5-2016

Conversational Alignment: A Study of Neural Coherence and Speech Entrainment

Kristen M. Jensen
Utah State University

Stephanie A. Borrie
Utah State University

Breanna E. Studenka
Utah State University

Ronald B. Gillam
Utah State University

Recommended Citation

Jensen, Kristen M.; Borrie, Stephanie A.; Studenka, Breanna E.; and Gillam, Ronald B., "Conversational Alignment: A Study of Neural Coherence and Speech Entrainment" (2016). *All Graduate Plan B and other Reports*. 779.
https://digitalcommons.usu.edu/gradreports/779

EMMA ECCLES JONES
College of Education & Human Services
UtahStateUniversity

DEPARTMENT OF COMMUNICATIVE
DISORDERS AND DEAF EDUCATION

Conversational alignment: A study of neural coherence and speech entrainment

Kristen M. Jensen, Stephanie A. Borrie, Breanna E. Studenka, & Ronald B. Gillam

Utah State University
Communicative Disorders and Deaf Education

Master’s Plan B

Conversational alignment: A study of neural coherence and speech entrainment

Introduction

Conversational alignment refers to the tendency for communication partners to adjust their verbal and non-verbal behaviors to become more like one another during the course of human interaction. This alignment phenomenon has been observed in neural patterns, specifically in the prefrontal areas of the brain (Holper et al., 2013; Cui et al., 2012; Dommer et al., 2012; Holper et al., 2012; Funane et al., 2011; Jiang et al., 2012); verbal behaviors such acoustic speech features (e.g., Borrie & Liss, 2014; Borrie et al., 2015; Lubold & Pon-Barry, 2014), phonological features (e.g., Babel, 2012; Pardo, 2006), lexical selection (e.g., Brennan & Clark, 1996; Garrod & Anderson, 1989), syntactic structure (e.g., Branigan, Pickering, & Cleland, 2000; Reitter, Moore, & Keller, 2006); and motor behaviors including body posture, facial expressions and breathing rate (e.g., Furuyama, Hayashi, & Mishima, 2005; Louwerse, Dale, Bard, & Jeuniaux, 2012; Richardson, March, & Schmit, 2005; Shockley, Santana, & Fowler, 2003; McFarland, 2001).

While conversational alignment in itself, is a largely physical phenomenon, it has been linked to significant functional value, both in the cognitive and social domains. Cognitively, conversational alignment facilitates spoken message comprehension, enabling listeners to share mental models (Garrod & Pickering, 2004) and generate temporal predictions about upcoming aspects of speech. From a social perspective, behavioral alignment has been linked with establishing turn-taking behaviors, and with increased feelings of rapport, empathy, and intimacy between conversational pairs (e.g., Lee et al. 2010; Nind, & Macrae, 2009; Smith, 2008; Bailenson & Yee, 2005; Chartrand & Barg, 1999; Miles, Putman & Street, 1984; Street & Giles, 1982). Benus (2014), for example, observed that individuals who align their speech features are perceived as more socially attractive and likeable, and have interactions that are more successful. These cognitive and social benefits, associated with conversational alignment, have been observed in both linguistic and neural data (e.g., Holper et al., 2012; 2013, Cui et al. 2012; Jiang et al., 2012; Egetemeir et al., 2011; Stephens et al. 2010).

The purpose of the current study was to examine conversational alignment as a multi-level communication phenomenon, by examining the relationship between neural and speech behaviors. To assess neural alignment, we used Near-Infrared Spectroscopy (NIRS), a non-invasive neuroimaging technology that detects cortical increases and decreases in the concentration of oxygenated and deoxygenated hemoglobin at multiple measurement sites to determine the rate that oxygen is being released and absorbed (Ferrari & Quaresima, 2012). While still considered a relatively new neural imaging technique, NIRS has been well established as an efficacious and effective data collection approach, particularly appropriate for social interaction research (e.g., Holper et al., 2013; Jiang et al., 2012; Holper et al., 2012; Suda et al., 2010). We utilized hyperscanning, a technique that allows for the quantitation of two simultaneous signals, allowing us to document neural alignment between two individuals (Babiloni & Astolfi, 2012). Recent studies have revealed neural alignment between two persons in cooperative states, including alignment in the right superior frontal cortices and medial prefrontal regions (Cui et al., 2012; Dommer et al., 2012; Funane et al., 2011). This increased prefrontal interbrain alignment has also been observed in other social interactions, including joint attention tasks (Dommer et al., 2012), imitation tasks (Holper et al., 2012), competitive games (Cheng et al., 2015, Duan et al., 2013), teaching-learning interactions (Holper et al., 2013), face-to-face communication (Jiang et al., 2012), mother-child interactions (Hirata et al., 2014), and during cooperative singing tasks (Osaka et al., 2015). Interestingly, Jiang et al. (2012) showed that increased neural alignment only occurred between conversational participants when they were speaking face-to-face, but not when participants had their backs facing one another. The authors speculated that the multi-sensory information, for example motor behaviors such as gestures, was required for neural alignment to occur.

Current Study

This study investigates if neural coherence drives alignment of behaviors in the speech domain. We used NIRS as a vehicle to examine this, specifically addressing the following research questions: (1) does neural alignment in the frontal cortex occur between dyads engaged in spoken dialogue; and (2) does neural alignment correlate with speech entrainment. Based on research implicating conversational alignment as a multi-level communication phenomenon and robust evidence of both neural coherence and speech entrainment in isolated studies, we hypothesize that evidence of neural alignment in the frontal cortex will be observed, and further that the degree of alignment will correlate with the degree of speech entrainment.

Methods

Participants

Data for this study was collected from 20 young, healthy female participants, with ages ranging from 19 to 46 years (*M* = 26.5, *SD* = 6.62). All participants were right handed, with normal or corrected-to-normal vision and hearing. Participants were all native speakers of American English with no neurological history or pathological speech patterns. Participants were randomly assigned as pairs; members of pairs were not well acquainted with each other before experiment. The study procedures were approved by the Institutional Review Board of Utah State University.

Procedure

Each dyad participated in a single experimental recording session in a quiet lab at Utah State University. Upon entering the lab, each individual participant was fitted with a “3x5” NIRS cap, that was aligned across the forehead according to standardized procedures. Central emitters and detectors were aligned to midline by orientating central optodes with the nasion, and aligned vertically by placing the bottom of the cap along the supraorbital ridge. In addition to the NIRS cap, each participant was fitted with a wireless CVL Lavalier microphone, synced with a Shure BLX188 DUAL Lavalier System connected to a Zoom H4N Portable Digital Recorder. Separate audio channels and standard settings (48 kHz; 16-bit sampling rate) were employed.

Participants were then seated in chairs, directly facing one another and were told that they would engage in two tasks, an initial resting state task followed by a spoken dialogue task. For the resting state task, standardized procedures were used where participants were instructed to close their eyes, rest their mind, and remain as still as possible (Lu et al., 2010). For the dialogue task, each participant was given one of a pair of pictures and was instructed to hold their picture at an angle at which it would not be visible to their conversational partner. The dyad was then told that their task was to work together, simply by speaking to one another, to identify the differences between the pair of pictures. They were also informed that individual conversational turns of each participant were to be ten seconds in length, with a beep signaling the end of one participant’s turn and the beginning of the other. The dyad was instructed to continue conversing like this until indicated to stop. This dialogue elicitation procedure continued for 5 minutes, consisting of 15 conversational speaking turns per participant. On completion of the dialogue task, the audio recording equipment was turned off and the NIRS caps were removed. Participants were debriefed and thanked before leaving the lab. The neural and audio data was then transferred to a computer for subsequent analysis.

Data Collection and Analysis

Neural data. NIRS measurements were conducted using an ETG-4000 optical topography at two wavelengths (695nm and 850nm), with a sampling rate of 0.1 seconds, to determine the concentration and absorption of oxygenated and deoxygenated hemoglobin during the tasks. Oxygenated hemoglobin has an increased sensitivity to changes in cerebral blood flow, therefore, the current study analyzed oxygenated hemoglobin concentrations (Lindenberger et al., 2009). To assess the relationship between the neural activation of each dialogue, coherence increase measures were implemented across the dialogue and resting state tasks. Wavelet transfer coherence (WTC) was used to assess the neural activation by measuring cross-correlation between the two time-series as a function of frequency and time by modifying and applying the Matlab package by Grindsted et al. (2004). We were then able to calculate the average interbrain coherence at our bands of interests (50 to 7 Hz) during the initial rest period and the dialogue task. The average coherence value of the dialogue was then subtracted from the initial resting state to reveal an index of neural synchronization increase, which was then converted to a z-value before statistical tests were performed.

Acoustic data. Acoustic analysis software, Praat (Boersma & Weenink, 2014) was used to manually annotate the audio recording files of the dialogue task, for each dyad, for individual speaking turns. The beginning of each speaking turn is identified as the moment that a participant began articulating an utterance and ends when articulation ceased. Utilizing the segmented .wav files, four acoustic features were computed for each speaking turn: average pitch, average intensity, pitch standard deviation, and intensity standard deviation. Using this data, we computed a synchrony score for each dyad. The synchrony analysis, used in a number of previous speech entrainment studies (Borrie, et al., 2015; Lubold & Pon-Barry, 2014), considers entrainment to be a local phenomenon occurring on a turn-by-turn basis and provides a conversation-level score that reflects the amount of entrainment for a single acoustic feature throughout the conversation. To illustrate, two speakers exhibit synchrony when they modulate their acoustic features in tandem (see also Levitan & Hirschberg, 2011). For example, two conversational participants may have very different raw feature values for their average pitch, but on a turn-by-turn basis as they engage in spoken dialogue, they adjust their pitch in the same direction as that of their partner. Thus, when one speaker increases their pitch, the other speaker reacts by also increasing their pitch. A synchrony score was computed for each dyad using Pearson's correlation coefficient with a two-tailed *t*-test on the speakers' raw feature values at each turn.

Results

Neural Data

Paired samples *t* tests were used to examine the difference between neural coherence between dyads during the resting state and during the spoken dialogue. This was done across all dyads and also at the level of the individual dyad. Across all dyads, there was a significant difference between neural coherence during resting state and neural coherence during spoken dialogue, $t(219) = 2.988$, $p < .001$. Thus, dyads exhibited significantly greater neural coherence when conversing with one another relative to the silent rest period. When the t tests were run at the level of the individual dyad, this relationship was observed in seven of the ten dyads—Dyad 1, $t(219) = 2.18$, $p = .04$, Dyad 2, $t(219) = 2.80$, $p = .01$, Dyad 5, $t(219) = 3.75$, $p < .001$, Dyad 7, $t(219) = 3.37$, $p = .003$, Dyad 8, $t(219) = 4.26$, $p < .001$, Dyad 9, $t(219) = 2.86$, $p = .009$, and Dyad 10, $t(219) = 5.14$, $p < .001$. This serves to indicate that neural activation patterns align more closely when conversational participants are engaged in spoken dialogue.

Acoustic Data

Synchrony scores, for all four acoustic features, were not significant. Thus, an investigation into the relationship between neural data and acoustic data is not justified.

Discussion

The purpose of the current study was to examine conversational alignment as a multi-level communication phenomenon, investigating if neural coherence drives entrainment in the speech domain. While we found evidence of neural coherence for participants engaging in spoken dialogue, we failed to find evidence of speech entrainment. These findings are discussed in more detail below.

In the present study, we found, by applying inter-brain neural coherence increase measures, robust neural alignment during a spoken dialogue over the frontal regions. These findings are consistent with previous research that has also employed hyperscanning to investigate neural coherence during social interactions (Osaka et al., 2015; Holper et al., 2013; Jiang et al. 2012). Using a similar methodology to the one reported in the current study, Jiang and colleagues investigated neural coherences in regards to activation with the frontopolar cortical regions during face-to-face conversations versus back-to-back conversations. The authors reported neural coherence between communication partners during face-to-face conversations but not during back-to-back conversations; and speculated that the addition of non-verbal behaviors, when communicating face-to-face, may be critical for neural alignment. Saito et al. (2007), also conducted a conversational alignment study, but used fMRI as opposed to NIRS. Here, the authors found neural coherence in the medial frontal cortex and bilateral anterior superior temporal gyrus. Thus, the current findings add further support to existing literature that affords evidence of neural coherence in the frontal cortex during spoken dialogue.

Communication is a diverse neural task that requires multi-modalities. Speech and language are recognized to occur in the posterior-superior temporal gyrus (Wernicke’s area) which is responsible for the comprehension of the message, and the superior inferior temporal gyrus (Broca’s area) responsible for movements of articulators. However, the frontal cortex plays a crucial role during social interactions integrating non-verbal messages into our understanding of communication (Duffy, 2005). The medial frontal cortex, in particular, plays an important role associated with social cognition (Amodio & Frith, 2006). The medial prefrontal cortex is responsible for action monitoring, self-knowledge, judgements, predicting others behaviors and outcome monitoring (Amodio & Frith, 2006). Therefore, neural coherence increase in the frontal cortex, as found in the current study and existing research, may relate to the role that the medial frontal cortex plays during successful communicative interactions, permitting the communication partner to mentalize, make judgements, internalize perceptions and to monitor the success of the conversations.

In contrast to neural alignment, we did not observe speech entrainment in the current data. This was not expected given the abundant amount of evidence of speech entrainment in the literature. However, methodological differences may account for the absence of entrainment in the acoustic data collected for the present study. One methodological feature of the current study was the unique nature of the spoken dialogue task. Here, participants were required to speak for a specified period of time (10 seconds) and then listen for the same amount of time, before repeating the cycle for a specified number of turns. This tightly-controlled, rapid-fire conversational turn-taking paradigm may recruit additional cognitive processes not needed for more typical conversational tasks in which conversational turns take a more organic course. Thus, the unique focus on quick, alternating speaking turns may have interfered with the mechanisms that underlies speech entrainment in typical conversations.

Another potential explanatory factor for the absence of speech entrainment in the current study was the short duration of the dialogue task. Previously, acoustic alignment during spoken dialogue has been found in research studies which elicited longer social interactions, ranging from ten minutes and up (Mason, 2013; Chartrand 1999). While, a specific length of time has not been identified as the duration required for acoustic alignment to transpire, the 15 conversational turns (5 minutes) in the current study appears to be insufficient to achieve significant acoustic entrainment, especially combined with the unique nature of the dialogue task. Future studies are warranted to investigate the time course of speech entrainment. That is, how long does it take for significant entrainment to occur between conversational partners.

A third reason that may account for why we did not observe acoustic entrainment may simply be the small number of dyads. While 20 participants is considered to be a relatively standard sample size in neural coherence studies (e.g., Duan et al., 2013; Dommer et al., 2012; Funane et al., 2012, Jiang et al., 2012;; Funane et al. 2011)—and the number we recruited for the current study—speech entrainment studies have employed much larger participant numbers to study, and provide evidence, of this behavioral alignment phenomenon. It is not uncommon, for example, for acoustic entrainment studies to have upward of 40 participants (e.g., Adank; Babel, 2010; Brennan & Chartrand, 1999). This may suggest that neural coherence is a relatively robust phenomenon and as such, a larger number of participants may not be required to reveal significant behavioral alignment. In contrast, speech entrainment may be a more precarious act and accordingly, requires a larger number of dyads to reveal significant alignment data.

Our chosen measure of acoustic alignment, speech synchrony scores, may be another compounding factor for the lack of speech entrainment observed. In the current study, a turn-by-turn basis correlation measure was employed across the entire dialogue task. Using more detailed analysis techniques may have revealed subtler changes in the data. Currently, there are multiple methods to measure and document behavioral alignment and no one measure has been identified as the gold standard. The acoustic properties chosen to identify alignment also play a significant role in the degree of entrainment. Levitan (2012), found significant acoustic alignment across mean intensity, maximum intensity, mean pitch, maximum pitch, jitter, shimmer, noise-to-harmonics ratio, and syllables per second. However, the degree of alignment found by Levitan et al. (2012) changed based on the gender ratio of the dyad. The greatest entrainment across all acoustic measures was found during male-female interactions, whereas, the same gendered dyads were only aligned across mean intensity, max intensity and syllables per second. The present study only had female-female dyads. It is evident that there are many variables that can affect the degree of alignment and multiple methods to analyzing this phenomenon. Research is needed to better understand these variables and methods.

Finally, while we did observe significant findings in our neural data, it is important to discuss some of the limitations of using NIRS as a tool to measure brain activation patterns.

Compared to fMRI, NIRS has lower spatial resolution which is approximately 3 cm (i.e., nearly equal to one gyrus) making it unable to detect deep structure. NIRS is also sensitive to non-activation blood flow, such as heart rate (Koike et al., 2015). In the present study, optodes were only placed on the frontal cortex. Language, speech and communication is a diverse neural task. Therefore, involvement of other cortical structures could not be determined. However, given these drawbacks, NIRS is still a suitable option for research regarding social communications, as it allows for the examination of activation patterns during more naturalistic conversational tasks.

Conclusion

In summary, we measured conversational alignment during a face-to-face spoken dialogue using NIRS and acoustic analysis techniques. We confirmed findings from previous studies, demonstrating increased neural coherence during spoken dialogue, however, we failed to detect the presence of speech entrainment. Accordingly, the link between neural and acoustic data could not be investigated.

References

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science*, 21, 1903–1909.

Amodio, D. M., & Frith, C. D. (2006). Meeting of Minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, 7, 268-277.

Babiloni, F., & Astolfi, L. (2014). Social neuroscience and hyperscanning techniques: Past, present and future, *Applied Neuroscience,* 44:76–93.

Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science*, 16, 814–819.3.

Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation, *Journal of Phonetics*, 40, 177–189.

Babel, M., & Bulatov, D. (2012). The role of fundamental frequency in phonetic accommodation, *Language and Speech*, 55(2), 231–248.

Benus, S. (2014). Social aspects of entrainment in spoken interaction. *Cognitive Computation*, 6(4), 802-813.

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer (Version 5.3.42) [Computer software]. Available from www.praat.org.

Borrie, S. A., & Liss, J. M. (2014). Rhythm as a coordinating device: Entrainment with disordered speech. *Journal of Speech Language Hearing Research,* 57(3), 815-824.

Borrie, S. A., & Schafer, M. C. M. (2015). The role of somatosensory information in speech perception: Imitation improves recognition of disordered speech. *Journal of Speech Language Hearing Research,* 58, 1708-1716.

Borrie, S. A., McAuliffe, M. J., Liss, J. M., O'Beirne, G. A., & Anderson, T. J. (2013). The role of linguistic and indexical information in improved recognition of dysarthric speech. *The Journal of the Acoustical Society of America*, 133, 474–482.

Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G. A., & Anderson, T. J. (2012a). Familiarization conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Language and Cognitive Processes*, 27, 1039–1055.

Borrie, S. A., McAuliffe, M. J., Liss, J. M., O'Beirne, G. A., & Anderson, T. J. (2012b). A follow-up investigation into the mechanisms that underlie improved recognition of dysarthric speech. *The Journal of the Acoustical Society of America,* 132, EL102–EL108.

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

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 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, 893–910.

Cui, X., Bryant, D. M., & Reiss, A.L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. , 59, 2430–2437.

Dommer, L., Jaeger N, Scholkmann, F., Wolf, M., & Holper, L. (2012). Between-brain coherence during joint n-back task performance: A two-person functional near-infrared spectroscopy study. *Behavior Brain Research*, 234, 212–222.

Duan, L., Liu, W. J., Dai, R,N., Li, R., Lu, C. M., Huang, Y. X., & Zhu, C. Z. (2013). Cross-brain neurofeedback: Scientific concept and experimental platform. *Plos One*, 8(5), 1-5.

Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management*. St. Louis, Mo: Elsevier Mosby.

Ferrari, M. & Quaresima, V. (2012). A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *Neuroimage*, 62(2), 921-935.

Egetemeir, J., Stenneken, P., Koehler, S., Fallgatter, A. J., & Herrmann, M. J. (2011). Exploring the neural basis of real-life joint action: Measuring brain activation during joint table setting with functional near-infrared spectroscopy. *Frontiers in Human Neuroscience*, 5(95), 1-9.

Funane, T., Kiguchi, M., Atsumori, H., Sato, H., Kubota, K., & Koizumi, H. (2011). Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near infrared spectroscopy. *Journal of Biomedical Optics,* 16(7), 1-10.

Furuyama, N., Hayashi, K., & Mishima, H. (2005). Interpersonal coordination among articulations, gesticulations, and breathing movements: A case of articulation of /a/ and flexion of the wrist. *Transcriptions of Japanese Society for Artificial Intelligence*, 20(3), 247-258.

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, 27, 181–218.

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes Geophys* 11, 561–566.

Hirata, M., Ikeda, T., Kikuchi, M., Kimura, T., Hiraishi, H., Yoshimura, Y., & Asada, M. (2014). Hyperscanning MEG for understanding mother-child cerebral interactions. *Frontiers in Human Neuroscience,* 8(118) 1-6.

Holper, L., Goldin, A., Shalom, D., Battroc, A., Wolf. M., & Sigman, M. (2013). The teaching and the learning brain: A cortical hemodynamic marker of teacher–student interactions in the Socratic dialog. *Internal Journal of Educational Research,* 59, 1–10.

Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fNIRS. *Neuroimage,* 63, 212–222.

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., & Lu, C. (2012). Neural synchronization during face-to-face communication. *Journal of Neuroscience,* 32, 16064–16069.

Koike, T., Tanabe, H., & Sadato, N. (2015). Hyperscanning neuroimaging technique to reveal the "two-in-one" system in social interactions. *Neuroscience Research,* 90, 25-32.

Lee, C., Black, M., Katsamanis, A., Lammert, A., Baucom, B., Christensen, A., . . . Narayanan, S. (2010). Quantification of prosodic entrainment in affective spontaneous spoken interactions of married couples. *Proceedings of Interspeech, Makuhari, Japan,* 793–796.

Levitan, R., Gravano, A., Wilson, L., Benus, S., Hirschberg, J., & Nenkova, A. (2012). Acoustic-prosodic entrainment and social behavior. Proceedings of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT '12), Montreal, Canada 11–19.

Levitan, R., & Hirschberg, J. (2011). Measuring acoustic-prosodic entrainment with respect to multiple levels and dimensions. *Proceedings of Interspeech, Florence, Italy,* 3081–3084.

Lindenberger, U., Li, S-C., Gruber, W., & Mueller, V. (2009). Brains swinging in concert: Cortical phase synchronization while playing guitar. *BMC Neuroscience*, 10-22.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, 36(8), 1404–1426

Lu, C. M., Zhang, B. B., Zang, Y. F., Peng, D. L., & Zhu, C. Z. (2010). Use of fNIRS to assess resting state functional connectivity. *J. Neuroscience Methods*, 186, 242-249.

Lubold, N., & Pon-Barry, H. (2014). Acoustic-prosodic entrainment and rapport in collaborative learning dialogues. *MLA 14,* 5-12.

Mason, J. H., Brynat, G., Gervais, M., & Kline, M. (2013). Convergence of speech rate in conversation predicts cooperation. *Evolution of Human Behavior*, 34, 419-426.

McFarland, D. H. (2001). Respiratory markers of conversational interaction. *Journal of Speech, Language, and Hearing Research*, 44, 128–14.

Miles, L. K., Nind, L. K., & Macrae, C. N. (2009). The rhythm of rapport: Interpersonal synchrony and social perception. *Journal of Experimental Social Psychology*, 45, 585–589.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *J. Acoustic. Soc. Am.*, 119(4), 2382-2393.

Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Minamoto, Y., & Osaka, M. (2015). How two brains make one synchronized mind in the inferior frontal cortex: fNIRS-based hyperscanning during cooperative singing. *Frontiers in Psychology*, 6(1811), 1-11.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogu. *Behavioral Brain Science*, 24(2), 169-190.

Putman, W. B., & Street, R. L., Jr. (1984). The conception and perception of noncontent speech performance: Implications for speech accommodation theory. *International Journal of Sociology of Language*, 46, 97–114.

Reitter, D., Moore, J. D., & Keller, F. (2006). Priming of syntactic rules in task-oriented dialogue and spontaneous conversation. In R. Sun & N. Miyake (Eds.), Proceedings of the 28th Annual Conference of the Cognitive Science Society (pp. 685–690).

Richardson, M. J., Marsh, K. L., & Schmit, R. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 62–79.

Saito, D N., Tanabe, H. C., Izuma, K., Hayashi, M. J., Morito, Y., Komeda, H., et al. (2010). Stay tuned: Inter-individual neural synchronization during mutual gaze and joint attention. *Frontier Integr. Neuroscie.* 4(127), 1-12.

Shockley, K., Santana, M. V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2), 326–332.

Smith, E. R. (2008). An embodied account of self-other “overlap” and its effects. In G. R. Semin & E. R. Smith (Eds.), Embodied grounding: Social, cognitive, affective and neuroscientific approaches (pp. 148–159). New York, NY: Cambridge University Press.

Stephens, G. J., Silbert, L. J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *PNAS,* 107(32), 14425-14430.

Street, R. L., Jr., & Giles, H. (1982). Speech accommodation theory: A social cognitive approach to language and speech behavior. In M. E. Roloff & C. R. Berger (Eds.), Social cognition and communication (pp. 193–226). Beverly Hills, CA: Sage.

Suda, M., Takei, Y., Aoyama, Y., Narita, K., Sato, T., Fukudo, M., & Mikuni, M. (Frontopolar activation during face-to-face conversation: An in situ study using near-infrared spectroscopy. *Neuropsychologia,* 48, 441-447.