*, 549-571.

Ireland, M.E., Slatcher, R.B., Eastwick, P.W., Scissors, L.E., Finkel, E.J., & Pennebaker, J.W. (2011). Language style matching predicts relationship initiation and stability. *Psychological Science*, *22*, 39-44.

29

Jungers, M.K., Palmer, C., & Speer, S.R. (2002). Time after time: the coordinating influence of tempo in music and speech. *Cognitive Processing*, *1-2*, 21-35.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*, 354-364.

Kiyonari, T., Tanida, S., & Yamagishi, T. (2000). Social exchange and reciprocity: confusion or a heuristic? *Evolution and Human Behavior*, *21*, 411-427.

Kurzban, R. (2001). The social psychophysics of cooperation: Nonverbal communication in a public goods game. *Journal of Nonverbal Behavior*, *25*, 241-259.

Lakin, J.L., Jefferis, V.E., Cheng, C.M., & Chartrand, T.L. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, *27*, 145-162.

Levenson, M.R., Kiehl, K.A., & Fitzpatrick, C.M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of Personality and Social Psychology, 68*, 151-158.

Lynch, R. (2011). It's funny because we think it's true: Laughter is augmented by implicit preferences. *Evolution and Human Behavior, 31*, 141–148.

Marsh, K.L., Richardson, M.J., Baron, R.M., & Schmidt, R.C. (2006). Contrasting approaches to perceiving and acting with others. *Ecological Psychology*, *18*, 1-38.

Mehu, M., & Dunbar, R. I. M. (2008). Naturalistic observations of smiling and laughter in human group interactions. *Behaviour, 145*, 1747–1780.

Meyer, A.S, & Bock, K. (1999). Representations and processes in the production of pronouns: some perspectives from Dutch. *Journal of Memory and Language*, *41*, 281-301.

Miles, L. K., Griffiths, J. L., Richardson, M. J., & Macrae, C. N. (2009). Too late to coordinate: Contextual influences on behavioral synchrony. *European Journal of Social Psychology, 40*, 52–60.

Miles, L. K., Lumsden, J., Richardson, M. J., & Macrae, N. C. (2011). Do birds of a feather move together? Group membership and behavioral synchrony. *Experimental Brain Research, 211*, 495–503.

Newman, M.L., Groom, C.J., Handelman, L.D., & Pennebaker, J.W. (2008). Gender differences in language use: an analysis of 14,000 text samples. *Discourse Processes*, 45, 211-236.

Niederhoffer, K.G., & Pennebaker, J.W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology, 21*, 337-360.

Owren, M. J., & Bachorowski, J. A. (2003). Reconsidering the evolution of nonlinguistic communication: the case of laughter. *Journal of Nonverbal Behavior, 27*(3), 183–200.

Pardo, J. S., Cajori Jay, I., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics, 72*, 2254-2264

Pennebaker, J.W., Booth, R.J., & Francis, M.E. (2007). *Linguistic Inquiry and Word Count (LIWC 2007): A computer-based text analysis program*. Austin, TX: LIWC.net.

Pennebaker, J.W., Francis, M.E., & Booth, R.J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC 2001*. Mahwah, NJ: Erlbaum.

692 Pennebaker, J.W., Mehl, M.R., & Niederhoffer, K.G. (2003). Psychological aspects of
693 natural language use: our words, or selves. *Annual Review of Psychology*, *54*,
694 547-577.
695 Phillips-Silver, J., Aktipis, C.A., & Bryant, G.A. (2010). The ecology of entrainment:
696 foundations of coordinated rhythmic movement. *Music Perception*, *28*, 3-14.
697 Pickering, M.J., & Garrod, S. (2004). Towards a mechanistic psychology of dialogue.
698 *Behavioral and Brain Sciences, 27*, 169-226.
699 Platow, M. J., Haslam, S. A., Both, A., Chew, I., Cuddon, M., & Goharpey, N. (2005).
700 "It's not funny if they're laughing": Self-categorization, social influence, and
701 responses to canned laughter. *Journal of Experimental Social Psychology, 41*,
702 542–550.
703 Rucas, S., Gurven, M., Kaplan, H., & Winking, J. (2010). The social strategy game:
704 Resource competition within female social networks among small-scale forager-
705 horticulturalists. *Human Nature, 21*(3), 1–18.
706 Smoski, M. J., & Bachorowski, J. A. (2003a). Antiphonal laughter between friends and
707 strangers. *Cognition & Emotion, 17*, 327–340.
708 Smoski, M.J. & Bachorowski, J.A. (2003b) Antiphonal laughter in developing
709 friendships. *Annuals of the New York Academy of Sciences, 1000*, 300–303.
710 Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann,
711 G., Rossano, F., de Ruiter, J.P., Yoon K-E, & Levinson, S.C. (2009). Universals
712 and cultural variation in turn-taking in conversation. *Proceedings of the National*
713 *Academy of Sciences*, *106*, 10587-10592.

Street, R.L. (1984). Speech convergence and speech evaluation in fact-finding interviews. *Human Communication Research, 11,* 139-169.

Taylor, P.J., & Thomas, S. (2008). Linguistic style matching and negotiation outcome. *Negotiation and Conflict Management Research, 1,* 263-281.

Valdesolo, P., & DeSteno, D. (2011). Synchrony and the social tuning of compassion. *Emotion*, *11*, 262-266.

Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: synchrony promotes cooperative ability. *Journal of Experimental Social Psychology*, *46*, 693-695.

van Baaren, R.B., Holland, R.W., Steenaert, B., & van Knippenberg, A. (2003). Mimicry for money: behavioral consequences of imitation. *Journal of Experimental Social Psychology*, *39*, 393-398.

Van Huyck, J.B., Battalio, R.C. & Beil, R.O. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, *80*, 234-248.

Vigil, J. M. (2007). Asymmetries in the social styles and friendship preferences of men and women. *Human Nature, 18,* 143-161.

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

Wiltermuth, S.S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science*, 20, 1-5.

33