Failure to (Mis)communicate: Linguistic Convergence, Lexical Choice, and Communicative Success in Dyadic Problem Solving

Alexandra Paxton (paxton.alexandra@gmail.com)
Cognitive and Information Sciences, University of California, Merced
Merced, CA 95343 USA

Jennifer M. Roche (jroche@bcs.rochester.edu) **Alyssa Ibarra (aibarra@bcs.rochester.edu)**
Michael K. Tanenhaus (mtan@bcs.rochester.edu)
Brain and Cognitive Sciences, University of Rochester
Rochester, NY 14627 USA

Abstract

The current study evaluates how lexical choice impacts task performance in dyads tasked with building an object together without a shared visual environment. Our analyses suggest that, while interpersonal lexical convergence in target linguistic categories promotes successful communication, success does not require convergence in all categories. In the absence of shared visual workspaces and face-to-face communication, success increases when interlocutors converge in establishing common ground but decreases with increased co-occurrence of knowledge-state words, perhaps due to mutual hedging and uncertainty. Finally, miscommunication increases with use of ambiguous spatial terms and with markers of confusion, pointing to unique lexical signatures for successful and unsuccessful communication.

Keywords: communication; convergence; coordination; joint action; miscommunication; psycholinguistics

Introduction

Joint action, broadly defined, is our ability to work cooperatively with others to achieve a shared goal (Clark, 1992). While many philosophers have focused on higher-level processes in achieving joint action (e.g., Bratman, 1992), cognitive scientists have framed much of their work around the ways in which lower-level processes can contribute to joint action (e.g., Tollefsen & Dale, 2012). From dancing the tango to preparing a meal, we often work with others to achieve mutual goals, and we rely heavily on our shared environments and one another to coordinate our actions (e.g., Sebanz, Bekkering, & Knoblich, 2006).

However, with the use of mobile technologies and computer-mediated communication on the rise, we increasingly find ourselves coordinating with others who do not share our physical environment. As technology breaks down barriers to collaboration imposed by long distance, it creates new questions about the ways in which we communicate during joint action. A major goal of the present project is to examine how individuals work together to achieve a mutual goal without the benefit of a shared visual field or face-to-face communication. Here, we focus on the impact of miscommunication on task performance from a framework heavily influenced by work on interpersonal convergence and joint action.

Interpersonal Convergence

Research on interpersonal interaction over the last several decades demonstrates that individuals tend to become increasingly attuned to one another over the course of an interaction across a variety of communication channels. This phenomenon can be generally referred to as *interpersonal synchrony* or *convergence*. Evidence for interpersonal convergence has been found across a number of dimensions: Interacting individuals begin to have more similar linguistic choices (e.g., Niederhoffer & Pennebaker, 2002; Pickering & Garrod, 2004), body movements (e.g., Richardson, Marsh, Goodman, & Schmidt, 2007), gaze patterns (e.g., Dale, Kirkham, & Richardson, 2011), and even physiological responses (e.g., Helm, Sbarra, & Ferrer, 2011). A prevailing view of interpersonal convergence suggests that it serves to facilitate interaction and strengthen social bonds (e.g., Lakin, Jefferis, Cheng, & Chartrand, 2003). Greater degrees of convergence have been linked to increased rapport (Hove & Risen, 2009) and liking (Chartrand & Bargh, 1999).

While shared environment, goals, and stimuli facilitate convergence (e.g., Shockley, Richardson, & Dale, 2009), interpersonal convergence has also been shown in the absence of face-to-face communication. Individuals who hear recorded speeches will exhibit similar gaze (Richardson & Dale, 2005) and neural patterns (Stephens et al., 2010), and producing speech with similar stress patterns leads to synchronized postural sway, even when partners cannot see one another (Shockley, Richardson, & Dale, 2009). Other work has found that individuals rate computer-mediated conversations more highly when they linguistically converge with their partners (Niederhoffer & Pennebaker, 2002), and coordination and task performance improve during computer-mediated communication when individuals share virtual work environments (Introne & Alterman, 2006).

Interpersonal convergence appears to be a robust phenomenon that can occur even in relatively impoverished communication. Interacting individuals may influence one another even when they have only a single communication channel available (e.g., text; Niederhoffer & Pennebaker, 2002) or when speaker and

listeners are separated temporally and spatially (e.g., Richardson & Dale, 2005; Stephens et al., 2010). Additionally, increasing shared resources and common ground between individuals facilitates task performance (e.g., Introne & Alterman, 2006).

Miscommunication

As reviewed above, interpersonal convergence promotes cooperation and perspective taking (e.g., Lakin et al., 2003). Therefore, we believe convergence may provide a salient point of comparison between successful communication and miscommunication. Miscommunication has been relatively understudied compared to successful communication, but we use existing work to guide our expectations.

Miscommunication has traditionally been conceptualized as uninformative noise in the system (cf. Keysar, 2007), but investigations of dialogue posit that this "noise" may lead to more precise communicative representations (e.g., Clark; 1996; Healy, 1997). Coupland, Giles, and Weimann (1991) suggest that miscommunication can actually provide rich information about how interlocutors come to communicate successfully. Successful communication necessarily requires interlocutors to coordinate and update mutual knowledge, experiences, beliefs, and assumptions (see Clark & Marshall, 1981). However, the process of regularly updating this information may be ridden with unsuccessful attempts that may ultimately help interlocutors reach their common goal. McTear (2008) suggests that a common root of these communication failures is related to misaligned mental states or perspectives on shared visual contexts. Given the difficulty of assessing the former during active communication, we here address communication relative to shared visual context.

The Present Study

The present study uses linguistic convergence as a lens through which to examine dyads' collaborative task performance and miscommunication when deprived of face-to-face interaction. More specifically, convergence may be optimal along some dimensions but not all (e.g., Riley, Richardson, Shockley, & Ramenzoni, 2011). For instance, recent work by Fusaroli et al. (2012) reveals that convergence of task-related word choice may be more important for successful dyadic task performance than generalized linguistic convergence. Rather than focus on overall linguistic convergence, therefore, we will focus on a handful of linguistic factors identified from previous analyses of this corpus (Roche, Paxton, Ibarra, & Tanenhaus, 2013) as being significantly related to states of successful and unsuccessful communication.

We analyzed transcripts with linguistic categories provided by LIWC (Linguistic Inquiry and Word Count; Pennebaker, Booth, & Francis, 2007). LIWC, a well-established text analysis tool, calculated the degree to which interlocutors use a number of language categories. Such categories were not meant to map onto any specific linguistic forms or stages of language processing. Rather, LIWC served as a useful tool to provide a descriptive picture of discourse as it unfolded.

LIWC distinguishes among a number of linguistic categories (e.g., some predicate classes, pronouns, verbs). The categories selected for the analysis were determined by an initial evaluation of the bloco^{tm} corpus (Roche et al., 2013) and included the following categories: *Cognitive Mechanism* (e.g., "think," "know"); *Perceptual* (e.g., "see," "hear"); *Spatial* (e.g., "top," "bottom"); *Assent* (e.g., "yes," "mmhmm"); *First Person Pronouns* (e.g., "I," "we"); and *Second Person Pronouns* (e.g., "you"). Roche et al.'s findings indicated that a subset of linguistic categories differentiated successful from unsuccessful communication. In order to relate the present analyses to previous ones, the categories most relevant to predicting communication states in previous analyses were integrated into the current analysis to determine if lexical convergence affected the congruence of dyads' visual environments.

From our earlier results and other work from the miscommunication literature (e.g., Coupland et al., 1991; McTear, 2008), we approach the study with several hypotheses about positive and negative predictors of communicative success and the shift from successful to unsuccessful communication. First, we predict that the convergence of grounding, assent words, and personal pronouns will be positively predictive of effective performance. As participants offer new information to one another, assent words may increase when the information offered is correct (i.e., indicative of grounding), but also while listeners verbally track speakers (especially without the benefit of nonverbal tracking). Increased use of personal pronouns may be associated with speakers' attempts to relate or take one another's perspective. In all of these cases, increased understanding of one another's perspective may allow participants to act more effectively as a dyad.

Second, we ask whether convergence on other dimensions might predict miscommunication. It seems plausible that convergence of negative words would be diagnostic of miscommunication. However, it is less clear what to expect for other categories – particularly words related to cognitive processes, spatial words, and words related to perceptual processes. Moreover, whereas convergence along these dimensions could be correlated with temporary miscommunication, convergence might alternatively predict overall task success, because use of these words is consistent with effective use of repair strategies. Increased co-occurrence of negation terms may suggest that participants are acknowledging problems with the task itself or in understanding one another, and use of cognitive words (e.g., "think") may indicate hedging, ambiguity, or uncertainty. Without the benefit of a shared visual field, spatial terms and perceptual words

may also prove to be a point of confusion (e.g., "up").

Finally, in addition to studying the aforementioned lexical items in the context of interpersonal convergence, we explore how different lexical choices may be more predictive of miscommunication. This last analysis will not focus on the degree to which individuals are converging in their use of terms. Instead, we will examine how use of these words may predict task failure when used by either participant.

Method

The current project analyzed part of a larger corpus aimed at capturing the linguistic and behavioral dynamics of dyadic task performance with and without shared visual fields. In the present subset of the corpus, participants were engaged in a turn taking task that required them to build three-dimensional puzzles based on pictorial instructions cards. Participants were unable to see their partner, their partner's workspace, or their partner's instruction cards during the interaction and were only able to coordinate the building through spoken language exchanges. The dyadic interactions were transcribed and annotated for relevant linguistic and behavioral measures discussed below.

Participants

Participants included 20 dyads of paid undergraduate students from the University of Rochester (N = 40; females = 26; mean age = 19 years). All participants were native speakers of American English with normal to corrected vision. None reported speech or hearing impairments.

Stimuli and Procedure

Stimuli included two bloco^{tm} objects, three-dimensional animal puzzles consisting of approximately 27 unique pieces each. The building process was divided into steps (M = 14; range = 13-15). Unique pictorial instruction cards guided participants through each step.

Participants were seated at identical workspaces separated by a partition. Participants were asked to work together to build the figure using the instruction cards provided to them. They were told to take turns providing the instructions but that both participants could otherwise speak freely. Once they completed the final instruction, the researcher informed the dyad whether they had built the object correctly. (Only about 2 dyads had mistakes after completing the figure, and both were minor, e.g., with grasshopper legs upside-down.) If both participants did not construct the figure completely correctly, the participants were told that something did not match and to identify and fix their mistake.

During the experiment, each dyad was video-recorded from three angles to get full views of each participant's workspace and to capture both participants together in profile. This aided in coding behavioral and performance data through the course of the interaction, based on how well the dyad's visual environments matched during the interaction (described below). The video recordings also included audio information, from which we fully transcribed the verbal exchanges between participants.

Grounding was determined by a similar procedure described by Nakatani and Traum (1999) in their discussion of grounding units. At each turn, T^A (Talker A) presented a new piece of information. It was not until T^B (Talker B) accepted this information that the linguistic exchange was coded as grounded. For example:

T^A: Take the big green piece and put the holes facing up.
T^B: ***Okay.***

This example was counted as an attempt to ground the information presented. It is also important to note that we are not considering other forms of grounding. We only evaluate this form of explicit verbal grounding in the current paper but intend to consider other forms of grounding in the future.

Visual Congruence

Task success was operationalized as the visual congruence of the interlocutor's workspaces. An undergraduate research assistant (RA) coded the dyads' workspaces as either matching or mismatching on a turn-by-turn basis. There were a total 8493 turns (M = 425 turns; standard deviation = 176) across the 20 dyads. The RA coded the visual environment at the end of each turn (i.e., the state of the environment when Talker A finished talking and Talker B began talking).

We determined the reliability of the coding by having two additional blind coders (with no prior knowledge of the experiment) evaluate 5% of the visual congruence codes (425 turns) from the original RA codes. These blind coders were asked to code whether they agreed or disagreed with the first RA's visual congruence codes. These codes were then subjected to an inter-rater reliability analysis, indicating high agreement with the primary coder (kappa = .96).

Example of the Visual Congruence Coding Procedure
Often, a speaker (T^A) was required to describe a spatial orientation to his or her partner (T^B). If T^B physically moved the object to the correct orientation (as intended by T^A based on by T^A's workspace and instruction card), the current turn was marked as visually congruent. However, if T^B failed to put the object in the correct orientation, the turn was marked as visually incongruent. Figure 1 was created to demonstrate what an incongruent turn might have looked like. In this turn, T^A instructed T^B to orient the holes in an upward fashion, but the ambiguous use of "up" resulted in a visually incongruent turn.

Results

We explored the relationship between task performance

Figure 1. Visually incongruent orientation for instruction: "Take the big green piece and put the holes facing up."

and communication patterns using two statistical models with distinct goals: (1) predicting communicative success (visual congruence) with lexical convergence and (2) uncovering lexical contribution to localized task failure (visual incongruence). We expected that aligning at the lexical level should contribute to the congruence of the visual workspace across participants, thus resulting in task success. Additionally, we expected that moments of imperfect communication (i.e., miscommunication) should also be uniquely identified within the lexical context. Models were run as a linear mixed-effects model and a mixed logit model, respectively.

Conditional Probability of Task Success

We were interested in overall task success; therefore, we calculated the overall conditional probability of *Visual Congruence* given *Grounding*. This provided a preliminary test of the anticipated positive relation between grounding and congruence: If participants were grounding, their workspaces should be visually congruent. Consistent with these expectations, participants were, in fact, grounding 64% (∂ = .16) of the time when their workspaces were visually congruent.

Task Success: Linear Mixed-Effects Model

Congruence was measured as cross-correlation coefficients between the target time series of each participant (e.g., Paxton & Dale, 2013), calculated at time lags of +/- 10 turns (see Figure 2). Because we were primarily interested in temporal structure of the convergence, the predictors in this model were not main effects but rather were interaction terms between the target variables and lag (+/-10 turns). These terms allowed us to measure how performance is affected by linguistic convergence as it occurs *in time*, a common framework within the convergence literature. We chose to analyze congruence of success (rather than congruence of miscommunication) due to the necessary nature of shared success to complete a joint task. In this sense, miscommunication can be alternatively conceptualized as a sort of decoupling of success.

Using the cross-correlation coefficients as continuous predictor and outcome variables, we created a linear mixed-effects model predicting *Visual Congruence* with the LIWC variables and *Grounding*, with fully specified random slopes and dyad as the random intercept. All terms were centered and standardized prior to the analysis, allowing us to interpret the estimates as effect sizes. After removing *Perceptual* and *Spatial* categories due to high covariation (*r*s > .35), the final model included the following predictor variables (as interaction effects with lag): *First* and *Second Person Pronouns*, *Assent*, *Cognitive Mechanism*, and *Grounding*. This resulted in significant effects for the predictor variables *Cognitive Mechanism* ($p < .05$, ß = 0.13) and *Grounding* ($p < .05$, ß = 0.09; see Figure 2).

The results from the mixed model indicate that an increase in the temporal convergence of *Grounding* is positively related to an increase in the convergence of physical workspace environments. However, the convergence of *Cognitive Mechanisms* was negatively related to visual congruence, indicating that as interlocutors became more aligned in their use of *Cognitive Mechanism* words, the less they aligned in their respective physical workspaces. This could suggest the tendency to hedge: Until potential ambiguities may be clarified, T^B may have waited until the information was clear to accept or reject a statement from T^A.

Additionally, the convergence of the workspace may not be contingent upon the convergence of other linguistic categories, such as *Personal Pronoun* and *Assent*. This does not necessarily mean that other linguistic categories do not contribute to task success; rather, they may not need to be aligned to result in the convergence of the visual workspace between interlocutors. Crucially, the convergence of *Grounding* did improve task success, which could indicate a mutual agreement to accept information in a felicitous manner, consistent with existing theories of joint action (e.g., Brennan & Clark, 1996).

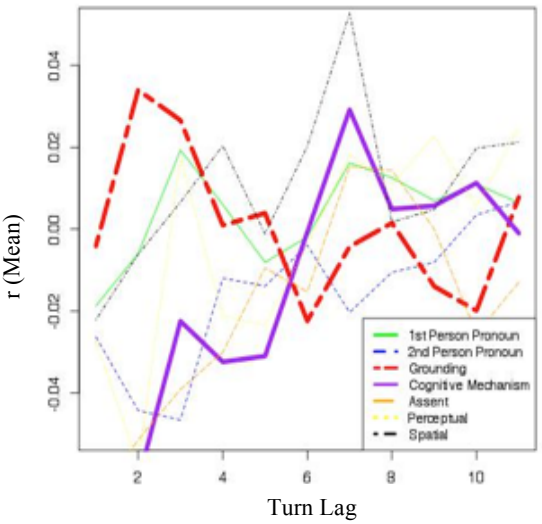

Figure 2. Cross-correlation coefficients (*r*) for all model predictors by time lag (in turns), with significant predictors (*Grounding* and *Cognitive Mechanism*) plotted in thicker lines (dashed and solid, respectively).

Task Failure: Mixed Logit Model

We next analyzed the distribution of visual incongruence according to lexical choice across dyads using a mixed logit model, predicting instances of miscommunication with LIWC categories. Miscommunication, in this case, was a binary (i.e., categorical) dependent variable. The predictors included in the model were *First* and *Second Person Pronoun*, *Cognitive Mechanism*, *Assent*, *Perceptual*, *Negation* and *Spatial* terms. This analysis provided us with an opportunity to explore whether miscommunication has a linguistic "signature" distinct from successful communication. For this model, we used the original time series (centered and standardized) for all variables, not cross-correlation coefficients or the variables' interaction terms with time lag.

The results find significant main effects for *Assent* ($z = -5.18$, $ß = .21$), *Perceptual* ($z = -2.01$, $ß = .09$), *Negation* ($z = 3.87$, $ß = .12$), and *Spatial* ($z = 3.74$, $ß = .12$) terms. As expected, visual incongruence is negatively related to the use of *Assent* ($p < .001$), suggesting that *Grounding* or *Assent* improved task success, as their omission signal poorer task performance. Visual incongruence also increases when *Perceptual* terms decrease ($p < .05$) and when *Negation* and *Spatial* terms increase (both $p < .001$). Using *Negation* words may indicate the interlocutor has recognized a miscommunication and may be attempting to resolve the issue. Miscommunication may have been exacerbated by an increase use of *Spatial* terms. Without shared visual information, spatial references may often be ambiguous (e.g., Figure 1), which may be a common reason for miscommunication generally (e.g., McTear, 2008).

Discussion

Rooted in work on interpersonal convergence, the current study approaches successful communication and miscommunication by comparing changes in linguistic markers as dyads work together to achieve a collaborative goal while overcoming their lack of shared visual field. With this project, we hope to add to these literatures by demonstrating that interpersonal convergence promotes successful communication – but also to support the recent view that perhaps not all communication channels or linguistic choices must be aligned to promote success (e.g., Fusaroli et al., 2012; Riley et al., 2011). It seems that under certain communicative contexts not all aspects of communication are aligned equally and that the temporal convergence of some communicative structures carries more weight than others. While several of our hypothesized variables do not seem to affect communicative success in these analyses, our first model demonstrates that the convergence of *Grounding* and *Cognitive Mechanisms* differentially influence task success: The convergence of *Grounding* promotes the congruence of the visual environment, while the convergence of *Cognitive Mechanisms* may indicate that interlocutors are strategic in their grounding behaviors, hesitating until they receive the information necessary to commit to grounding.

Our second model confirms our expectation that word choice might provide a window into ongoing communicative success. In our analyses, the proportion of incongruence in the visual environment is predicted by specific lexical categories. For example, the use of *Negation* or *Spatial* terms is more likely to signal likely miscommunication. However, additional work should be done in more naturalistic settings to determine whether miscommunication also behaves in these specific ways in the "wild," in addition to experimental settings.

Future Directions

Roche et al.'s (2013) findings and the current results provide an initial look into the mechanisms responsible for successful communication and miscommunication during problem-solving tasks. As shown here, *Grounding* and *Assent* significantly predict successful communication. However, in our observation of the interactions, we found that participants use assent words in at least two different ways: for grounding and for verbal tracking. In the present study, we do not distinguish between these two uses of assent words, but we imagine that each may differentially impact successful communication. In future work, we hope to explore how assent may separate into distinct behavioral patterns.

In addition, we are currently working towards expanding beyond simple descriptions of the communicative system and intend to delve deeper into the system's dynamics. We hope to do this by evaluating state changes between successful and unsuccessful communication to model miscommunication and repair as they unfold in time. While not inconsistent with other explanations, we believe that our current results are highly consistent with the idea of communicative state as a dynamical system: Interacting individuals who begin to fall into miscommunication may push themselves out of that state and into a successful communicative state by increasing their understanding of one another and the situation at large. In addition to modeling communication state changes, we hope to explore ideas of miscommunication and successful communication as attractor states.

Conclusion

We have argued that miscommunication behaves in interesting ways that are distinct from successful communication. Interpersonal convergence along some linguistic dimensions – notably, mutual grounding – gives dyads an advantage in task performance by promoting successful communication. Additionally, the evaluation of lexical choices of interlocutors provides insight into a speaker's current cognitive state and thus predicts successful and unsuccessful communication. Though the results here provide only a preliminary look at these connections, we believe that the present study contributes

towards a more comprehensive framework of communication in general and provides direction for future comparisons of communication across multiple settings.

Acknowledgments

Special thanks go to Rick Dale (University of California, Merced) for feedback and advice on statistical modeling. We also wish to thank to our undergraduate research assistants at the University of Rochester (Chelsea Marsh, Eric Bigelow, Derek Murphy, Melanie Graber, Anthony Germani, Olga Nikolayeva, and Madeleine Salisbury) and the University of California, Merced (Chelsea Coe and J.P. Gonzales). Preparation of this manuscript was supported by grants from the National Institute of Health (RO1 HD027206) to Michael Tanenhaus.

References

Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review, 101*(2), 327-341.

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

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910.

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

Clark, H. H. & Marshall, C. (1981). Definite reference and mutual knowledge. In A. Joshi, B. Webber & I. Sag (Eds.), *Elements of Discourse Understanding*.

Coupland, N., Giles, H. & Wiemann, J. M. (1991). *Miscommunication and problem talk*. Newbury Park: Sage.

Dale, R., Kirkham, N., & Richardson, D. (2011). The dynamics of reference and shared visual attention. *Frontiers in Cognition*, *2*.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylen, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science*, *23*(8), 931–939.

Healey, P. (1997). Expertise or expert-ese: The emergence of task-oriented sub-languages. In M. Shafto & P. Langley (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Helm, J. L., Sbarra, D., & Ferrer, E. (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion*, *12*(4), 748.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, *27*(6), 949–960.

Introne, J., & Alterman, R. (2006). Using shared representations to improve coordination and intent inference. *User Modeling and User-Adapted Interaction*, *16*(3-4), 249–280.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics, 4*, 71-84.

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*(3), 145–162.

McTear, M. (2008). Handling miscommunication: Why bother? *Recent Trends in Discourse and Dialogue, 39*, 101-122.

Nakatani, C. & Traum, D. (1999). Coding discourse structure in dialogue. *University of Maryland Institute for Advanced Computer Studies Technical Report, 1*, 1-42.

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, *21*(4), 337–360.

Paxton, A. & Dale, R. (2013). Argument disrupts interpersonal synchrony. *Quarterly Journal of Experimental Psychology, 66*(11), 2092-2102.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic Inquiry and Word Count: LIWC* [Computer software]. Austin, TX: LIWC.net.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(02), 169–190.

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R. L., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, *26*(6), 867–891.

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*(6), 1045–1060.

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

Roche, J. M., Paxton, A., Ibarra, A., & Tanenhaus, M. (2013). From minor mishap to major catastrophe: Lexical choice in miscommunication. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Conference of the Cognitive Science Society.* Austin, TX: Cognitive Science Society.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76.

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

Tollefsen, D., & Dale, R. (2012). Naturalizing joint action. *Philosophical Psychology*, *25*(3), 385-407.