ABSTRACT

Title of Dissertation: THE NEURAL MECHANISMS SUPPORTING STRUCTURE AND INTER-BRAIN CONNECTIVITY IN NATURAL CONVERSATION.

Nuria Yasmin AbdulSabur, Doctor of Philosophy, 2014

Dissertation directed by: Professor William Idsardi
Department of Linguistics

Conversation is the height of human communication and social interaction, yet little is known about the neural mechanisms supporting it. To date, there have been no ecologically valid neuroimaging studies of conversation, and for good reason. Until recently, imaging techniques were hindered by artifact related to speech production. Now that we can circumvent this problem, I attempt to uncover the neural correlates of multiple aspects of conversation, including coordinating speaker change, the effect of conversation type (e.g. cooperative or argumentative) on inter-brain coupling, and the relationship between this coupling and social coherence. Pairs of individuals underwent simultaneous fMRI brain scans while they engaged in a series of unscripted conversations, for a total of 40 pairs (80 individuals).

The first two studies in this dissertation lay a foundation by outlining brain regions supporting comprehension and production in both narrative and conversation – two aspects of discourse level communication. The subsequent studies focus on two unique features of conversation: alternating turns-at-talk and establishing inter-brain coherence through speech. The results show that at the moment of speaker change, both people are engaging attentional and mentalizing systems – which likely support orienting toward implicit cues signaling speaker change as well as anticipating the other person's intention to either begin or end his turn. Four networks were identified that are significantly predicted by a novel measure of social coherence; they include the posterior parietal cortex, medial prefrontal cortex, and right angular gyrus.

Taken together, the findings reveal that natural conversation relies on multiple cognitive networks besides language to coordinate or enhance social interaction.

THE NEURAL MECHANISMS SUPPORTING STRUCTURE AND INTER-BRAIN CONNECTIVITY IN NATURAL CONVERSATION

by

Nuria Yasmin AbdulSabur

Dissertation submitted to the Faculty of the Graduate School of the

University of Maryland, College Park in partial fulfillment

Of the requirements for the degree of

Doctor of Philosophy

2014

Advisory Committee:

Professor William Idsardi, Chair

Dr. Allen R. Braun

Professor Ellen Lau

Professor Luiz Pessoa

Professor Elizabeth Redcay

Dedication

To Zakaria, the world’s best conversationalist.

Acknowledgements

I would like to thank and acknowledge Yisheng Xu, Siyuan Liu, and Ho Ming Chow for their guidance and mentorship. And thanks again to them and Meghan Healey for spending so many Saturday mornings in the NMR center collecting data with me. Many thanks to Joel Kuipers, Jacqueline Hazen, Briel Kobak, and Emma Kelsey for lending their expertise. I am also grateful to the summer interns and volunteers who helped sift through, organize, or score data from this study: Samantha Perugini, Vikram Siberry, and India Bane. Finally, tremendous appreciation goes to Yusuf Abdul-Salaam and my parents for their support and encouragement.

Table of Contents

The general aim of this thesis is to use a hyperscanning fMRI paradigm to explore the neural networks supporting natural and interactive conversation, a completely novel study. An initial goal is to characterize the brain regions underlying comprehension and production and compare them to narrative, the other element of discourse level communication. Most importantly, I aim to go beyond this and examine the neural mechanisms underpinning phenomena unique to conversation, such as alternating turns-at-talk and establishing inter-brain coupling through speech.

Background

Conversation, a critical element of discourse level language, is the cornerstone of human communication. This is easily observable in our daily lives, as we engage in possibly dozens of conversations face-to-face, on the phone, and even via text (such as text messaging). Cognitive scientists propose that human culture is built upon the ability to identify with other humans (Tomasello, 1999). After developing a sense of his intentionality, a child begins to recognize intentionality in others (Tomasello, 1999). This is facilitated through social interactions. While not all social interactions are verbal, conversation is undoubtedly a crucial component. Language and, more specifically, conversation require the coordination of both meaning and understanding and are both social and cognitive in nature (Clark, 1996).

Additionally, many conversational features appear to be universal, leading many to believe in a biological basis to this complex behavior (Stivers et al., 2009). Because of its importance and ubiquity, any ecologically valid and complete neuro-cognitive model of language needs to take conversation into account. Garrod and

Pickering (2004) argue that conversation takes advantage of the innate propensity towards "interactive alignment", the process through which interlocutors synchronize their mental representations, which may not be specific to language. They posit that interactive alignment addresses the mirroring phenomenon that has been observed in conversation (and other behaviors), which is characterized by a gradual adoption of another speaker's phrases, intonation, vocal intensity, and posture among other features (Garrod and Anderson, 1987; Giles et al., 1991; LaFrance, 1985). Another way of framing this desire for social alignment is the consideration that language evolved from early primate gesture and was spurred on by collaborative activities that result from communal living arrangements (Corballis, 2003; Tomasello, 2008).

Involvement of Extralinguistic Cognitive and Psychological Functions

It is given that discourse necessarily entails multiple levels of language processing, e.g. phonological, lexical, morphological, syntactic, etc. However, at the discourse level, and especially in conversation, paralinguistic and extralinguistic features can enhance or completely alter the construal of spoken language (Bryant and Fox Tree, 2002; Kelly et al., 2010; Weber et al., 2006). Paralinguistic attributes of discourse include intonation, stress (emphasis), and volume, cues that are carried on the same signal of language but are not in and of themselves linguistic. Extralinguistic features can include eye-gaze, body positioning, facial expressions, and gesture – features beyond language and carried on a separate modality.

Presumably, one must recruit cognitive systems other than language, such as working memory, attention, and response inhibition, to maintain and organize the many aspects of conversation. For example, one must attend to and interpret the other person’s speech, incorporate para- and extralinguistic information, formulate a response, and hold that response for an appropriate time. However, it remains unclear how these systems interact.

Some suggest the primary goal of language is to influence the attention of others (Tomasello, 1999). Eye-gaze, a common indicator of attention, is also important in initiating and directing joint attention (JA). In children, joint attention supports word learning and the development of communication and social skills (Carpenter et al., 1998; Tomasello and Farrar, 1986). In adults, it is one of the major elements in social interaction. There are many examples of people engaged in joint attention involving external and concrete targets, for example pointing to direct one’s attention to an object. Other examples of joint attention can include focusing on a shared task or performance. Joint attention is also critical in conversation. Indeed without it, it would be impossible to coordinate behavior. But JA in conversation, which necessarily calls for drawing others’ attention to the actions, objects, or ideas one is trying to convey, may include both external concrete objects or stimuli and internal and/or abstract targets. In the latter type of dyadic joint attention, the target of gaze often becomes the other person and his facial expressions, gestures, etc. It should then come as no surprise that eye-gaze, again as a metric of JA, is also critical in coordinating conversational behaviors, such as turn-taking (Wiemann and Knapp, 1975). Additionally, in natural conversation, joint attention can be ascertained in other ways, such as back channel responses from listener (Fries, 1952; Yngve, 1970).

Conversation, like other social interactions, also requires mentalizing, i.e. understanding the beliefs, feelings, desires, and intentions of oneself and others. For a speaker to successfully communicate his thoughts he must make assumptions about the listener's knowledge and opinions of the world, thereby essentially inferring the other's perspective (Tomasello, 1999) while firmly grounded in his own. Listeners, on the other hand, assume the speaker's intentions and infer his meaning. These inferences (for both speaker and listener) are most apparent when negotiating implicit statements, which include (but are certainly not limited to) the use of metaphor and irony. Very few statements are perfectly explicit. Instead, most draw upon (to varying degrees) the context of the conversation and above-mentioned assumptions (Grice, 1975). Consider the following exchange: "Will you be the concert on Saturday?", "I have to babysit". Although it is assumed the answer to the questions is 'I cannot go to the concert because I will be babysitting elsewhere at that time and cannot possibly be in both locations at the same time', the second speaker does not need to say that. Producing and successfully interpreting such implicit statements are ecologically advantageous in that they expedite communication of complex concepts, but they are also undoubtedly cognitively and computationally weighty. Beyond implicit statements, mentalizing is certainly essential to assessing how the other person receives one's message. A speaker is constantly evaluating whether the other person is listening, whether he understands, and predicting how he might respond. It is this evaluation that allows one to deliver a message appropriate to his audience.

Clinical Relevance of Conversation

For most, conversations are performed with little effort. However, some neurological conditions impact one's ability to perform or understand discourse. Traumatic brain injury (TBI) has been linked to impairments in both narrative

(Coelho et al., 2005; Tucker and Hanlon, 1998) and conversation (McDonald and Flanagan, 2004; Snow et al., 1998). Considering lexical, semantic, and syntactic elements of language use are often intact, discourse impairments (particularly those pertaining to conversation) may be related to difficulties in social communication (Dahlberg et al., 2006; McDonald and Flanagan, 2004), specifically the ability to mentalize, recognize emotion in others, and follow cultural norms within conversation (Turkstra et al., 2001).

Patients with right hemisphere brain damage (RHD), i.e., damage to right cortical tissue resulting from various etiologies, exhibit a range of conversational impairments. RHD is associated with difficulty with pragmatic elements of language, and particularly conversation. They are less likely to use facial expressions (Blonder et al., 1993) and more likely to stray off topic (Lehman Blake, 2006). Those with RHD also tend to have difficulty interpreting non-literal text (Kaplan et al., 1990). This condition has also been linked to relatively poor performance on mentalizing (or theory-of-mind, TOM) tasks (Siegal et al., 1996; Winner et al., 1998). Although some suggest poor TOM performance may relate to impaired understanding of verbal presentation of the task (Surian and Siegal, 2001; Tompkins et al., 2008a).

Similarly, those with autism spectrum disorder (ASD) suffer from impaired social communication in addition to cognitive impairments. Even people with high functioning autism (who may have average or above-average intelligence and language skills) have difficulty with social interactions. Atypical conversational behaviors in this group include failure to respond to questions, offering fewer contributions and less sharing of personal experience, as compared to typically developing peers (Capps et al., 1998). This can lead to severe psychosocial impairments, such as an inability to establish and maintain relationships (Whitehouse et al., 2009).

Each of these conditions, in which lower-level language skills are relatively preserved while conversational skills are severely impaired, underscores the importance of extralinguistic cognition in successful naturalistic communication.

Review of Neuroimaging Studies of Conversation

To date, there have been very few neuroimaging studies involving conversation at all. This is not altogether without reason. Conversation presents unique technological and methodological challenges. The only way to actually assess conversation is with a naturalistic design. Yet, conversation, like all discourse level communication, is complex and difficult to control. Turns-at-talk are extemporaneous and of variable length, which contributes to this difficulty. Moreover, natural communication entails continuous, overt speech (i.e. longer than one or two seconds). Although functional magnetic resonance imaging (fMRI) provides the best balance of temporal and spatial resolution and a low level of invasiveness, continuous speech creates dramatic noise in fMRI data. Head motion and consequently motion-related artifact are obviously exacerbated by overt speech production. However, the most significant hurdle is susceptibility artifact, the changes in magnetic field resulting from boundaries between tissues with varying magnetic susceptibility. Susceptibility artifact is significantly aggravated by continuous speech which necessitates movement of air-tissue boundaries around the mouth, jaw, and tongue (Birn et al., 1998). Nevertheless, a few researchers have attempted to shed light on the neural correlates of conversation using fMRI and other neuroimaging techniques.

Caplan and Dapretto (2001) asked subjects to listen to conversations that contained implicit topic shifts while undergoing fMRI brain scans. The authors experimentally manipulated whether a final sentence in the conversation logically followed the context, as established by the rest of the conversation. The participants were asked to make explicit judgments about the whether the last sentence was contextually congruent with the rest of the narrative. However, the topic shifts remained implicit. They found that implicit shifts in topics elicited activation in the right hemisphere homologues of language areas, while explicit judgments of context involved left-lateralized areas. While this study is informative, a potential limitation is that the subjects only listened to the conversations of others, essentially simulating eavesdropping or overhearing. One cannot assume that such activation patterns will remain during participation in a conversation, which is certainly more engaging and likely also more cognitively demanding. Additionally, the authors miss the opportunity to delve into features that are unique to conversation, such as exchanging turns at talk.

In a more recent study, Suda et al. (2010) used near-infrared spectroscopy (NIRS) to scan subjects while they spoke with an interviewer in 15-second segments. As a control task, participants repeated consonant-vowel (CV) syllables. The authors found that inferior frontal and superior temporal channels were more active during the conversation condition. While this study takes a step in the right direction by allowing face-to-face interaction, its scope is limited by both technical restrictions and study design. The authors' array of NIRS sensors did not cover the entire head and provided a very limited view of the brain activity. Moreover, the lack of high-resolution imaging makes it impossible to know specifically which parts of the frontal and temporal cortices were more engaged by the experimental task. An additional concern is that the authors used CV as a baseline task, although CV has no linguistic content and can only control for the motor and acoustic properties of conversation, at best. Such a study design makes it impossible to demonstrate whether the observed changes in inferior frontal and superior temporal brain regions are due to emergent features of conversation itself or another level of language processing. Further, separating the conversations into 15-second intervals is certainly a departure from the natural flow of conversation.

The same research group conducted two follow-up studies, one comparing schizophrenic patients and normal controls (Takei et al., 2013) and the other in typically developed adults who were assessed with the Autism Spectrum Quotient (AQ, Suda et al., 2011). They found that as AQ scores increased, patients demonstrated decreased activation of channels over the left superior temporal sulcus. Those with schizophrenia, on the other hand, exhibited decreased activation of channels over the right inferior frontal gyrus and bilateral temporal lobes. However, they repeated the study design and again imposed 15-second alternating turns, which means that these studies suffer from the same departures from natural conversation.

Stephens et al. (2010) used fMRI to look at the brains of speaker and listeners. The authors designed a study in which, using fMRI, they scanned on person narrating in a natural way events from her life. They then played back the narrative to a cohort while they individually underwent fMRI brain scans. The authors outline brain regions that demonstrated coupling between speaker and listeners, such as the left inferior frontal gyrus, medial prefrontal cortex, superior temporal gyrus, and precuneus. Their approach is important, in that, as a tool, inter-subject coupling could potentially reveal key aspects of conversation. However, their design has not captured conversation itself. Instead, it is a study on narrative comprehension and production, and one can only hypothesize in how their results may apply to natural conversation.

Overall, these studies point to diffuse regions in the frontal, temporal, and medial parietal cortices. While these findings contribute to our understanding of the neural correlates of communication, methodological limitations and a lack of ecological validity have made it difficult to apply these results to natural conversation.

To date, no study has comprehensively characterized the neural correlates of any of the numerous aspects of conversation. Yet, these neural substrates are important to uncover for several reasons. First is a purely scientific approach, as such as study would shed light on the biological basis of a critical feature of human behavior. Uncovering these neural correlates can lead to a better understanding of the cognitive processes that subserve conversation, such as memory, language, and social cognition. It is a first step toward exploring how such cognitive functions and associated brain networks interact. Moreover, once a model of brain activity during conversation is established in healthy individuals, we may be able to make more informed hypotheses about the underlying mechanisms affecting patient groups and how best to treat or mitigate their conditions.

A Glance at Imaging Studies of Joint Attention and TOM

Studies on joint attention and mentalizing can shed some light on the neural mechanisms supporting conversation. As discussed above, conversation relies heavily on these processes and may draw upon similar brain structures. Fortunately, imaging research on mentalizing and JA have been much more fruitful than studies of conversation itself.

Bristow et al. (2007) compared brain activation related to gaze shifts in both social and non-social contexts, i.e., when the person in the video appears to either look directly at the participant (social) or elsewhere (non-social) and then shifts his or her gaze to a target (correct) or another location (incorrect). They found a main effect of gaze shift in the posterior superior temporal sulcus and left middle frontal gyrus. Interestingly, they also found that correct, social eye gaze recruited the medial prefrontal cortex and precuneus, when compared to incorrect, non-social eye gaze respectively. In contrast, perceiving non-social and incorrect gaze shift was associated with increased activation of a fronto-parietal attention network and posterior superior temporal sulcus. The findings suggest all these brain regions may be related to shifting gaze or, more likely, shifting joint attention as a function of eye gaze. However, this study also suggests the medial prefrontal cortex and precuneus hold specialized roles in social interactions where the participant is directly involved.

Rather than contrasting social and nonsocial shifts in eye gaze, Schilbach et al. (2010) scanned participants while following or leading the gaze of another person, similar to initiating and responding to joint attention. They found that following gaze recruited the anterior medial prefrontal cortex, while directing gaze engaged the ventral striatum. This study is one of many to find differences depending on the role one plays in joint attention, follower or leader. This has important implications for dyadic interaction or any relationships where one may be likely to lead (e.g. parent/child). Interestingly, the researchers found that following another's gaze recruited the medial prefrontal cortex (MPFC), while directing gaze was limited to subcortical structures. The MPFC is consistently linked with social interaction and mentalizing (Amodio and Frith, 2006; Spreng et al., 2009), and one might expect its involvement in directing another's behavior.

In fact, in their study of initiating and responding to joint attention Redcay et al. (2012) found that both engage dorsomedial prefrontal cortex, as well as the right poster superior temporal sulcus (pSTS). Interestingly, they also demonstrate a role for the intraparietal sulcus and middle frontal gyrus, regions associated with other types of attention (Pessoa et al., 2009). Also, the authors find that responding to bids for joint attention recruited ventral medial prefrontal cortex. Involvement of the pSTS is also critical because this region shows a preference for human stimuli, both auditory and visual, demonstrating sensitivity for human interaction (Belin et al., 2000; Grossman et al., 2000). Another study (Laube et al., 2011) examined participants only responding to bids for joint attention (through eye gaze and head movement) and found increased activation of the right poster superior temporal sulcus, as Redcay et al. also did. However, in this study the fusiform gyrus is also implicated.

Clearly, there is not consensus between studies on joint attention, which may be in part to differing task of joint attention. Still, consistent patterns are emerging that might be informative for natural conversation. Together, these studies link the medial prefrontal cortex and posterior superior temporal sulcus (likely right-lateralized) to either responding to or initiating joint attention (or both).

Thus far, the majority of studies of joint attention have focused on eye gaze, rather than spoken language, perhaps in part because of technical limitations on imaging connected speech explained earlier, but now we can. Examining the neural correlates of joint attention in diverse ecologically valid social settings can lead to a deeper understanding of the brain regions supporting joint attention and how they may be influenced by changes in modality (e.g. visual or auditory) or target (e.g. concrete or abstract targets).

Research on theory-of-mind has also been very productive and should be helpful in forming expectations of networks supporting conversation.

Saxe and Kanwisher (2003) show that the temporoparietal junction (TPJ, which anatomically overlaps with portions of the posterior superior temporal, angular and supramarginal gyri) subserves building representations of others' minds. They demonstrate that the TPJ was significantly engaged by stories detailing the mental states of others, rather than their physical characteristics. They also show that the TPJ did not respond to stories lacking social interaction, indicating specificity to social stimuli. In a related study, the authors provided evidence that the TPJ and posterior cingulate cortex are recruited when reading about the thoughts of others, but not when reading about other "socially relevant information", such as one's appearance (Saxe and Powell, 2006). They also found that the medial prefrontal cortex was engaged in all stories, regardless of content, suggesting it plays a more general role in social cognition. Taken together, these two studies demonstrate that, at least when as it relates to reading stories, activation of the TPJ discriminates between content pertaining to others' mental states and other types of information.

In an older study, with a similar design, the participants underwent PET scanning while listening to stimuli from three story conditions: stories requiring mentalizing, stories dominated by physical information, and unlinked sentences (Fletcher et al., 1995b). The authors found that both types of stories (mentalizing and physical) recruited the posterior cingulate cortex, bilateral temporal poles, and the left STG. However, the middle frontal gyrus (BA8) and posterior cingulate cortex were engaged in mentalizing stories (which required the attributing mental states to the characters) as compared to physical stories. Unlike the previous studies, the TPJ is not implicated here. However, they all find increased activation of the posterior cingulate cortex,

Overall, this body of research implicates the TPJ, posterior cingulate cortex, medial prefrontal cortex, and posterior superior temporal sulcus support modeling mental states of others – an essential part of conversation. Rather than face-to-face interaction, many studies still entail passive interpretation of others' social interactions, through stories or false-belief tasks. Social interaction, in which the person is directly engaged, should also engage these regions, but it is important to extend our understanding of the role of this network in other settings, particularly those more closely mirroring daily experiences.

Outline of this Dissertation

I have set out to fill in the holes in the current literature by exploring the neural correlates of the quintessential social act: conversation. In order to do so, I conducted an fMRI study in which pairs of people were scanned simultaneously in separate MRI scanners, while engaged in a series of conversations. While they were naturalistic, unscripted, unrehearsed conversations, they were designed such that the contents of each conversation would differ. During one conversation, participants worked together to develop a detailed solution to a hypothetical problem. In another, they shared their experiences from their lives. In a third type of conversation, they informally debated issues surrounding immigration policies. Participants also engaged in a control task in which each person reported on unrelated factual topics and interrupted one another to insert turns-at-talk.

This study is the first of its kind, no other study have been able to clearly outline the neural mechanisms supporting conversation. Due to technical innovations, we can now undertake this endeavor. Although there is no neuroimaging study for close comparison, based on what is known behaviorally about conversation and imaging studies of related tasks, some predictions can be made.

I expect mentalizing, joint attention, and of course language production and comprehension to be engaged throughout conversation. However, I also predict that the involvement of all of these cognitive processes (either in activation or functional connectivity) can be modulated by several behavioral, cognitive, and psychosocial factors.

For example, based on literature on conversation analysis (discussed in more detail in Study 2), I believe that the cognitive processes underlying coordinating turn transitions are inherently different from those that support producing or listening to a turn-at-talk, likely involving regions related to joint attention.

Additionally, behavioral scientists have shown that the context and content of a conversation can alter particular behavioral patterns (also presented in more detail in Study 2), captured in the conversation types we use. Similarly, functional imaging studies of narrative have proven that content can modulate neural correlates of narrative comprehension (Chow et al., 2013; Saxe and Powell, 2006). I predict that different types of conversations can influence the neural correlates of conversation. Depending on the goal, content, and context, some may rely more heavily on distinct cognitive skills, and that this difference will be reflected in the both conversation behavior and the underlying neural mechanisms. Specifically, I predict that conversation type (and relatedly, content) will influence the manner and degree to which participants build up mutual understanding and common ground.

Importantly, inter-brain coupling has been demonstrated in other discourse-level tasks (Stephens et al., 2010), and I expect that particular regions or networks will demonstrate brain-to-brain coupling in conversation. Moreover, conversational coherence (the establishment and maintenance of common ground) will influence inter-brain synchronicity, likely in a positive relationship.

Further, conversation shares some features with narrative, the other element of discourse level language. And I expect there to be neuroanatomical overlap between the brain regions supporting these two high-order language tasks. However, to accurately assess this there needs to be a comprehensive description of the brain regions related to both narrative production and comprehension.

My aim is to systematically test these predictions. I will begin this dissertation by outlining the neural correlates of discourse level comprehension and production in a study that examines narrative production and comprehension in the same group of subjects. In Study 2, I examine the brain regions supporting both comprehension and production in conversation, and I highlight similarities and differences between narrative and conversation. The third and fourth studies delve into two features unique to conversation, i.e., cannot be studied at any other level of language. The first is turn-transitions, the act of alternating speakers that is one of the more essential hallmarks of conversation. The second is inter-brain coupling during spoken language, particularly as it relates to conversational coherence. In the latter study, I will also explore how inter-brain coupling is influenced by conversation type and the psychosocial factor of personality.

Study 1: Neural Correlates of Narrative Comprehension and Production: a combined fMRI and PET study

Why Start With Narrative?

Discourse level communication includes both narratives and conversation. Most imaging studies have discourse have involved only narratives and expository texts and focused on narrative comprehension for many reasons that were outlined in the previous sections, like problems with susceptibility artifacts and controlling stimuli. Still, we know that narrative processing shares features with conversation. Both are language tasks that involve (perhaps to differing degrees) extralinguistic cognitive processes. Moreover, both narrative and conversation are complex, requiring multiple levels of linguistic processing in parallel. Also, in most typical settings, both are unrehearsed and generated on the spot. Even in the case of retelling stories or sharing information, seldom are these accounts memorized and related verbatim. Importantly, both require establishing and constantly updating in real-time coherence or connectedness (Schiffrin, 1987), a feature that differentiates narratives from a random collection of events and conversation from scattered statements or turns-at-talk. A comprehensive outline of the neural correlates of narrative processing may be useful in making predictions for conversation. Also, like all communication, both narrative and conversation require a speaker and listener.

However, some differences also exist. One essential difference is that during narrative, one person is the speaker, the other is the listener, and these roles do not changes. Clearly, this is not the case in conversation, where alternating speakers is one of the defining characteristics. Relatedly, conversation is truly cooperative in that both (or all) parties must participate in establishing shared understanding and moving the conversation forward. As a result, it is reasonable to expect some brain regions to support both tasks while others might be unique to one or the other.

Introduction

The two sides of natural language – language as it is heard and as it is spoken – together constitute a cornerstone of human culture. However, the scientific investigation of language comprehension and production, for purposes of simplicity and experimental control, has most often been confined to the level of sentences and words. This is a significant oversight because, during the vast majority of real-world interactions, language is used at the level of discourse (which includes both narratives and conversations). This is the context in which the pragmatic properties of language naturally emerge: language, as it is produced and understood at this level, is characterized by distinctive features that are not manifest in words or sentences alone, and likely relies upon complex interactions between the language system and other cognitive domains. For example, discourse processing involves the construction of a situation model (or mental representation) of the narrative by drawing upon one's experience, memory, and world knowledge, representing both the narrative macro- and microstructure (van Dijk and Kintsch, 1983).

There are a number of ways in which language operates at the discourse level – it is, for example, used to communicate plans, to instruct, persuade or convey other sorts of expository detail. Stories are critical elements of human society. They are frequently the means by which people learn about the world (for example, engaging cognitive processes that facilitate understanding social interactions) and the context in which they make sense of it (that is, by organizing and imposing an event structure).

Stories are also clinically relevant: production and comprehension of normed, well-controlled narratives has proven valuable in testing patients with a variety of disorders that affect communication (Barnes and Baron-Cohen, 2012; Coelho et al., 2012; Crinion et al., 2006; Davis and Coelho, 2004; Norbury and Bishop, 2003; Spalletta et al., 2010; Thompson et al., 2012). Stories are useful in this context because the symptoms of many of these disorders only emerge at this level; comprehension or production deficits that are not apparent when patients process words or sentences in isolation are clearly manifest when they process narrative. Understanding the brain mechanisms responsible for this may lead to a fuller understanding of the pathophysiology of these disorders (why, for example, some disorders have a selective impact on production or comprehension while others typically affect both) and the prospects for their treatment.

How then does the brain organize the elements of a story in order to tell it, and deconstruct these when a story is heard? Intuitively, it seems clear that there must be both similarities and differences between comprehension and production in this context. Both engage language mechanisms as well as cognitive faculties that must interact with language to support discourse level understanding. Unfortunately, the relationships between language comprehension and production in this important context, the features shared by them and the ways in which they differ, remain poorly characterized. Understanding the neural circuits that support them – specifically the degree to which these systems overlap or remain anatomically discrete – should be integral to any model of natural language use.

These processes are likely to be engaged in different ways during storytelling (when a narrative is formulated and produced), and story comprehension (during which subjects process and decode incoming narrative information). These issues will be addressed in the present study.

Historically, the earliest theories describing the anatomical and functional correlates of language comprehension and production were precise but oversimplified. The earliest neurological models attributed speech production to anterior and comprehension to posterior perisylvian brain areas of the left hemisphere (Geschwind, 1972). It is now known that the anatomical foundations of speech comprehension and production and their interrelation are more complex (Hickok and Poeppel, 2007; Indefrey, 2011). For example, clinical studies have made it clear that lesions to anterior perisylvian areas can lead to comprehension deficits, while lesions to posterior perisylvian areas can result in deficits in speech production (Blumstein et al., 1977; Dronkers et al., 2004). Moreover, it became clear that language processing is not strictly lateralized to the left and that the right hemisphere may play a greater role than originally proposed, particularly at higher levels of language performance (Marini, 2012; Tompkins et al., 2008b; Vigneau et al., 2011).

The growing accessibility of neuroimaging methods has expedited attempts to clarify the distinctive features of language comprehension and production. The use of these methods, including both functional magnetic resonance imaging (fMRI) and positron emission tomography (PET), revealed both similarities (Okada and Hickok, 2006; Papathanassiou et al., 2000) and differences (Indefrey et al., 2004; Wise et al., 1991), between production and comprehension. But, as pointed out earlier, these studies, for the most part, remained confined to the level of sentences and words.

Neuroimaging studies that have evaluated discourse comprehension or production demonstrate that language processing at this level engages an array of brain regions that extend beyond the left perisylvian areas typically associated with the processing of words and sentences (Mar, 2004; Martin-Loeches et al., 2008; Mazoyer et al., 1993; Xu et al., 2005), reinforcing the idea that narrative discourse is characterized by emergent properties involving the interaction of language with other cognitive domains. These studies however, have almost always examined discourse comprehension or production in isolation. Further, they have used different tasks and, in doing so, have made comparisons between comprehension and production difficult to interpret (Blank et al., 2002; Braun et al., 2001; Martin-Loeches et al., 2008; Troiani et al., 2008; Xu et al., 2005). Examining them both in a single experiment – i.e., using the same comprehension and production tasks and a controlled, within-subjects design that permits direct comparisons between comprehension and production in the same subjects – would remove potential confounds presented by differences between subject populations, equipment, or study design.

Thus far, a single study has directly compared discourse level production and comprehension using a within-subjects design (Awad et al., 2007). This study used positron emission tomography (PET) however, which is subject to a number of technical limitations. In addition to relatively poor temporal and spatial resolution, the dose restrictions that accompany the use of radionuclides limit the number of data points that can be collected in an experiment, leading to decreased statistical power. These characteristics make fMRI a potentially superior method for comparing narrative comprehension and production. Another important advantage of BOLD (blood oxygenation level dependent) fMRI over PET is that the absence of dose-limitations and the ability to collect many more data points over shorter periods of time makes it possible to employ statistically robust functional network connectivity methods. The use of these methods should be of particular advantage in studying discourse: the integration of language and other cognitive systems that emerges at this level would clearly benefit from the capacity to investigate large scale network interactions, complementing and extending the information that is available through the use of conventional GLM methods.

But there are special problems with the use of BOLD fMRI in imaging speech production - specifically the production of overt continuous speech, which is subject to the generation of susceptibility artifacts that have proven very difficult to circumvent or correct (Barch et al., 1999; Birn et al., 1998; Kemeny et al., 2005). PET, however, is impervious to these artifacts and until now has remained the gold standard in imaging continuous speech production. Here we use an innovative denoising method for processing continuous speech data that has been shown to effectively remove susceptibility as well as other physiological artifacts without requiring changes to task design that would compromise naturalistic speech (Liu et al., 2012).

Using fMRI and conventional within-subject boxcar design we compare brain activity during overt storytelling and story comprehension with low-level baseline (recitation and listening) tasks. We also collected PET data using an equivalent experimental protocol to validate this method and provide converging evidence that should strengthen the reliability of our findings. Capitalizing on the ability to use BOLD fMRI under these circumstances, we use spatial independent component analysis to assess functional network connectivity (Allen et al., 2011; Doucet et al., 2011).

Together GLM contrast and functional connectivity methods provide complementary types of information that make it possible to better characterize cognitive processes of interest. In this study, we combine these methods to address a number of unresolved questions related to discourse level comprehension and production: Of the neural patterns that unambiguously differentiate the processing of spoken narrative from lower levels of language use, which are seen for both comprehension and production and which are unique for either process? Which of these features are found in classical perisylvian language or language-related areas? Which are detected in extrasylvian areas, particularly those that play a role in higher cognitive functions (inference making, mentalizing, situation modeling)? How do these extrasylvian areas interact with areas known to be involved in language processing?

Methods

Participants

Eighteen healthy volunteers (7 males, 11 females; aged 20-32 years) participated in this study. All participants were right-handed, native speakers of American English and were free of neurological or psychiatric illnesses. All eighteen participants were scanned in an fMRI experiment and seventeen of them participated in the PET experiment. Written informed consent was obtained for all participants under a protocol approved by the National Institutes of Health CNS Institutional Review Board (NIH 92-DC-0178).

Task Paradigm

Each subject performed speech production and comprehension tasks during two conditions: narrative stories (NA) and nursery rhymes (NR). The NA stimuli were twelve pre-trained stories, each depicted in a series of three picture cards taken from a standardized stimulus set (Helm-Estabrooks and Nicholas, PRO-ED Inc., 2003), each card corresponding to the beginning, middle, and end of the story. The NR stimuli were memorized American nursery rhymes (e.g., Mary Had a Little Lamb) in traditional verses that all participants had been exposed to earlier in life. In each production task, the subject was required to retell a story or repeat a nursery rhyme at a natural and relatively constant speaking rate and volume and with prosodic intonation commensurate with narrative production. In the comprehension tasks, the subject was instructed to attend to pre-recorded auditory stimuli from the same set of stories and nursery rhymes. The speaker for the recorded stimuli was an adult, male native-speaker of American English.

Training Paradigm

Participants were exposed to all twelve of the narratives during training. A training paradigm was presented on a Dell desktop computer. Participants saw the narrative title (e.g. “The Softhearted Lobsterman”), followed by digitized picture cards (e.g., Figure 1B) for each story, presented one at a time. Underneath each picture was text that corresponded to the depicted events. The experimenter read this text out loud once and advanced to the next picture until all three story cards were read. At that point, the experimenter again presented each card for the same story. At this time, the participant was asked to retell the story in his own words, using only the picture cues. Once the participant was able to retell the entire narrative without prompts from the experimenter and without missing essential components (e.g. main characters and events, changes in scene or time, etc.), he advanced to the next narrative. Training was deemed complete when the participant could retell all twelve narratives when provided only the title. Training most often took place within 48 hours of scanning. Immediately before collecting scans (i.e., less than 30 minutes), participants' ability to remember narratives was again assessed with standardized questions.

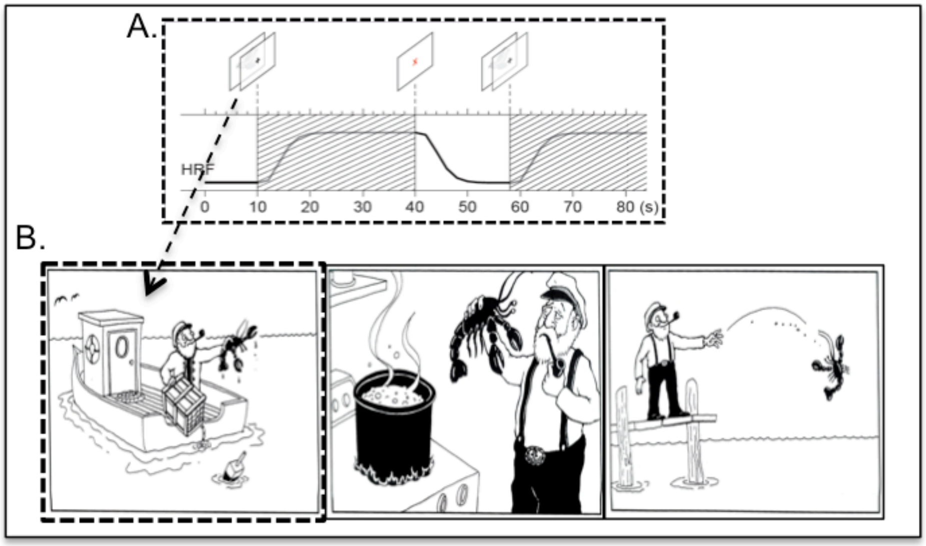


Figure 1. Task Design

A. Timing for fMRI task. The shaded area represents when a task (either experimental or baseline) was being performed. The clear area represents periods of rest. B. Picture panels representing the beginning, middle, and end of narratives. These panels were only used during training. The first panel was used a visual cue during the experiment, as indicated by the dotted arrow.

Experimental Design

fMRI Experiment

BOLD fMRI scans (Figure 1) consisted of 30s blocks for each of four task conditions: Narrative comprehension (NAc), Narrative production (NAp), Nursery Rhyme comprehension (NRc), and Nursery Rhyme production (NRp). Each block was cued by a picture for one-second duration, which was either the first story card for the NA conditions or a line drawing for the NR conditions; this was followed by a one-second written instruction containing the task condition and a title of the NA story or NR verse for the subsequent block. A white fixation cross ("+") was presented during each task block; and a red fixation "x" signaled the end of the task block and remained on display throughout the 16 second rest period between blocks. The order of tasks was randomized within each run. All four conditions were presented six times, with a different story or nursery rhyme in each instance. Of the total twelve narratives, six were randomly selected and used for production tasks. The remaining stories were used during narrative comprehension tasks. As a result, during the experiment, participants never listened to and produced the same narrative.

An experimenter transcribed each narrative and nursery rhyme to confirm compliance to the task, which was assessed by one's ability to correctly recall and retell (in his own words) the macrostructure of the narrative, particularly the characters, location, introduction of conflict, and resolution.

PET Experiment

In the PET experiment, three scans were acquired for each of the four conditions. Three additional scans were acquired during resting fixation. The order of scans was counterbalanced across subjects.

An instruction slide indicating the task and condition of each upcoming scan was presented prior to injection of radioisotope. Task cues consisted of the same pictures and line drawings used in the fMRI task, along with the title of the task, and remained on screen for two seconds. During each scan, a white fixation cross (“+”) was displayed for 60 seconds until an instruction screen cued the subject to stop. Scans were automatically initiated by detection of radiotracer in the brain and continued for 60 seconds. The delay period between injection and scan onset was calibrated every seven scans to estimate the vein-to-brain time (delay between injection and scan onset). These values were used to adjust the task cue, which was presented on average 8 seconds prior to the onset of scans.

Presentation and Recording Devices

All stimuli (for both fMRI and PET) were presented using *E-Prime* software 1.2 (Psychology Software Tools, 2002). The visual cues were projected to the subject by a mirror reflection system using a DLP projector in the fMRI study. During PET, images were displayed on a computer monitor placed in the center of the participants’ field of view. For fMRI, the auditory stimuli in the comprehension tasks were delivered to the subject through a pair of Silent Scan™ 3100 pneumatic headphones (Avotec, Stuart, FL, USA). For the PET portion of the study, auditory stimuli were presented through free-field speakers positioned near the scanner gantry. The subjects’ speech was recorded by a FOMRI™ II noise canceling optical microphone (Optoacoustics, Or Yehuda, Israel) in the fMRI scans. During PET scans, the subjects’ voice was recorded using an SM11 dynamic lavalier microphone (Shure, Niles, IL, USA). Pitch variation (standard deviation of pitch, used as an index of prosody) was calculated for the nursery rhyme and narrative tasks. No significant differences between tasks were detected for either comprehension (two-sample t-test, p = 0.399) or production (two-sample t-test, p = 0.603).

Data Acquisition

MRI/fMRI Image Acquisition

T2-weighted BOLD images were acquired on a General Electric (GE) Signa HDxt 3.0 Tesla scanner (GE Healthcare, Waukesha, WI, USA) with an 8-channel HR Brain Coil. A single-shot gradient-echo echo-planar (EPI) sequence with ASSET parallel imaging was used. The detailed scanning parameters were as follows: TR = 2000 ms, TE = 30 ms, flip-angle = 90º; 64×64 matrix, FOV = 227 mm, ASSET factor = 2. Whole brain coverage was achieved using 40 interleaved sagittal slices with a thickness of 4 mm. In addition to the functional data, T1-weighted high-resolution structural images were acquired sagittally using a magnetization-prepared rapid gradient-echo (MPRAGE) sequence. Sagittal acquisition facilitated offsetting head motion produced during speech tasks, which is more likely to be in a pitch plane.

PET Image Acquisition

PET images (scans) were acquired for each subject on a GE Advance scanner. The axial FOV (153 mm), including 35 slices separated by 4.25 mm, covered the whole brain with a reconstructed resolution of 6.5 mm (FWHM) in x-, y- and z-axes. In order to correct for attenuation, a transmission scan was performed. For each of the scans, 10 mCi of H215O was injected intravenously. Injections and scans were separated by five-minute intervals.

Data Analysis

fMRI Data Preprocessing

Head motion in the fMRI time series was corrected by both in-plane and rigid-body image registration algorithms in AFNI (Cox, 1996). The former was applied before the slice-time correction and the latter after. In addition, the structural image was co-registered to the fMRI data using a mutual-information based algorithm (Studholme et al., 1999). The segmentation and normalization of the structural image was then computed in a unified framework based on the tissue probability maps provided by SPM5 (Ashburner and Friston, 2005). The segmented grey matter and white matter tissue class images in their native (un-normalized) space were added up, thresholded, and then resampled to the fMRI data grid to create a brain mask for denoising and further analysis.

In order to remove the imaging artifacts generated by continuous overt speech production (Kemeny et al., 2005), a denoising procedure based on spatial independent component analysis (sICA) was applied on the motion-corrected fMRI time series after temporal concatenation. This procedure was implemented using the infomax ICA algorithm (Bell and Sejnowski, 1995) provided by the GIFT (Group ICA fMRI Toolbox) software (Calhoun et al., 2001). Prior to sICA decompositions, dimensionality estimation and additional preprocessing steps were performed on the input fMRI time series. The order of source dimensionality (i.e., the number of independent components) was estimated by the minimum description length (MDL) criterion (Li et al., 2007). The preprocessing steps, including centering, whitening and dimensionality reduction, were applied to reduce the complexity of ICA (Hyvarinen and Oja, 2000). Whitening and dimensionality reduction were achieved with principal component analysis (PCA).

The resulting component maps from sICA were used for signal-noise classification by five human raters. A set of concrete operational criteria based on the spatial characteristics of component maps was provided for classification. The criteria determined whether a component belonged to the "noise" category and included: 1) low degree of spatial clustering; 2) major clusters fall outside the brain; 3) major clusters surround the edge of the brain; 4) high degree of neighborhood connectedness between major positive and negative clusters; and 5) high degree of slice-wise variation. Inter-rater reliability was assessed by Fleiss' kappa statistics (Fleiss, 1971), which yielded a value (kappa = 0.9696) indicating perfect agreement between the raters (Landis and Koch, 1977). For each component, the classification scores of the five human raters were synthesized so that the component would be labeled as "noise" only with a consensus score reached by at least three raters. This final set of scores was used for reconstructing a denoised dataset, during which the variance of the selected noise components was subtracted from the input fMRI time series.

The denoised fMRI datasets were then transformed into a standard brain space by applying the affine plus nonlinear spatial normalization parameters of the structural image. The resulting datasets, with a resampled voxel size of 3 × 3 × 3 mm^3, were spatially smoothed in a stepwise fashion to a target FWHM (Full Width at Half Maximum) of 8 mm in x-, y-, and z-directions using AFNI. This procedure was confined within the brain mask, which prevented the denoised voxels from being contaminated by artifacts outside the mask during the Gaussian blur.

fMRI General Linear Model Analysis

Subject level general linear modeling (GLM) was computed in SPM using classical restricted maximum likelihood (REML) estimation based on a canonical hemodynamic response function (HRF). Whole-brain mean signal was used as a covariate in order to reduce the effect of global BOLD signal changes caused by the fluctuations in arterial PCO2 (Birn et al., 2006) that result from continuous speech production (Hoit and Lohmeier, 2000).

For the group analysis, in order to determine the neural correlates of narrative production, the NAp condition was modeled and directly contrasted with NRp using t-tests on a voxel by voxel basis. The same contrast was performed for NAc and NRc, in order to determine the neural correlates of narrative comprehension.   The resulting student-*t* maps were thresholded with a voxel *p*-threshold of 0.01 and a cluster-size threshold of 67 voxels, corresponding to a family-wise error of 0.05 based on Monte Carlo simulations (Forman et al., 1995).   The conjoint activations between these two contrasts were identified as voxels met the above thresholds for both contrasts (i.e. NAp > NRp and NAc > NRc), distinguishing these from voxels that met thresholds for only one of these (i.e. either NAp > NRp or NAc > NRc respectively), which are also identified. Additionally, we directly compared these contrasts (i.e., NAp > NRp directly contrasted with NAc > NRc) in order to identify statistical differences between task conditions.

fMRI Functional Network Connectivity Analysis

Task-related modifications in functional connectivity were investigated across the entire brain. This analysis started from the residual time-series of the subject level GLM analysis with boxcar-shaped task effects removed. A finite impulse response (FIR) band-pass filter was subsequently applied to remove low-frequency fluctuations below 0.03 Hz and high-frequency noises above 0.08 Hz (Cordes et al., 2001). To account for the delay of hemodynamic response, the data within each block were shifted forward for three images and concatenated for each task condition respectively, resulting in 90 data points per subject and condition.

To reduce the dimensionality of the search space, group-level sICA (again using GIFT MATLAB toolbox) was used to derive 60 spatially independent components, each representing a self-organized functional unit (or network) with homogenous temporal dynamics. Prior to the sICA, data underwent two steps of reduction at the time domain using principal component analysis (PCA): one within each subject and condition, and the other at the group level after concatenating the principal components across all subjects and conditions. The group-level dimensionality for PCA and ICA decompositions (i.e., the selection of 60 components rather than another number) was estimated by identifying the minimal number of principal components that can capture all variances in the first data reduction step and approximating the maximal true degree of the freedom among all input datasets. The major purpose for this procedure was to use a high-order decomposition to maximize the observable effects while avoiding possible over-fitting errors (Sarela and Vigario, 2003).

After group ICA decomposition, nuisance components with spatial patterns clearly localized in major cerebral arteries, ventricles, or dural vein sinuses were rejected from further analysis, leaving 57 of the original 60 independent components. The time courses of remaining components for each subject and condition, which were computed from the group ICA time courses by a PCA-based back-reconstruction method (Erhardt et al., 2011), were used for functional network connectivity (FNC) analysis (Allen et al., 2011; Doucet et al., 2011; Jafri et al., 2008).

In FNC, for each subject, Pearson's correlation coefficients and their Fisher's z' transformations were computed between each component and every other component to indicate the strength of connectivity between functional networks. The resulting N x N matrices (where N equals the number of components, in this case 57) were averaged together across all subjects, resulting in a mean correlation matrix was for each experimental condition (i.e., NAp and NAc). The correlation matrix, R, was converted to a distance or linkage matrix, D = 1-R, indicating dissimilarity between each pair of components. Agglomerative hierarchical clustering on these distance values was done for sorting the components in a data-driven way, so that those with similar temporal dynamics were placed together in a cluster (Doucet et al., 2011). The distance between two clusters was the average distance between all pairs of their elements. For each condition, a dendrogram plot was generated to illustrate the hierarchical, binary cluster tree, in which leaves (or end points) represent components and the height of paths between leaves represents the distances between components (see Appendices I and II).

Independent Component 47 (IC-47), which demonstrated spatial overlap with the GLM conjunction results (demonstrated in Appendix III), was selected as an anchor for functional connectivity analysis (i.e., used to identify the component cluster of interest, refer to Appendices II and II). Clusters consisted of components within three degrees (i.e., steps within a dendrogram) from IC-47 for NAp and NAc.

PET Image Preprocessing and General Linear Modeling Analysis

We used the FSL linear image registration tool (FLIRT) to align each subjects PET scans to the first scan acquired for that subject. An average image of these aligned PET images was then computed and co-registered with subjects' structural MRI images using a mutual-information based algorithm (Studholme et al., 1999) provided in the SPM package. The nonlinear warping steps of the MRI image were utilized in the PET spatial normalization. This combined normalization scheme was used to transform the aligned PET images into MNI space. The resulting images were spatially smoothed with an 8mm FWHM Gaussian kernel. At the group level, an analysis of covariance (ANCOVA) was used to complete the general linear modeling for each task. NAp was directly contrasted with NRp. The same was performed for the comprehension tasks (NAc and NRc), resulting in t-maps. Because of relatively lower sensitivity of the PET and the importance of comparing fMRI and PET, it was necessary to set thresholds for PET were not as stringent as those used for fMRI. For comprehension data, a threshold cut off of $p < 0.05$ was used, without multiple comparison correction. A cut off of $p < 0.02$ was used for production, again without correction. To facilitate the comparison of fMRI and PET activation maps, the threshold for PET was relaxed (relative to the $p < 0.05$, corrected threshold used for fMRI). Due to dose limitations, PET allows for collection of significantly fewer data points and, as a result, suffers reduced statistical power when compared to fMRI.

Measuring the effect of denoising method: Comparison between fMRI and PET

In order to quantify the effect of the fMRI data denoising method on each participant, we performed Pearson's correlations between all within-mask voxels in the PET dataset and both the denoised and non-denoised fMRI datasets. This resulted in two sets of correlation coefficients (i.e., the correlation between PET and non-denoised fMRI and the correlation between PET and denoised fMRI) for each condition. The correlation coefficients were converted to z-scores using Fisher's z-transformation. The z-scores were then entered into separate paired t-tests for NAp and NAc.

Results

Task Performance

During the fMRI task, the participants typically produced narratives for the entirety of the 30 s block (duration mean ± standard deviation, 29.3 s ± 0.7 s). The narratives contained a mean of 88 words (± 12) and 7.7 t-units (± 2.1).

Positron Emission Tomography – Comparison with fMRI

The correlation between PET and denoised fMRI was significantly higher than the correlation between PET and non-denoised fMRI data for both the narrative production (two sample t-tests, $p = 4.8 \times 10^{-6}$) and narrative comprehension (two sample t-tests, NAc: $p = 0.005$) tasks. There was also a significant interaction between task (i.e., NAp or NAc) and the application of the denoising method ($F(1,16) = 27.23$, $p = 0.0001$), indicating increased efficacy during the production task.

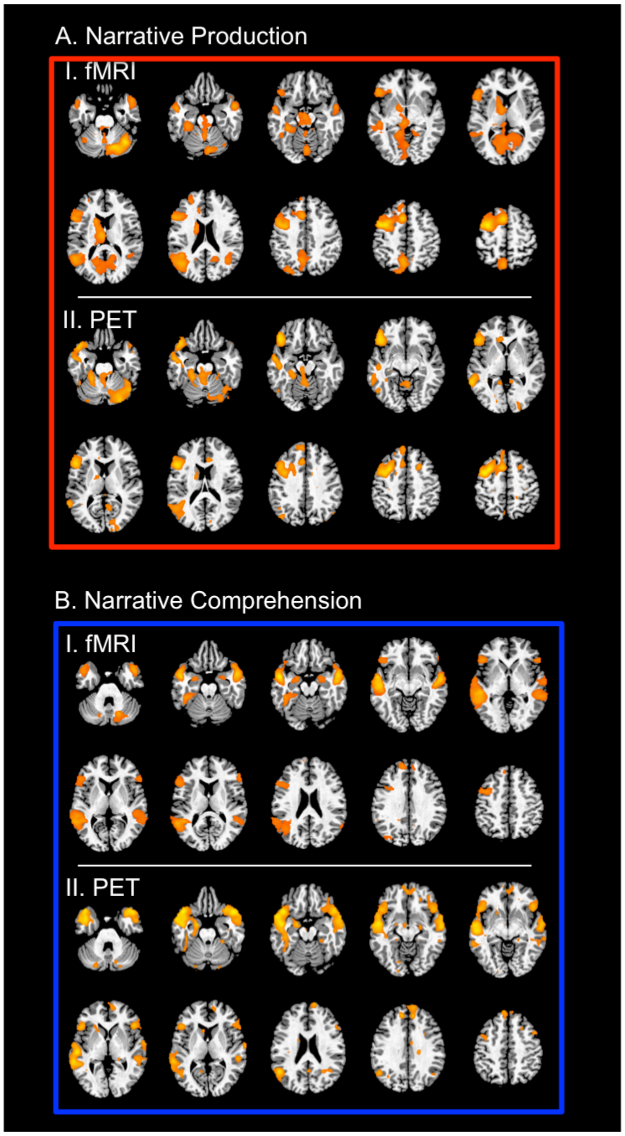


Figure 2. Axial images of results from contrast between either (A) Narrative Production and Nursery Rhyme production (NAp > NRp) or (B) Narrative Comprehension and Nursery Rhyme Comprehension in MRI (I) and PET (II). fMRI data are thresholded at level of $p < 0.05$, FWE corrected. The threshold for PET data presented in 2A,II is $p < 0.02$, uncorrected. The threshold for PET data presented in 2B,II is $p < 0.05$, uncorrected.

Figure 2A illustrates the patterns of activation detected during the narrative production task (versus the nursery rhyme production baseline) for both fMRI and PET. Both methods revealed virtually identical activation patterns that included perisylvian (left IFG, temporal pole and STS, left MTG), extrasylvian (left angular gyrus, dmPFC, and precuneus), visual (left and right lingual gyri) and motor related areas (left SMA, pre-SMA, and right cerebellar hemisphere). While common activation patterns predominated, some differences were found (e.g., activation of left lateral orbital frontal cortex was selectively detected by PET; right STS and angular gyrus by fMRI).

Figure 2B similarly illustrates patterns of activation detected during narrative comprehension (versus the nursery rhyme comprehension baseline) for both imaging methods. Similar to narrative production, common activation patterns predominated. Perisylvian regions were engaged bilaterally (left and right IFG, STG, MTG, temporal poles and anterior STS) during narrative comprehension for both methods, as were extrasylvian regions, also bilaterally, including the dmPFC, angular gyri and temporal poles, and cerebellar hemispheres. Again, some differences were found: activation of the orbital cortex was detected only in PET; both left and right amygdalae were found in fMRI, whereas only the left amygdala was found to be active in PET.

fMRI: Neural Correlates of Narrative Production and Comprehension

Contrasts between the narrative and respective baseline conditions (NAp-NRp; NAc-NRc) for production and comprehension and the conjunctions between these contrasts are summarized in Figure 3 and Table 1A-C. Figure 3 depicts clusters of significant activations unique to either production (red) or comprehension (blue), with yellow representing voxels in which activations are significant for both contrasts. Table 1 provides the corresponding t-values and coordinates of the voxel with peak activation within each significant cluster.

Activations Common to Narrative Production and Comprehension

Activations in perisylvian brain regions common to both production and comprehension were predominantly left lateralized and included the left superior temporal (STG), middle temporal (MTG), and inferior frontal gyri (IFG; BA 45,47). Common activations in the temporal poles and anterior superior temporal sulci (STS), in contrast, were bilateral.

Common activations, also left lateralized, were also detected in an extrasylvian network comprised of the left dorsal medial prefrontal cortex (dmPFC), precuneus and angular gyri (in which activations were stronger and more widespread in the left hemisphere), and in the left parahippocampal gyrus

Lastly, shared activations were found in motor-related areas, including the left and right pre-supplementary motor area (pre-SMA), the left dorsal premotor cortex (PMd) and in the right posterior cerebellar hemisphere.

Activations Unique to Narrative Production

There were no activations in perisylvian regions uniquely related to narrative production (that is, activations detected in these regions during production were also found during narrative comprehension, as outlined above).

In extrasylvian cortices, the cluster of activation of the left dmPFC seen for production extended beyond that identified in the conjunction analysis, laterally into the dorsal lateral prefrontal cortex (dlPFC, BA10) and dorsally into the superior portions of the dorsomedial prefrontal cortex (BA8). Similarly, activation of both the left precuneus and left angular gyrus during production extended beyond the cluster identified in the conjunction analysis. In addition, the left anterior cingulate cortex and both the left and right cuneus and lingual gyri were selectively activated during narrative production.

Activation of the pre-SMA and PMd during narrative production extended beyond the clusters shared with comprehension. Further, left hemisphere subcortical regions uniquely activated during narrative production included the dorsal caudate and anterior and dorsomedial thalamus. Narrative production was also uniquely associated with right lateralized activation of the lateral cerebellum.

These results were verified by direct statistical comparisons of our experimental conditions (i.e., NAp and NAc), which showed that these same left-lateralized regions were significantly more engaged during NAp, including the dlPFC (peak voxel in cluster: -27, 47, 21; $t = 5.92$; $p < 0.05$), rostral PFC (-27, 47, 21; $t = 5.92$; $p < 0.05$), a portion of the left precuneus (-6, -67, 48; $t = 6.23$; $p < 0.05$), and angular gyrus (-42, -70, 24; $t = 4.25$; $p < 0.05$).

Additionally, a cluster with a peak voxel in the pre-SMA (-6, 17, 45; $t = 9.52$; $p < 0.05$) extended into the cingulate cortex and PMd. The thalamus (-6, -13, 12; $t = 6.46$; $p < 0.05$) was also significantly more engaged during NAp, as was a relatively large cluster with a peak voxel in the cerebellum (36, -58, -33; $t = 10.69$; $p < 0.05$) and extended to the cuneus and lingual gyri.

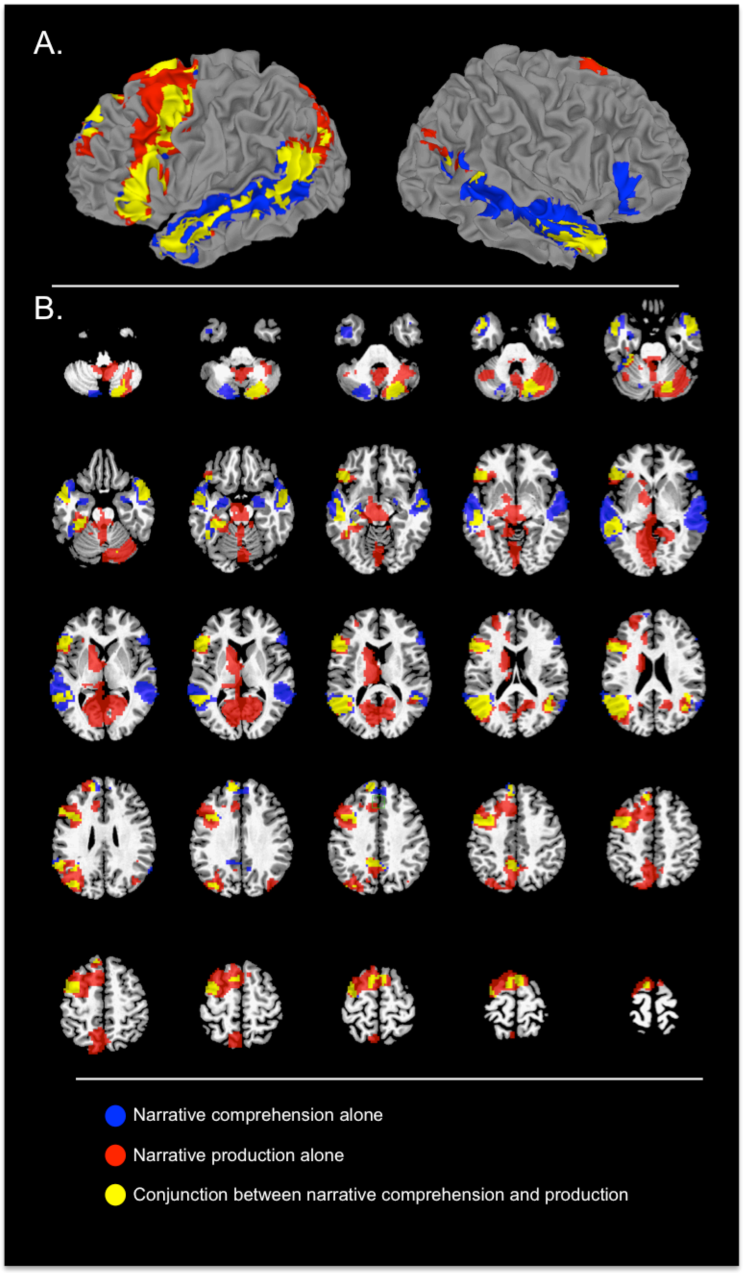


Figure 3. Conjunction analyses presenting MRI results in axial brain images. The color red indicates voxels significantly activated for narrative production above baseline (NAp > NRp). Blue represents voxels recruited for narrative comprehension over its respective baseline (NAc > NRc). The overlap, in

yellow, consists of voxels that are activated above threshold in both contrasts (NAp > NRp and NAc>NRc). Images thresholded at $p < 0.05$, FWE corrected. Section A depicts rendered images of the left and right surfaces of a template brain. Section B illustrates axial slices progressing from inferior (top left) to superior (bottom right) brain regions.

Activations Unique to Narrative Comprehension

The hallmark of activation associated with narrative comprehension appeared to be strong bilaterality. Comprehension was uniquely associated with activation of right hemisphere homologues of the left perisylvian areas that were identified in the conjunction analysis, including the right IFG and a wide extent of the right superior and middle temporal gyri, extending from the pole to the temporoparietal occipital junction and into the angular gyrus (where it encompassed a larger area than that identified in the conjunction analysis).

Comprehension was also associated with activation of contralateral homologues of other regions identified in the conjunction analysis, including a small portion of the right posterior dmPFC and the left cerebellar hemisphere.

Finally, comprehension was uniquely associated with robust bilateral activation of the amygdalae.

Direct comparisons between NAc and NAp confirmed that these regions were significantly more engaged during NAc, notably in the right IFG (peak voxel, 48, 41, 3; $t = 5.61$; $p < 0.05$) and the right (57, -1, -6; $t = 7.42$; $p < 0.05$), as well as the left (-60, -16, -6; $t = 7.73$; $p < 0.05$), superior temporal cortex and middle temporal cortices.

Table 1A

FMRI REGIONS OF CONJUNCTION: ACTIVATED DURING BOTH NARRATIVE PRODUCTION AND COMPREHENSION

BRAIN REGION	**BA**	**X**	**Y**	**Z**	**T-score**
LEFT HEMISPHERE					
Cortical					
Perisylvian Areas					
Inferior frontal gyrus, pars opercularis	45	-54	20	15	5.94
Inferior frontal gyrus, pars triangularis	47	-48	26	-3	6.43
Superior temporal sulcus, anterior	21	-51	-1	-21	7.43
Middle temporal gyrus, posterior	21	-60	-49	0	4.34
Extrasylvian Areas					
Pre-supplementary motor area	6	-9	14	63	4.16
Dorsal premotor cortex	6	-39	5	51	6.29
Angular gyrus	39	-45	-64	18	6.71
Dorsal-medial prefrontal cortex	8	-12	50	33	4.94
Precuneus	7	-3	-52	39	3.97
Parahippocampal gyrus	36	-24	-37	-18	3.85
Fusiform gyrus	20	-33	-37	-21	5.10
RIGHT HEMISPHERE					
Cortical					
Perisylvian Areas					
Superior temporal sulcus, anterior	21	54	2	-21	7.66
Extrasylvian Areas					
Angular gyrus	39	42	-61	18	4.77
Pre-supplementary motor area	6	9	17	63	3.53
Subcortical					
Cerebellum, posterior	--	18	-79	-45	8.13

Table 1B

FMRI REGIONS ACTIVATED DURING NARRATIVE PRODUCTION ALONE

REGION OF INTEREST	**BA**	**X**	**Y**	**Z**	**T-score**
LEFT HEMISPHERE					
Cortical					
Perisylvian Areas					
Inferior frontal gyrus, pars triangularis	47	-30	29	-3	4.30
Extrasylvian Areas					
Dorsal-medial prefrontal cortex	8	-12	35	48	4.74
Dorsal-lateral prefrontal cortex	10	-24	47	21	4.96
Dorsal premotor cortex	9	-42	20	33	6.72
Supplementary motor area	6	-6	14	54	10.19
Cingulate cortex	32	-12	29	24	5.21
Precuneus	7	-6	-58	45	7.65
Cuneus	18	-3	-61	6	5.39
Subcortical					
Dorsal caudate	--	-15	8	15	7.47
Dorsal thalamus	--	-3	-16	12	7.56
Cerebellum, lateral	--	-36	-61	-30	5.33
RIGHT HEMISPHERE					
Cortical					
Posterior cingulate cortex	23	9	-57	12	5.02
Precuneus	7	12	-55	45	3.20
Angular gyrus	39	42	-76	30	4.22
Lingual gyrus	18	24	-55	0	4.92
Subcortical					
Cerebellum, lateral	--	39	-58	-33	13.14
Cerebellar vermis	--	3	-76	-18	6.97

Table 1C
FMRI REGIONS ACTIVATED DURING NARRATIVE COMPREHENSION ALONE

REGION OF INTEREST		**BA**	**X**	**Y**	**Z**	**T-score**
LEFT HEMISPHERE						
Cortical						
	Temporal pole	20	-42	2	-42	5.01
	Superior temporal gyrus, anterior	22	-57	-4	-9	9.82
	Superior temporal gyrus, posterior	22	-33	-46	18	3.95
Subcortical						
	Amygdala	34	-21	-4	-21	5.61
	Hippocampus	28	-27	-16	-15	5.37
	Cerebellum, posterior	--	-21	-76	-39	6.55
RIGHT HEMISPHERE						
Cortical						
Perisylvian areas						
	Inferior frontal gyrus	47	51	26	-3	5.35
	Superior temporal gyrus, anterior	22	57	-10	-9	9.31
	Middle temporal gyrus	21	57	-40	-3	5.70
Extrasylvian areas						
	Angular gyrus	39	60	-58	21	4.07
	Dorsal-medial prefrontal cortex	9	6	44	36	4.74
Subcortical						
	Amygdala	34	24	-4	-21	4.95

Table 1 presents the regions of interest from the GLM analyses alongside corresponding Brodmann Areas, XYZ values, and t-scores. 1A) Regions significantly active for both narrative production (NAp > NRp) and narrative comprehension (NAc > NRc). 1B) Regions uniquely activated for narrative production (NAp>NRp). 1C) Regions uniquely engaged by narrative comprehension (NAc>NRc).

Functional Connectivity in Narrative Production and Comprehension

Conjunction Component

IC-47 appeared to exhibit spatial overlap with the regions shared by NAp and NAc, presented earlier. To quantify this observation, a voxel-wise Boolean mask was applied to identify voxels that are above threshold in both IC-47 and the regions shared by NAp and NAc. The resulting image (see Appendix III) demonstrates considerable spatial overlap in the left IFG, superior temporal sulcus, inferior parietal lobule, dorsal premotor cortex, medial prefrontal cortex, and right cerebellum. On this basis, IC-47 is here on referred to as the Conjunction Component.

Network Connectivity during both Narrative Comprehension and Production

During both NAp and NAc the Conjunction Component was most closely connected to IC-49, which contains bilateral superior and middle temporal cortices (Figures 4 and 5).

Network Connectivity during Narrative Production

During NAp, the Conjunction Component also connects to IC-46, which is characterized by the left lateralized dlPFC. At the third degree, the Conjunction Component connects to four components: IC-41, IC-51, IC-13, and IC-36. This group of components is distinguished by cortical and subcortical regions related to motor planning and coordination (Bohland et al., 2010; Manto et al., 2012; Paus et al., 1993; Schulz et al., 2005). See Figure 4.

Network Connectivity during Narrative Comprehension

The Conjunction Component is closely linked to IC-34 (Figure 5), which consists of t he right-lateralized perisylvian regions. There are third degree connections to three components: IC-52, IC-32, and IC-54. Each of these third-degree components contains medial and lateral regions that are key elements of the theory-of-mind network (Saxe et al., 2004; Spreng et al., 2009).

More detailed discussion of FNC results is available in Appendix IV.

Discussion

The current study is the first to use BOLD fMRI methods to examine naturalistic speech production and comprehension within the same cohort of subjects using a reliable artifact reduction method. We focused on language as it is used at the level of discourse, specifically in processing narrative fiction. Using a well-established story-based paradigm, we used both GLM contrast and ICA-based network connectivity methods to pinpoint what is common and what is unique in storytelling and story comprehension, paying attention to the roles of perisylvian, extrasylvian and sensorimotor systems in both left and right hemispheres.

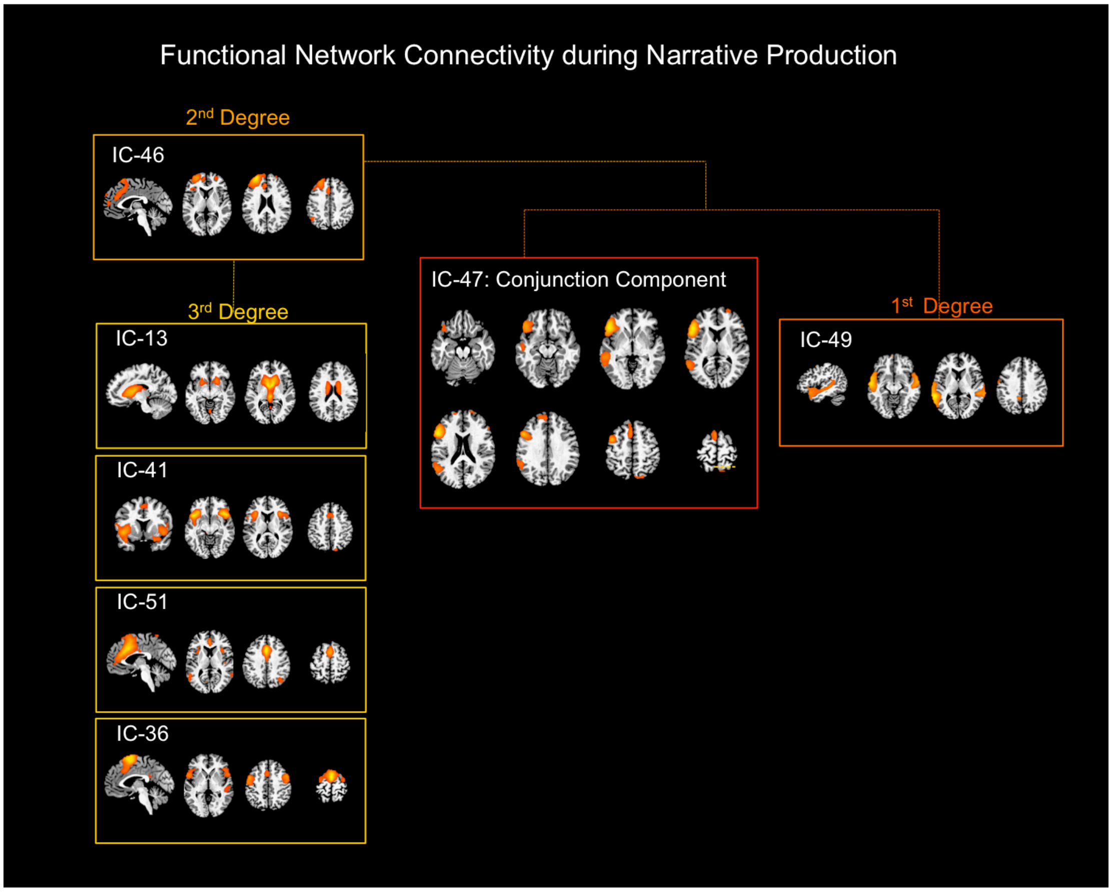


Figure 4. Brain images from independent components connected to the anchor component (IC-47) during narrative production. Again, IC-47 (the Conjunction Component; red box, center) is most closely connected to IC-49 (right side, dark orange box). A second-degree connection to IC-46 (left side, light orange box) is next. Last, there are third degree connections to three components (IC-54, IC-32, IC-52; yellow boxes).

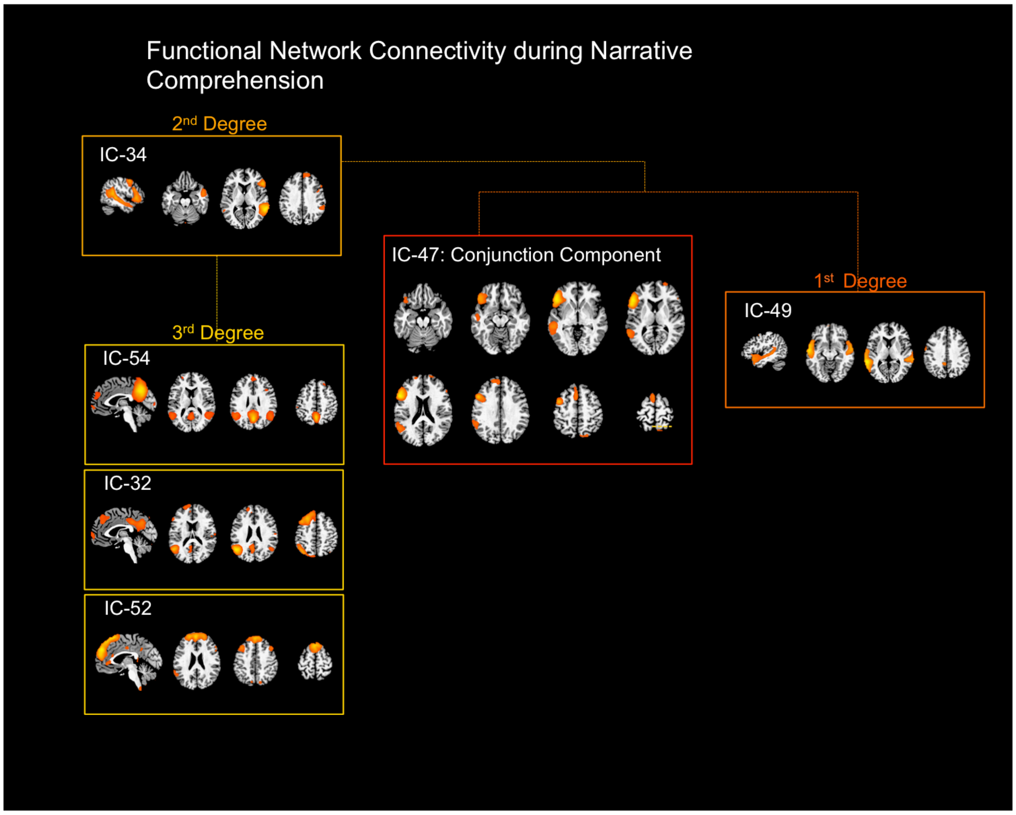


Figure 5. Brain images from independent components connected to the anchor component (IC-47) during narrative comprehension. IC-47 (the Conjunction Component; red box, center) is most closely connected to IC-49 (right side, dark orange box). IC-46 (left side, light orange box) demonstrates the next closest link. Four components (yellow boxes) demonstrate third degree connections.

To do this, we introduced a novel method that rectifies deep-rooted difficulties that have precluded the imaging of continuous speech production using BOLD fMRI. Comparison of both PET and fMRI contrast results suggest the method is valid and strengthens the validity of our findings, which have shed light on a number of the questions we raised at the outset and have raised additional questions in turn.

PET: Comparison with fMRI

The striking similarities between results obtained with PET and MRI are illustrated in Figure 2. Both methods detected activation of perisylvian regions (IFG, STG, STS, MTG) as well as temporal poles/anterior STS), premotor areas (PMd and pre-SMA), and extrasylvian regions (left dmPFC, angular gyrus, parahippocampal gyrus) for both narrative production and comprehension.

Moreover, both MRI and PET demonstrated the same selective activations for production (ACC, superior dmPFC, rostral PFC, visual association areas) and comprehension (activation of right inferior frontal and temporal cortices and amygdala).

While the similarities between PET and fMRI results are evident, disparities were also observed, some of which may be related to technical differences between these methods. Some of these – e.g. strong activation of subcortical regions during narrative production (see Figure 2A) and activation of both amygdalae (see Figure 2B) during comprehension – were detected only with fMRI, and could simply be ascribed to the increased sensitivity of this method. It is also possible that auditory scanner noise (which is present in fMRI but absent in PET) and its potential effect on attention (e.g. through auditory interference) could lead to modality specific differences. Other differences –e.g. the fact that activation of the orbital frontal cortex was detected by PET, but not MRI – may reflect technical shortcomings of fMRI, attributable to proximity of this region to air/tissue boundaries that are vulnerable to fMRI signal loss (Ojemann et al., 1997).

In a larger sense, the primarily uniform correspondence between PET and fMRI results – particularly in inferior fronto-temporal areas subject to susceptibility artifact (Barch et al., 1999; Birn et al., 1998; Kemeny et al., 2005) underscores the reliability of the ICA-based denoising methods used here in post-processing the narrative production data, and indicates that the present findings can be reported and interpreted with confidence.

fMRI: Complementary GLM and Functional Network Connectivity Analyses

Traditional GLM underscores task-related differences in brain activity. ICA Connectivity, on the other hand, highlights the within-condition correlation between networks without the need for contrasts. As such, functional network connectivity (FNC) can shed light on the functional relationships between regions identified in GLM. When taken together, they present a more comprehensive view of the neural grounding of cognitive functions. The results of both approaches are discussed below.

fMRI: Neural Correlates of Narrative Production and Comprehension – Shared Responses

Shared activations, i.e., increases in BOLD signal above baseline that are common to both story comprehension and production, were identified using GLM based conjunction methods.

Perisylvian areas

As expected, brain regions long associated with language processing – classical perisylvian areas including the left IFG, STG and MTG, as well as regions more recently associated with language processing such as the temporal poles/anterior STS – were active, and functionally coupled to one another, during both language production and comprehension (see Figure 3, Table 1A). It is well-established that these brain regions support basic speech and language processes at the phonological, lexical and sentential levels (Acheson et al., 2010; Binder et al., 1997; Friederici et al., 2006; Moro et al., 2001; Okada and Hickok, 2006; Rissman et al., 2003; Visser et al., 2010), yet it is important to note that these regions are active following subtraction of the propositional speech baseline tasks.

It is possible that the additional activity in these regions is due to the more extemporaneous nature of the narrative tasks – the demands associated with on-line syntactic or phonological processing, both in formulation and comprehension of stories - that are not manifest during the execution of the nursery rhyme tasks.

It is also possible that increased activity may in some cases reflect higher-level cognitive or linguistic functions that are essential to storytelling or story comprehension. For example, activation of the temporal poles may be linked to processing more complex discourse-level semantic knowledge (Visser et al., 2010) or tying together connected sentences (Mar, 2004), functions that characterize both production and comprehension tasks. Similarly, the IFG and posterior MTG may play a role in higher-level semantic retrieval and integration processes (Hagoort, 2005) that emerge at the level of discourse. Indeed, the IFG has been linked to semantically appropriate lexical selection during narrative (Marini and Urgesi, 2012). Moreover, both the left IFG and posterior MTG are engaged in processing what are unambiguously discourse level features such as metaphor, and sarcasm (Eviatar and Just, 2006; Shibata et al., 2010; Uchiyama et al., 2006).

It is important to note that essentially the same set of regions was identified as an integrated network common to both conditions (referred to as the “Conjunction Component”) using ICA based connectivity methods. These regions included these same classical left hemisphere perisylvian language cortices, such as the left IFG and MTG. Similar to earlier research that underscored structural and functional connections between the left IFG and MTG (Turken and Dronkers, 2011), our results identify them as elements of the same functional network. However, the network also extended beyond these to include extrasylvian areas that appear to interact with the language system during the production and comprehension of natural language.

Interestingly, while the GLM contrasts show that a large extent of superior and middle temporal gyri is more active bilaterally only during narrative comprehension (discussed further below in the section on fMRI response unique to narrative comprehension), the FNC results show that during both narrative comprehension and production the Conjunction Component is significantly coupled to a component that encompasses these temporal regions. One possible interpretation of this finding is that this production-related increase in connectivity between the Conjunction Component and bilateral superior temporal cortex is due to auditory-motor interactions – e.g. self monitoring mediated either by direct auditory feedback or through an internally modeled or dynamic representation of the vocal tract and its auditory output (Hickok et al., 2011). In the case of narrative comprehension, processing incoming auditory/linguistic information generated from an external source may place greater demands on the auditory system, resulting in stronger activation, more robust changes in local field potentials, and concomitant increases in the BOLD signal in addition to the increased functional connectivity between the systems.

Left lateralized perisylvian regions were also reported in an earlier study of narrative production and comprehension (Awad et al., 2007), but an important distinction between the present findings and those of Awad et al. relates to the robust activation of the left IFG we detected for both narrative production and comprehension. Because the left IFG was present in both our PET and fMRI results this discrepancy is not due to differences between the imaging modalities employed, but may result from two significant differences in study design. The first may have to do with the baseline tasks used. While we used matched, relatively simple language tasks for both comprehension and production, the tasks used by Awad et al. – in particular processing of spectrally rotated speech – may have placed functional demands upon the IFG that could have induced activations that obscured any discourse-related increases (as the authors themselves suggest). Another distinction lies in the experimental task itself, which may similarly place additional demands on the IFG. While subjects in the Awad et al. study produced personal narratives from memory (e.g. how one spent a weekend) these may have been less likely to contain a standard narrative structure – an introduction followed by the development and resolution of a plot – like those used in the present study. For example, the left IFG could be sensitive to this sort of internal narrative structure (e.g. causal event structure or "story grammar"), which may engage the unification or integrative processes mediated by that region, while recitation of more unstructured autobiographical events may not.

Motor-related areas

Both narrative tasks were associated with activation of motor-related areas including the PMd and pre-SMA, and these areas were functionally coupled to the perisylvian language cortices during both comprehension and production. While the processes carried out by these premotor regions may be intuitively connected with production (e.g. rapid selection and organization of narrative elements, as discussed in the next section) there is evidence that the PMd and pre-SMA also play a role in narrative comprehension. For example, in addition to its involvement at lower levels of speech perception (Canessa et al., 2008; Meister et al., 2007; Saur et al., 2008), the PMd is more active during narrative comprehension than during the processing of disconnected words or sentences (Mano et al., 2009; Yarkoni et al., 2008) and the pre-SMA appears to play a general role in reconciling conflicting or inconsistent information (Wittfoth et al., 2006) that might account for its involvement in detecting sarcasm (Uchiyama et al., 2006). The pre-SMA is also implicated in cognitive processes that should be engaged during both narrative comprehension and production, such as executive control of working memory (Marvel and Desmond, 2010), monitoring of behavioral sequences (Shima and Tanji, 2000) or the cerebral representation of space and time (Beudel et al., 2009). Strengthening this assertion, networks containing the pre-SMA were functionally linked to the Conjunction Component during both narrative production and comprehension (see Figures 4 and 5).

Extrasylvian areas

While many of the perisylvian and motor-related areas have been linked to speech and language processing at lexical and sentential levels, we also observed activations in a set of regions – dorsomedial prefrontal cortex, precuneus and inferior parietal lobules – that may emerge only at the level of discourse. Awad et al. (2007) also found overlap in a number of these areas. In the present study, these regions were engaged during both narrative tasks (Figure 3, Table 1A), likely supporting interactions between language and other cognitive systems that play a crucial role during storytelling or story comprehension.

A relationship between one of these regions – the medial prefrontal cortex – and discourse-related processes has been described for some time. For example, an early study (Fletcher et al., 1995b) found that this area is engaged when assessing the motives of characters in a narrative, but not when evaluating mechanical or physical properties of the narrative. Willems et al. (2010) showed that this region is sensitive to the communicative intent of utterances rather than to their linguistic complexity. The medial prefrontal cortex is in fact a higher order heteromodal area that plays many roles and integrates multiple cognitive operations (Ramnani and Owen, 2004) – including motivation (Stuss and Levine, 2002), orientation and allocation of attentional resources (Burgess et al., 2007), and recognition of intentional actions (Chaminade et al., 2011) – all of which come into play when language is used to communicate information at the discourse level. Among the more specific functions that have been described, there are many that are explicitly engaged during the processing of stories, e.g. "source monitoring" (retrieving information about when, where and how events occurred) (Turner et al., 2008), extraction of a thematic message (the “moral”) of a story (Nichelli et al., 1995), and – importantly – the use of self-referential information to understand the mental states of others (Gilbert et al., 2007; Mitchell et al., 2005).

Indeed, the medial prefrontal cortex, together with the precuneus, inferior parietal lobules, and parahippocampal gyri – all of which were activated during both narrative tasks – constitutes a network, that has been linked to social cognition and to “mentalizing”, also known as theory of mind (Amodio and Frith, 2006; Saxe et al., 2004; Spreng et al., 2009) – the ability to intuit the goals, beliefs, and intentions of others. The process of inferring beliefs, goals and intentions may be directly involved in narrative comprehension and production in at least two ways. First, for both hearer and storyteller, mentalizing should play a crucial role in understanding the thoughts, feelings and motives of the stories’ characters and the complex social interactions between them. Second, as a “second order” process, the speaker needs to infer the expectations of his or her audience and the hearer must similarly infer the intentions of the storyteller. Further, the ability to process many of the extratextual or implicit elements of discourse, such as irony or sarcasm, relies on mentalizing and activates the same network of regions (Shibata et al., 2010; Uchiyama et al., 2006).

While storytelling or story comprehension clearly involve this sort of social or emotional inference making, it is also the case that discourse processing is mediated by more general inferencing mechanisms, of which mentalizing is but one example. For example, during both comprehension and production, many non-social – e.g. temporal, spatial, causal – relationships between situations, events or other elements of a plot must frequently be inferred (by both listener and storyteller) when these relationships are not explicitly formulated and encoded in the story. In fact, the same set of regions plays a role in more generic inference making (Ferstl and von Cramon, 2001; Kuperberg et al., 2006; Mason and Just, 2011; Sieborger et al., 2007). In a broader sense, inference making – which crucially requires incorporation of the speaker's or hearer's own world knowledge – is key in building narrative coherence (Graesser and Kreuz, 1993; Graesser et al., 1994) and the same network clearly supports coherence building (Ferstl and von Cramon, 2001; Kuperberg et al., 2006; Mason and Just, 2011; Sieborger et al., 2007) as well.

In addition, the same regions have been implicated in the retrieval of episodic memory (Sestieri et al., 2011). In fact, this is the principal interpretation made by Awad et al. regarding similar observations reported in their study (during which subjects processed autobiographical materials). Although our stimuli consisted of fictional stories, the events in these stories likely elicited implicit retrieval of subjects' episodic memories during both production and comprehension of these stories. These memories may be related to one's own experience with similar situations. Yet, it is also possible that these regions are engaged in retrieving the macrostructure of each narrative's situation model.

It should be stressed that this same network of regions is called by many names (e.g. default mode, mentalizing, task-negative - largely related to the experimental context in which it has been described) and subserves multiple functions, in addition to the long list cited above (Legrand and Ruby, 2009; Spreng et al., 2009). The heteromodal brain regions that make up this network are connected not only with each other, but are connected to and process information from a wide array of cortical and subcortical areas distributed throughout the brain, and likely perform neural computations - higher order integration and orchestration of information, whether exteroceptive or internally generated - that underlie the growing array of cognitive functions with which it has been associated. Nevertheless, while domain general and certainly not dedicated to processing discourse, this network is clearly engaged in reading or telling stories and may, in this context represent a system that tethers together language and other cognitive and sensorimotor domains.

What may unify all of these processes is that both social and nonsocial inferencing, and the incorporation of critical elements of world knowledge (derived from both semantic and episodic memory) all represent key mechanisms by which an ongoing multidimensional representation of the story – a mental model (Johnson-Laird, 1980) or situation model (van Dijk and Kintsch, 1983) – is created. Indeed, a number of studies have found that these same regions play a role in building, manipulating and updating such models (Ferstl et al., 2005; Martin-Loeches et al., 2008; Whitney et al., 2009; Yarkoni et al., 2008). Taken together, whether they are hearing or telling stories, our participants may use this network to infer and understand mental states, action consequences, social interactions and the relationships between story elements, drawing upon their knowledge of the world in order to create a situation model that represents the narratives and the characters and events within them.

It should also be noted however, that the activations we report here involve only a subset of regions that have been associated with this network. For example, while it is commonly described as bilateral, the conjunctions we report were markedly left lateralized, e.g. activations of the medial prefrontal cortex, precuneus and parahippocampal gyrus are confined to the left hemisphere. Additionally, we did not observe activation of some regions that have frequently been described as elements of this network – e.g. the posterior cingulate cortex.

It is possible that these differences are characteristic of narrative processing *per se*. On the other hand, they might also be attributable to the fact that subjects had already been exposed to the stories' content (incorporated into our design to provide a greater measure of control). That is, it is possible that if subjects were hearing the stories for the first time or were generating them spontaneously, we may have detected activation of wider array of regions within this network, including homologues within the right hemisphere. Future studies should clarify this issue.

fMRI: Responses Unique to Narrative Production

The GLM analyses (Figure 3, Table 1B) showed that narrative production was associated with strong bilateral activation of visual association cortices. Activation of these regions may support participants' representation of visual features of the stories being told (Chen et al., 1998; Ganis et al., 2004; Lambert et al., 2002). It is unclear why the same pattern was not found for narrative comprehension, during which one might also expect similar generation of visual images.

Importantly, the majority of brain regions that were activated solely during narrative production play a role in action selection, speech motor sequencing, phonation and articulation (Bohland et al., 2010; Manto et al., 2012; Paus et al., 1993; Schulz et al., 2005), including the dorsal ACC (including the cingulate motor area), the caudate nucleus and anteromedial thalamus, a broader extent (i.e., beyond the voxels shared with narrative comprehension) of the pre-SMA and left dorsal premotor area and right lateral cerebellar hemisphere.

While all of these cognitive and motor processes are certainly engaged to some degree during the baseline task, the narrative production task involves more rapid, on-line language formulation – e.g. lexical selection, syntactic construction that occurs in real time and may place greater demands upon cortical and subcortical regions that support on-line selection processes (Forstmann et al., 2008) leading to their selective activation during narrative production.

The caudate nucleus and dorsal thalamus may also support higher-level functions such as conceptual sequencing (Chan et al., 2011) and, in concert with other regions selectively activated during production – e.g. dorsolateral prefrontal areas and pre-SMA (Kennerley et al., 2004) - may play a role in discourse level processes such the organization and sequencing of narrative elements into event structures, in a manner similar to the organization of action sequences in motor planning. Taken together, these interpretations suggest the possibility that a novel, discourse-specific role for cortical and subcortical motor systems may emerge during the formulation and production of a narrative.

These functional results are bolstered by the finding that during narrative production, the Conjunction Component was selectively coupled to a set of components encompassing premotor regions, as well as basal ganglia and thalamus. This evidence suggests that motor-related regions are not only more active, but that their activity is strongly correlated with that in perisylvian and extrasylvian regions, including the dorsolateral prefrontal cortex, during the production of stories. Together all of these regions may function in a cognitive-motor cascade that orchestrates more complex language formulation – in which information flows from cognitive and language related areas through the prefrontal cortex to premotor and subcortical areas that organize articulation.

fMRI: Responses Unique to Narrative Comprehension

Narrative comprehension was, in general, associated with strong bilateral activations - differentiating it from production (Figure 3, Table 1C). This was most apparent in perisylvian areas where comprehension was associated with robust activation of right hemisphere homologues of left hemisphere regions that were activated during both conditions. These activations included the right IFG, STG, and MTG, the latter extending dorsally into the right angular gyrus (where they extended beyond the smaller cluster of activation shared with production). In addition, even within the left hemisphere, comprehension elicited stronger and more widespread activations of the superior and middle temporal gyri, accentuating the bilateral nature of the response.

Speech perception is unambiguously associated with bilateral activation of the temporal cortices (Binder et al., 2000; Hickok and Poeppel, 2007). Therefore, a bilateral response would be expected during the nursery rhyme comprehension task. The fact that bilateral activity in auditory association cortices was greater for stories than for the baseline task suggests that this activity is specifically related to discourse processing. This is consistent with previous imaging studies that have reported selective bilateral superior temporal activation for narrative comprehension (Awad et al., 2007; Robertson et al., 2000; Stephens et al., 2010) and might in this case be related to a variety of features present in discourse – for example, the emergence of more complex discourse level semantics (Jung-Beeman, 2005).

Consistent with our findings, previous studies have also demonstrated that discourse comprehension elicits activation of the right IFG, while sentence level processing does not (Robertson et al., 2000). Additionally, the right IFG has been found to be sensitive to discourse level manipulation of context (Menenti et al., 2009) and is activated in making causal inferences between ambiguously related sentence pairs (Kuperberg et al., 2006; Mason and Just, 2011) – all processes that could be more essential in processing stories than in telling them.

In addition, connectivity analyses of narrative comprehension demonstrate that the Conjunction Component is selectively coupled to a network entirely comprised of its perisylvian homologues in the right hemisphere.

The network connectivity analysis of narrative comprehension also revealed strong connections between the Conjunction Component and a more widespread set of the "mentalizing" regions in both left and right hemispheres, particularly the medial prefrontal, precuneus and bilateral IPL. This finding suggests that the mentalizing network may not support inference making for narrative production and comprehension equally. Instead, it might be argued that intuition and inference are more critical when characters, their motivations, relationships and intentions need to be processed by the hearer (rather than the storyteller who already has created a model of these relationships) or when gaps in the textual information (ostensibly available to teller) must be filled in.

Limitations

Keeping our tasks as naturalistic as possible was a primary consideration in this study. For example, we encouraged participants to retell the stories in their own words. Nevertheless, as a measure of control, they had been exposed to the stories' content prior to both production and comprehension blocks. It will be important in the future to identify patterns of brain activity in participants as they create fictional stories in a truly spontaneous fashion and compare these to patterns induced as they listen to stories they had never heard. These spontaneous conditions are arguably closer to the way language is used in the real world, and we might expect critical differences to emerge that were not detected in the present study. Moreover, the fact that we separated comprehension and production into separate blocks, again to maintain experimental control, is another departure from natural language use.

Summary and Future Directions

In this study, we sought to detect and contrast patterns of brain activity associated with language comprehension and production as these operate in a naturalistic context. Using both contrast and connectivity methods, we have shown that both comprehension and production of narrative fiction engage not only perisylvian areas, but extrasylvian systems that appear to interact with language at this level. We have argued that these patterns support cognitive processes that are related to both storytelling and story comprehension – e.g. mentalizing, inference-making, construction of a situation model, or model of the narrative world.

Implications for Conversation

As mentioned earlier, narrative and conversation share many features. Most basically, both entail comprehension and production of complex, largely extemporaneous linguistic stimuli. Yet, beyond that both also depend one's understanding of both context and others' mental states to make sense of rapidly changing and, at times, abbreviated information. Some of the same underlying cognitive processes mentioned above, such as mentalizing, are essential to successfully engaging in conversation. For example, just as inference-making supports pulling together information and building a coherent narrative, it is also involved in integrating world knowledge to make sense of implicit statements, which abound in natural conversation (Grice, 1975). Similarly, building a situation model may also be important during conversation, although it is constructed by multiple people, instead of only one as in narrative processing.

Due to these established and potential similarities between narrative and conversation, I expect to find similar overlapping regions during comprehension and production of conversation, including peri- and extrasylvian regions, pertaining to high-order language processing and social cognition.

However, some differences exist between these two elements of discourse. Other than comprehending and producing complex speech, conversation requires relatively immediate alternation between these states (Stivers et al., 2009). Undoubtedly, the anticipation and coordination of conversational turn taking adds a dimension that cannot be brought to light with any other sort of language use.

Study 2: Comprehension and Production During Conversation

Introduction

In the previous chapter, several peri- and extrasylvian regions were found to support both narrative production and comprehension. Virtually nothing is known of the neural correlates of production and comprehension during natural conversation. Consequently, it is important to explore the basic features. The first step is to examine comprehension and production. Because of the similarities discussed earlier, it is reasonable to expect narrative and conversation to share some neural correlates. Yet, it will also be important to explore the differences.

In contrast to narrative, conversation requires fluid and repeated alternation between speaker and listener. In conversation, speaker roles are more flexible. Listeners in conversation are often communicative, while speakers may "listen" by interpreting cues (either verbal or nonverbal) to assess the listener's understanding, attention, and emotions (Yngve, 1970). Moreover, simultaneous speech, e.g., interruptions, active listening, or terminal overlaps, occurs frequently in natural conversation, although it is usually short-lived, (Sacks et al., 1974). For these reasons, the focus should be activations in *both* speaker and listener at this time, rather than attempting to separate them.

In addition to comparing narrative and conversation, another important step is to develop a comparable baseline task for use in conversation studies. In the previous study, narrative was compared with overlearned nursery rhymes, which may be an appropriate comparison task for conversation as well. Yet, nursery rhyme recitation differs from conversation on a great number of levels. A conversation-specific baseline task should closer match conversation by including interaction and a degree of extemporaneity (rather than being completely memorized).

To address this issue, we have devised a comparison task in which participants share unrelated expository information on something they are familiar with, but have not memorized. Participants insert turns-at-talk by interrupting at intervals of their choosing. Also, participants are not to address one another directly or respond anything the other person says. Such a task is devoid of true social interaction. And while it retains the overall "shape" of conversation (by consisting of alternating turns-at-talk), it does not require the same degree of collaboration and interdependence.

Group ICA is used to reduce the data and provide self-organizing networks with unique time courses and spatial patterns. These networks are used to reveal neural correlates of conversation, instead of functional connectivity. This method differs from the traditional GLM approach used in Study 1, which identifies clusters of voxels instead of networks. As such, direct comparisons between narrative and comprehension cannot be done. Nevertheless, qualitative comparisons can shed light on basic similarities and differences between the two tasks. Clearly, regions fundamentally related to language should be engaged during conversation, as they were in narrative. Also, extrasylvian regions will be recruited, particularly networks pertaining to mentalizing and inference making. There should also be some differences, possibly related to coordinating behavior between two individuals.

Methods

Participants

Study participants were eighty right-handed, adult, native speakers of American English (44 female, 36 male; aged 21 to 33 years), arranged into 40 gender-matched pairs of interlocutors. Subjects were free of neurological and psychological disorders. All provided written consent in accordance with National Institutes of Health Institutional Review Board under protocol NIH 92-DC-0178.

Experiment Design and Training

The subjects engaged in four completely unscripted conversations, each of which was designed to model differing types of conversation.

During Autobiographical (AB) conversation, participants discussed their personal experiences in college. In order to maintain a certain level of novelty throughout all pairs and to ensure that none of the pairs were recounting shared experiences, friends who knew each other well in college discussed vacations they took separately. The AB conversation served to model a typical social interaction in which participants share information pertaining to themselves and their personal lives.

In the Debate (DB) conversation, participants informally argued the pros and cons of immigration policies in the United States. The participants assumed opposite positions, which did not always reflect their personal opinions. They were instructed to maintain their positions for the duration of the conversation. Unlike AB conversations, DB is goal oriented in that each presents his argument and counters those raised by one's opponent.

In Problem Solving (PS) conversation, the subjects collaborated on developing a plan to survive being stranded on a tropical Island. They were given a hypothetical situation (provided as part of training materials in Appendix V) minutes before undergoing scanning and were instructed to discuss the situation as if it were real. This conversation forced the interlocutors to cooperate, plan, and think creatively in order to reach a mutually beneficial conclusion. Importantly, in this conversation, it was more critical that speakers align themselves to the same mental representation, essentially "staying on the same page".

The fourth conditional was the Conversational Control (CC). In this task, each participant spoke on completely separate expository topics. They were given the choice of only one of four possible topics: the American Civil War, the Solar System and Space, Human Physiology, and Earth Science. Participants were instructed to not directly address one another at any point. In order to insert "turns", they interrupted one another whenever they chose, rather than follow any implicit cues typically associated with exchanging turns-at-talk. CC was designed to mimic the alternating, complex speech of conversation. Yet, it was devoid of the cooperative and social features that typically characterize natural conversation.

As an additional control speech task, subjects were also asked to recite overlearned nursery rhymes (NR) and were able to select three NRs with which they were most familiar during the training session.

Each subject underwent training at least one day before, and typically within a week of scanning. At training, they were given verbal instructions on the task. They also selected their topic for CC and a position (pro/con) for DB during the training session. Training materials are provided in Appendix V.

Paradigm Design

Each scanning series, or run, lasted a total of 12 minutes and 6 seconds. Each subject had a turn reciting one NR for 30 seconds with a 16 second interval between. After another 16-second interval, there were of 8 minutes allowed for conversation. Following the conversation, each subject again had a turn reciting one NR, again separated by 16 seconds of rest. Before NR a text page was presented for 1 second, which displayed the name of one of the participants and the NR that was randomly selected for him. For conversations, the visual cue consisted of text revealing the topic. In the case of CC and DB, subjects were reminded during the cue which topic or position, respectively, they had chosen. There were a total of 4 runs, one for each conversation type, the order of which was randomly selected. A white fixation point ("+") was presented at the center of the screen during the entire 8 minutes of conversation and for each 30-second block of NR. Participants were cued to stop speaking with a the word "STOP" in red bold font, which was presented for 500 ms. During rest periods, participants were provided a red "X" as a fixation point. *E-Prime* software 1.2 (Psychology Software Tools, 2002) was used to present the stimuli. A mirror reflection system and a Digital Light Processing (DLP, Texas Instruments, Dallas, TX) projector were used to project the visual cues to the participants.

Data Acquisition

T2-weighted BOLD images were acquired on two General Electric (GE) Signa HDxt 3.0 Tesla scanners (GE Healthcare, Waukesha, WI, USA) with an 8-channel HR Brain Coil. A single-shot gradient echo echoplanar (EPI) sequence was used. The scanning parameters were as follows: TR = 2000 ms, TE = 30 ms, flip-angle = 90°; 64×64 matrix, FOV = 227 mm. Forty interleaved sagittal slices with a thickness of 4 mm were used to cover the whole brain. In addition to the functional data, sagittal T1-weighted high-resolution structural images were acquired using a magnetization prepared rapid gradient echo (MPRAGE) sequence. Participants were scanned simultaneously in two separate MRI scanners (here on referred to as 3T-A and 3T-B). Each scanner was outfitted with equipment that allowed the participants to hear one another. 3T-A had a STAX electrostatic earbud audio system (STAX Limited, Saitama, Japan) and Optoacoustic head mounted fiber optic noise canceling microphone (Optoacoustics, Or Yehuda, Israel). In 3T-B, sound was presented through a pair of Silent Scan 3100 pneumatic headphones (Avotec, Stuart, FL, USA) and voice was recorded by a FOMRI II noise canceling optical microphone (Optoacoustics, Or Yehuda, Israel), which was mounted to the scanner table and positioned near the participant's mouth. Audio was recorded using Audition (Adobe Systems Incorporated, San Jose, CA). Unfortunately, because of technical limitations, participants were only able to hear one another. As a result, extralinguistic cues (e.g. eye-gaze and gesture) cannot be factored into this study.

Data Analysis

Conversation Transcription and Timing

For each pair, individual subjects were recorded in separate audio channels and transcribed using CLAN software and CHAT transcription method. The times for each turn onset, duration, and transition event marker were automatically measured using a MATLAB script that utilized the audio files from both speakers to delineate events. Turns-at-talk were defined as any continuous speech by a speaker.

Functional MRI Data

First, the structural image of each subject was segmented and normalized into MNI space using tissue probability maps in SPM (Ashburner and Friston, 2005). AFNI was used to perform in-plane registration, slice-time correction and volumetric rigid-body registration to the functional datasets using AFNI (Cox, 1996). Traditional motion correction algorithms can correct misalignments due to whole head movements, but not motion-related susceptibility artifacts caused by continuous and overt speech production (Birn et al., 1998). As in Study 1, in order to correct for susceptibility artifact, spatial independent component analysis (sICA, (McKeown et al., 1998) was applied to the motion and slice-time corrected functional data on each subject level using GIFT (Group ICA for fMRI Toolbox), a MATLAB (Mathworks, Natick, MA, USA) toolbox. In sICA, each BOLD image was treated as a mixture of multiple spatially independent signal and noise sources. The number of components in each dataset was estimated by minimum description length (MDL) criterion (Li et al., 2007). The systematic classification of artifactual and neuronal ICA components was performed by the same human raters in Study 1, based on the same criteria. The noise components and their variances were subtracted from the original dataset. The remaining components were added together to construct the denoised BOLD data. After denoising with sICA, data were normalized into MNI space at a voxel size of 3 x 3 x 3 mm by applying the transforms derived from the structural image normalization, and smoothed to a target full-width-half-max (FWHM) of 8 mm. Turns-at-talk were modeled as variable-length blocks, either production (p) or comprehension (c). NR was modeled as blocks of a fixed duration.

Group Level Independent Component Analysis

As in Study 1, group-level ICA was applied to these data. The exact methods were followed to acquire these data with two exceptions. Three PCA steps were used to reduce the data, instead of two as in the previous study. Also, 30 independent components were yielded using GIFT, each of which has a unique time series. Of these, four components were identified as noise, based on the same criteria presented in Study 1, and removed from the reconstructed BOLD signal. These criteria included spatial patterns localized to ventricles, dural vein sinuses, and cerebral arteries.

Applying GICA allows for statistical separation of these sources into independent components (ICs) *before* running more traditional general linear modeling (GLM) analyses. Additionally, GICA reduces the data to fewer components (i.e., 30 self-organizing networks, as opposed to thousands of voxels), which may be more manageable. It is important to point out that GICA produces components from the entire group (i.e., aggregate data). However, through back-reconstruction, components for the individual subject can be computed. The back-reconstruction uses the input BOLD data, aggregate data, and data reduction steps to estimate the spatial and temporal characteristics of components at the individual level.

To demonstrate the reliability of back-reconstructed data in this task, Pearson's correlation coefficient was calculated between 1) the back-reconstructed time series of the component spatially corresponding to bilateral motor cortex (see Appendix VI) and 2) a predictor of BOLD activity related to speech production onset and offset for a single randomly selected subject. The results demonstrate a strong correlation between the two time courses ($r(352) = 0.7485$, $p < 0.0001$). For further analyses, we use back-reconstructed ICA time series (i.e., time courses for each participant that indicate temporal variations of mixing weights across time and are derived from both the aggregate ICA data and the individual subject's BOLD data). So, while the components do not reflect direct measures of brain activity or changes in blood flow, they are derived from and strongly correlated with BOLD signal.

ICA-Based GLM

Subject-level general linear modeling (GLM) was computed using back-reconstructed ICA time courses, again using REML estimation based on a canonical HRF in SPM. Using SPM and MATLAB, a "dummy" header was created, providing standard spatial information like image size, dimensions, offset, etc. The selected variables were identical to real functional MRI datasets, with the exception of the image size, which was set to 2 x 3 x 5 voxels (a total of 30 voxels, one for each independent component). The image file (.img) consisted of the time series of each of the 30 components. The result was a four-dimensional dataset (i.e., 2 x 3 x 5 voxels x number of time points), which was used for traditional regression analyses. Each voxel represented the time course for a single independent component. The spatial characteristics of the components were not taken into account at this point. Unlike typical fMRI datasets, the position and size of each voxel is arbitrarily (but not randomly) set. To identify the "location" of each component was compared on a voxel-by-voxel basis with the original time course from each component until each voxel was matched to a particular component. Voxels representing noise components were masked, and not considered in the analysis.

Group Level GLM

For ICA-based GLM, production periods were modeled and contrasted first with NR baseline then with production periods during CC, both using t-tests on a voxel by voxel basis. The same contrasts were performed for comprehension periods (i.e., listening when the other speaker has the floor or when the other person is reciting an overlearned nursery rhyme). The resulting t-maps were thresholded at p = 0.05, with no correction for multiple comparisons because of the limited number of voxels.

Results

Turns at talk: Shared Activations, Comprehension and Production vs. NR

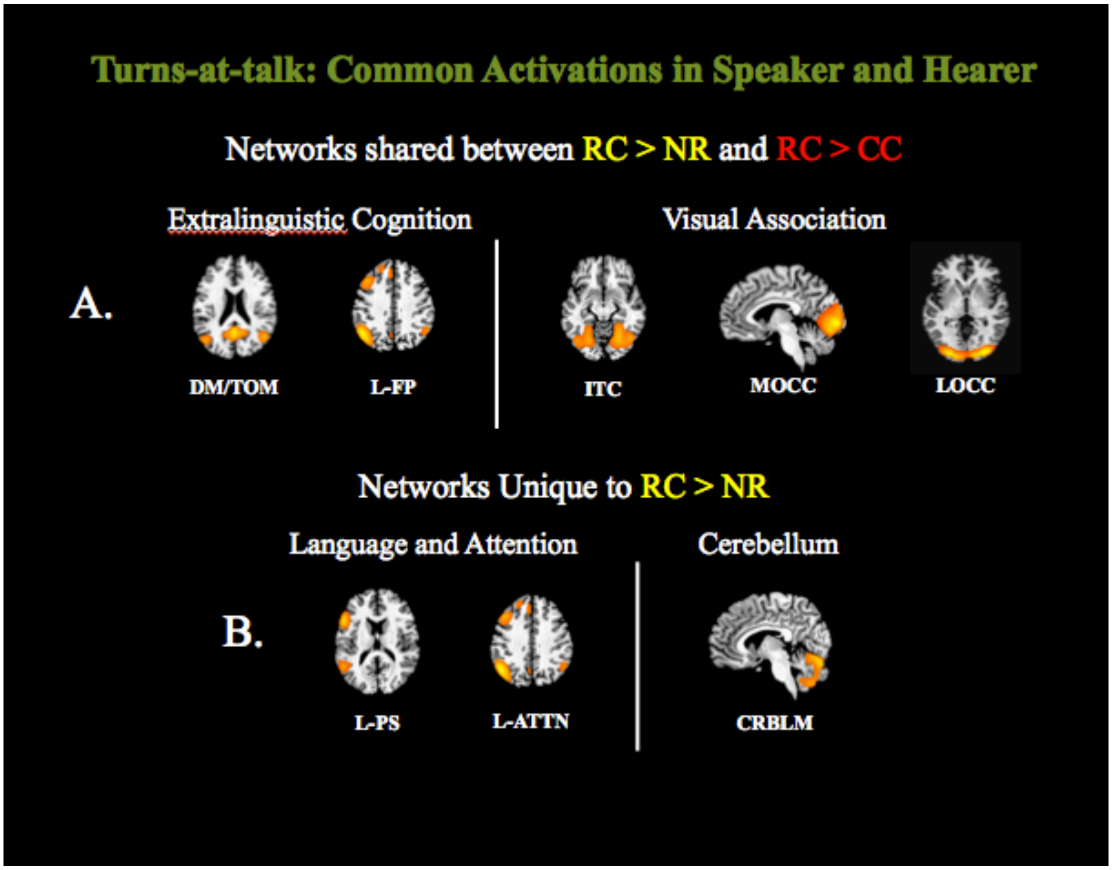
Comprehension and production during all the conversations (i.e., PS, AB, DB) – referred to as "real conversation" comprehension (RCc) and production (RCp) – was contrasted with either nursery rhyme comprehension (NRc) or production (NRp) respectively. Eight ICs were shared between these contrasts, demonstrating greater activity during *both* comprehension and production during conversation than nursery rhymes (RCc > NRc and RCp > NRp), presented in Figure 6 and Table 2.

These included and left lateralized perisylvian network (L-PS), which consisted of the inferior frontal gyrus (IFG), as well as a cluster with a peak in the middle temporal gyrus (MTG) and extended into the superior temporal sulcus (STS) and posterior STG, all regions traditionally associated with language processing.

Several extrasylvian, cognitive networks were also activated. One network (DM/TOM) had a maximally active voxel in the posterior cingulate cortex (PCC) and included bilateral inferior parietal lobule (IPL), dorsal medial prefrontal cortex (dmPFC) and parahippocampal gyrus – all regions typically associated with both the default mode and theory-of-mind networks (Spreng et al., 2009). Additionally, a left fronto-parietal network (L-FP) was more engaged and was comprised of the left angular gyrus, left dorsolateral prefrontal cortex (dlPFC), precuneus, and right angular gyrus. Another network (L-ATT) consisting of the left middle frontal gyrus and intraparietal sulcus (IPS) was also significantly engaged.

In addition to these, RCp and RCc significantly activated three networks in either the inferior temporal or occipital cortex, likely related to vision. The first of these (ITC) was comprised of the bilateral inferior temporal cortices, with primary activation in the fusiform gyrus. The two other ICs consisted of regions in the occipital cortex, specifically the medial (MOCC) and posterior-lateral (LOCC) occipital cortex.

Additionally another network consisting of the cerebellum was significantly engaged.



Turns-at-talk: Common Activations in Speaker and Hearer

Networks shared between RC > NR and RC > CC

Networks Unique to RC > NR

Figure 6. Axial and sagittal brain slices representing A) the networks that are significantly engaged during comprehension and production after both contrasts (RC > NR and RC > CC) and B) the networks that were only found when conversations were compared to nursery rhyme (RC > NR).

Table 2.
Conjunction: Comprehension and Production In Conversation

Network	Anatomical Description	X	Y	Z	T-score
RC > NR					
Language Related Network					
L-PS	Left inferior frontal gyrus, middle temporal gyrus, superior temporal sulcus	-47	19	-2	5.64
Extralinguistic Cognitive Networks					
DM/TOM	Medial prefrontal cortex, precuneus, bilateral IPL, parahippocampal gyrus	0	-53	17	2.45
L-FP	Left and right angular gyrus, left dorsolateral prefrontal cortex, precuneus	-47	-63	34	4.56
L-ATT	Left middle frontal gyrus and intraparietal sulcus	-44	10	33	2.95
Visual Networks					
ITC	Bilateral medial inferior temporal lobe	26	-72	-17	1.05
MOCC	Medial occipital lobe	-6	-90	-1	3.44
LOCC	Lateral, posterior occipital lobe	32	-87	-1	2.39
Other Regions					
CBLM	Cerebellum	29	-68	-23	2.19
RC > CC					
Extralinguistic Cognitive Networks					
DM/TOM	Medial prefrontal cortex, precuneus, bilateral IPL, parahippocampal gyrus	0	-53	17	6.97
L-FP	Left and right angular gyrus, left dorsolateral prefrontal cortex, precuneus	-47	-63	34	7.20
Visual Networks					
ITC	Bilateral medial inferior temporal lobe	26	-72	-17	4.09
MOCC	Medial occipital lobe	-6	-90	-1	3.57
LOCC	Lateral, posterior occipital lobe	32	-87	-1	3.86

Table 2 (see above) Table contains the networks significantly activated during *both* all comprehension and production when compared to either NR (RCp > NRp *and* RCc > NRp, indicated in the table as **RC > NR**) or CC (RCp > CCp *and* RCc > CCp, indicated in the table as **RC > CC**). Also reported are the XYZ coordinates of the peak voxel for each component (taken from the aggregate data), as well as the minimum t-score (that is, between NAp and NAc) for each network.

Turns at talk: Shared Activations, Comprehension and Production vs. CC

To explore how CC compared to NR as a baseline task for a conversational task, RCp and RCc were contrasted with production (CCp) and (CCc) comprehension respectively. This resulted in five significant components, all of which were also significantly more engaged when compared to NR (i.e., RC > NR, see Table 2). These networks were the DM/TOM, L-FP, ITC, LOCC, and MOCC networks, illustrated in Figure 6 and Table 2.

Discussion

Narrative and Conversation: Shared Findings

In order to compare comprehension and production during narrative to that during conversation, conversational turns-at-talk were compared with a nursery rhyme baseline task – the same baseline used in the narrative study. However, these are obviously separate analyses, using different methods. The previous study on narrative utilized traditional voxel-wise GLM methods, while the conversation data are ICA based. As a result, for now, any comparisons can only be qualitative. Despite these obstacles, there were striking similarities in the neural correlates of narrative and conversation that could be quantified with future studies.

Language Related Neural Correlates of Conversation

Importantly, as expected, conversation recruits a left lateralized network consisting of the inferior frontal gyrus, superior temporal sulcus, and middle temporal gyrus. As discussed in the previous chapter, all of these regions are commonly observed in languages studies, from single word processing to narrative and now, conversation. These regions support multiple facets of phonological, lexical, and sentential language processing, all of which are essential in producing and understanding language during conversation (Acheson et al., 2010; Binder et al., 1997; Friederici et al., 2006; Moro et al., 2001; Okada and Hickok, 2006; Rissman et al., 2003; Visser et al., 2010). In the previous chapter, I mentioned the possibility that the activation of these left perisylvian regions above NR – another linguistic task – is due to additional discourse-specific processing related to higher-level semantic retrieval and integration, semantically appropriate lexical retrieval (Eviatar and Just, 2006; Hagoort, 2005; Marini and Urgesi, 2012; Shibata et al., 2010; Uchiyama et al., 2006). In addition, it was earlier suggested that increased activation of these regions was in part due to extemporaneity inherent to discourse, particularly when compared to overlearned speech.

Here it is important to note that the left perisylvian network (L-PS) did not significantly differentiate natural conversations and CC, also a discourse-level, spontaneously generated task. When taken together with increased activation of the L-PS network when compared to NR, this finding supports the view that traditional language regions play an additional role in discourse processing and that this role may not be specific to conversation.

Involvement of Social and Cognitive Networks

The DM/TOM network consisted of the medial prefrontal cortex, retrosplenial cortex, bilateral parahippocampal gyrus, bilateral superior frontal gyrus, and bilateral IPL, while the L-FP network included left fronto-parietal regions with peak values in the angular gyrus. The majority of these regions comprise the default mode network (Saxe and Kanwisher, 2003; Spreng et al., 2009). This network of regions is often associated with periods of "rest", yet a growing body of research, including data presented in Study 1, demonstrates a role for this network in discourse processing (Fletcher et al., 1995b; Mar, 2011).

Importantly, there is notable anatomical overlap between the default mode network and regions supporting theory-of-mind (Spreng et al., 2009). It is likely that in dyadic conversation, these networks support establishing and maintaining representations of the mental states of oneself and others, undoubtedly a crucial element of coherence building in social interaction. With either contrast (NR or CC) there is an increase in the both the DM/TOM and L-FP networks during conversation. Although one could argue both comparison tasks were discourse-level, they were both void of direct social interaction. Even though at first glance CC, which involved alternating turns-at-talk, may appear interactive, it important to note that exchanges was interruptive – unilateral, rather than cooperative. Additionally, during CC all participants shared impersonal information that was unrelated to what the other person said.

As discussed in the previous chapter, some of these regions in the DM/TOM network (particularly the medial prefrontal cortex, precuneus, and bilateral IPL) may participate in a more general inference-making mechanism. In natural conversation, the ability to draw inference is vital. Some common phenomena require drawing inferences, such as the use of sarcasm, metaphors, and idioms. Yet it is also important to note that very few statements are perfectly explicit. Instead, they require integration of verbal, prosodic, and contextual cues. Also, like all implied statements, they are grounded in assumptions of the other's mental framework (Grice, 1975). But these heteromodal regions may also support building a multidimensional mental model or representation of situational context created within each conversation (Ferstl et al., 2005; Martin-Loeches et al., 2008; Whitney et al., 2009; Yarkoni et al., 2008). Although this phenomenon is traditionally applied to narrative (Johnson-Laird, 1980; van Dijk and Kintsch, 1983), it may apply to conversation which also requires coherence, connectedness between events and sentences (as well as turns-at-talk) – albeit these are constructed by at least two people.

At this time, it is difficult to tease apart all the possible contributions of these networks, i.e. whether they are related to social cognition, the default mode, inference-making, or situation-model building. Yet, future studies can help. For example, taking a closer look at conversational topic shift and management may help tap into neural mechanisms supporting the maintenance of a mental model in conversation. As another example, if one or both of these networks were involved in inference making, one would reasonably expect increases in activation during more implicit statements. However, the DM/TOM and L-FP networks include such functionally diverse regions that they may support all or any combination of these critical functions.

Narrative and Conversation: Divergent Findings

While there is considerable anatomical overlap between narrative comprehension/production and listening/speaking during a conversation, there are also some differences.

When compared to NR, participating in conversations engaged the L-ATT network, which included the MFG and IPS. The MFG cluster in the L-ATT network does overlap with a cluster found in narrative production and comprehension, extending from a peak in the dorsal premotor cortex. However, the IPS was not engaged during narrative comprehension and production. Both these regions are often recruited in tasks requiring directed attention, particularly switching attention (Pessoa et al., 2009; Rossi et al., 2009; Salmi et al., 2009; Yamasaki et al., 2002). Activation of this network suggests that the fluid and rapidly changing nature of conversation may require more shifts in attention. The target of this attention may be abstract phenomena, like topic shifts. But this network could also facilitate attending to the nonverbal social cues constantly produced and interpreted during conversation and transmitted through one's voice, facial expressions, gestures, etc.

Yet another finding unique to conversation is the involvement of the visual cortex for both comprehension and production, as opposed to just production, as was the case in the previous chapter. Although the occipital cortex clearly mediates visual processing, some researchers find a more cognitive and social role for it. Some have suggested a role for the visual cortex in directed attention, even in the absence of visual stimuli (Kastner et al., 1999). Others notice involvement of the occipital cortex in social situations. Altered functional connectivity between the occipital and frontal cortices has been found in those with social anxiety disorder (Ding et al., 2011). The occipital lobe, specifically the cuneus, has also been linked to processing emotional prosody (Sander et al., 2005) and is increased when reading emotional words (Fossati et al., 2003). Additionally, there is evidence that the occipital cortex is engaged in inference-making (Mason and Just, 2011).

However, the precise role of the occipital cortex in these non-visual social, cognitive tasks is unclear. Mason and Just (2011) speculate engagement of the occipital cortex relates to embodied cognition and that drawing inferences about occurrences may draw upon one's experiences of similar situations or sensations. In this study, in which participants are engaging in social interaction without visual input, the occipital cortex may relate to mental imagery and visual imagination rather than drawing upon or reliving experiences – although, undoubtedly mental imagery is rooted in experience.

That we also found increased activation of a component made up of the fusiform gyrus, including the fusiform face area, lends support to this suggestion. Facial expression and gestures are important elements of interpersonal communication. In conversations for which there is no visual input (such as telephone calls), interlocutors may compensate by visualizing paralinguistic information that typically accompanies and enriches conversation. Additionally, others suggest the fusiform gyrus is involved in social cognition, beyond facial recognition and expression (Schultz et al., 2003).

Unfortunately, the exact involvement of the visual and inferior temporal cortices cannot be fully explained with this study. Future studies of naturalistic conversation in which the presence or absence of facial expression is experimentally manipulated are necessary.

Use of Conversation Control Task (CC)

For this study, we designed a new task to use specifically as a comparison to natural conversation. This task was designed to share particular aspects of conversation, e.g. be spontaneous generated (as opposed to overlearned), complex, and language based. Additionally, we wanted this task to consist of participant-driven turns-at-talk. However, CC was also designed to be a task in which, unlike natural conversation, there is no need (or opportunity) to maintain a shared thread of consciousness, to establish and build common ground, or to consider the other person's degree of understanding – all of which are critical to communication. During this task, participants were to neither address nor respond to the other person. In addition, participants were asked to insert turns by interrupting the other person, i.e., they were instructed not to wait for a transition relevance place (described further in following chapter).

As a baseline task, CC resulted in similar findings as NR when contrasted with conversation periods, such as relative de-activation of social and visual networks. Yet, there were a few exceptions. One exception is L-PS, the network consisting of left perisylvian regions typically associated with language. Considering that CC is extemporaneous and linguistically varied (as opposed to repeated like NR), it is not surprising that there is left perisylvian regions do not differ between conversation and CC. Another exception is that when compared to conversation, CC did not differ in L-ATT, the left-lateralized attention network. This suggests that either CC also engages attentional systems or, at least, CC does not involve de-activate them. Still, this study seems to demonstrate that, as compared to conversation, NR and CC lead to very similar results.

What is gained by CC is that, unlike NR, it has a similar structure of natural conversation. Both necessarily entail variable length turns that conclude with speaker-change. That means that events within conversation, such as turn-transitions, can be compared between conversation and CC but not NR.

Summary

There are many similarities between narrative and conversation, two components of discourse-level communication. For example, they engage both traditional language network and regions beyond that, particularly pertaining to social cognition. However, differences are important to point out. Conversation engaged attentional and bilateral inferior temporal networks not seen in narrative. Also, the occipital cortex was engaged during both comprehension and production during narrative processing, but not during narrative. It is important to note when considering these qualitative comparisons between narrative and conversation that one cannot rule out the possibility that some of these differences are associated with methodological differences (i.e., voxel-wise GLM vs. ICA-based GLM). For better comparisons, ICA-based GLM should be conducted with narrative comprehension and production.

In this study, I also introduce a new baseline tasks CC to be used in conversation studies. This task resulted in comparable findings when compared to conversation as NR did. Yet, CC is structurally closer to conversation, making it a better comparison task.

In this chapter, I examined the brain regions supporting both comprehension and production during narrative and conversation. However, it is difficult to paint a full picture of natural conversation without looking at some of the features that are unique to this type of communication. In the next chapters, I will explore conversational turn taking and inter-brain connectivity.

Study 3: Conversational turn-taking

Introduction

Conversation is typically unscripted and unpredictable. However, after years of observational research, conversation analysts have found that conversation is actually very structured and governed by a complex set of (largely) implicit rules (Sacks et al., 1974). Some rules pertain to the structure of the conversation (e.g., coordinating speaker change). Others relate to preserving the coherence or logical flow of the conversation (e.g., when and how to shift from one topic to another).

Conversational Turn-taking

Perhaps not surprisingly, many of these unspoken rules relate to conversational turn taking, which is a hallmark of conversation and does not exist at any other level of communication. Successful turn taking, like the rest of conversation, requires the drawing of another's attention to a task by the speaker, similar to joint-attention tasks. Additionally, conversation necessitates coordinating speaker change. Although this often takes place without much conscious effort, turn-taking demands predicting another's behavior (e.g., ending a turn) and planning one's own turn either simultaneously or in very short succession. The above-mentioned rules facilitate this synchronization of behavior.

Social interaction occurs largely because of and through turn taking, which also provides much of the structure of conversation. Clearly, turn taking is a ubiquitous element of human communication. So much so that some argue that it is an innate human characteristic to look to align ourselves with others through speech (Garrod and Pickering, 2004). Simply from our experiences with communication, turn taking is undeniably crucial to the daily experience of almost every human. As such, it is important to understand the biological basis of it.

Sacks et al. (1974) are the first to systematically characterize behaviors supporting conversational turn-taking and list “rules” governing conversation, which are more like principles based on observed norms.

One of the most essential principles is that speaker-change reoccurs and takes place at transitional relevance places (TRPs), which are periods, implicitly agreed upon by both speakers, at which the likelihood of speaker change increases. That is to say, a turn-transition does not typically happen at random places in a conversation. Rather, there is a set of features that characterize TRPs. Sack et al. also pointed out that turn order is not specified. Instead, in conversations with more than two interlocutors, any speaker can claim a turn at a TRP (unless explicitly selected by the current speaker). Also, turn length is not fixed. While this is an obvious fact, it is worth noting that this feature adds a level of complexity to the predictive nature of conversational turn taking.

Another apparent but important feature outlined by Sacks et al. (1974) is that ordinarily there is only one speaker at a time. It is this one-speaker-at-a-time rule that forces a structure based on turn taking. When a speaker has the floor, he maintains it until a TRP at which another speaker may claim the floor. While there are often instances in which more than one person is speaks at the same time, these periods are usually short-lived. Additionally, overlapping speech is usually managed by another set of rules.

Mechanisms of Conversational Turn-taking

Other than interruptions, turn transitions take place at Transition Relevance Places (TRPs). Importantly, TRPs can re-occur within a turn-at-talk and are present whether or not a switch in turns actually takes place (Sacks et al., 1974). At a TRP either another speaker can claim the floor or the current speaker can continue his turn. In other words, according to Sacks et al. (1974), all transitions take place at TRPs, but not all TRPs result in turn transition.

While all conversational analysts agree in their existence, TRPs have yet to be clearly characterized. There is still much debate on exactly what constitutes a TRP, how long they last, and, importantly, which cues carry the most information (i.e., are most likely to influence prediction of speaker change).

Some argue that prosodic cues are most essential. For example, Wells and Macfarlane (1998) identify accent patterns that are only present where turns transitions take place and argue that these acoustic modulations are integral to predicting the end of a turn. However, this result does not fit well with the notion that not all TRPs result in speaker change (Sacks et al., 1974). In a more recent study, de Ruiter et al. (2006) found that with only lexicosemantic (or textual) information, participants were able to make similar predictions about turn transitions as those who heard the conversation as it naturally occurred. The authors concluded that lexicosemantic information change was the most important cue. Many others have suggested that a combination of linguistic and paralinguistic factors is most essential to signaling turn-switching (Caspers, 2003; Duncan, 1974; Selting, 1996). Still others suggest that the important features of turn-taking change developmentally (Keitel et al., 2013).

All this goes to show how complex this issue is, although it may appear simple at first glance. Although the exact mechanisms remain unresolved, it is agreed that interlocutors are using primarily implicit cues to signal, anticipate, and synchronize turn transitions. As evidence of this, a study by Stivers et al. (2009) sampled ten languages from cultures with differing social norms and linguistic characteristics and measured the time between turns. They found no significant differences between the amount of time between the offset of one turn-at-talk and the onset of the next, with an average for each language ranging between 0 and 200 ms and with a mode of 0 ms. Such a small time window obviously does not allot for formulation and execution of the linguistic and articulatory-motor planning necessary to claim a turn-at-talk (Levelt et al., 1991). Therefore, it follows that, during natural conversation, interlocutors are anticipating transitions (likely facilitated by cues provided during or just before TRPs), shifting their attention to the cues preceding a TRP, and preparing for their turn while the other speaker still has the floor.

Predictions

Due to the social nature of conversation and the predictive quality of turn-taking in particular, one can reasonably expect the involvement of cognitive functions that support intuiting the intentions and mental states of oneself and others. Consistently, regions in the medial prefrontal cortex (MPFC), bilateral temporoparietal junction (TPJ), posterior cingulate cortex, precuneus and anterior superior temporal sulcus (STS) are active during mentalizing tasks (Gallagher et al., 2000; Saxe and Kanwisher, 2003; Spreng et al., 2009). I hypothesize that regions related to social cognition, which should be engaged by conversation and may also be recruited during TRPs, which require drawing upon information to infer (or predict) when one will release a turn. TRPs may also require attention to the implicit cues that accompany them, in which case attentional systems should also emerge at this time period.

Methods

The participants and data collection and processing methods are identical to the previous study. The same ICs are entered into ICA-based GLM. The difference here is that transitions are being modeled rather than turns-at-talk. Transition points were defined as instantaneous events (duration of 0.0 s) at the end and beginnings of turns, i.e. when one either claims or releases the floor. Although TRPs can occur within a turn-at-talk (Sacks et al., 1974), they are still exceedingly difficult to identify reliably. Because of this, only TRPs preceding a completed turn-transition were identified. Transition periods during all of the natural conversations (RCtrans) were averaged together and compared with transitions during CC (CCtrans) with a voxel-by-voxel t-test. The resulting t-map was thresholded at $p = 0.05$.

Results

All transition periods during the real conversations were compared to those during CC, resulting in 10 components. Five of these were the same components found to differentiate real conversations from CC during comprehension and production turns, suggesting a more generalized difference between the real conversations and CC. They are DM/TOM (0, -53, 17; $t = 8.71$), L-FP (-47, -63, 34; $t = 10.90$), ITC (26, -72, -17; $t = 7.20$), MOCC (-6, -90, -1; $t = 4.14$), and LOCC (32, -87, -1; $t = 2.62$).

However, there were five components that were found to differentiate the real conversations from CC uniquely during transition periods, and these may underscore brain regions related to coordinating transition points (see Figure 7A). The maximum voxel in the first component was located in the dlPFC (-44, 10, 33; *t* = 10.28). The component also included the left intraparietal sulcus (IPS). Another IC consisted of bilateral IPS (23, -72, 51; *t* = 5.83). Another IC consisted of the dorsal precuneus (0, -66, 57; *t* = 3.07). There is also component consisted of bilateral caudate and putamen (23, 4, -9; *t* = 5.26). The last component exhibited a peak voxel in the right dorsal precentral gyrus (35, -23, 67; *t* = 3.24).

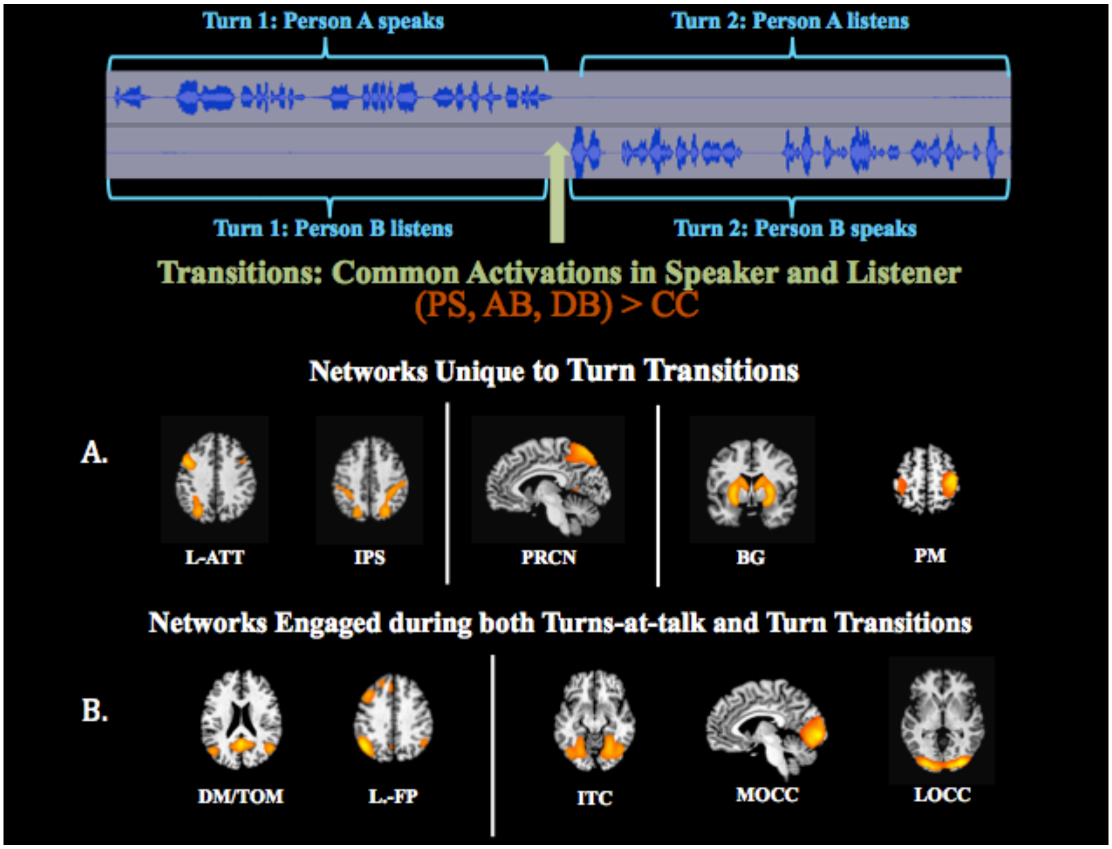


Figure 7. Results from ICA-based GLM of turn-transitions during real conversation compared to those from CC (RCtrans > CCtrans). Five networks (A) are unique to turn transitions. Five networks were identical to those found to be active during comprehension and production (B).

Discussion

The goal of this study was to shed light on the brain regions supporting turn taking, a task unique to interactive communication. To do so, we contrasted transitions (that is, the periods when a speaker either begins or ends his turn-at-talk) during natural conversations and a control task in which participants interrupted one another to insert turns rather than wait for cues indicating a transition relevance place.

Activations Shared with Comprehension and Production

Of the ten networks that were significantly engaged during transitions, half of them were the exact networks that differentiated speaking and listening during actual conversations from CC. As discussed in the previous chapter, these networks may be involved in multiple aspects of conversation, from mentalizing to building a coherent framework or mental imagery. That comparisons of both turns-at-talk and transitions result in these components suggests that they reflect general and sustained differences in the tasks themselves. These results demonstrate that throughout natural conversations (whether speaking, listening, or coordinating speaker change) these five networks underpin communication and social interaction, rather than language formulation or comprehension, or even discourse processing itself.

Activation Unique to Turn Transitions

Turn taking is the predominant feature of conversation and one of its most intriguing phenomena. Across cultures and languages, turn transitions are usually well-coordinated and extremely short (Stivers et al., 2009). Conversation analysts agree that cues – implicitly transmitted by the speaker and interpreted by the listener – facilitate the prediction of potential periods for speaker change (i.e., transition relevance places). Although it remains unknown which features are most essential, these implicit cues can be based in intonation, volume, and syntax. Nevertheless, until now, we could only speculate on the neural mechanisms supporting this process.

We found that transition periods uniquely engaged networks consisting of the bilateral intraparietal sulcus (IPS) and dorsolateral prefrontal cortex. These regions contribute to the dorsal attention network (DAN), which mediates top-down orienting toward external stimuli (Corbetta and Shulman, 2002; Vanhaudenhuyse et al., 2011). The DAN (also known as Task Positive Network) is engaged in tasks requiring directed attention, and it typically co-occurs with deactivation of the medial prefrontal cortex, bilateral IPL, and precuneus, regions associated with the default mode network (Fox et al., 2005; Raichle et al., 2001). Some proffer this as evidence of an intrinsic internal-external (or self-other) dichotomy in neuronal networks (Vanhaudenhuyse et al., 2011). Yet, during turn-transitions these networks are co-activated, demonstrating that these systems are neither completely exclusive nor necessarily anticorrelated. Instead, social and attentional systems are simultaneously activated.

In turn-transitions, activation of the attention network might reflect attending to the implicit cues preceding turn transitions. However, some of the same regions (bilateral frontal cortex, intraparietal sulcus) are also associated with joint attention (i.e., directing or meeting another's focus), the foundation of social interaction (Callejas et al., 2013; Redcay et al., 2012; Williams et al., 2005). Indeed, brain regions supporting mentalizing and other elements of social cognition (e.g. medial prefrontal cortex, precuneus, posterior STS, IPL) are also linked to either initiating or responding to joint attention (Bristow et al., 2007; Laube et al., 2011; Redcay et al., 2012; Schilbach et al., 2010).

At transition periods, the attention networks may work in tandem with other regions related to both joint attention and mentalizing (e.g. medial prefrontal cortex, precuneus, posterior superior temporal sulcus, and inferior parietal lobule) to either initiate or respond to (depending on conversational role) bids for joint attention transmitted through implicit signals.

The precuneus was also engaged during turn transitions. This region is consistently recruited by both theory-of-mind and self-referential tasks, establishing its importance in social cognition (Fletcher et al., 1995b; Gusnard et al., 2001; Saxe and Kanwisher, 2003; Saxe et al., 2006). It may be in this capacity that the precuneus supports coordinating turn transitions. The precuneus is, however, also closely linked to several other complex, integrative cognitive functions, such as visuospatial imagery, episodic memory retrieval, and consciousness (Cavanna, 2007; Cavanna and Trimble, 2006; Fletcher et al., 1995a; Ghaem et al., 1997; Krause et al., 1999). Importantly, studies involving visual and auditory stimuli have shown that the precuneus also supports voluntary shifts of attention (Le et al., 1998; Nagahama et al., 1999; Shomstein and Yantis, 2004).

The precuneus is both structurally and functionally connected to the intraparietal sulcus and superior and middle gyri (Cavanna and Trimble, 2006; Margulies et al., 2009). The caudate and putamen, which subserve sequencing of speech and cognitive planning (Bohland et al., 2010; Monchi et al., 2006; Schulz et al., 2005), are structurally connected to the precuneus through afferent projections

(Cavanna and Trimble, 2006), as well as functionally connected, as demonstrated through resting state studies (Di Martino et al., 2008; Zhang and Li, 2012).

Given the functional centrality of the precuneus, it may play multiple roles throughout a conversation. Specifically, when one performs turn transitions, the precuneus likely cooperates with attentional and mentalizing networks and the basal ganglia to orient one's focus to verbal and nonverbal signals and coordinate the execution of a turn-transition. The interaction of these networks possibly allows a listener to accurately predict a turn-transition and synchronize his response with the speaker's releasing of the floor. Likewise, the speaker may also be assessing the listener's willingness or readiness to claim a turn, also through implicit cues.

Summary

In this study, we demonstrated that in addition to networks related to social cognition that are also engaged during comprehension and production, coordinating speaker change recruits attentional networks that may facilitate orienting toward signals preceding a turn transition and interpreting the other person's intention to either release or claim the floor. Additionally, we suggest that the precuneus interacts with fronto-parietal regions linked to directed attention and with the basal ganglia to coordinate speaker change.

This first look helps illuminate the cognitive systems supporting coordinating turn exchanges, it also leaves many other questions. For example, what difference, if any, might there be between releasing and claiming a turn? Also, provided the regions that are activated during transitions, it may be easier to test hypotheses regarding the presence of within-turn transition relevance places and which linguistic or paralinguistic cues accompany them. Another important next step would be to examine how these neural correlates may change with different types of transitions, for example, interruptions, questions, or unusually long inter-turn pauses – which occur infrequently and may reflect a breakdown in the conversation (Sidnell, 2010).

Study 4: Inter-brain connectivity during conversation

Introduction

Significance of Inter-brain Connectivity

Imaging studies of social interaction often involve individuals usually observing others' interaction or interacting with a computer or experimenter in highly controlled situations. Such studies fail to capture two important elements of social interaction. The first is conversation's extemporaneity. It is made up as it progresses, and none of the participants can predict exactly how the conversation will go, which in large part is due to the second factor: conversation is created by at least two people, who influence one another's actions. Exploring inter-brain connectivity is essential to outlining the manner in which parties work together to produce a unique social experience.

To date, there have been several hyperscanning studies on inter-brain connectedness, most of which use electroencephalography (EEG). Despite employing differing methodologies, many of these studies converge on a similar finding: social interaction increases inter-brain synchronicity (Chatel-Goldman et al., 2013). Some of these studies examined interaction limited to coordinated motion. For example, Dumas et al. (2010) collected EEG from pairs of participants as they spontaneously imitated one another's hand movements. They found that behavioral synchrony correlated with an inter-brain network in the right centroparietal cortex. Another study of musicians playing guitar duets found increased inter-brain coherence in frontal and central parietal electrodes during periods calling for increased musical coordination (Sanger et al., 2012). Yun et al. (2012) reported that after taking part in cooperative interactions (in which one person mimicked the other's finger movements), participants unconsciously continued to match finger movements and demonstrated increases neural synchrony. Interestingly, another EEG hyperscanning study of participants engaging in cooperative games supports the theory that cooperation itself increases inter-brain synchronization (De Vico Fallani et al., 2010). They found that pairs who defect (as opposed to pairs who cooperate with one another) demonstrate fewer inter-brain connections and greater modularity in the theta band. A relatively recent fMRI study of inter-brain coherence during narrative production and comprehension showed that the better participants understood a story, the greater the brain-to-brain coupling between listener and speaker (Stephens et al., 2010)

This research seems to consistently find that social interaction (and particularly, more cooperative interaction) increases inter-brain coupling. We can extrapolate from this that conversational coherence – the building up of common ground, mutual understanding, and shared intention – may also be reflected in increased inter-brain coupling. This ability to establish interpersonal coherence is essential to aligning oneself to the other person and communicating on the same figurative wavelength. Yet, it remains unseen whether this effect of inter-brain connectedness will be observed in fMRI, particularly during natural conversation. Also, other questions remain, such as which brain regions are modulated by behavioral coherence? And how might psychosocial factors influence brain-to-brain connectedness between partners?

The Relevance of Conversation Type

Thus far, we have discussed conversation generically, but as we can attest from our own experiences, conversations vary across many dimensions. The content of a conversation can either facilitate or hinder the development of discourse coherence and a feeling of social closeness. Research has demonstrated that interlocutors report increased positive affect and social connectedness, including feelings of closeness and liking, after conversations containing self-relevant information, i.e., when the interlocutors discuss themselves and their personal experiences (Aron et al., 1997; Sprecher et al., 2013; Vittengl and Holt, 2000). Vittengl and Holt (2000) asked participants to engage in casual, dyadic conversations for ten minutes. It is worth noting that participants were not acquainted before the experiment. Before and after the conversation, participants filled affect questionnaires. After conversations, participants also filled self-disclosure rating scales. The authors found that participants who shared more about themselves and their life experiences were more likely to report positive feelings overall and, importantly, feelings of "social attraction" (i.e. liking the other person and wanting to be friends). Other studies have found that sharing positive emotions extends the self-other overlap between individuals (Waugh and Fredrickson, 2006) and is closely related to developing new and strengthening old relationships (Baumeister and Leary, 1995). In fact, agreement and disagreement are typically performed differently (Goodwin and Heritage, 1990). While one (agreement) is usually instant and emphatic, disagreement usually comes after delays, mitigations, prefaces, etc. associated with minimizing their occurrence. These studies imply a preference toward agreement and affiliation. Moreover, conversations filled with personal content increase affiliation and connectedness in a way that impersonal conversations do not.

In additional to self-disclosure, collaboration may play a role in either developing or maintaining social relatedness. As mentioned earlier, De Vico Fallani et al. (2010) found that cooperative interactions were characterized by increased inter-brain connectivity. When engaged in cooperative conversations (often set around collaborative tasks like solving a puzzle), people tend to subconsciously mirror their partner's body posture more frequently than when engaged in other types of conversation (Shockley et al., 2003). Additionally, Wilkes-Gibbs (1995), who argued that "coherence of language depends on coherence of activity", found that as conversation coherence increased (measured by spoken constructs like co-constructed phrases and shared sentence completion), so did performance on a collaborative task. Specifically, partners who were more aligned took less time to come to consensus on a joint labeling task. Taken together, these studies demonstrate a close relationship between task collaboration and conversation (or social) coherence.

Overall, these studies demonstrate that conversations requiring cooperation, joint goal setting, and affiliation may foster social connectedness, while conversations defined by opposition (like Debate) may not have the same effects.

Each of the three conversation types (PS, AB, DB) alters the nature of social interaction. In AB, interlocutors are purely engaging in self-disclosure. PS, on the other hand, mandates that the interlocutors collaborate to achieve a shared goal. During debate, speakers must refute each other at every turn. Considering previous research, these various types of conversation should exhibit differing patterns of inter-brain connectivity, particularly as they relate to discourse coherence.

Psychosocial Influences on Conversation

Perhaps not surprisingly, behavioral research suggests one's personal disposition can influence his social interactions. Personality traits such as neuroticism (i.e. propensity towards expressing or feeling anxiety, frustration, and anger) and extraversion influence conversational behaviors.

One would expect that anxious individuals, particularly those suffering from social anxiety, tend to participate less in conversation and other social interactions. Multiple studies have shown that, in addition to avoiding social situations altogether, socially anxious people are less likely to initiate conversations, they speak for less time than others, are more likely to avert eye gaze, and are less likely to claim a turn-at-talk (Cheek and Buss, 1981; Schlenker and Leary, 1985). Another study (Leary et al., 1987) asked participants to fill self-report questionnaires then take part in a five-minute face-to-face interaction with a stranger of the same sex. They found that participants who reported being more anxious were more likely to ask questions, produce more acknowledgments and confirmation of the other speakers, and tended to produce fewer of their own informational statements (or edification). A more recent study (Lopes et al., 2003) showed that those who scored high in neuroticism on the NEO Five Factor Inventory (NEO-FFI) were significantly less likely to report having positive social interactions with others, although it remains unclear what underlies this correlation.

In contrast to anxious individuals, extraverts reported significantly more positive social relationships (Lopes et al., 2003). Additionally, compared to pairs of introverts or mismatched pairs, conversations of pairs of extraverts typically cover greater range of topics and are characterized by more indications of common ground (Cuperman and Ickes, 2009; Thorne, 1987). In fact, some evidence suggests extraverts are more able to interpret nonverbal communication cues during social interactions (Akert and Panter, 1988), which may be attributed to more experience.

Literature on brain correlates of personality and social interaction

There is evidence for a biological basis for trait neuroticism. Some suggest that decreased connectivity between the amygdala and anterior cingulate cortex may lead to impaired inhibition of the amygdala's anxiety response (Ormel et al., 2013). However, neuroticism may also be associated with reduced connectivity between prefrontal cortex and both the posterior cingulate cortex and angular gyrus, adjacent to TPJ (Servaas et al., 2013a).

Similarly, imaging studies have identified neurological differences between extraverts and introverts. Extraverts demonstrate a higher degree of functional clustering during resting state connectivity (Gao et al., 2013). Others argue that extraversion is associated with decreased cortical thickness of the inferior frontal cortex and right fusiform gyrus and distinct activation patterns in the amygdala (Canli et al., 2002; Wright et al., 2006). Extraversion is also associated with increased cerebral glucose metabolism in the orbital frontal cortex (Deckersbach et al., 2006).

Clearly, there is not yet agreement on the neural correlates of personality traits. This is due to several confounding factors including varying personality assessment tools, imaging methods, and behavioral tasks. However, there is a growing body of imaging research involving personality traits. What these studies demonstrate is a neurological basis for personality traits known to influence social interaction, and more specifically performance in conversation.

Aims

The goal of this study to shed light on inter-brain connectivity by answering three questions. First, in which networks is inter-brain connectivity modulated by behavioral coherence and whether psychosocial factors (specifically neuroticism and extraversion) influence connectedness between partners? Second, how is this relationship affected by conversation type? To get to the heart of these first questions, we collaborated with linguistic anthropologists at George Washington University to develop a novel measure that is a composite of dozens of behaviors that either contribute to or detract from establishing common ground. The resulting score quantifies the degree of social coherence in conversation. The last question is whether psychosocial factors (specifically, the personality factors neuroticism and extraversion) influence connectedness between partners.

I have used fMRI and group ICA to identify components and calculated the temporal correlation between components' time series. As discussed earlier, most studies of brain-to-brain coupling utilize EEG because of its superior temporal resolution, which allows EEG to identify coupling in frequency bands that are unavailable to fMRI. However, for this study of inter-brain connectivity during conversation, we use fMRI, which has better spatial resolution, and for which there is a method to circumvent artifacts caused by continuous speech. Such a solution has not yet been found for EEG, which is affected by electrical potentials from the mouth and jaw that propagate onto the scalp and interfere with real signal.

Methods

This study was conducted using the same participants as in Studies 2 and 3, 40 gender-matched pairs of healthy, right-handed native English speakers.

Task Design is identical to Chapter 2

Data Acquisition:

FMRI data acquisition is identical to methods described in Chapters 2 and 3.

Personality Inventory: NEO-FFI

Twenty-six pairs (52 total participants) completed the NEO Five Factor Inventory (NEO-FFI, PAR, Inc., Odessa, FL), a self-report questionnaire, to assess personality traits. Although scores were collected for all five factors, only neuroticism scores and extraversion scores were used.

Data Analysis

Personality Inventory

Subjects were separated into two tiers (here on referred to as ‘high’ or ‘low’) in each factor based on K-means clustering. As a result, sorting individuals in high and low groups for each personality factor are relative to our sample. Pairs of interlocutors were then categorized into High (H, two individuals with high scores on a particular factor), Low (L, two individuals with low scores on a particular factor), and Mismatch (M, one person scored high while the other scored low). The inter-brain connectivity was calculated only for High and Low pairs for each factor. There were a total of 14 pairs for Neuroticism (5 High, 9 Low) and 15 pairs for Extraversion (10 High, and 5 Low).

Discourse Coherence Composite Score

Conversations of 19 pairs (limited due to time restraints) were coded by trained raters at George Washington University, resulting in over 45 measures of behaviors known to affect discourse coherence, defined as the building up of common ground (mutual knowledge, beliefs, and assumptions), which is achieved by working together to integrate and interpret conversational cues (Clark and Brennan, 1991; Schiffrin, 1987). The definition of each score is included in Appendix VII. Each score was assigned a valence between -2 and 2 to indicate the degree to which it either contributed to or detracted from overall discourse coherence (Appendix VIII). Measures contributing to the building of common ground included continuers (akin to active listening) and bids for joint action. Disadvantageous measures included trouble-sources preceding conversational repair, i.e. utterances indicating there was a miscommunication or misunderstanding of any sort. The overall discourse composite score (DCS) is the sum of the weighted scores. There is one DCS per conversation type per pair (i.e., AB, PS, DB), a total of 45.

To assess the effect of discourse coherence on inter-brain connectivity in 19 pairs (38 total subjects), an analysis of variance (ANOVA) was calculated between each z-transformed correlation coefficient and composite DCS. The ANOVA includes data across the three conversations.

Group Independent Components Analysis

As in Study 3, group level ICA was applied to BOLD data. However, in this study, subject-level input datasets were shortened to only include the eight minutes of conversation before ICA to ensure the resulting datasets and subsequent analyses were influenced only by conversations themselves (and not nursery rhyme or rest periods). This was performed by removing fMRI images collected before and after the conversation and only retaining the 240 images collected during the conversation. The shortened datasets were entered into GIFT (MATLAB toolbox) with the same parameters as Study 3. Again, 30 independent components (ICs) were collected. Of these, two were identified as noise, based on the same criteria presented in Study 1.

Inter Brain Connectivity

Brain-to-brain connectivity was calculated between the independent components obtained through GIFT. Pearson's correlation coefficients were calculated across the time course for every back-reconstructed IC of each participant and every back-reconstructed IC *of his partner*, resulting again in four (one for each conversation type and CC) 30-by-30 correlation matrices for each of the 40 pairs of interlocutors. In order to prevent possible effects of scanner variability (Montague et al., 2002), the correlation matrices for each pair were then duplicated and transposed (i.e., rotated 90 degrees so that the x and y axes are switched). The correlation coefficients were then converted to z-scores using Fisher's transformation. Further, z-transformed correlation matrices were also entered into an analysis of variance to test the effect of the DCS on inter-brain correlations. All p-values were corrected for multiple comparisons with FDR and a threshold of $q < 0.05$. At this time, I focus on correlations on the diagonal of the matrix, that is correlations between identical networks (e.g. Network A in one person's brain correlated with Network B in his partner's brain). Paired t-tests (restricted to the inter-brain connections on the diagonal) were used to directly compare conversation types as (specifically, PS vs. DB and AB vs. DB), as well as to compare all conversation and CC to a null baseline (empty matrix).

Results

Inter-brain Connectivity: All Conversations

Four components were found to be significantly influenced by conversational coherence. See Figure 8 for a scatter plot of these data.

The first was a posterior medial network (PMN), made up of the restroplenial cortex and precuneus in one contiguous cluster (MNI coordinates of peak voxel, 0, -72, 37). This network also included small clusters in the left and right superior parietal lobule.

The second of these components consisted of right lateralized fronto-parietal (R-FP) network with peak activation in the angular gyrus (AG; 38, -66, 51). Frontal regions of this network included the right middle frontal gyrus (MFG).

A third component was comprised of a large portion of the medial prefrontal cortex (MPFC; 0, 56, 22). The last components (L-PS) consisted of left lateralized perisylvian areas typically associated with language processing. Its peak voxel was in the left IFG (-47, 19, -2), and it included the superior temporal sulcus and middle temporal gyrus, as well as small cluster in the dorsomedial prefrontal cortex.

Not only were all these inter-brain correlations modulated by DCS, but they were also greater than the null condition for each conversation type – with one exception, inter-brain correlations for L-PS during DB.

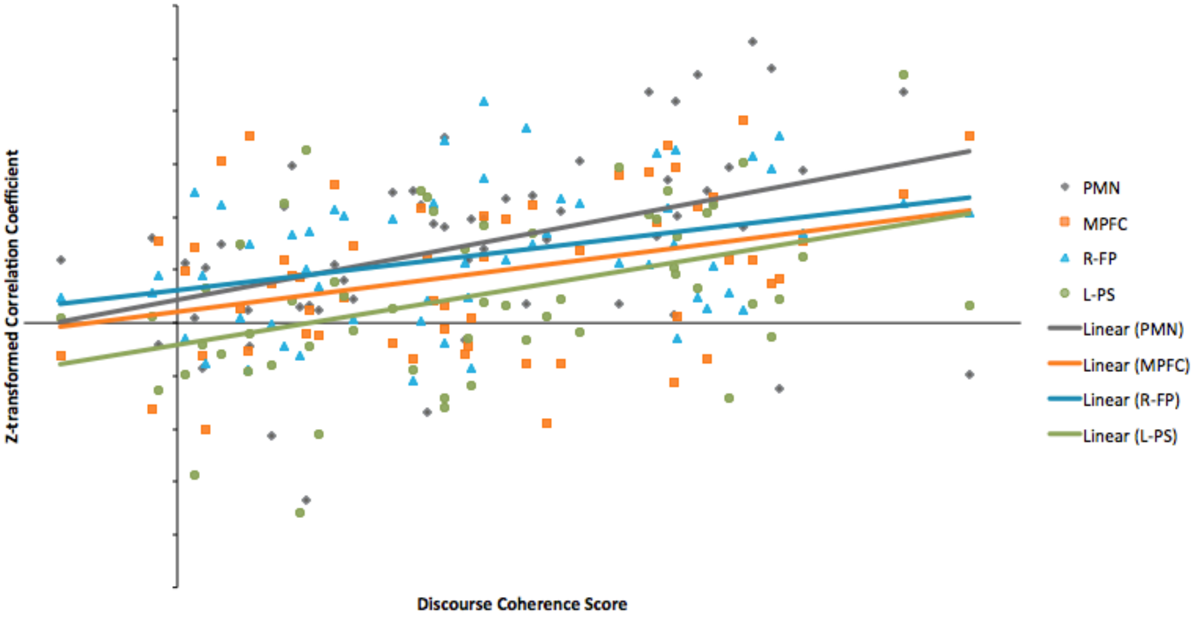


Figure 8. Scatterplot and linear regression for DCS vs. z-transformed correlation for each of the four networks (PMN, R-FP, MPFC, and L-PS)

Interbrain Connectivity and Conversation Type

Each of the four networks mentioned above was tested for differences between conversation types.

The inter-brain connection at the PMN was significantly greater during PS than either of the other conversations or control task. In turn, inter-brain correlation during AB was significantly greater than during either or CC. Last, DB was significantly greater than CC. See Figure 9 and Table 3.

Also demonstrated in Figure 9 and Table 3, at the R-FP network, there was no significant different between PS and AB. However, connectivity at R-FP was greater for both PS and AB than DB. R-FP connectivity was greater during all three conversations when compared to CC.

For the network consisting of the MPFC, inter-brain correlation was significantly greater during PS than all other conditions (Figure 10 and Table 3). On the other hand, connections during AB did not differ from either DB or CC. Also, there was no significant different in DB and CC.

Lastly, brain-to-brain correlation in the L-PS network was greater during PS than any other condition (also presented in Figure 10 and Table 3). Also, connectivity was higher during AB than DB, but did not significantly differ from CC. DB was significantly *lower* than CC.

Table 3. Differences in Inter-Brain Connectivity by Conversation Type

Posterior Medial Network (PMN)

	PS	AB	DB	CC
PS	–	*3.05***	*7.03****	*6.03****
AB	–	–	*3.85****	*3.71 ****
DB	–	–	–	*ns*

Right Fronto-parietal Network (R-FP)

	PS	AB	DB	CC
PS	–	*ns*	*4.13****	*5.22****
AB	–	–	*3.09***	*5.21****
DB	–	–	–	*3.35***

Medial Prefontal Cortex (MPFC)

	PS	AB	DB	CC
PS	–	*3.45***	*2.73**	*4.08****
AB	–	–	*ns*	*ns*
DB	–	–	–	*ns*

Left Perisylvian Network (L-PS)

	PS	AB	DB	CC
PS	–	*4.54****	*6.35****	*3.69****
AB	–	–	*3.07***	*ns*
DB	–	–	–	*-2.42**

Table 3. Data table containing the t-scores and significance resulting from two-tailed t-tests between z-transformed correlation coefficients between each conversation type. This was repeated for each of the four connections (PMN, R-FP, MPFC, L-PS). ***, $p < 0.0005$; **, $p < 0.005$; *, $p < 0.05$; *ns*, not significant.

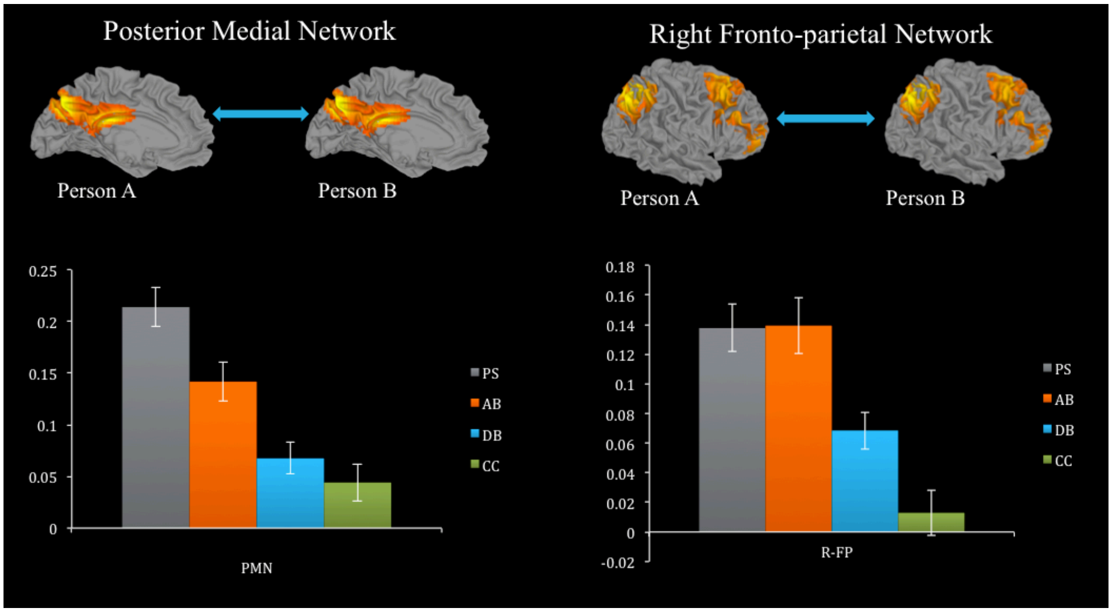


Figure 9. Surface rendering of the aggregates spatial maps for two networks (PMN and R-FP) as well as z-transformed inter-brain correlation coefficients for each conversation type (i.e., PS, AB, and DB) and the control task (CC).

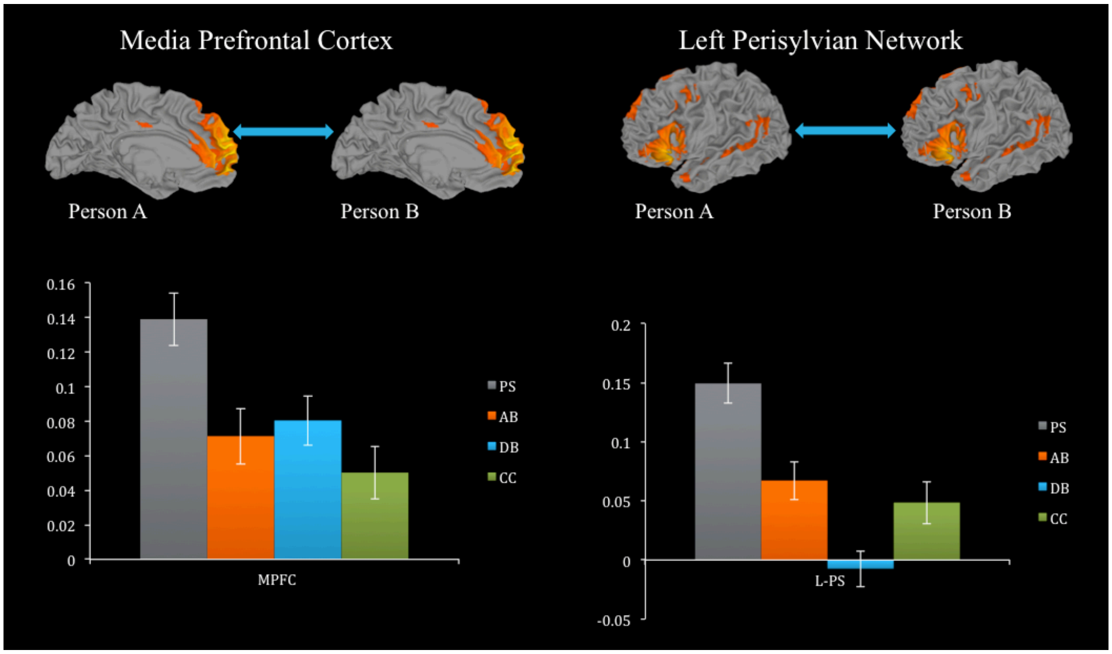


Figure 10. Surface rendering of the aggregate spatial maps for two networks (MPFC and L-PS) as well as z-transformed inter-brain correlation coefficients for each conversation type (i.e., PS, AB, and DB) and the control task (CC).

Overall Inter-brain Connectedness based on Personality Type and Relationship

Pairs of participants on the higher end of neuroticism scale demonstrated significantly fewer brain-to-brain correlations than pairs on the lower end ($t(2) = -11.80$, $p < 0.001$). Extraversion trait demonstrated the opposite effect, where extraverts demonstrated a greater number of inter-brain correlations than more introverted participants ($t(2) = 5.04$, $p < 0.05$).

Discussion

We examined the temporal brain-to-brain correlations between fMRI-based independent components and found that during natural conversations, several networks are functionally connected in both brains. Importantly, these connections are predicted by our behavior-based discourse coherence measure. We also show that inter-brain functional connectivity can be influenced by the type of conversation. Interestingly, we also show that personality traits influence global inter-brain connectivity.

Inter-brain Coupling during Natural Conversation

Four networks were significantly connected to identical networks in the other person's brain and modulated by participants' conversational coherence score. Such inter-brain connections are important in that they indicate regions in which both brains are temporally synchronized through an unknown mechanism, possibly due to the interlocutors orienting themselves and each other to an abstract signal.

Right Fronto-parietal Network (R-FP)

The right lateralized fronto-parietal network, consisting of the angular gyrus (where the strongest signal was located) and the middle frontal gyrus (MFG). The angular gyrus, which benefits from numerous anatomical connections, is intricately involved in several cognitive functions including spatial cognition, semantic processing, and arithmetic processing (Grabner et al., 2009; Mort et al., 2003; Seghier et al., 2010; Uddin et al., 2010). Similarly, the MFG is recruited by cognitive tasks, like spatial working memory and attention (Leung et al., 2002; McCarthy et al., 1996; Pessoa et al., 2009; Yamasaki et al., 2002). Importantly, both regions are closely associated with multiple aspects of social cognition. Studies of participants observing social acts (e.g. interpersonal interaction, expression of emotion, reading or listening to stories, or false belief theory-of-mind tasks) frequently engage the MFG and angular gyrus (Lawrence et al., 2006; Leube et al., 2012; Mar, 2011). The angular gyrus supports self-reflection, autobiographical reasoning, and awareness of one's actions (D'Argembeau et al., 2013; Farrer et al., 2008; Kjaer et al., 2002). It is also engaged by empathy and social reasoning (Decety and Lamm, 2007; Kubit and Jack, 2013), demonstrating a role in social cognition pertaining to both one's self and others. Perhaps it is in this capacity that the R-FP network is closely connected to conversational coherence.

This network distinguished the two most affiliative conversations from debate and the control task but did not differentiate between them, even though the collaborative problem-solving task necessarily entailed cooperation while the autobiography moistly did not. Inter-brain connectivity of this network may indicate that both subjects are engaged in social evaluation, likely playing a role in considering the other person's mental state – certainly a task that is important to PS and AB, and not so much with DB and CC.

Left Perisylvian Network (L-PS)

As discussed in depth in previous chapters, the regions in L-PS support multiple aspects of linguistic processing, e.g. inferior frontal gyrus, superior temporal sulcus, posterior middle temporal gyrus. Considering that this network supports both comprehension and production, it is not completely surprising that is correlated across brains. Not only is this network significantly connected within both people, the inter-brain connectivity increases with social coherence. The regions in this network – commonly associated with more fundamental aspects of language production and comprehension, like articulation and lexical-phonological integration (Hickok and Poeppel, 2007; Okada and Hickok, 2006) – might contribute to coherence because of intelligibility, which may be particularly pertinent in the relatively noisy environment of MRI scanner. Stephens et al. (2010) found inter-brain coupling in speakers and listeners in multiple regions, including the inferior frontal gyrus. Additionally, they found that this coupling increased as subject reported understanding each other more. Yet this does not answer why the score is so low in debate, where understanding and intelligibility are still critical elements. It becomes even more difficult to interpret its role in coherence, because the inter-brain connections of this network are significantly lower in DB than the control task. Additional studies can help clarify this issue, perhaps by manipulating intelligibility in conversation.

Medial Prefrontal Cortex (MPFC)

A large portion of the dorsal medial prefrontal cortex was identified as a single component, covering functionally distinct regions (Amodio and Frith, 2006). The MPFC is a multimodal region, yet it is frequently associated with internal monitoring, self-assessment, and action monitoring and evaluation (Amodio and Frith, 2006; Gusnard et al., 2001). All of these skills will ensure one's behavior is fitting with the context, certainly a vital component of conversation. Those who report self-monitoring (i.e., adjusting their behavior to fit the context) are more likely to initiate conversation and tend to communicate more easily than others (Dabbs et al., 1980; Ickes and Barnes, 1977). A preliminary study found that training people with high functioning autism in self-monitoring techniques improves their social interactions (Ganz et al., 2013).

All conversations and the control task, exhibit significant (above zero) inter-brain correlation in the MPFC. Moreover, the MPFC is significantly modulated by discourse coherence. This suggests that as both subjects engage in self-monitoring, discourse coherence increases, lending more evidence to behavioral studies on the role of self-monitoring. Interestingly, inter-brain correlation during the collaborative problem-solving task stood apart from the other conditions. It is possible that the dorsal MPFC, which is engaged to a degree with all conversations, is particularly sensitive to cooperation, as has been suggested (McCabe et al., 2001). Yet other imaging studies suggest that cooperative interactions engage the ventral MPFC, which is not included in this network (Decety et al., 2004; Rilling et al., 2002).

Posterior Medial Network

Last is the posterior medial network, consisting primarily of the restroplenial posterior cingulate cortex and precuneus. The posterior cingulate cortex is integral to developing and maintaining a sense of self (Northoff and Bermpohl, 2004), while the precuneus is linked to consciousness, self-assessment, and autobiographical memory (Cavanna, 2007; Lou et al., 2005; Saxe et al., 2006). Additionally, both regions contribute to self-reflective and self-referential tasks (Johnson et al., 2006; Northoff et al., 2006). Together these regions make up a network supporting self-evaluation. This network is not only correlated between both partners, but, like the other networks discussed here, it is strongly predicted by behavioral measures of coherence. In addition, this network distinguished all conversations. Not only does the evaluation of and reference to oneself synchronize over time, but this synchronicity is also modulated by conversation type. Self-referencing is critical to developing coherence, which entails interactive alignment – lining up the understanding of both interlocutors (Menenti et al., 2012). And an essential element of establishing interactive alignment or mutual understanding, is the ability to assess one's own understanding, intentions, and expectations. That is to say, that social alignment is still very much grounded in the self.

Conversation Type and Social Coherence

We can support previous research that the content of a conversation can influence the degree to which people demonstrate feelings of social connectedness (Aron et al., 1997; Sprecher et al., 2013; Vittengl and Holt, 2000). Our composite discourse coherence score, based in dozens of submeasures of social interaction in conversation, reliably differentiated personally relevant conversations from debate, which is less personal and more intellectual. Our coherence measure also indicated a greater degree of coherence during the collaborative task that the autobiographical task. We can now add that cooperative conversations that may entail less self-disclosure also contribute to behaviors consistent with establishing mutual understanding and connectedness. In fact, affiliation and cooperation, rather than self-disclosure, may be the key elements to developing feelings of closeness and social attraction. Intuitively, one would expect that (for example) affirmative and affiliative responses to self-disclosure would increase both the likelihood of future disclosure and feelings of social connectedness. Moreover, one could predict that negative and dissenting responses to self-disclosure would have the opposite effect. However, to our knowledge, such a study has not yet been done.

Personality Type

While there was no difference in the overall strength of inter-brain correlation in either the angular gyrus or posterior medial network, we found that people who reported high neuroticism demonstrated significantly fewer connections overall than pairs where both partners reported low levels of neuroticism. If inter-brain connectivity underlies developing rich social interaction, our result sheds some light on behavioral studies reporting the tendency of neurotic people to report having less satisfying personal relationships (Lopes et al., 2003). In fact, it may be the relative decrease in inter-brain correlation that leads to this disparity. The exact neural mechanisms of this difference are unknown. Previous neuroimaging studies have highlighted distinct patterns of activity and within-brain functional connectivity in more neurotic populations (Ormel et al., 2013; Servaas et al., 2013a; Servaas et al., 2013b).

Interestingly, we also found that pairs of extraverts demonstrated a greater number of inter-brain correlations than pairs of introverts. Again, this corresponds with earlier findings that extraverts also tend to have more engaging and varied conversations and more satisfying interpersonal relationships (Lopes et al., 2003; Thorne, 1987).

How does inter-brain synchronicity take place?

So far, we have been able to demonstrate functional brain-to-brain connectivity in natural conversation using fMRI. Additionally, this connectivity is reliably influenced by discourse coherence and personality traits. These findings provide a critical step towards uncovering the neural correlates of common ground in natural conversation. Still, it remains unclear exactly *how* this happens. How are individuals synchronizing brain activation of specific brain regions? There are two theories on this phenomenon.

Hasson et al. (2012) suggest that inter-brain correlations are due to both (or all) interlocutors responding to environmental signals. The authors go on to compare this exchange of information with wireless communication. Perhaps, during conversations, there are linguistic and paralinguistic cues that both people receive and their brain respond to. In the case of our study, the signal would need to be auditory, since no other information is available. Yet this raises additional questions, such as whether the more information, such as extralinguistic cues like facial expressions and gestures (thereby, essentially increasing the communication bandwidth), would alter inter-brain coherence. Importantly, this particular theory does not address the absence of the inter-brain correlations discussed above from the control task, which presumably contains the same acoustic cues.

The second theory of inter-brain synchronicity claims that social interaction is characterized by self-organizing rhythm behaviors (Neda et al., 2000). This is supported by observations that people tend to implicitly match one other in simple movements like clapping, eye movements, and rocking chairs (Neda et al., 2000; Richardson et al., 2007a; Richardson et al., 2007b). Specific to conversation, Wilson and Wilson (2005) puts for an oscillator model of conversational turn-taking. Specifically, they suggest that, based on the speaker's speech rate, both speaker and listener synchronize to internal oscillators. Of course, this rhythm is implicitly established and sustained by interlocutors. This theory leads to very specific, testable predictions about the relationship between speech rate and turn duration or length of silence between turns. However, theory also cannot explain the effect of conversation types and the control task on inter-brain correlation.

A combination of the external/internal theories would be more accurate. For example, Pickering and Garrod (2013) suggest speakers' covert imitation of implicit cues and forward models to predict a turn-transition. I propose that during conversation, participants pick up on implicit cues of social alignment, signals that the other person either agrees with, supports, or is interested in the current interaction. These cues may encourage the generation of more, and this continues for the duration of the conversation. At some point, interlocutors may come to a point where they are behaviorally synchronized, and the conversation becomes more fluid, probably characterized by fewer dysfluencies and pauses and less silence between turns.

Of course, future research is needed to tease these issues apart. However, with this study, we lay the groundwork by identifying brain regions and specific inter-brain connections to look for.

Summary

In this study, I examined inter-brain coupling during conversation, specifically how it is modulated by a novel measure of social coherence. We found that networks known to support self-monitoring, social cognition, and self-awareness were synchronized across brains. Moreover, this synchronization increased with measured social coherence. We demonstrated a strong relationship between conversation type and social coherence, highlighting the importance of conversation goals and content. Lastly, in this study we were able to demonstrate that personality traits associated with improved or impaired social interaction also exhibited increased or decreased inter-brain connectivity, respectively.

Off-diagonal inter-brain correlations (i.e. networks connected to different, as opposed to identical, networks in the partner's brain), which may reveal more about inter-brain correlations. For example, the primary motor regions in one person are always strongly connected with the superior temporal gyrus in his partner and vice versa, which is likely driven by the exchanging roles of speaker and listener. A subsequent study would identify which inter-brain correlations are depending on alternating roles, perhaps by examining correlations during comprehension and production separately.

Conclusions

Summary of Findings

I set out to uncover the neural correlates related to comprehension, production, and turn-taking in natural conversation. My aim was also to explore inter-brain coherence and reveal brain regions supporting discourse coherence. While previous studies have examined conversation, inter-brain connectivity, or comprehension or production, none have combined them in a naturalistic paradigm. As such, the studies just presented are the first of their kind.

The first study, a well-controlled examination of narrative production and comprehension in the same cohort, provided a comprehensive description of the neural correlates of narrative processing and outlined differences and similarities in functional connectivity. In that study, we argued that despite having some unique neural substrates, both narrative comprehension and production (and likely discourse more generally) recruit extrasylvian regions that support drawing inferences from context and building a situation model. The first study clearly demonstrates that processing connected speech relies are regions far beyond traditional left-lateralized language areas. Still, just as narrative demonstrates emergent features, we predicted then that conversation should do the same.

The other studies involved fMRI hyperscanning of participants while they engaged in a series of conversations and conducted group-level independent components analysis (ICA), applying a data-driven method to identify self-organizing networks. The resulting networks were used in the subsequent studies. As one might expect, both listening and speaking during conversation engages left-lateralized languages areas. It was also shown that, similar to narrative, conversation engages extrasylvian brain regions related to social cognition and inference-making.

Interestingly, in addition to mentalizing networks, turn-transitions engaged networks related to directed attention. They also engaged the precuneus and basal ganglia. Turn-transitions are typically highly coordinated behaviors that involve predictive on the part of both speakers. It is argued that the attention networks work in tandem with mentalizing networks and the precuneus to both draw and gauge the other speaker's attention to the implicit cues signaling an impending speaker change. The basal ganglia may interact with the precuneus and eventually pre-motor regions to coordinate the act of either claiming or releasing the floor.

In another chapter, brain-to-brain correlations that are modulated by a behavioral measure of social coherence are explored. These networks, in addition to being associated with social coherence, are strongly correlated to identical networks in the other's brain. Two of these networks, the poster medial network (comprised of the precuneus and posterior cingulate cortex) and medial prefrontal cortex, relate to self-reference and self-monitoring. Another network, consisting of fronto-parietal regions with the strongest loadings on the angular gyrus, is closely linked to social cognition, in particular considering the mental states of others. These results demonstrate that social coherence is establish and sustained between two people when they both align representations of themselves and the other.

Lastly, we were able to show that personality type affects the degree of inter-brain connectivity between speakers. Extraverts engage a greater number of networks, while neurotic people engage fewer. Although how this is achieved is unclear, previous studies also show that both extraverts and neurotics exhibit brain patterns distinct from others. Additionally, extraverts report having more satisfying relations; and the opposite is true of neurotic people (Lopes et al., 2003). The variations in overall inter-brain correlation may be related to such feelings. Alternatively, this difference could be a consequence of the personality differences themselves. For example, perhaps extraverts engage in more conversation with more partners and essentially have more "practice" at establishing common ground in conversation. However, at this time, we can only speculate.

Future Directions

These studies have laid a foundation for exploring the neural underpinnings of naturalistic conversation. They are the first to identify important features of conversation, such as the cognitive systems employed during conversational turn-taking, network level brain-to-brain connectivity, and the networks influenced by social coherence. Now that these mechanisms are revealed in healthy adult participants, the study can be repeated in other populations like those with autism, traumatic brain injury (TBI), schizophrenia, or other populations with communication impairments. For example, examining inter-brain connectivity and social coherence in patients on the autism spectrum or studying turn-transitions in TBI patients may shed light on the neural underpinnings of impairments specific to their conditions.

Like much scientific research, these studies answered some very important questions while raising many others. Although, I discussed qualitative similarities and differences between narrative and conversation, it would be best to design a study in which the two could be quantitatively compared. Such a study should employ ICA-based analyses on both, comparing network activation.

So much more work remains to be done on turn-transitions. I was able to successfully demonstrate a role for attentional networks in coordinating turn transitions, but my analyses are limited to transition relevance places that resulted in a successful speaker change. According to Sacks et al. (1974), transition relevance places can reoccur throughout a turn-at-talk and do not always culminate in speaker change. Future studies can identify TRPs, particularly those not ending in speaker change and test whether the neural correlates are similar. Relatedly, we can look at the time of transitions. In the current study, transitions are instantaneous moments at the beginning or end of turn. However, it is known that (at least the beginning of) turns-at-talk are planned before one begins speaking. Future studies of turn-transitions can attempt to identify exactly how long before a speaker exchange an interlocutor begins to prepare. Another future study should examine varying kinds of transitions, such as interruptions, questions or transitions consisting of long pauses.

Inter-brain connections also need to be explored further. A logical next step would be to delve into off-diagonal connections. These are interesting because they may entail some kind of directionality (or causality) or be related to alternating speaker roles. Lastly, the analyses I performed were simple temporal correlations. A future study should examine whether any of these inter-brain networks are temporally lagged from one another, either in inter- or intra-brain connectivity. Such an analysis might reveal the involvement of multiple networks, in cascading activation.

Methodological Advancements

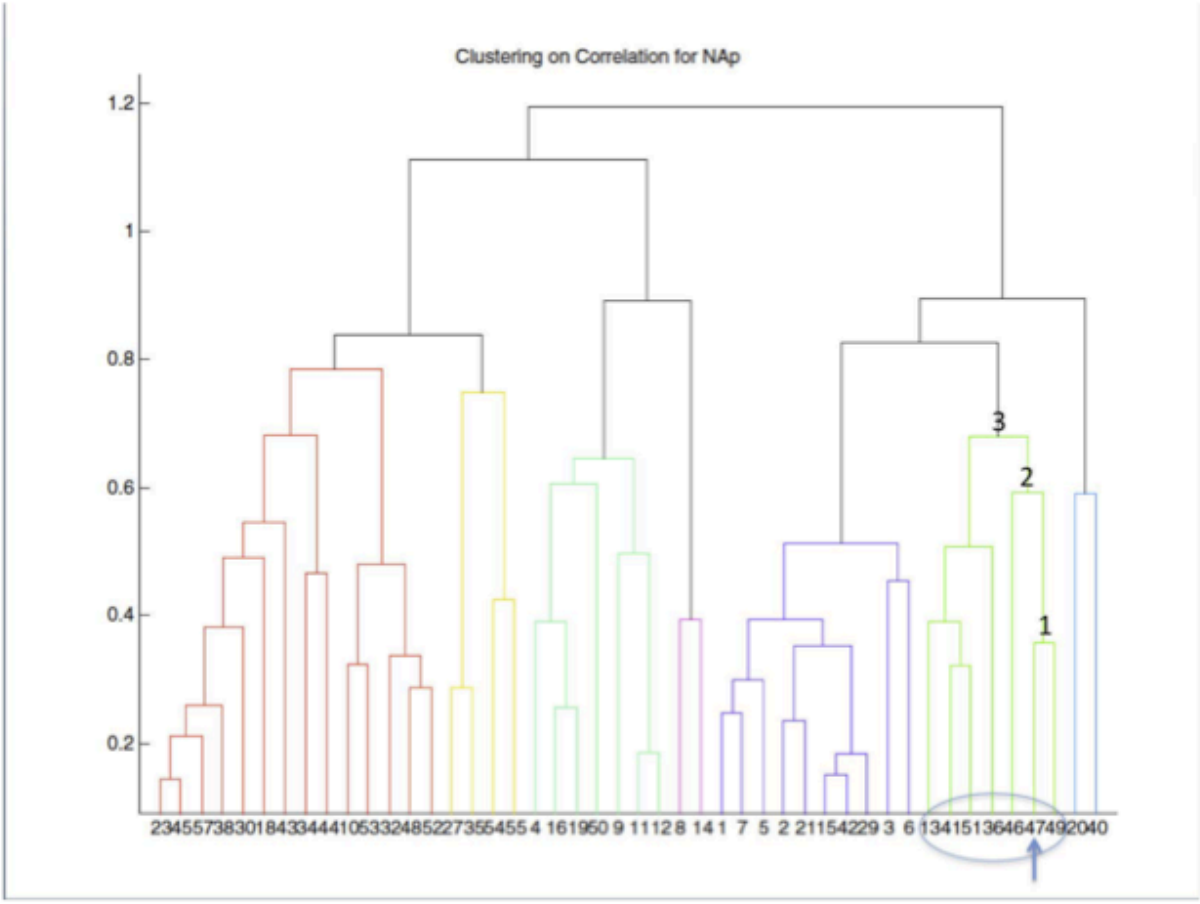
Another critical contribution of this thesis is the tools for future research. Due to the novelty of this study, several materials had to be developed specifically for use in conversational imaging research. The first is a method for reliably measuring social coherence established within conversation. This measure, a composite of dozens of behavioral submeasures, is robust and able to discriminate between types of conversation. Importantly, the definitions for all these measures are carefully outlined to facilitate repeatability. This measure could be useful for research in both social and natural science.

I have also designed and implemented a complex linguistic task that is a reasonable comparison condition for natural conversation. The Conversation Control (CC) task requires participants to take self-paced turns sharing unrelated expository information. Because the content is not overlearned, participants have to construct their speech in a manner similar to natural conversation. However, because they do not address one another and they are not even discussing the same topic, social interaction and communication are markedly attenuated – making this an appropriate comparison to assess the interactive nature of conversation.

Another methodological improvement has been the utilization of ICA-based GLM. Group ICA allows one reduce data from thousands of voxels to dozens of stable networks. Moreover, ICA has the ability to unmix signals from the same brain regions, which traditional GLM is unable to do (Xu et al., 2013). With ICA-base GLM, one can unmix the signals prior to running GLM and presumably. As a result, spatial group ICA is an ideal tool for conversation, which inherently entails concurrent activation of multiple networks.

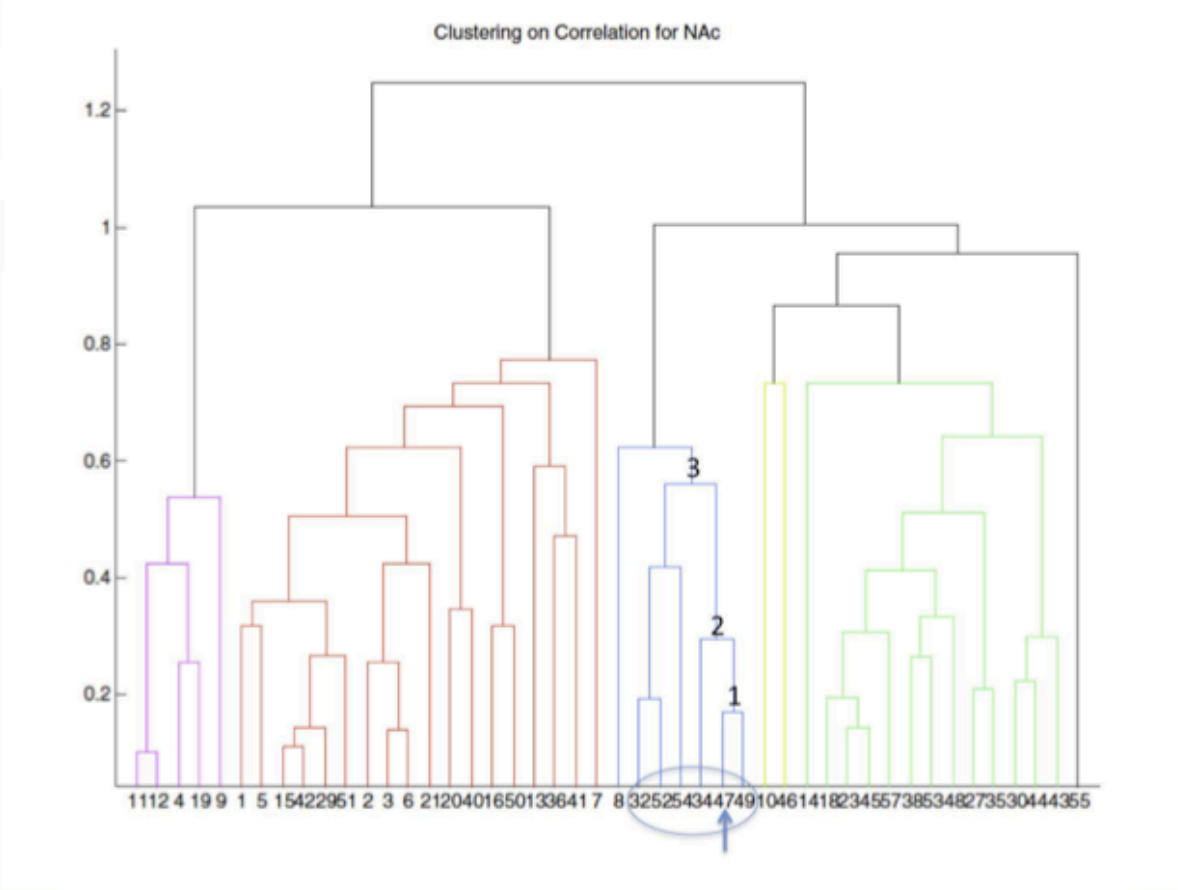
This collection of studies is the first step towards unraveling the neural substrates of ecologically valid conversation.

Appendix I. Narrative Production Dendrogram



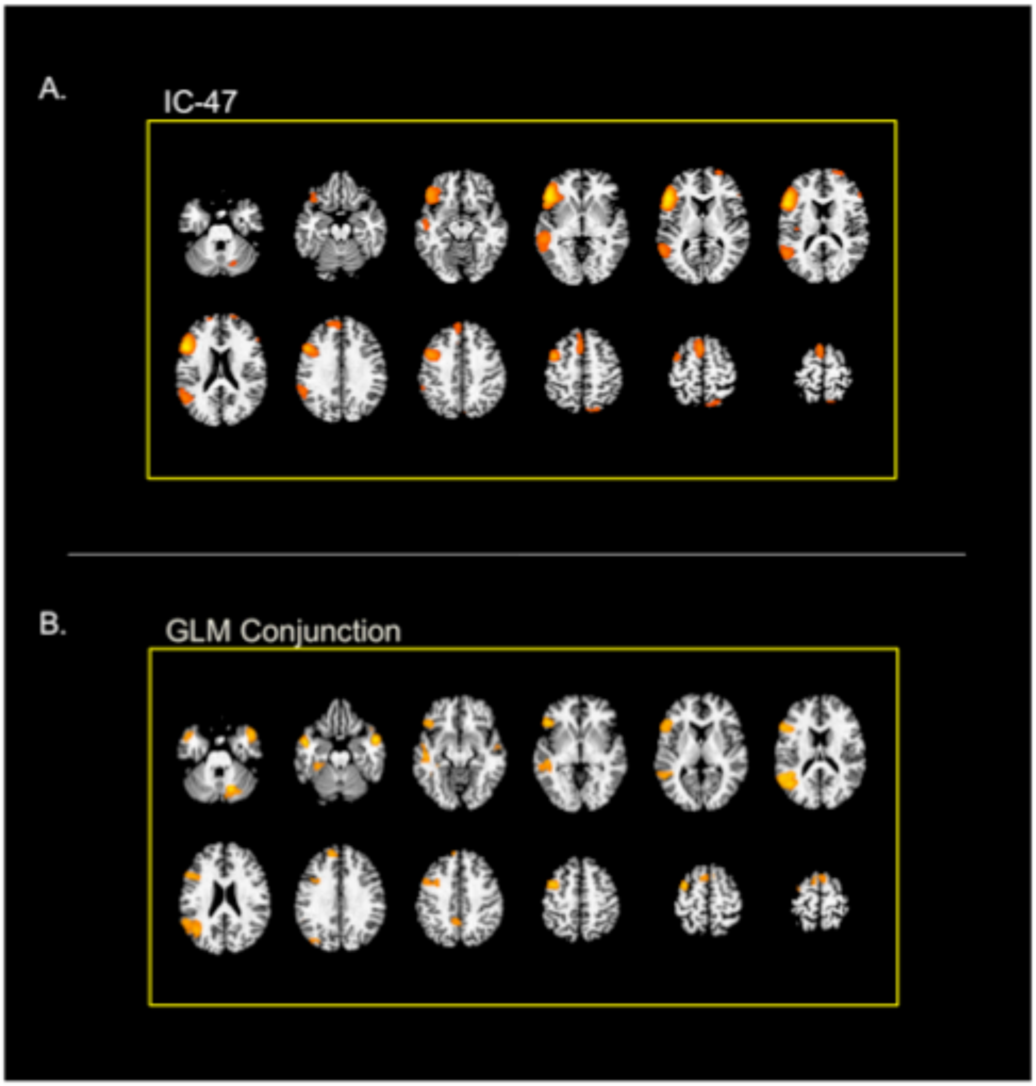
Appendix I. Dendrogram depicting correlations between components during Narrative Production (NAp). Color assignment indicates different nodes as defined by the 80% threshold discussed in Methods section. The node containing the Conjunction Component is marked in Green and circled. The Conjunction Component itself is located at the blue arrow. The cluster cut-off of three levels for the Conjunction Component is also indicated.

Appendix II. Narrative Comprehension Dendrogram



Appendix II. Dendrogram depicting correlations between components during Narrative Comprehension (NAc). Color assignment indicates different nodes as defined by the 80% threshold discussed in Methods section. The node containing the Conjunction Component is marked in Blue and circled. The Conjunction Component itself is located at the blue arrow. The cluster cut-off of three levels for the Conjunction Component is also indicated.

Appendix III. Visual Comparison between GLM Conjunction and IC – 47.



Appendix III. A) Axial brain slices representing IC-47 (the Conjunction Component), thresholded at z=3.00. B) Axial brain slices depicting the conjunction from GLM analyses, thresholded at p=0.05, corrected.

Appendix IV. Detailed FNC Results

Functional Connectivity in Narrative Production and Comprehension

Conjunction Component

The conjunction component (IC-47) depicts a network virtually identical to the GLM conjunction analyses reported above (see Figure 2 and Table 1). This component included left lateralized perisylvian areas [i.e., inferior frontal gyrus (IFG), middle temporal gyrus (MTG), and superior temporal gyrus (STG), as well as the left anterior superior temporal sulcus (STS)], the presupplementary motor area (pre-SMA), dorsal premotor cortex (PMd), and dorsal medial prefrontal cortex (dmPFC). Supplementary Figure 3 highlights the similarity between IC-47 and the GLM conjunctions.

Network Connectivity during Narrative Production

At the first level, the Conjunction Component is functionally connected to IC- 49, which consists of bilateral superior temporal gyrus and STS. IC-49 also extends to the middle temporal gyrus, particularly in the left hemisphere.

The left dorsolateral prefrontal cortex (dlPFC) dominates IC-46, the second-degree connection to the Conjunction Component. IC-46 also contains significantly smaller clusters in the left dorsal anterior cingulate cortex (ACC), pre-SMA, and anterior insula. Another small cluster exists in the right dlPFC.

At the third level, the Conjunction Component is connected to IC-13, IC-41, IC-51, and IC-36. IC-13 consists of the caudate nucleus bilaterally and portions of the medial thalami. Both IC-41 and IC-51 contain the bilateral anterior insula and dorsal ACC. However, IC-41 contains a larger portion of the insular cortex, while IC-51 consists of a greater extent of the dorsal ACC and pre-SMA, as well as smaller clusters of the bilateral dlPFC and posterior MTG.

Network Connectivity during Narrative Comprehension

As in NAp, the Conjunction Component is most closely connected to IC-49, which contains bilateral superior temporal cortex during NAc.

IC-34 is the only second-degree connection to the Conjunction Component. IC-34 consists of the right IFG, MTG, and anterior STG. This component is identical to brain regions found to be active uniquely for NAc by way of GLM analysis (see Figure 2 of the main text).

The third degree connections consist of three components: IC-32, IC-52, and IC-54. IC-32 includes the left dlPFC, superior parietal lobule (SPL), inferior parietal lobule (IPL), the posterior cingulate cortex (PCC), and a portion of the precuneus. IC-52 consists primarily of the dmPFC but includes the dlPFC, pre-SMA, ACC, and right IFG. Lastly, a large cluster encompassing the precuneus and retrosplenial PCC characterizes IC-54. IC-54 is also distinguished by bilateral IPL and a smaller portion of the dmPFC.

Appendix V: Conversation Paradigm Training Materials

--------------------- TRAINING SHEET USED BY EXPERIMENTER ---------------------

Training _ Two Way Conversation

Materials to have reviewed before training:

- Website and video of immigration issue
- Expository texts

Administer questionnaire

Instructions:

For this study, you will participate in a conversation. Before and after each conversation, you will have a turn reciting a NR. NR – conversation – NR. We will repeat that 4 times. Each of the conversation will have a different topic.

For the most part, what you should remember is that your conversations should be as natural as possible. Feel free to ask each other questions, cut each other off, tell jokes, whatever is natural for you.

Things to avoid:

- Try not to totally dominate a conversation, or allow the conversation to be dominated.
- Ask questions, if that's natural, but do not turn it into an interview.
- Make sure to stay on topic!

The 4 conversations will be:

College experience – you can have a normal conversation where you share your college experiences. Be careful not to fall into interviewer/interviewee pattern. You should both share your experience and ask questions in response to the other person's experience throughout the conversation.

Debate – You need to take opposing sides and stick to them throughout this debate. Respond to each as best you can, make rebuttals and responses, and present your points. Let's review the pros and cons of the argument. Who is taking which side? Do you have personal opinions or preferences?

Problem solving – We will put you in a hypothetical perilous situation. You two should work together to try to make sure you both make it out alive. Although the situation is not real, you should discuss it as if it were.

Parallel Speech – This conversation will be a little unnatural. In natural speech there is back and forth and what a person says influences what you will say next. In fact, this isn't really going to be a conversation. You will each discuss separate topics. You are to just stick to the facts of your topic, although you are not limited to the facts I have presented to you. Do not respond to anything the other person says, although you should listen to them when they are talking. You should make sure that both of you get equal speaking time, even if that means interrupting the other person.

Practice:

We're going to practice having a conversation. Feel free to talk about your latest vacations.

---- *listen and give them pointers*

Now we're going to review the Nursery Rhymes. Which of the 5 do you both know best? We only need three. Let's practice saying them as if we were in the scanner.

Possible Nursery Rhymes: Mary Had a Little Lamb, Jack Be Nimble, Humpty Dumpty, Hey Diddle Diddle, Hickory Dickory Dock

Problem solving situation: TO BE PROVIDED IMMEDIATELY BEFORE SCANNING

You were on a cruise ship, which docked for a couple of hours at a tiny and seemingly uninhabited island in the middle of the ocean, which appears the size of NIH's main campus. You both stayed a little too long, and the cruise ship left without. All you have are the clothes you are wearing, whatever is in normally in your purse/backpack in addition to a book of matches, a pocketknife, and enough food/water to stretch one or two days. The island is a stereotypical tropical landscape with densely packed trees (some of which bear unfamiliar fruit), rocks, sandy beaches, fish in the reef off shore, small land animals, but there may also be dangerous animals on and off shore. There is a small mountain peak in the center of the island, but it appears to be a smoldering volcano. You know that another cruise ship will pass by a half mile off shore in 90 days. What will you do? In what order? And how will you do it?

Do not discuss until cued to do so in the scanner.

----- TOPICS USED FOR Conversational Control TASK AND SAMPLE TEXT------

SOLAR SYSTEM

The solar system consists of the Sun and 8 planets that orbit it: Mercury, Venus, Mars, Earth, Jupiter, Saturn, Uranus, and Neptune. The Sun is by far the largest body in the solar system, with a mass equivalent to almost 333 thousand Earths. The Sun, which is essentially a star, has 8 planet bound by its gravitational pull. The four planets closest to Sun: Mercury Venus, Mars, and Earth, are primarily made up of rock and metal. The four farthest from the sun (Jupiter, Saturn, Uranus, and Neptune) are made primarily of gas and tend to be much larger than the closer planets. Each planet has distinct properties.

Mercury is the closest to the Sun. With a size that is less than 6% the mass of Earth, it is the smallest planet. It has a large iron core and thin mantle.

Venus is second farthest from the sun. It is only slightly smaller than Earth. The atmosphere on Venus is 90 times as dense as Earth's atmosphere. Venus is tremendously hot with a surface temperature of over 400 degrees Celsius.

Earth is the largest and densest of the rocky planets. It is the only planet definitely known to contain water and life. It has one moon.

Mars is only a tenth the mass of Earth. Its atmosphere consists primarily of carbon dioxide. Mars is characterized by its red color, which is caused by iron oxide (also known as rust) in its surface soil.

Jupiter is that largest planet in our solar system. It is 318 times the mass of Earth, and is 2.5 times larger than all the other planets combined. Of the gaseous planets, Jupiter is the closest to the Sun, but is the fifth closest planet overall. It is orbited by 63 moons, the largest of which is bigger than Mercury.

Saturn is distinguished by its system of rings. It is the least dense planet in the solar system, but has a mass 95 times that of Earth. It has 62 moons, one of which, Titan, is the only moon with a substantial atmosphere.

Uranus has a mass 14 times that of Earth. It radiates little heat, and is thought to be cooler at its core. Uranus has 27 known moons orbiting it.

Neptune is only slightly smaller than Uranus, but it is much more dense, which means it has a great mass, equivalent to 17 Earths. Neptune has 13 moons.

(From Wikipedia)

EARTH SCIENCE

The Earth is considered made up of four interactive spheres. The hydrosphere is the network of rivers, lakes, oceans and all bodies of water on earth. The biosphere refers to all life on Earth, plant and animal. The geosphere is the Earth's crust and the solid parts of earth, made of rock. The atmosphere consists of all the gases that surround Earth and make up its air. The land on Earth that is above water is broken into seven continents: Africa, Antarctica, Asia, Australia, Europe, North America, South America. The largest bodies of water (oceans) are: the Pacific, Atlantic, Indian, and

Arctic oceans. The atmosphere consists of 78% nitrogen and 21% oxygen. The Earth is a sphere. The bulge in the middle is due to its rotation. At it's widest point, the Equator, the Earth measures 24,860.2 miles around. Two magnetic poles characterize the Earth, corresponding the North and South poles. Earth rotates on its axis as it orbits the sun. The rotation is equivalent to 1 day. It takes 1 year to complete an orbit around the sun. (From Wikipedia)

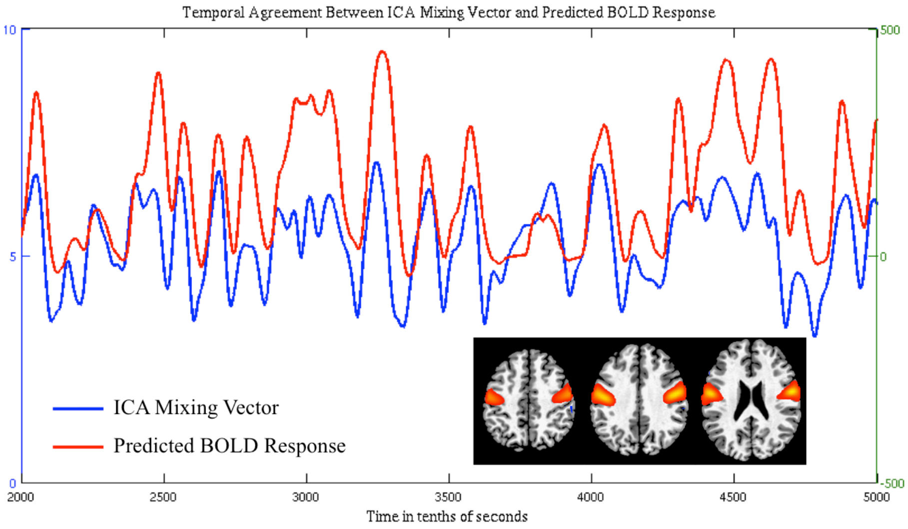
CIVIL WAR

Fought 1861-1865, the American Civil War was the result of decades of sectional tensions between the North and South. Focused on slavery and states rights, these issues came to a head following the election in 1860 of Abraham Lincoln, who was against the spread of slavery to newly acquired American territories. To the Southern states, this was threat to their economic stability. Over the next several months, eleven southern states seceded and formed the Confederate States of America, including Alabama, Georgia, Louisiana, and Texas. During the first two years of the war, Southern troops won numerous victories under the direction of Confederate General Robert E. Lee. However, they saw their fortunes turn after losses at Gettysburg and Vicksburg in 1863. It was by winning these battles that Union General Ulysses S. Grant gained a strategic advantage, which was to claim control of the Mississippi River and essentially cut the South in two. From then on, Northern forces worked to conqueror the South, forcing them to surrender in April 1865. After the Civil War, three Amendments were made the Constitution: 13th, 14th and 15th amendments which respectively abolished slavery, extended legal protections to all citizens regardless of race, and removed racial restrictions on voting rights.
(From About.com)

HUMAN PHYSIOLOGY

Traditionally, human physiology is seen as a collection of interacting systems. Each has it's own function and purpose, but they work together and are heavily dependent on one another. The nervous system consists of the central nervous system (CNS, the brain and spinal cord) and the peripheral nervous system (PNS, the network of nerves throughout the body). The nervous system interacts with organs and muscles to control their functions. Additionally, it receives information from organs and muscles, such as receiving sensory information. The musculoskeletal system consists of the skeleton (bones, tendons, ligaments, cartilage) and the network of muscles attached to the bones. This system gives the body its basic structure and facilitates movement. The circulatory system is made of the heart and blood vessels. It's the body's transport system. It delivers nutrients and oxygen to cells throughout the body and removes waste. The respiratory system consists of the nose, trachea, pharynx, and lungs. It takes oxygen from the air, delivers it to the blood system. It excretes carbon dioxide and water back into the air. The gastrointestinal system is made up of the mouth, stomach, intestines and other organs. It is responsible for breaking down food to nutritional molecules and delivering them to the circulatory system to be distributed throughout the body. It also excretes and unused bits of waste from food. The immune system is comprised of white blood cells and the lymphatic system has the job of distinguishing its own cells from alien and potentially dangerous cells and then, destroying or neutralizing those alien cells. There are other systems, which include the endocrine system and the reproductive system, which are integral to human physiology.

Appendix VI. Comparison of Motor Component and HRF Convolved Turns-at-talk.



Data from six minutes of a specific conversation. The raw, back-reconstructed ICA mixing vector for independent component (IC) consisting of bilateral motor cortex is shown in blue. The spatial pattern of this particular IC is presented in inset. In red is the predicted BOLD signal resulting purely from the onset and offset of speaking, i.e. this particular person's periods of speech convolved with canonical hemodynamic response function (HRF). These time course are significantly correlated ($r(352) = 0.7485$, $p < 0.0001$), indicating that group ICA back-reconstructed time courses faithfully reproduce behavior induced signal.

Appendix VII: Operational Definitions For Each DCS Submeasure

Final Codes and their Operational Definitions

Produced by Jacqueline Hazen and Briel Kobak in collaboration with Nuria AbdulSabur

Intra-turn Codes

Turn construction units (TCUs)

Each turn is composed of unit-types, which form comprehensive units of turns--single words, phrases, or whole sentences. These allow the speaker and recipient to project which unit-type is under way (Sacks, Schegloff and Jefferson 1974:702). A single turn-at-talk may be built of several turn construction units. Each turn at talk will be coded as the following:

Lexical (No Valence)

The verbal content of a turn is composed of a single word or words with no syntactical structure (e.g. Right, right, right. Twenty-four.) This includes turns solely composed of laughter.

Phrasal (No Valence)

The verbal content of a turn is limited to words that do not form a complete English sentence (i.e. dependent clauses, phrases), thereby not possessing a verb

Sentential (No Valence)

The verbal content of a turn includes a finite verb and forms a complete English sentence. Even if the sentence is incomplete, if it contains a verb and has syntactic coherence, it should be coded as sentential.

Multi-Sentential (No Valence)

The verbal content of a turn includes two or more sentences (sentences will be demarcated by a pause or completed thought)

Narrative (No Valence)

The verbal content of a turn contains a 1) *disjunct marker*—utterance that signals the talk to follow is not topically coherent with the adjacent prior talk, such as “oh” or “incidentally”--and 2) *embedded repetition*—which locates, but does not explicitly cite, the element of prior talk which triggered the story (Jefferson 1978) and 3) also qualifies as multi-sentential. The narrative should depict a single event, and not just a general description of the past.

Presupposition

Presupposition refers to the idea that context is invoked and managed in conversation, and that "institutional imperatives originating from outside the interaction are evidenced" (Heritage 1998:4). The ways in which speakers design their words and invoke contexts with the recipient in mind reflects the degree to which the two interlocutors are in fact able to draw upon shared, mutual knowledge. Ways in which speakers can do so include:

Recognitional forms (+2)

- The use of a content word that produces a singular referent for both speaker and hearer, and which attempt to elicit the presupposition of specific people/objects/places.
- This can include proper nouns or recognitional descriptors (e.g. Nuria, or the neuro-imager who trained us), or a mutually recognized category that one can only know from personal experience (e.g. those fMRIs in the basement of Building 10). From the context, the extent to which personal experience informs the use of a recognitional term, the extent to which it is meant to be informationally salient, will determine whether or not people/objects/places are coded as recognitional forms.
- Recognitional forms have to be shared and understood by both speaker and hearer, which can be noticed by the subsequent remark (i.e. unmarked continuers indicate non-sharedness; epistemic stance markers can highlight sharedness; etc.). They can be set off by the use of a demonstrative pronoun.
- Ex: PA2: Neuroscience.
  PA 1: Neuroscience. [That's] uh that's the same as /Nuria/.
  PA2: [Yeah]
- The naming of colleges, which isn't exactly an emic ability, will only count when the school is nicknamed or shortened in some way.
- Has to be used to defend knowledge

Epistemic stance marker: External (+2)

"*Epistemic stance* refers to a person's knowledge or belief, including sources of knowledge and degrees of commitment to truth and certainty of propositions (Chafe and Nichols 1986)...

Typically, these actions and stance displays relate to common or similar topics and goals" (Ochs 2004:109). The context of knowledge and the process of learning is invoked here, and discussed explicitly. Evidentials (Schieffelin 1996:440) refer to the ways in which speakers convey experience and/or visual, verbal, or sonic information as proof of something. Looking for evidentials will help point to the existence of epistemic stance markers since they include linguistic phenomena, such as reported speech.

- Epistemic Stance Marker: External constitutes an utterance in which the speaker explicitly refers to his own means for or extent of the knowledge he has shared, or is about to share, in order to explain or reinforce the use of a referent that exists prior to and externally to the conversation; "markers that indicate something about the source of the information in the proposition" (Bybee 1985: 184); this can be constructed within a single turn, or across turns at talk.
- Ex: PA2: Yeah I got pretty cold. I_don't_know (.) You_know I've actually been to Cleveland a whole bunch my family's from there.
  PA1: Oh no way.
  PA2: Yeah but u:m (.) Yeah actually my uh (.) my grandfather was Director_of_Alumni_Relations for like twenty years.=
  PA1: =at Case?
  PA2: At Case.
  PA1: But then so you know all about like the whole Case_Western like versus like (1) like I guess Case like branding versus like Case_Western_Reserve_University branding.
  PA2: Right yeah.
  PA1: Ye:a:h [I]-
  PA2: [He] complained about that a lot.
  PA1: @@ (1) Yeah I-the_ol-the old timers were really bad about it. Um (.) I guess I did like this tele-telephone (.) thing. I don't know like you worked in the the call center and you got like a student_wage.

Entailment

Entailment acknowledges that conversation inevitably produces new data and invokes new contexts on a moment-to-moment basis, providing the grounds for future presupposition. "The assumption is that it is fundamentally through interaction that context is built" and the words and ideas produced therein are “made real and enforceable for the participants” as a conversation proceeds (Heritage 1998: 4).

Creation and **Repetition of co-constructed phrase (+1)**

- **Creation of co-constructed phrase** marks the first use of an original or emic term/phrase that indexes something meaningful in the current state of talk
- **Repetition of co-constructed phrase** marks the continued use of an original or emic term/phrase that indexes something meaningful in the current state of talk after the original utterance
    - *The repetition of the co-constructed phrase will have to be noticed first by the coder, who can then go back in the transcript and look for the creation of the co-constructed phrase being repeated.*
- Ex: PA1: =at Case?
  PA2: At Case.
  PA1: But then so you know all about like the whole Case_Western like versus like
  (1) l like I guess Case like branding versus like
  Case_Western_Reserve_University branding.
- Note: This also gets linked to **Shared Use of Prounouns** in a way didn’t anticipate re: Whole Case Western verus like Case Western Reserve University branding and Connecticut; they’re pretty unwieldy (the ‘you guys’ and the Jewish student group)

Shared use of pronouns (+1)

Pronouns signal that their referents have been previously mentioned, or are readily identifiable in the context of communication or on the basis of the speaker and hearer’s mutual knowledge (Gee 1999:120).

- Shared use of pronouns for topic referent- the use of a pronoun across more than one turn at talk by both speakers, which is not merely anaphoric, but rather indexes shared knowledge and understanding of a referent being used by both speaker and recipient in pronomial form.
- The pronoun becomes more than a stand-in as each speaker instead uses it as a connective thread to establish and fortify intersubjectivity within the conversation.
- The use of these shared, demonstrative pronouns across turns do not necessarily have to be the same exact pronoun (i.e. ‘that’ and ‘it’) as long as they both index the same referent.

- Ex: PA1: So uh when you cooked at school did, it was-that was (.) like did you have that all four years or did you just like get into and it expanded or you just do it for yourself after a while or--
  PA2: Well yeah. You_know I did it for like and a half and then like (.) passed it on to someone else and (.) you_know did other things. (1) You_know I_don't_know I focused a_lot more on my work.

Epistemic stance marker: Internal (0)

- Epistemic stance marker: Internal constitutes an utterance in which the speaker explicitly refers to the means for or extent of the mutual knowledge he has shared, or is about to share, in order to explain or reinforce the use of a referent that was produced within, and therefore must reference, the conversation at hand; this can happen within a single turn or across turns at talk. How the speaker knows what he knows, or chooses to explain what he does NOT know, has to come from data provided from the conversation being had.
- This will often take a negative form, wherein the speaker references prior talk due to his lack of knowledge.
- Ex: PA2: = just like a whole bunch of guys that came through and I would (.) you_know try and go to the concert [whenever I could yeah]
  PA1: [that's awesome] That sounds like a >good ti-yeah the only one I didn't know is< Das_Racist?
  PA2: (0.7) Yeah man.

Feeling of nonseriousness

Wallace Chafe believes laughter to express emotion that belongs in the same categorization as joy and anger; this emotion would be called 'the feeling of nonseriousness." Laughter is a mechanism we use to distract ourselves and others from serious thought (2007).

All forms of laughter to be coded as either of these two:

Shared laughter (+2)

- The produced, audible of laughter of both recipients either 1) overlaps for the entire segment of laughter or a portion of it OR 2) the audible laughter of one participant is followed by the audible laughter of the other participant with no discernible silence between them

Single-sided laughter (0)

- The produced, audible laughter of one recipient is not followed by or overlapping with audible laughter of the other participant

The laughter codes will be further distinguished as:

Corrective Laughter (+1)

- This constitutes laughter that is meant to counterbalance the negative qualities attached to prior or surrounding language, which can range from discomfort to embarrassment, and the laughter then functions as a mechanism to ameliorate contextual abnormalities in the conversation; this can be produced by the either the speaker or listener of the negative speech.

Non-corrective laughter (-2)

- Laughter that occurs in response to undesirable situations that can be characterized as contemptible, cruel, rude, or threatening (Chafe 2007: 79) and is NOT meant to counterbalance the negative qualities of surrounding talk. Such undesirable utterances can be determined from context, and cues such as silences and markedly drawn out speech.

Non-serious Laughter (+1)

- This is the genuine stuff- laughter produced when there is no observable surrounding talk that contains negative or abnormal qualities and thereby represents the audible expression of a desirable emotion; to be looked for especially after *spoken invitations to laughter*.

Invitation to Spoken Laughter (+1)

Sidnell describes the elicitation of laughter as co-implicating recipients in a perverse hearing of prior talk (2010).

- This code must involve 1) a *non-seriousness disjunct marker*—this utterance will break with the preceding frames of talk (i.e. through intonation, pitch, duration of vowels, use of a curse word, or pragmatic expressions like ‘it’s so funny to me, ‘I mean let’s be honest’) so as to signal that its content is not to be interpreted in the same manner as preceding talk and 2) a spoken invitation to laughter has to involve a non-serious statement that is notably sarcastic, ironic, or silly given the context of the utterance.

- Note: Spoken invitations to laughter do not have to actually elicit laughter in order to be coded.

Inter-turn Codes

Speaker-change

The moment when a different speaker begins a turn. These moments are often negotiated through *Transition-relevant places,* which are moments in a turn where continuing or concluding a turn becomes relevant--moments of possible completion of a thought (Sacks, Schegloff and Jefferson 1974:707). In writing, these are often marked by punctuation like periods, commas, question marks, etc. In discourse they are marked by pauses, completion of phrases, or a "pitch peaks" of noticeably higher and louder pitch than the preceding syllables (Schegloff 1998).

The codes for speaker change are marked on the transcripts as the words through the first period in a TCU OR a pause of longer than 1 second.

Speaker-selected (+1)

The turn preceding the moment of speaker-change contains some discourse or behavior asking the other speaker to speak (Sacks, Schegloff, and Jefferson 1974:704-706).

- Markers of speaker-selected change might include:
    - A question directed at the other speaker: What about you? How was your xxx?
    - Tag questions directed at the other speaker: Bill? xxxx, right? (719)
    - Requests for comment or clarification: Tell me about xxxx.
    - An extended pause at a transition-relevant place (706-707; Sidnell 2010: 48)
    - An uprise in intonation that turns the middle or end of a turn into a question: We went to the car /and/?

Self-initiated Speaker Change (0)

The turn preceding the moment of speaker-change contained no markers asking the other speaker to speak. Self-initiated speaker-change may occur at transition-relevant places or at any other point during the current speaker's turn.

Interruption (-1)

- A special subset of self-initiated speaker-change that occurs at a non-transition-relevant place with overlap, e.g. in the middle of a word or grammatical phrase.[1]

Except for the initial opening TCU in a conversation, every TCU in our transcripts is coded either as a kind of speaker-change AND/OR a type of continuer. If PA1 begins speaking, PA2 produces a TCU as a continuer, then PA1 continues speaking through another TCU, only the initial TCU by PA1 is coded with speaker-change. As long as PA2's TCUs only consist of continuers, PA2 has not claimed the floor and a speaker-change has not taken place.

Infrequently, TCUs coded as Continuers also receive a code of Speaker-Selected Speaker Change. This occurs per the example below:
00:05:14-5 PA2: Alright should we also-- should we build a /spear/ or something to--to kill some of these small local animals?=

00:05:19-2 PA1: = /Heck yeah!/

00:05:21-5 PA2: How do we figure out if the fruit is poisonous or not, I've never seen this bef[ore]

00:05:24-2 PA1: [Um] look for it in fecal matter(.) Usually if you [look at]--

PA1's "/Heck yeah!/" acts a marked contingent phrasal continuer as she is not claiming a turn, but voicing her continued engagement with PA2's turn. However, PA2 has queried her in the middle of the her own turn, so PA1 has also been Speaker-Selected to participate in the conversation.

---

[1] It's important to note that as defined and used here, an interruption is not necessarily a negative or positive behavior in relation to the level of intersubjectivity shared between participants. Partly it will depend on the referential and sequential content.

Topic shifts

The movement in content from discourse centering around one thing to discourse about another. By coding for topic shifts, it's important to remember that we are coding for *how* speakers move between discourse centering around one thing to discourse about another thing. Centering discourse around one thing is co-created and organized (or disorganized) by participants (Sidnell 2010:223-224): topic is not a thing that exists outside the particular context of a conversation.

Topic closure, followed by topic generation (+1)

- A shift in topic in which one or both speakers close the previous topic, then one or both introduce a new related or unrelated center of discourse. Both speakers must acknowledge the topic closure by a) joining in discourse related to the new center and/or b) using one of the markers described in the next bullet point. If one speaker continues to speak around the previous topic center, then the topic is not closed.
- Markers of topic closure include *reciprocal confirmations* following a turn (e.g. P1: Okay, that's good. P2: Right.) or a *summarizing assessment* (e.g. That sounds perfect. Now, what about the xxx?) (Sidnell 2010:231, 234), prior to the TCU introducing the new topic.
- Topic generators include asking a question or making a statement that contextually moves the center of the discourse. Speakers respond to topic generators with discourse on that subject or their own topic generators (Sidnell 2010:231-233).
    - Segments of the data marked with this code include both the turns that close the previous topic and turns that introduce the new topic.
    - Codes include: a) **Topic closure, followed by topic generation that returns to a previous topic;** b) **Topic closure, followed by topic generation that moves to a co-class member** (see below for definition of co-class member); and c) **Topic closure, followed by topic generation that centers around a new topic.**

Gradual shift (0)

(aka Sacks 1972 *stepwise transitions)*

- Changes in topic in which no new topic is announced in a single TCU, or preceded by a topic close moving to a topic beginning. Marked by *topical pivots* (Sacks 1995), which are utterances in which the first part is relevant to the topic, and the second part is connected to the first part, but not the topic, e.g. as part of a discussion about rain boots shifting to a discussion of a party: “Yes **mine** are made of rubber and I wore **them** to a party last night.” Topical pivots would generally be contextually defined.

  - Codes include: a) **Gradual shift, via topical pivot;** and b) **Gradual shift to a co-class member** (see below for definition of co-class member).

- Other markers could be shifts in topic between contextually defined *co-class members*, e.g. in a discussion about bananas as a potential food source moving to a discussion about other fruits OR other potential food sources. Another co-class member example would be shifting from discussion Participant 1’s major to Participant 2’s major. If these co-class members are shared categories between P1 and P2, the center of the conversation discourse can switch gradually between them. Co-class markers often occur near discourse that speakers use to label something co-class: words like “also,” “as well,” or “another.”

Topic Jump (-2)

- A single TCU that requests an abrupt topic change while speakers are discussing the current topic.
- A topic jump isn’t the initiation of a new topic when a previous topic has been closed.
- A topic jump could be marked by the absence of makers of topic closure, like *reciprocal confirmations* following a turn or a *summarizing assessment* prior to the TCU introducing the new topic.
- Topic jumps often involve pragmatic dimensions in their movements to new centers of topic in the conversation (e.g. in A02711 problem dialogue, suddenly questioning the extent of the information in the descriptive paragraph). TCUs immediately responding to a topic jump may or may not begin with filler words, other-initiated repair asking about the new topic, or pauses.

Other contextual markers following shifts via topic closure, a gradual shift, or topic jumps would be change in the system of shared reference marked by pronominal coreference (e.g. **mine** and **them** are understood to mean rain boots as marked above). We are already coding for ‘Shared-system of pronouns’ in the Intra-turn code family in Presupposition and Entailment, so we will look for the boundaries of shared use.

Repair

An explicit acknowledgement by one or more speaker that intersubjective understanding has been lost or threatened by trouble, followed by an action to fix the intersubjective understanding (Sacks, Schegloff and Jefferson 1974:723-724). A repair generally concludes with reciprocal confirmations and a return to the prior topic of conversation and an unmarked flow of discourse.

Trouble Sources (-2)

"The segment of talk to which the repair is addressed" (Sidnell 2010:110).

- Trouble sources precede the repair initiators. All are contextually defined as trouble sources by the surrounding discourse. Trouble sources can be:
    - **Phonological** (trouble with sounds)
    - **Semantic** (trouble with the meaning of a word or phrase)
    - **Syntactic** misunderstandings (trouble with conventional grammar)
    - **Pragmatic** (trouble with the movement of the conversation like prolonged overlap, false starts, interrupting; trouble with diverging topic centers) Speakers perceive trouble as a compromise to intersubjective understanding, therefore on a basic level, all trouble sources are in some way pragmatic. *Pragmatics* in the linguistic sense encompasses the "inter-relation of language structure and the principles of language use," (Levinson 1983:9). Using language to build a conversation creates intersubjectivity, so compromised intersubjectivity *for any reason* is a pragmatic threat to conversation. However, trouble has an extra pragmatic dimension when it is based on non-shared topics and problems with speaker-change.

Whether a trouble source may be announcing phonological, pragmatic, semantic, or syntactic trouble may not be clear from the context. In the initial four full runs that we coded, it was found that the majority of trouble sources do not fit definitively into one category. Therefore, all trouble sources will be simply coded as a 'trouble source.' If a single repair had more than one trouble source, we code the multiple trouble sources separately.

After a repair is identified, we examine previous discourse for the trouble source. Trouble sources are coded as any and all discourse leading to the repair. In most instances for other-initiated repair, this is the TCU preceding the repair initiator.

However, in cases where prolonged overlap or false starts are the trouble sources, both participants' overlapped words are coded as the trouble source. For self-initiated repair, the trouble source was coded as the words that the participant repeated and modified or substituted other words for with the repair. For example, PA1: "Um otherwise we should probably (1.05) uh split up and find like other stuff around the campus uh around the island." "Around the campus uh" in green is the trouble source; "Around the island" in yellow is the repair.

Repair initiators mark a possible disjunction with the immediately preceding talk and attempt to resolve the trouble source preceding them. Different forms of repair initiators are described below. Please see the code family "Adverse Effects on Discourse" for more details about the codes for open and closed repair initiators.

Self-initiated repair (+2)

Repair actions enacted by the speaker who said the trouble source (Sidnell 2010:114-115).

- Repair sequences are contextually defined, but often include repetition, paraphrasing, adding information, and restarting the previous utterance. Self-initiated repair can be distinguished from typical pauses and repetition in TCUs by marked changes in emphasis, intonation patterns, and rhythm, co-occurring with the following indicators:

    1) repetition of one's own words
    2) ending the pronunciation of a word half-way through
    OR 3) stopping in the middle of a turn at a non-transition-relevant place.

However, if a participant simply repeats words *without* changing the words at all, inserting a filler word, or changing intonation, this repetition is not considered a self-initiated repair. If the participant changes intonation markedly OR repeats some words, but changes others or inserts a filler word between the repetition, this is coded as self-repair.

For example, PA1: "Um otherwise we should probably (1.05) uh split up and find like other stuff around the campus uh around the island." "Around the island" would be coded as a self-initiated repair, since PA1 changes his 'campus' to 'island' and inserts a filler word. But PA1: "I think--I think we should have our raft ready by now" would not be coded for self-initiated repair as there is no marked change in intonation, no significant pause or filler words in between the repetition, nor does PA1 changes his words; he just simply repeats them.

Other-initiated repair over 2 turns (+1)

Repair actions enacted as a reaction to another speaker saying a trouble marker that consist of 1 turn for each speaker. Repair sequences are contextually defined, but often include repetition, paraphrasing, adding information, and restarting the previous utterance.

Other-initiated repair sustained over multi-turns (+1)

- Repair actions enacted as a reaction to another speaker saying a trouble marker that consist more than 1 turn for one or both speakers. This sequence could consist of multiple trouble markers indicating that repair has not been achieved and multiple attempts at repair.
- Repair sequences are contextually defined, but often include repetition, paraphrasing, adding information, and restarting the previous utterance.
- When a repair takes multiple TCUs from one or both speakers, we will highlight and examine the content of the full repair and code for multiple trouble sources (if they exist).

Non-acknowledgement of repair initiator (-1)

- A self- or other-repair initiator that does not get acknowledged by the recipient of the repair-initiator.
  This code is marked on the TCU following the repair initiator, and would include the entire TCU.

Continuers

Utterances that vary from 1 word to a short sentence that encourage the current speaker to continue the turn. These TCUs almost always occur at transition-relevant places with little pause in the conversation (Sacks, Schegloff, and Jefferson 1974)--otherwise, they could be classified as interrupting devices or trouble sources. They do not initiate turns meant to create speaker-change, but effectively pass a possible place of speaker-change (Sidnell 2010:135). They can indicate positive alignment with the other speaker's discourse.

Both sets of *marked* continuers would suggest more engagement with the other speaker's words--a non-minimal response, while both sets of *unmarked* continuers suggest the minimal response that encourage the speaker to continue the turn.

Marked non-lexical and lexical (+1)

- A sound or word that does not specifically relate to the TCU, e.g. Huh. Mhm. Great. Nice. Yeah. The tone is more intense or varied, and/or the pitch is higher. Repetition of a single word without significant pauses, e.g. Right, right, right is considered lexical.

Unmarked non-lexical and lexical (0)

- A sound or word that does not specifically relate to the TCU, e.g. Huh. Mhm. Great. Nice. Yeah. The tone is flat or even. Repetition of a single word without significant pauses, e.g. Right, right, right, is considered lexical.

Marked contingent lexical/phrasal (+2)

- One or more words that specifically relate to the content of the TCU via contingent phrases or exact repetition of words from the TCU, e.g. Who won the game? or /Right,/ Castaway. These presuppose the hearer's understanding of the speaker's TCU. We would look for both
    - Semantic contingency
    - Marked intonation change (pitch, intensity): The tone is more intense and/or the pitch is higher.

Unmarked contingent lexical/phrasal (0)

- One or more words that specifically relate to the content of the TCU via contingent phrases or exact repetition of words from the TCU, e.g. Mhm, soccer. These don't necessarily presuppose the hearer's understanding, as they could be a simple repetition of heard phrases.

If a participant simply repeats a word, i.e. "Right, right" or "Yeah, yeah," this is not considered phrasal, but simply lexical. If a participant says two different words, "Oh, right," that is coded as phrasal.

Codes for Joint Shift in Frames

Linguistic anthropologist Michael Agar describes *frames* in discourse as dynamic sets of expectations that inform participants' interpretations and perceptions of specific speech acts (1996:130-139, 141-145, 165-166). Participants draw implicitly on multiple frames to interpret discourse, build new frames and shift expectations dialectically based throughout a speech act. The following codes highlight contextual cues used by individual participants to mark a shift in frames. *Contextualization cues* are "signs that help speakers hint at, or clarify" inferences about others' discourse, and cues include "prosodic features... paralinguistic features... choice of code, and particular lexical expressions" (Gumperz 1992:229). For all of these codes, we will be considering the semantic content, as well as how the discourse is constructed prosodically.

Irony Markers (0+2)

Irony, as defined by Sidnell, is "the expression of one's meaning by using language that normally signifies the opposite" (2010:70). Irony markers are a stance for a joint alignment in an ironic frame of joint play. Markers of irony include semantic content that violates conversational norms, and intonation that is unusually flat or unusually affected.

- Like laughter as an expression for **feelings of nonseriousness** (Chafe 2007), we will be coding irony markers as **single-sided** or **shared**. Sidnell notes that to maintain intersubjectivity after an ironic marker, the recipient must "show that they understood not only what the words mean, but, moreover, what the speaker meant in using those words" with laughter or by joining the ironic frame with further ironic discourse (70).*
    - **Single-sided irony (0):** Discourse in a TCU that differs markedly in intonation and semantic content in such a way that based on context, it means the opposite of its literal meaning. It is neither preceded by, nor followed by, other ironic markers or laughter.
    - Like single-sided laughter, single-sided irony does not mean that the other participant did not respond positively or negatively to the irony with a facial expression. By itself, it has a valence of 0 for discourse coherence. The audible responses of the other participant--either positively by **sharing the irony (+2)** as defined below or negatively as an **Unreturned bid for play (-1)** reflect the weight of irony in discourse coherence.
    - **Shared irony (+2):** Discourse in a TCU that differs markedly in intonation and semantic content in such a way that based on context, it means the

opposite of its literal meaning. It is either preceded by other ironic markers, or followed by other ironic markers or laughter. We will highlight and code both instances of irony.

    - Like shared laughter, shared irony shows that both participants are on the same page and jointly participating in creating an ironic frame to the discourse. Joining the ironic frame is an explicit action that includes marked intonation changes (+2).

*Responses to irony that demonstrate an understanding of the literal meaning of ironic discourse, but explicitly refute the ironic frame, will be coded with **Unreturned bid for play (-1).**

Shared Performativity (+2 or 0)

The concept of performance in discourse analysis explores *how* humans use language in combination with *what* our use of language does and creates in particular cultural matrices (Dent 2009:54). One aspect of performance is a speaker's orientation towards the audience as a group that can evaluate and respond to discourse (Dent 2009: 46, 239). Another aspect is the speaker's attention to the meanings carried by the lexical expressions and intonation of the discourse itself, or its poetic features (45; Jakobson 1960:358).

- Performativity in a specific piece of discourse can be analyzed by its degree of presence or absence. Marked performativity demarcates an explicit or overt orientation towards an audiences' expected or actual reaction to discourse, as well as towards the discourse itself.
- The code for shared **performativity** will highlight contextualization cues that overtly move the conversation to a highly performative frame, or create a highly performative frame (e.g. pretending that participants are not in fMRI tubes, but somewhere else). Markers include semantic content or turns constructed as if participants are located in another space or time (e.g. "Dude, did you see the size of that fucking shark from the island scenario?"), or if participants use semantic content and intonation to assume a different identity (e.g. in the vacation conversation, PA1 uses reported speech to mime a vendor in the Grand Bazaar: "Pashmina for $15 dolla").
- Like irony markers, if **shared performativity** leads another participant responding by joining in with parallel discourse content or laughter, it will be given a valence of +2 as an explicit and marked indication of jointly created discursive frame. **Single-**

sided performativity will be measured as 0, and an **Unreturned bid for play** would be measured as -1.

Bids for joint action (+2)

Discourse in which a participant expressly requests that both participants be joined together in a common action or experience. They explicitly shift, or ratify, a frame of mutual participation. Bids will be defined by their use of pronouns and verbs that include both participants (we, you and I, let's) constructed as an imperative injunctive to do something together (e.g. "Let's go at it" in the debate.) A bid for joint action can also be constructed as a question that includes pronouns and verbs that include both participants (we, you and I, let's).

- A bid for joint action rates a +2 valence as it overtly requests and projects shared participation in future actions and discourse, along with its more subtle references to the shared nature of the conversation with pronouns and verb conjugations.

Interrogative Reciprocity (+2)

Questions that a) create a speaker-selected speaker change while b) requesting a shift to a topic that parallels the current topic, but from the other participants' perspective.

- For example, if PA1 has been describing their college major, then closes their turn with "What did you major in?" or "What about you?" or "What did you study in college?" Another example would be PA2 expressing an opinion about specific immigration legislation, then asking for PA1's opinion: "What's your opinion on X?" or "How do you feel about X?" These parallel topic centers can be considered co-class members. Interrogative reciprocity can occur at moments we already code for **Topic closure, followed by topic generation that shifts to a co-class member.** Interrogative reciprocity ranks a valence of +2 as it explicitly requests that the other participant join in creating the topic center of a conversation. It asks for a more balanced and shared production vs. PA1 expounding and PA2 using continuers.

Metapragmatics (+2)[2]

Words, phrases, or questions that overtly signal and comment on the dialogic nature of the participants' conversation. *Pragmatics* in the linguistic sense encompasses the "inter-relation of language structure and the principles of language use," (Levinson 1983:9) i.e. the social

---

[2] Given our definition of metapragmatic discourse and our knowledge of the participants, when comparing the frequency of metapragmatic comments between kinds of conversation, we should remember that because we have the exact problem description that the participants read as well as heard from the staff, so we can mark instances of references to that description as metapragmatic indices of a shared outside discourse. For the debate and the college/biographical discussions, we don't have an equivalent corpus to work from.

ideologies that participants use to evaluate and create discourse. Metapragmatics is reflexive discourse that explicitly addresses or discusses how language is used (Lucy 1993:2). Categories highlighted by this code include:

1. Pronouns whose referents index the conversation itself, e.g. "Let's go at **it**" example from the A02711 debate. The code would encompass the relevant piece of discourse in which the pronoun occurs.
2. Questions whose semantic content overtly requests to continue the conversation by referring to the current topic center or a previous topic center, e.g. "Where else should we go?" from the vacation conversation, or "What about some protein sources?" from the problem conversation.
3. TCUs during a repair that request a repetition of a previous participant's TCU, e.g."What did you say?" or "I'm sorry?" when discourse hasn't been heard.
4. Metalanguage that reframes or ratifies the common ground of the participants in these experiments. *Common ground* can include shared general beliefs and language from culture; co-present physical experiences and setting while conversing face-to-face; any previous shared conversation/experiences; and current shared conversation (Clark 1996:12). As common ground can be difficult to determine operationally, we will highlight easily defendable pieces from the shared contexts of the experiment and conversations themselves.
    - ○ Specifically we will mark any a) references to and comments about the paragraph read by the participants before the problem conversation;* b) references to the nursery rhymes, visual signals for stopping and starting the conversation, or pre-experiment training with the NIH staff; c) references to previous discourse that the participants have shared in this conversation (e.g "As I said earlier," or more direct references).
    - ○ This is a specific subset of **External Epistemic Stance Markers** that refers to experiences that we know the participants have <u>shared</u> based on the experiment protocol. In the Discourse Coherence Scale, any quote double-coded as **Metapragmatics** and **External Epistemic Stance Marker** will count as a single +2 valence.

      *Indications that mentions of matches, volcanoes, bags, clothes etc are references to the paragraph will include use of definite pronouns (e.g. the, that) with the noun or a possessive pronoun; and demarcations of remembrance: "I remember we have..." or "They said we've got..." and might include context of marked intonation in responses from the other participant on the noun.

  - All other forms of metapragmatics will be weighed with a +2 valence as they explicitly frame or ratify the conversation as a shared production of the participants, and/or explicitly reference the participants' shared common ground of the experiment's protocol and processes.

Discourse markers (+1)

"Sequentially dependent elements that bracket units of talk" and frame the units' pragmatic meaning in the conversation for participants (Schiffrin 2008:57). Fraser notes that discourse markers are a type of commentary pragmatic marker: "a class of expressions, each of which signals how the speaker intends the basic message that follows to relate to the prior discourse," (Fraser 1990:387).

- We will specifically be looking for a select set of discourse markers that frame TCUs as pieces of a dialogic conversation. Not all occurrences of these words will act as pragmatic discourse markers. Therefore, we will search and find their instances of use, and only code occurrences where the words function as defined below based on the context and surrounding text.
  - a) *'Well'* as a discourse marker that signals a responsive utterance and which "displays a speaker in a particular participation status--respondent." (Schiffrin 1987:103). Used at the beginning of a turn or sentence, *well* frames the discourse following as a response to another participant or to the speaker themselves.
  - b) In certain usages, *'so'* as a discourse marker of a participant's stance. Used as the beginning of a turn, phrase, or sentence, *so* frames the discourse following as containing a stance or opinion that results from prior statements or is being framed as resulting from prior knowledge/understanding, e.g. "So, I mean..." "So, that could cause..." (Schiffrin 1987:223-225).[3] Used to start

---

[3] Schiffrin notes that 'so' is always used at transition-relevant spaces, where it can sometimes pragmatically offer a space for a speaker change to another participant (217-222). If the other participant does not claim a turn, the original speaker can continue talking as it is not a direct request for a speaker-change, e.g. "I think that sounds good. So. Eh. Then I think we should look for coconuts." We will not be coding 'so' in these instances. We will also not code instances where 'so' is used to mark a consequence and not a participant's opinion, e.g. "And they did a deferred rush, so they didn't actually join." or they didn't have

or within a question, ‘so’ can imply that a participant is asking for the other participant’s opinion.

  - c) *‘You know’* as a discourse marker is used by one speaker as a metapragmatic way of checking the other participant's knowledge in general conversation and arguments (271; 280) and "marker of consensual truths" (i.e. We're not all perfect, y'know.) (276). It frames TCUs as dialogic by asking the other participant to focus on a specific piece of information as shared or by checking its sharedness). Occasionally, ‘you know’ is used to begin TCUs in which the other participant cannot have prior knowledge of what follows in the TCU; however, this is still coded as a discourse marker because it cues the other participant that the following discourse is relevant knowledge for both participants to share. Instances of ‘you know’ as a discourse marker generally have a marked intonation which differentiates them from instances of ‘you know’ used as a filler word. [4]

- Discourse markers will weigh +1 on the discourse coherence scale, as they are contextual cues to the participants about how the speaker believes his/her discourse fits into the larger conversation and can positively shape a shared frame of understanding. However, they are more implicit than explicit.

Shared Greetings (+1) An adjacency pair with some combination of ‘Hi’, ‘Hey’, ‘Good xxx’ reflects a pragmatic ideology of one polite/proper/normal way to open a conversation, as well as a way to check that the other person is ready to begin the conversation. We will code the specific adjacency pairs of greetings if one participant opens with ‘Hi/Hey’ plus the other participant’s name or ‘Hi/Hey’ singularly, and then the other person joins the greeting frame with a ‘Hi’ or ‘Hey.’

Codes for Mirroring

Word Repetition (+1 or -1)

---

[3]anyone join sororities until winter. ‘So’ used to begin summarizing or transitioning phrases, even when containing an opinion, will not be coded as a discourse marker (eg. “So that was nice”).

[4] We have a separate code for filler words like 'um' and the transcribers at UMD coded for all filled pauses, so the discourse marker code represents a new significance for specific different occurrences of ‘you know’ as described above.

By looking at word repetition, we will be examining the choice of lexical forms and formulaic expressions that can help to distinguish between main points and qualifying information or side sequences (Gumperz 1992:213-232). The re-iteration of a phrase by another speaker implies that that segment of talk is being foregrounded, which has a significant impact on both speakers' inferential processes.

- The repetition of words and/or phrases from prior turns by the other participant. For this specific code, the repetition will have to reference specific words (more than 1 or more than one generic referent that isn't a proper noun--i.e., marching band counts as one word--or grammatical normality, i.e. 'wanted to') from the turn preceding the utterance being coded, and include some specific word or phrase from the preceding turn. We will only code the segmented repetition, and not the original utterance that is being reproduced; the repetition has to occur in the following line to count.
- This includes when a participant does not use a pronoun to express a referent following its use by another participant, but instead repeats the same referent or substitutes another phrase as an equivalent. This particular repetition represents a lexical choice that has the effect of "separating shared or known items from new information" (Gumperz 1992:232). Since one speaker is choosing not to condense information into a pronoun, he/she is in effect highlighting the known-ness and sharedness of the referent.
- The valence will depend on the purpose of the repetition: affiliative (+1) or agonistic (-1).
- Agonistic turns will be determined by the content of the utterance following the repetition. If the next speaker feels it necessary to defend a prior utterance, then the repetition will be coded as agonistic. If the word repetition does not require a remedy, it will be coded as affiliative.
- (-1) is assigned since the agonistic word repetitions do not threaten intersubjectivity, so much as mark disagreement.

Ex**:** *PA1: "As a anti_immigratio-- to discourage immigrants from."*

PA2: "To discourage immigrants?!"

Mirroring Pauses between Turns (+1) **To be done with script**

Conversation is inherently built upon the taking of turns (see ‘Grossly Apparent Facts’ by Schegloff in the GWU “Conversation Analysis” presentation). This means that speakers are always organizing their talk so as to minimize gaps between talk--so that there is no silence--and make sure there are not too many overlaps--so that both speakers aren’t talking at the same time (Sidnell 2010:37). “The placement and timing of pauses in spoken discourse conveys significant information about the speaker’s discourse production process (Chafe, 1980b) and orientation toward the ongoing conversational interaction (e.g., Goodwin, 1981)” (Edwards and Lampert 1993:61).

- Using the timestamps provided by UMD and NIH, we will calculate the pauses between turns at talk, and we will look for speakers pausing for similar durations in between turns. ‘Similar’ durations will be defined as those that fall within the same standard deviation of the distribution of all pauses between the pair.
- The marking of pauses as ‘similar’ or non-similar can also be thought of in terms of marked or unmarked in regards to length; if someone picks up on it as a bid for common ground or as a trouble marker.
- This code receives a (+1) valence since it is a subconscious affiliation on the part of the speakers, and never gets deliberately address or invoked throughout the conversation

Mirroring of Syntax (+1)

Mirroring of syntax refers to the tendency of speakers to match turn lengths, in this case through the syntactic structure of their utterance: Lexical, Phrasal, Sentential, Multi-Sentential, or Narrative. Some scholars call this phenomenon *syntactic priming*, which “...refers to the increased probability of re-using recently pro-cessed syntactic structures” (Jaeger and Snider 2013:57).

- For this code, we will compare the TCU construction between two turns, and if they match, we will give the second of these turns a +1 valence.
- A score will be assigned to each subsequent turn, unless they do not match, in which case we will not give the turn any valence.
- This code receives a (+1) valence since it is a subconscious affiliation on the part of the speakers, and never gets deliberately address or invoked throughout the conversations.

Codes for Indices of Joint Participation

First person plural (+1)

The basic communicative function of deictic references "... is to individuate or single out objects of reference or address in terms of their relation to the current interactions context in which the utterance occurs" (Hanks 1992:47). When one speaker uses a deictic referent that indexes the two speakers as a shared unit, he/she is in effect saying something about his/her relation to the other and their roles in the current context.

- We will apply this code when an interlocutor indexes the two speakers as some sort of combined unit, or implies their shared agentive or affiliative relations through a deictic reference.
- For example, the use of 'we' (where 'we' indexes the two interlocutors), as well as 'Let's,' 'you and I' and 'our/ours' fit into this. (Note: Referencing a single speaker or a 'we' that indexes a group outside of the dyadic conversation at hand will not be counted).
- This code receives a (+1) valence since it is a subconscious affiliation on the part of the speakers, contributing to a shared state without doing so explicitly.

  Ex: *"We gotta go there"* in vacation conversations

Agreement/Affirmation/Ratification (+1 or +2)

Assessment is an integral aspect of conversation: "One activity that both speakers and recipients perform within the turn at talk is evaluating in some fashion persons and events being described within their talk" (Goodwin and Goodwin 1992:154). In most cases, interlocutors have a high preference for assessments that suggest agreement, and not disagreement; these "...assessments are the vehicles for action..." (Sidnell 2010:82). In this case the action would be to ratify, bolster, or affiliate with the utterances of the other speaker.

- We will code moments in the conversation as being **Agreement** when the utterance of one speaker marks some sort of confirmation, support, or encouragement towards the other interlocutor. These assessments will be based both on word choice and context.
- The quantitative score depends on the markedness of the utterance in terms of intonation, duration, intensity, etc. Marked utterances will receive a valence of +2 and unmarked utterances of Agreement will receive a +1. Specifically, we will look

for intonation and words of the turn following it to see if it's agreement, or just a non-jarring, socially acceptable way of claiming a turn.

- These moments include **Topic Closure Followed by Topic Generation** that include reciprocal confirmations/summarizing assessment, as well as **Marked** and **Unmarked contingent continuers**.
- Ex: Remarks like "I know right" or "same;" Commiseration, e.g. "Oh that sucks;" Turns that begin with "Yeah" or "No, me either"

Codes for Adverse Effect on Discourse Coherence

Open-class Repair Initiators (-2)

Within a repair episode, repair initiators are the reaction to a trouble source. Specifically, open-class repair initiators indicate a problem with the prior turn but don't get particular (Sidnell 2010:119).

- These utterances will be coded by looking for a trouble source and seeing the response; if it is generic and does not reference anything particular in the prior utterance, then it will be open-class.
- Ex: "What?" "Huh?"
- These typically deal with problems that have to do with what the speaker means to accomplish through his/her utterance.
- Since they are so vague, and indicate an inability to articulate the exact trouble source, they will receive a valence of -2. They must be resolved and remedied, engaging both speakers.

Class-specific Repair Initiators (-1)

Within a repair episode, repair initiators are the reaction to a trouble source. In the case of class-specific repair initiators, they identify a particular kind of item in the prior talk as in need of repair (Sidnell 2010:124).

- These utterances will be coded by looking for a trouble source and seeing the response; if it is specific and references something particular in the prior utterance, then it will be class-specific.
- Ex: "Who?" "Where?" "When?'"

- Since one interlocutor would be able to specify the trouble source he/she encountered, the intersubjectivity would appear stronger here compared to the open-class repair initiators. Therefore we will assign these codes a -1 valence.

Unreturned bid for play (-1)

When one speaker makes a bid for play (see **Irony** or **Performativity**), the recipient can fail to accept the bid by either failing to understand entirely, or refusing to carry on within a playful frame. In these cases, the recipient of an ironic or playful utterances fails to "show that they understood not only what the words mean, but, moreover, what the speaker meant in using those words" by NOT responding with laughter or by joining the ironic frame with further ironic discourse (Chafe 2007:70).

- When one speaker does NOT return another speaker's bid for sarcastic or ironic play either by abstaining from laughter or failing to continue the playful frame. We will mark the first sentence of the next turn.
- This code receives a (-1) since the unreturned bid for play--absence of laughter or discontinued play--marks a moment of discontinuity in the affiliation of the two interlocutors, and the breaking of a frame, but does not necessarily threaten shared understanding on the whole.

  Ex: PA2: "What's it like to be rich and go to private school?"

  PA1: "I wouldn't know. I had scholarships..."

Disagreement/Denial/Dissent (-1)

Assessment is an integral aspect of conversation: "One activity that both speakers and recipients perform within the turn at talk is evaluating in some fashion persons and events being described within their talk" (Goodwin and Goodwin 1992:154). However, speakers highly disprefer choosing an assessment that is disagreement (Sidnell 2010:82).

- These utterances are contrarian or dissenting responses to a prior turn at talk. We will determine if the content counts as disagreement by checking context and content to avoid including ironic disagreement or other playful statements.
- Ex: Remarks that express disagreement "No," "I disagree;" Negative assessments of the content of prior speech "That's not true!"

- For the debate, the turn counts as disagreement only if it refers back to a specific point/topic made in the prior turn. So bringing up an entirely new issue in the next turn or diverting the topic does not count as disagreement in these terms.
- This is only a -1 because the two speakers can understand each other, but they just disagree on the topic at hand, which can require re-establishing common ground but does not necessarily threaten intersubjectivity.

Lengthy pauses between turns (-1) **To be done with script**

Interlocutors strive to maintain 'normal' gaps between utterances, however, there are inevitable moments of breakdown. "In terms of positioning, dispreferred responses are often delayed both by inter-turn gap and turn-initial delay." (Sidnell 2010:78). These overly-long pauses mark discontinuity, hence their dispreferred status, which affects "participants' perception of discourse-level coherence, thus influencing interpretation as such" (Gumperz 1992:231).

- We will accrue all the pauses between turns and if one of them falls outside of two standard deviations within the distribution, we will mark the first word of the turn that follows the pause as 'lengthy.'
- Since these pauses are dispreferred, they receive a valence of (-1), which indicates that they break from the expected trajectory of conversation.

NOTE: NIH team (with their definition of turns) will use script for onset and offset of turns in order to look for pauses between turns that are outliers according to some statistical measure.

Filler Words (-1)

Discontinuities within a single speaker's turn often takes the form of filler words, or placeholders. "When speakers in a conversation experience difficulty remembering a word, they may engage in a search for that word...Languages typically offer a range of devices or accomplishing that delay...and these can take...non-lexical but nonetheless conventionalized sounds, such as English uh/uhm, or Hebrew *e* (Schegloff 1979; Clark and Fox Tree 2002)..." (Fox 2010:1)

- Filler words are those with no significant semantic content, or syntactic/pragmatic implications for subsequent speech, such as 'uh' or 'um.' We will code each individual filler word (as many as there are within a turn) and each will be assigned a valence of (-1).

- Note: This code will not be applied to include continuers.
- These codes will receive a negative valence of (-1) since they mark a discontinuity, due to the hesitation that results from searching for a word or idea, but they do not threaten the overall common ground between the interlocutors since they themselves do not require remedy so much as delayed completion.

Works Cited

Agar, Michael (1996). *Language Shock: Understanding the Culture of Conversation*. New York: Harper.

Clark, Herbert (1996). "Common Ground." *In* Using Language. Cambridge: Cambridge University Press. Pp 92-121.

Dent, Alexander (2009). *River of Tears: Country Music, Memory, and Modernity in Brazil*. Durham, NC: Duke University Press.

Fox, Barbara (2010). "Introduction." In Amiridze, N., Davis, B. H., & Maclagan, M. (eds.) *Fillers, Pauses and Placeholders*. Philadelphia: John Benjamins Publishing Co. 1-10.

Gumperz, John (1992). "Contextualization and Understanding." In *Rethinking Context: Language as an Interactive Phenomenon*. Alexander Duranti and Charles Goodwin, eds. Cambridge: Cambridge University Press. 229-252.

Goodwin and Goodwin (1992). "Assessments and the construction of context." In *Rethinking Context: Language as an Interactive Phenomenon*. Alexander Duranti and Charles Goodwin, eds. Cambridge: Cambridge University Press. 147-189.

Hanks, W. (1992). "The indexical ground of deictic reference." In *Rethinking Context: Language as an Interactive Phenomenon*. Alexander Duranti and Charles Goodwin, eds. Cambridge: Cambridge University Press. 43-76.

Jakobson, Roman (1960). "Closing Statement: Linguistics and Poetics." In *Style in Language*. Pp. 350-377.

Lucy, John, ed. (1993). *Reflexive Language: Reported Speech and Metapragmatics*. Cambridge: Cambridge University Press.

Schiffrin, Deborah (2001). "Discourse Markers: Language, Meaning, and Context." In *The Handbook of Discourse Analysis*. Schiffrin, Deborah, Deborah Tannen, and Heidi E. Hamilton, eds. Oxford:Blackwell Publishers, Ltd. 54-75.

-----(1987). *Discourse Markers*. Cambridge: Cambridge University Press.

Sidnell, Jack (2010). *Conversation Analysis: An Introduction*. Malden, MA: Wiley-Blackwell Publishing.

T. Florian Jaeger, Neal E. Snider (2013). "Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience." In *Cognition*, Volume 127, Issue 1: 57–83.

Appendix VIII. Valence Scale for Discourse Coherence Score

Produced by Briel Kobak and Jacqueline Hazen in collaboration with Nuria AbdulSabur

(-2) Very adversely affects Discourse Coherence

The utterance threatens the intersubjectivity of speakers to an extent that a remedy or repair episode must follow, which engages both speakers; a remedy could simply be a significant pause

"Dispreferred responses typically contain explanations or justifications indicating why a dispreferred response is being produced..." (Sidnell 2010: 79).

(-1) Somewhat adversely affects Discourse Coherence

The utterance breaks a frame, rhythm, or trajectory of surrounding speech but does not threaten intersubjectivity to such a degree that a remedy or repair is given substantial attention by both participants

(0) Has neutral effect on Discourse Coherence

The utterance does not mark any significant break in the surrounding speech or intersubjective standing of either participant

(+1) Somewhat improves Discourse Coherence

The utterance contributes to a shared understanding, i.e. shared epistemic or affective stance, without doing so on explicit terms

(+2)Vastly improves Discourse Coherence

The utterance contributes to a shared understanding, i.e. shared epistemic or affective stance, either by doing so overtly/explicitly or by doing so in a marked way (intonation, intensity, etc.)

-2 **Very adversely affects discourse coherence**	**-1** **Somewhat adversely affects discourse coherence**	**0** **Has neutral effect on discourse coherence**	**1** **Somewhat improves discourse coherence**	**2** **Vastly improves discourse coherence**
Non-corrective Laughter	Class-specific Repair Initiator	Epistemic Stance Marker: Internal	Corrective Laughter	Bids for Joint Action
Open-class Repair Initiators	Disagreement/Denial/Dissent	Gradual Shift Topic Shift	Creation and Repetition of Co-constructed Phrase	Epistemic Stance Marker: External
Topic Jump	Filler Words	Self-initiated Speaker Change: Relevant	Discourse Markers	Interrogative Reciprocity
Trouble Source Codes followed by Repair Sequences	Interruption	Single-sided Irony Markers	First Person Plural	Marked Agreement
	Lengthy Pauses between Turns	Single-sided Laughter	Invitation to Spoken Laughter	Marked Contingent Lexical/Phrasal Continuers
	Non-acknowledgement of Repair Initiator	Unmarked Non-lexical and Lexical Continuers	Marked Non-lexical and lexical continuers	Metapragmatics
	Opposite of Pronomial Co-reference: Agonistic	Unmarked Contingent Lexical/Phrasal Continuers	Mirroring of Syntax	Performativity
	Self-initiated Speaker Change: Non-relevant		Mirroring Pauses between Turns	Recognitional Forms

	Unreturned Bid for Play		Non-serious Laughter	Self-initiated Repair
	Word Repetition: Agonistic		Opposite of Pronomial Co-Reference	Shared Irony Markers
			Other-initiated Repair (2 turns or multi-turn)	Shared Laughter
			Self-initiated Speaker Change	
			Shared Use of Pronouns	
			Speaker-selected Speaker Change	
			Topic closure, followed by a topic generation	
			Unmarked Agreement	
			Word Repetition: Affiliative	

References

Acheson, D.J., Hamidi, M., Binder, J.R., and Postle, B.R. (2010). A common neural substrate for language production and verbal working memory. J Cogn Neurosci *23*, 1358-1367.

Akert, R.M., and Panter, A.T. (1988). Extraversion and the ability to decode nonverbal communication. Personality and Individual Differences *9*, 965-972.

Allen, E.A., Erhardt, E.B., Damaraju, E., Gruner, W., Segall, J.M., Silva, R.F., Havlicek, M., Rachakonda, S., Fries, J., Kalyanam, R., *et al.* (2011). A baseline for the multivariate comparison of resting-state networks. Front Syst Neurosci *5*, 2.

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

Aron, A., Melinat, E., Aron, E., Vallone, R., and Bator, R. (1997). The experimental generation of interpersonal closeness: a procedure and some preliminary findings. Personality and Social Psychology Bulletin *23*, 363-377.

Ashburner, J., and Friston, K.J. (2005). Unified segmentation. Neuroimage *26*, 839-851.

Awad, M., Warren, J.E., Scott, S.K., Turkheimer, F.E., and Wise, R.J. (2007). A common system for the comprehension and production of narrative speech. J Neurosci *27*, 11455-11464.

Barch, D.M., Sabb, F.W., Carter, C.S., Braver, T.S., Noll, D.C., and Cohen, J.D. (1999). Overt verbal responding during fMRI scanning: empirical investigations of problems and potential solutions. Neuroimage *10*, 642-657.

Barnes, J.L., and Baron-Cohen, S. (2012). The big picture: storytelling ability in adults with autism spectrum conditions. Journal of Autism and Developmental Disorders *42*, 1557-1565.

Baumeister, R.F., and Leary, M.R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. Psychol Bull *117*, 497-529.

Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. Nature *403*, 309-312.

Bell, A.J., and Sejnowski, T.J. (1995). An information-maximization approach to blind separation and blind deconvolution. Neural Comput *7*, 1129-1159.

Beudel, M., Renken, R., Leenders, K.L., and de Jong, B.M. (2009). Cerebral representations of space and time. Neuroimage *44*, 1032-1040.

Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., and Possing, E.T. (2000). Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex *10*, 512-528.

Binder, J.R., Frost, J.A., Hammeke, T.A., Cox, R.W., Rao, S.M., and Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. J Neurosci *17*, 353-362.

Birn, R.M., Bandettini, P.A., Cox, R.W., Jesmanowicz, A., and Shaker, R. (1998). Magnetic field changes in the human brain due to swallowing or speaking. Magn Reson Med *40*, 55-60.

Birn, R.M., Diamond, J.B., Smith, M.A., and Bandettini, P.A. (2006). Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage *31*, 1536-1548.

Blank, S.C., Scott, S.K., Murphy, K., Warburton, E., and Wise, R.J. (2002). Speech production: Wernicke, Broca and beyond. Brain *125*, 1829-1838.

Blonder, L.X., Burns, A.F., Bowers, D., Moore, R.W., and Heilman, K.M. (1993). Right hemisphere facial expressivity during natural conversation. Brain Cogn *21*, 44-56.

Blumstein, S.E., Baker, E., and Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. Neuropsychologia *15*, 19-30.

Bohland, J.W., Bullock, D., and Guenther, F.H. (2010). Neural representations and mechanisms for the performance of simple speech sequences. J Cogn Neurosci *22*, 1504-1529.

Braun, A.R., Guillemin, A., Hosey, L., and Varga, M. (2001). The neural organization of discourse: an H2 15O-PET study of narrative production in English and American sign language. Brain *124*, 2028-2044.

Bristow, D., Rees, G., and Frith, C.D. (2007). Social interaction modifies neural response to gaze shifts. Soc Cogn Affect Neurosci *2*, 52-61.

Bryant, G.A., and Fox Tree, J.E. (2002). Recognizing verbal irony in spontaneous speech. Metaphor and Symbol *17*, 99-119.

Burgess, P.W., Dumontheil, I., and Gilbert, S.J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. Trends Cogn Sci *11*, 290-298.

Calhoun, V.D., Adali, T., Pearlson, G.D., and Pekar, J.J. (2001). A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp *14*, 140-151.

Callejas, A., Shulman, G.L., and Corbetta, M. (2013). Dorsal and Ventral Attention Systems Underlie Social and Symbolic Cueing. J Cogn Neurosci.

Canessa, N., Borgo, F., Cappa, S.F., Perani, D., Falini, A., Buccino, G., Tettamanti, M., and Shallice, T. (2008). The different neural correlates of action and functional knowledge in semantic memory: an FMRI study. Cereb Cortex *18*, 740-751.

Canli, T., Sivers, H., Whitfield, S.L., Gotlib, I.H., and Gabrieli, J.D. (2002). Amygdala response to happy faces as a function of extraversion. Science *296*, 2191.

Caplan, R., and Dapretto, M. (2001). Making sense during conversation: an fMRI study. Neuroreport *12*, 3625-3632.

Capps, L., Kehres, J., and Sigman, M. (1998). Conversation abilities among children with autism and children with developmental delays. Autism *2*, 325-344.

Carpenter, M., Nagell, K., and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. Monogr Soc Res Child Dev *63*, i-vi, 1-143.

Caspers, J. (2003). Local speech melody as a limiting factor in the turn-taking system in Dutch. Journal of Phonetics *31*, 251-276.

Cavanna, A.E. (2007). The precuneus and consciousness. CNS Spectrums *12*, 545-552.

Cavanna, A.E., and Trimble, M.R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. Brain *129*, 564-583.

Chaminade, T., Kawato, M., and Frith, C. (2011). Individuals' and groups' intentions in the medial prefrontal cortex. Neuroreport *22*, 814-818.

Chan, S.H., Ryan, L., and Bever, T.G. (2011). Role of the striatum in language: Syntactic and conceptual sequencing. Brain Lang.

Chatel-Goldman, J., Schwartz, J.L., Jutten, C., and Congedo, M. (2013). Non-local mind from the perspective of social cognition. Front Hum Neurosci *7*, 107.

Cheek, J.M., and Buss, A.H. (1981). Shyness and sociability. Journal of Personality and Social Psychology *4*, 330-339.

Chen, W., Kato, T., Zhu, X.H., Ogawa, S., Tank, D.W., and Ugurbil, K. (1998). Human primary visual cortex and lateral geniculate nucleus activation during visual imagery. Neuroreport *9*, 3669-3674.

Chow, H.M., Mar, R.A., Xu, Y., Liu, S., Wagage, S., and Braun, A.R. (2013). Embodied Comprehension of Stories: Interactions between Language Regions and Modality-specific Neural Systems. J Cogn Neurosci.

Clark, H.H. (1996). Using language (Cambridge [England] ; New York: Cambridge University Press).

Clark, H.H., and Brennan, S.E. (1991). Grounding in communication. In Perspectivies on Socially Shared Cognition, L.B. Resnick, J.M. Levine, and S.D. Teasley, eds. (Washington, DC: American Psychological Association).

Coelho, C., Le, K., Mozeiko, J., Krueger, F., and Grafman, J. (2012). Discourse production following injury to the dorsolateral prefrontal cortex. Neuropsychologia *50*, 3564-3572.

Coelho, C.A., Grela, B., Corso, M., Gamble, A., and Feinn, R. (2005). Microlinguistic deficits in the narrative discourse of adults with traumatic brain injury. Brain Inj *19*, 1139-1145.

Corballis, M.C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. Behav Brain Sci *26*, 199-208; discussion 208-160.

Corbetta, M., and Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci *3*, 201-215.

Cordes, D., Haughton, V.M., Arfanakis, K., Carew, J.D., Turski, P.A., Moritz, C.H., Quigley, M.A., and Meyerand, M.E. (2001). Frequencies contributing to functional connectivity in the cerebral cortex in "resting-state" data. AJNR Am J Neuroradiol *22*, 1326-1333.

Cox, R.W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res *29*, 162-173.

Crinion, J.T., Warburton, E.A., Lambon-Ralph, M.A., Howard, D., and Wise, R.J. (2006). Listening to narrative speech after aphasic stroke: the role of the left anterior temporal lobe. Cereb Cortex *16*, 1116-1125.

Cuperman, R., and Ickes, W. (2009). Big Five predictors of behavior and perceptions in initial dyadic interactions: personality similarity helps extraverts and introverts, but hurts "disagreeables". J Pers Soc Psychol *97*, 667-684.

D'Argembeau, A., Cassol, H., Phillips, C., Balteau, E., Salmon, E., and Van der Linden, M. (2013). Brains creating stories of selves: the neural basis of autobiographical reasoning. Soc Cogn Affect Neurosci.

Dabbs, J.M., Evans, M.S., Hopper, C.H., and Purvis, J.A. (1980). Self-monitors in conversation: what do they monitor? Journal of Personality and Social Psychology *39*, 278-284.

Dahlberg, C., Hawley, L., Morey, C., Newman, J., Cusick, C.P., and Harrison-Felix, C. (2006). Social communication skills in persons with post-acute traumatic brain injury: three perspectives. Brain Inj *20*, 425-435.

Davis, G.A., and Coelho, C.A. (2004). Referential cohesion and logical coherence of narration after closed head injury. Brain and Language *89*, 508-523.

de Ruiter, J.P., Holger, M., and Enfield, N.H. (2006). Projecting the end of a speakers's turn: a cognitive cornerstone of conversation. Language *82*, 515-535.

De Vico Fallani, F., Nicosia, V., Sinatra, R., Astolfi, L., Cincotti, F., Mattia, D., Wilke, C., Doud, A., Latora, V., He, B.*, et al.* (2010). Defecting or not defecting: how to "read" human behavior during cooperative games by EEG measurements. PLoS One *5*, e14187.

Decety, J., Jackson, P.L., Sommerville, J.A., Chaminade, T., and Meltzoff, A.N. (2004). The neural bases of cooperation and competition: an fMRI investigation. Neuroimage *23*, 744-751.

Decety, J., and Lamm, C. (2007). The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. Neuroscientist *13*, 580-593.

Deckersbach, T., Miller, K.K., Klibanski, A., Fischman, A., Dougherty, D.D., Blais, M.A., Herzog, D.B., and Rauch, S.L. (2006). Regional cerebral brain metabolism correlates of neuroticism and extraversion. Depress Anxiety *23*, 133-138.

Di Martino, A., Scheres, A., Margulies, D.S., Kelly, A.M., Uddin, L.Q., Shehzad, Z., Biswal, B., Walters, J.R., Castellanos, F.X., and Milham, M.P. (2008). Functional connectivity of human striatum: a resting state FMRI study. Cereb Cortex *18*, 2735-2747.

Ding, J., Chen, H., Qiu, C., Liao, W., Warwick, J.M., Duan, X., Zhang, W., and Gong, Q. (2011). Disrupted functional connectivity in social anxiety disorder: a resting-state fMRI study. Magn Reson Imaging *29*, 701-711.

Doucet, G., Naveau, M., Petit, L., Delcroix, N., Zago, L., Crivello, F., Jobard, G., Tzourio-Mazoyer, N., Mazoyer, B., Mellet, E.*, et al.* (2011). Brain activity at rest: a multiscale hierarchical functional organization. J Neurophysiol *105*, 2753-2763.

Dronkers, N.F., Wilkins, D.P., Van Valin, R.D., Jr., Redfern, B.B., and Jaeger, J.J. (2004). Lesion analysis of the brain areas involved in language comprehension. Cognition *92*, 145-177.

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. PLoS One *5*, e12166.

Duncan, S. (1974). On the structure of speaker-auditor interaction during speaking turns. Language in Society *3*, 161-180.

Erhardt, E.B., Rachakonda, S., Bedrick, E.J., Allen, E.A., Adali, T., and Calhoun, V.D. (2011). Comparison of multi-subject ICA methods for analysis of fMRI data. Hum Brain Mapp *32*, 2075-2095.

Eviatar, Z., and Just, M.A. (2006). Brain correlates of discourse processing: an fMRI investigation of irony and conventional metaphor comprehension. Neuropsychologia *44*, 2348-2359.

Farrer, C., Frey, S.H., Van Horn, J.D., Tunik, E., Turk, D., Inati, S., and Grafton, S.T. (2008). The angular gyrus computes action awareness representations. Cereb Cortex *18*, 254-261.

Ferstl, E.C., Rinck, M., and von Cramon, D.Y. (2005). Emotional and temporal aspects of situation model processing during text comprehension: an event-related fMRI study. J Cogn Neurosci *17*, 724-739.

Ferstl, E.C., and von Cramon, D.Y. (2001). The role of coherence and cohesion in text comprehension: an event-related fMRI study. Brain Res Cogn Brain Res *11*, 325-340.

Fleiss, J.L. (1971). Measuring nomincal scale agreement among many raters. Psychological Bulletin *76*, 378-382.

Fletcher, P.C., Frith, C.D., Baker, S.C., Shallice, T., Frackowiak, R.S., and Dolan, R.J. (1995a). The mind's eye--precuneus activation in memory-related imagery. Neuroimage *2*, 195-200.

Fletcher, P.C., Happe, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S., and Frith, C.D. (1995b). Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition *57*, 109-128.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., and Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med *33*, 636-647.

Forstmann, B.U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D.Y., Ridderinkhof, K.R., and Wagenmakers, E.J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. Proc Natl Acad Sci U S A *105*, 17538-17542.

Fossati, P., Hevenor, S.J., Graham, S.J., Grady, C., Keightley, M.L., Craik, F., and Mayberg, H. (2003). In search of the emotional self: an fMRI study using positive and negative emotional words. Am J Psychiatry *160*, 1938-1945.

Fox, M.D., Snyder, A.Z., Vincent, J.L., Corbetta, M., Van Essen, D.C., and Raichle, M.E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci U S A *102*, 9673-9678.

Friederici, A.D., Fiebach, C.J., Schlesewsky, M., Bornkessel, I.D., and von Cramon, D.Y. (2006). Processing linguistic complexity and grammaticality in the left frontal cortex. Cereb Cortex *16*, 1709-1717.

Fries, C.G. (1952). The structure of English (New York: Harcourt Brace).

Ganis, G., Thompson, W.L., and Kosslyn, S.M. (2004). Brain areas underlying visual mental imagery and visual perception: an fMRI study. Brain Res Cogn Brain Res *20*, 226-241.

Ganz, J.B., Heath, A.K., Davis, J.L., and Vannest, K.J. (2013). Effects of a self-monitoring device on socially relevant behaviors in adolescents with Asperger disorder: a pilot study. Assist Technol *25*, 149-157.

Gao, Q., Xu, Q., Duan, X., Liao, W., Ding, J., Zhang, Z., Li, Y., Lu, G., and Chen, H. (2013). Extraversion and neuroticism relate to topological properties of resting-state brain networks. Front Hum Neurosci *7*, 257.

Garrod, S., and Anderson, A. (1987). Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. Cognition *27*, 191-218.

Garrod, S., and Pickering, M.J. (2004). Why is conversation so easy? Trends Cogn Sci *8*, 8-11.

Geschwind, N. (1972). Language and the brain. Sci Am *226*, 76-83.

Ghaem, O., Mellet, E., Crivello, F., Tzourio, N., Mazoyer, B., Berthoz, A., and Denis, M. (1997). Mental navigation along memorized routes activates the hippocampus, precuneus, and insula. Neuroreport *8*, 739-744.

Gilbert, S.J., Williamson, I.D., Dumontheil, I., Simons, J.S., Frith, C.D., and Burgess, P.W. (2007). Distinct regions of medial rostral prefrontal cortex supporting social and nonsocial functions. Soc Cogn Affect Neurosci *2*, 217-226.

Giles, H., Coupland, N., and Coupland, J. (1991). Accommodation theory: communication, context, and consequence. (Cambridge, MA.).

Goodwin, C., and Heritage, J. (1990). Conversation Analysis. Annual Review of Anthropology *19*, 283-307.

Grabner, R.H., Ischebeck, A., Reishofer, G., Koschutnig, K., Delazer, M., Ebner, F., and Neuper, C. (2009). Fact learning in complex arithmetic and figural-spatial tasks: the role of the angular gyrus and its relation to mathematical competence. Hum Brain Mapp *30*, 2936-2952.

Graesser, A.C., and Kreuz, R.J. (1993). A theory of inference generation during text comprehension. Discourse Processes *16*, 145-160.

Graesser, A.C., Singer, M., and Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychol Rev *101*, 371-395.

Grice, H.P. (1975). Logic and Converation. In Syntax and Semantics, Speech Acts, P. Cole, and J.L. Morgan, eds. (New York: Academic Press).

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., and Blake, R. (2000). Brain areas involved in perception of biological motion. J Cogn Neurosci *12*, 711-720.

Gusnard, D.A., Akbudak, E., Shulman, G.L., and Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. Proc Natl Acad Sci U S A *98*, 4259-4264.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. Trends Cogn Sci *9*, 416-423.

Hasson, U., Ghazanfar, A.A., Galantucci, B., Garrod, S., and Keysers, C. (2012). Brain-to-brain coupling: a mechanism for creating and sharing a social world. Trends Cogn Sci *16*, 114-121.

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. Neuron *69*, 407-422.

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. Nat Rev Neurosci *8*, 393-402.

Hoit, J.D., and Lohmeier, H.L. (2000). Influence of continuous speaking on ventilation. J Speech Lang Hear Res *43*, 1240-1251.

Hyvarinen, A., and Oja, E. (2000). Independent component analysis: algorithms and applications. Neural Netw *13*, 411-430.

Ickes, W., and Barnes, R.D. (1977). The role of sex and self-monitoring in unstructured dyadic interactions. Journal of Personality and Social Psychology *35*, 315-330.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. Front Psychol *2*, 255.

Indefrey, P., Hellwig, F., Herzog, H., Seitz, R.J., and Hagoort, P. (2004). Neural responses to the production and comprehension of syntax in identical utterances. Brain Lang *89*, 312-319.

Jafri, M.J., Pearlson, G.D., Stevens, M., and Calhoun, V.D. (2008). A method for functional network connectivity among spatially independent resting-state components in schizophrenia. Neuroimage *39*, 1666-1681.

Johnson, M.K., Raye, C.L., Mitchell, K.J., Touryan, S.R., Greene, E.J., and Nolen-Hoeksema, S. (2006). Dissociating medial frontal and posterior cingulate activity during self-reflection. Soc Cogn Affect Neurosci *1*, 56-64.

Johnson-Laird, P.N. (1980). Mental Models in Cognitive Science. Cognitive Science *4*, 71-115.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. Trends Cogn Sci *9*, 512-518.

Kaplan, J.A., Brownell, H.H., Jacobs, J.R., and Gardner, H. (1990). The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks. Brain Lang *38*, 315-333.

Kastner, S., Pinsk, M.A., De Weerd, P., Desimone, R., and Ungerleider, L.G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron *22*, 751-761.

Keitel, A., Prinz, W., Friederici, A.D., Hofsten, C.V., and Daum, M.M. (2013). Perception of conversations: The importance of semantics and intonation in children's development. J Exp Child Psychol *116*, 264-277.

Kelly, S.D., Ozyurek, A., and Maris, E. (2010). Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. Psychol Sci *21*, 260-267.

Kemeny, S., Ye, F.Q., Birn, R., and Braun, A.R. (2005). Comparison of continuous overt speech fMRI using BOLD and arterial spin labeling. Hum Brain Mapp *24*, 173-183.

Kennerley, S.W., Sakai, K., and Rushworth, M.F. (2004). Organization of action sequences and the role of the pre-SMA. J Neurophysiol *91*, 978-993.

Kjaer, T.W., Nowak, M., and Lou, H.C. (2002). Reflective self-awareness and conscious states: PET evidence for a common midline parietofrontal core. Neuroimage *17*, 1080-1086.

Krause, B.J., Schmidt, D., Mottaghy, F.M., Taylor, J., Halsband, U., Herzog, H., Tellmann, L., and Muller-Gartner, H.W. (1999). Episodic retrieval activates the precuneus irrespective of the imagery content of word pair associates. A PET study. Brain *122 (Pt 2)*, 255-263.

Kubit, B., and Jack, A.I. (2013). Rethinking the role of the rTPJ in attention and social cognition in light of the opposing domains hypothesis: findings from an ALE-based meta-analysis and resting-state functional connectivity. Front Hum Neurosci *7*, 323.

Kuperberg, G.R., Lakshmanan, B.M., Caplan, D.N., and Holcomb, P.J. (2006). Making sense of discourse: an fMRI study of causal inferencing across sentences. Neuroimage *33*, 343-361.

LaFrance, M. (1985). Postural mirroring and intergroup relationships. Personality and Social Psychology Bulletin *11*, 207-217
.

Lambert, S., Sampaio, E., Scheiber, C., and Mauss, Y. (2002). Neural substrates of animal mental imagery: calcarine sulcus and dorsal pathway involvement--an fMRI study. Brain Res *924*, 176-183.

Landis, J.R., and Koch, G.G. (1977). The measurement of observer agreement for categorical data. Biometrics *33*, 159-174.

Laube, I., Kamphuis, S., Dicke, P.W., and Thier, P. (2011). Cortical processing of head- and eye-gaze cues guiding joint social attention. Neuroimage *54*, 1643-1653.

Lawrence, E.J., Shaw, P., Giampietro, V.P., Surguladze, S., Brammer, M.J., and David, A.S. (2006). The role of 'shared representations' in social perception and empathy: an fMRI study. Neuroimage *29*, 1173-1184.

Le, T.H., Pardo, J.V., and Hu, X. (1998). 4 T-fMRI study of nonspatial shifting of selective attention: cerebellar and parietal contributions. J Neurophysiol *79*, 1535-1548.

Leary, M.R., Knight, P.D., and Johnson, K.A. (1987). Social anxiety and dyadic conversation: a verbal response analysis. Journal of Social and Clinical Psychology *5*, 34-50.

Legrand, D., and Ruby, P. (2009). What is self-specific? Theoretical investigation and critical review of neuroimaging results. Psychol Rev *116*, 252-282.

Lehman Blake, M. (2006). Clinical relevance of discourse characteristics after right hemisphere brain damage. Am J Speech Lang Pathol *15*, 255-267.

Leube, D., Straube, B., Green, A., Blumel, I., Prinz, S., Schlotterbeck, P., and Kircher, T. (2012). A possible brain network for representation of cooperative behavior and its implications for the psychopathology of schizophrenia. Neuropsychobiology *66*, 24-32.

Leung, H.C., Gore, J.C., and Goldman-Rakic, P.S. (2002). Sustained mnemonic response in the human middle frontal gyrus during on-line storage of spatial memoranda. J Cogn Neurosci *14*, 659-671.

Li, Y.O., Adali, T., and Calhoun, V.D. (2007). Estimating the number of independent components for functional magnetic resonance imaging data. Hum Brain Mapp *28*, 1251-1266.

Liu, S., Chow, H.M., Xu, Y., Erkkinen, M.G., Swett, K.E., Eagle, M.W., Rizik-Baer, D.A., and Braun, A.R. (2012). Neural correlates of lyrical improvisation: an FMRI study of freestyle rap. Sci Rep *2*, 834.

Lopes, P.N., Salovey, P., and Straus, R. (2003). Emotional intelligence, personality, and the perceived quality of social relationships. Personality and Individual Differences *35*, 641-658.

Lou, H.C., Nowak, M., and Kjaer, T.W. (2005). The mental self. Prog Brain Res *150*, 197-204.

Mano, Y., Harada, T., Sugiura, M., Saito, D.N., and Sadato, N. (2009). Perspective-taking as part of narrative comprehension: a functional MRI study. Neuropsychologia *47*, 813-824.

Manto, M., Bower, J.M., Conforto, A.B., Delgado-Garcia, J.M., da Guarda, S.N., Gerwig, M., Habas, C., Hagura, N., Ivry, R.B., Marien, P.*, et al.* (2012). Consensus paper: roles of the cerebellum in motor control--the diversity of ideas on cerebellar involvement in movement. Cerebellum *11*, 457-487.

Mar, R.A. (2004). The neuropsychology of narrative: story comprehension, story production and their interrelation. Neuropsychologia *42*, 1414-1434.

Mar, R.A. (2011). The neural bases of social cognition and story comprehension. Annu Rev Psychol *62*, 103-134.

Margulies, D.S., Vincent, J.L., Kelly, C., Lohmann, G., Uddin, L.Q., Biswal, B.B., Villringer, A., Castellanos, F.X., Milham, M.P., and Petrides, M. (2009). Precuneus shares intrinsic functional architecture in humans and monkeys. Proc Natl Acad Sci U S A *106*, 20069-20074.

Marini, A. (2012). Characteristics of narrative discourse processing after damage to the right hemisphere. Semin Speech Lang *33*, 68-78.

Marini, A., and Urgesi, C. (2012). Please get to the point! A cortical correlate of linguistic informativeness. Journal of Cognitive Neuroscience *24*, 2211-2222.

Martin-Loeches, M., Casado, P., Hernandez-Tamames, J.A., and Alvarez-Linera, J. (2008). Brain activation in discourse comprehension: a 3t fMRI study. Neuroimage *41*, 614-622.

Marvel, C.L., and Desmond, J.E. (2010). The contributions of cerebro-cerebellar circuitry to executive verbal working memory. Cortex *46*, 880-895.

Mason, R.A., and Just, M.A. (2011). Differentiable cortical networks for inferences concerning people's intentions versus physical causality. Hum Brain Mapp *32*, 313-329.

Mazoyer, B., Tzourio, N., Frak, V., Syrota, S., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., and Mehler, J. (1993). The cortical representation of speech. Journal of Cognitive Neuroscience *5*, 467-479.

McCabe, K., Houser, D., Ryan, L., Smith, V., and Trouard, T. (2001). A functional imaging study of cooperation in two-person reciprocal exchange. Proc Natl Acad Sci U S A *98*, 11832-11835.

McCarthy, G., Puce, A., Constable, R.T., Krystal, J.H., Gore, J.C., and Goldman-Rakic, P. (1996). Activation of human prefrontal cortex during spatial and nonspatial working memory tasks measured by functional MRI. Cereb Cortex *6*, 600-611.

McDonald, S., and Flanagan, S. (2004). Social perception deficits after traumatic brain injury: interaction between emotion recognition, mentalizing ability, and social communication. Neuropsychology *18*, 572-579.

McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.P., Kindermann, S.S., Bell, A.J., and Sejnowski, T.J. (1998). Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp *6*, 160-188.

Meister, I.G., Wilson, S.M., Deblieck, C., Wu, A.D., and Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. Curr Biol *17*, 1692-1696.

Menenti, L., Petersson, K.M., Scheeringa, R., and Hagoort, P. (2009). When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. J Cogn Neurosci *21*, 2358-2368.

Menenti, L., Pickering, M.J., and Garrod, S.C. (2012). Toward a neural basis of interactive alignment in conversation. Front Hum Neurosci *6*, 185.

Mitchell, J.P., Banaji, M.R., and Macrae, C.N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. J Cogn Neurosci *17*, 1306-1315.

Monchi, O., Petrides, M., Strafella, A.P., Worsley, K.J., and Doyon, J. (2006). Functional role of the basal ganglia in the planning and execution of actions. Ann Neurol *59*, 257-264.

Montague, P.R., Berns, G.S., Cohen, J.D., McClure, S.M., Pagnoni, G., Dhamala, M., Wiest, M.C., Karpov, I., King, R.D., Apple, N.*, et al.* (2002). Hyperscanning: simultaneous fMRI during linked social interactions. Neuroimage *16*, 1159-1164.

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S.F., and Fazio, F. (2001). Syntax and the brain: disentangling grammar by selective anomalies. Neuroimage *13*, 110-118.

Mort, D.J., Malhotra, P., Mannan, S.K., Rorden, C., Pambakian, A., Kennard, C., and Husain, M. (2003). The anatomy of visual neglect. Brain *126*, 1986-1997.

Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Sawamoto, N., Toma, K., Nakamura, K., Hanakawa, T., Konishi, J*., et al.* (1999). Transient neural activity in the medial superior frontal gyrus and precuneus time locked with attention shift between object features. Neuroimage *10*, 193-199.

Neda, Z., Ravasz, E., Brechet, Y., Vicsek, T., and Barabasi, A.L. (2000). The sound of many hands clapping. Nature *403*, 849-850.

Nichelli, P., Grafman, J., Pietrini, P., Clark, K., Lee, K.Y., and Miletich, R. (1995). Where the brain appreciates the moral of a story. Neuroreport *6*, 2309-2313.

Norbury, C.F., and Bishop, D.V. (2003). Narrative skills of children with communication impairments. Int J Lang Commun Disord *38*, 287-313.

Northoff, G., and Bermpohl, F. (2004). Cortical midline structures and the self. Trends Cogn Sci *8*, 102-107.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain--a meta-analysis of imaging studies on the self. Neuroimage *31*, 440-457.

Ojemann, J.G., Akbudak, E., Snyder, A.Z., McKinstry, R.C., Raichle, M.E., and Conturo, T.E. (1997). Anatomic localization and quantitative analysis of gradient refocused echo-planar fMRI susceptibility artifacts. Neuroimage *6*, 156-167.

Okada, K., and Hickok, G. (2006). Identification of lexical-phonological networks in the superior temporal sulcus using functional magnetic resonance imaging. Neuroreport *17*, 1293-1296.

Ormel, J., Bastiaansen, A., Riese, H., Bos, E.H., Servaas, M., Ellenbogen, M., Rosmalen, J.G., and Aleman, A. (2013). The biological and psychological basis of neuroticism: current status and future directions. Neurosci Biobehav Rev *37*, 59-72.

Papathanassiou, D., Etard, O., Mellet, E., Zago, L., Mazoyer, B., and Tzourio-Mazoyer, N. (2000). A common language network for comprehension and production: a contribution to the definition of language epicenters with PET. Neuroimage *11*, 347-357.

Paus, T., Petrides, M., Evans, A.C., and Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: a positron emission tomography study. J Neurophysiol *70*, 453-469.

Pessoa, L., Rossi, A., Japee, S., Desimone, R., and Ungerleider, L.G. (2009). Attentional control during the transient updating of cue information. Brain Res *1247*, 149-158.

Pickering, M.J., and Garrod, S. (2013). An integrated theory of language production and comprehension. Behav Brain Sci *36*, 329-347.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., and Shulman, G.L. (2001). A default mode of brain function. Proc Natl Acad Sci U S A *98*, 676-682.

Ramnani, N., and Owen, A.M. (2004). Anterior prefrontal cortex: insights into function from anatomy and neuroimaging. Nat Rev Neurosci *5*, 184-194.

Redcay, E., Kleiner, M., and Saxe, R. (2012). Look at this: the neural correlates of initiating and responding to bids for joint attention. Front Hum Neurosci *6*, 169.

Richardson, D.C., Dale, R., and Kirkham, N.Z. (2007a). The art of conversation is coordination: common ground and the coupling of eye movements during dialogue. Psychol Sci *18*, 407-413.

Richardson, M.J., Marsh, K.L., Isenhower, R.W., Goodman, J.R., and Schmidt, R.C. (2007b). Rocking together: dynamics of intentional and unintentional interpersonal coordination. Hum Mov Sci *26*, 867-891.

Rilling, J., Gutman, D., Zeh, T., Pagnoni, G., Berns, G., and Kilts, C. (2002). A neural basis for social cooperation. Neuron *35*, 395-405.

Rissman, J., Eliassen, J.C., and Blumstein, S.E. (2003). An event-related FMRI investigation of implicit semantic priming. J Cogn Neurosci *15*, 1160-1175.

Robertson, D.A., Gernsbacher, M.A., Guidotti, S.J., Robertson, R.R., Irwin, W., Mock, B.J., and Campana, M.E. (2000). Functional neuroanatomy of the cognitive process of mapping during discourse comprehension. Psychol Sci *11*, 255-260.

Rossi, A.F., Pessoa, L., Desimone, R., and Ungerleider, L.G. (2009). The prefrontal cortex and the executive control of attention. Exp Brain Res *192*, 489-497.

Sacks, H., Schegloff, E.A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. Language *50*, 696-735.

Salmi, J., Rinne, T., Koistinen, S., Salonen, O., and Alho, K. (2009). Brain networks of bottom-up triggered and top-down controlled shifting of auditory attention. Brain Res *1286*, 155-164.

Sander, D., Grandjean, D., Pourtois, G., Schwartz, S., Seghier, M.L., Scherer, K.R., and Vuilleumier, P. (2005). Emotion and attention interactions in social cognition: brain regions involved in processing anger prosody. Neuroimage *28*, 848-858.

Sanger, J., Muller, V., and Lindenberger, U. (2012). Intra- and interbrain synchronization and network properties when playing guitar in duets. Front Hum Neurosci *6*, 312.

Sarela, J., and Vigario, R. (2003). Overlearning n marginal distribution-based ICA: analysis and solutions. Journal of Machine Learning Research *4*, 1447-1469.

Saur, D., Kreher, B.W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M.S., Umarova, R., Musso, M., Glauche, V., Abel, S*., et al.* (2008). Ventral and dorsal pathways for language. Proc Natl Acad Sci U S A *105*, 18035-18040.

Saxe, R., Carey, S., and Kanwisher, N. (2004). Understanding other minds: linking developmental psychology and functional neuroimaging. Annu Rev Psychol *55*, 87-124.

Saxe, R., and Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". Neuroimage *19*, 1835-1842.

Saxe, R., Moran, J.M., Scholz, J., and Gabrieli, J. (2006). Overlapping and non-overlapping brain regions for theory of mind and self reflection in individual subjects. Soc Cogn Affect Neurosci *1*, 229-234.

Saxe, R., and Powell, L.J. (2006). It's the thought that counts: specific brain regions for one component of theory of mind. Psychol Sci *17*, 692-699.

Schiffrin, D. (1987). Discourse Markers (Cambridge: Cambridge University Press).

Schilbach, L., Wilms, M., Eickhoff, S.B., Romanzetti, S., Tepest, R., Bente, G., Shah, N.J., Fink, G.R., and Vogeley, K. (2010). Minds made for sharing: initiating joint attention recruits reward-related neurocircuitry. J Cogn Neurosci *22*, 2702-2715.

Schlenker, B.R., and Leary, M.R. (1985). Social anxiety and communication about the self. Journal of Language and Social Psychology *4*, 171-192.

Schultz, R.T., Grelotti, D.J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., and Skudlarski, P. (2003). The role of the fusiform face area in social cognition: implications for the pathobiology of autism. Philos Trans R Soc Lond B Biol Sci *358*, 415-427.

Schulz, G.M., Varga, M., Jeffires, K., Ludlow, C.L., and Braun, A.R. (2005). Functional neuroanatomy of human vocalization: an H215O PET study. Cereb Cortex *15*, 1835-1847.

Seghier, M.L., Fagan, E., and Price, C.J. (2010). Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. J Neurosci *30*, 16809-16817.

Selting, M. (1996). On the interplay of syntax and prosody in the constitution of turn-constructional units and turns in conversation. Pragmatics *6*, 371-388.

Servaas, M.N., Riese, H., Renken, R.J., Marsman, J.B., Lambregs, J., Ormel, J., and Aleman, A. (2013a). The effect of criticism on functional brain connectivity and associations with neuroticism. PLoS One *8*, e69606.

Servaas, M.N., van der Velde, J., Costafreda, S.G., Horton, P., Ormel, J., Riese, H., and Aleman, A. (2013b). Neuroticism and the brain: A quantitative meta-analysis of neuroimaging studies investigating emotion processing. Neurosci Biobehav Rev *37*, 1518-1529.

Sestieri, C., Corbetta, M., Romani, G.L., and Shulman, G.L. (2011). Episodic memory retrieval, parietal cortex, and the default mode network: functional and topographic analyses. J Neurosci *31*, 4407-4420.

Shibata, M., Toyomura, A., Itoh, H., and Abe, J. (2010). Neural substrates of irony comprehension: A functional MRI study. Brain Res *1308*, 114-123.

Shima, K., and Tanji, J. (2000). Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. J Neurophysiol *84*, 2148-2160.

Shockley, K., Santana, M.V., and Fowler, C.A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. J Exp Psychol Hum Percept Perform *29*, 326-332.

Shomstein, S., and Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. J Neurosci *24*, 10702-10706.

Sidnell, J. (2010). Conversation Analysis: An Introduction (Malden, MA: Wiley-Blackwell Publishing).

Sieborger, F.T., Ferstl, E.C., and von Cramon, D.Y. (2007). Making sense of nonsense: an fMRI study of task induced inference processes during discourse comprehension. Brain Res *1166*, 77-91.

Siegal, M., Carrington, J., and Radel, M. (1996). Theory of mind and pragmatic understanding following right hemisphere damage. Brain Lang *53*, 40-50.

Snow, P., Douglas, J., and Ponsford, J. (1998). Conversational discourse abilities following severe traumatic brain injury: a follow-up study. Brain Inj *12*, 911-935.

Spalletta, G., Spoletini, I., Cherubini, A., Rubino, I.A., Siracusano, A., Piras, F., Caltagirone, C., and Marini, A. (2010). Cortico-subcortical underpinnings of narrative processing impairment in schizophrenia. Psychiatry Research *182*, 77-80.

Sprecher, S., Treger, S., and Wondra, J.D. (2013). Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions. Journal of Social and Personal Relationships *30*, 497-514.

Spreng, R.N., Mar, R.A., and Kim, A.S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J Cogn Neurosci *21*, 489-510.

Stephens, G.J., Silbert, L.J., and Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. Proceedings of the National Academy of Sciences of the United States of America.

Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J.P., Yoon, K.E., *et al.* (2009). Universals and cultural variation in turn-taking in conversation. Proc Natl Acad Sci U S A *106*, 10587-10592.

Studholme, C., Hill, D.L., and Hawkes, D.J. (1999). An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition *32*, 71-86.

Stuss, D.T., and Levine, B. (2002). Adult clinical neuropsychology: lessons from studies of the frontal lobes. Annu Rev Psychol *53*, 401-433.

Suda, M., Takei, Y., Aoyama, Y., Narita, K., Sakurai, N., Fukuda, M., and Mikuni, M. (2011). Autistic traits and brain activation during face-to-face conversations in typically developed adults. PLoS One *6*, e20021.

Suda, M., Takei, Y., Aoyama, Y., Narita, K., Sato, T., Fukuda, M., and Mikuni, M. (2010). Frontopolar activation during face-to-face conversation: an in situ study using near-infrared spectroscopy. Neuropsychologia *48*, 441-447.

Surian, L., and Siegal, M. (2001). Sources of performance on theory of mind tasks in right hemisphere-damaged patients. Brain Lang *78*, 224-232.

Takei, Y., Suda, M., Aoyama, Y., Yamaguchi, M., Sakurai, N., Narita, K., Fukuda, M., and Mikuni, M. (2013). Temporal lobe and inferior frontal gyrus dysfunction in patients with schizophrenia during face-to-face conversation: A near-infrared spectroscopy study. J Psychiatr Res *47*, 1581-1589.

Thompson, C.K., Cho, S., Hsu, C.J., Wieneke, C., Rademaker, A., Weitner, B.B., Mesulam, M.M., and Weintraub, S. (2012). Dissociations Between Fluency And Agrammatism In Primary Progressive Aphasia. Aphasiology *26*, 20-43.

Thorne, A. (1987). The press of personality: a study of conversations between introverts and extraverts. Journal of Personality and Social Psychology *53*, 718-726.

Tomasello, M. (1999). The Cultural Origins of Human Cognition (London, England: Harvard University Press).

Tomasello, M. (2008). Origins of Human Communication (Cambridge, MA: MIT Press).

Tomasello, M., and Farrar, M.J. (1986). Joint attention and early language. Child Dev *57*, 1454-1463.

Tompkins, C.A., Scharp, V.L., Fassbinder, W., Meigh, K.M., and Armstrong, E.M. (2008a). A different story on "Theory of Mind" deficit in adults with right hemisphere brain damage. Aphasiology *22*, 42-61.

Tompkins, C.A., Scharp, V.L., Meigh, K.M., and Fassbinder, W. (2008b). Coarse coding and discourse comprehension in adults with right hemisphere brain damage. Aphasiology *22*, 204-223.

Troiani, V., Fernandez-Seara, M.A., Wang, Z., Detre, J.A., Ash, S., and Grossman, M. (2008). Narrative speech production: an fMRI study using continuous arterial spin labeling. Neuroimage *40*, 932-939.

Tucker, F.M., and Hanlon, R.E. (1998). Effects of mild traumatic brain injury on narrative discourse production. Brain Inj *12*, 783-792.

Turken, A.U., and Dronkers, N.F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. Frontiers in Systems Neuroscience *5*, 1.

Turkstra, L.S., McDonald, S., and DePompei, R. (2001). Social information processing in adolescents: data from normally developing adolescents and preliminary data from their peers with traumatic brain injury. J Head Trauma Rehabil *16*, 469-483.

Turner, M.S., Simons, J.S., Gilbert, S.J., Frith, C.D., and Burgess, P.W. (2008). Distinct roles for lateral and medial rostral prefrontal cortex in source monitoring of perceived and imagined events. Neuropsychologia *46*, 1442-1453.

Uchiyama, H., Seki, A., Kageyama, H., Saito, D.N., Koeda, T., Ohno, K., and Sadato, N. (2006). Neural substrates of sarcasm: a functional magnetic-resonance imaging study. Brain Res *1124*, 100-110.

Uddin, L.Q., Supekar, K., Amin, H., Rykhlevskaia, E., Nguyen, D.A., Greicius, M.D., and Menon, V. (2010). Dissociable connectivity within human angular gyrus and intraparietal sulcus: evidence from functional and structural connectivity. Cereb Cortex *20*, 2636-2646.

van Dijk, T.A., and Kintsch, W. (1983). Strategies of discourse comprehension (New York: Academic Press).

Vanhaudenhuyse, A., Demertzi, A., Schabus, M., Noirhomme, Q., Bredart, S., Boly, M., Phillips, C., Soddu, A., Luxen, A., Moonen, G., *et al.* (2011). Two distinct neuronal networks mediate the awareness of environment and of self. J Cogn Neurosci *23*, 570-578.

Vigneau, M., Beaucousin, V., Herve, P.Y., Jobard, G., Petit, L., Crivello, F., Mellet, E., Zago, L., Mazoyer, B., and Tzourio-Mazoyer, N. (2011). What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing? Insights from a meta-analysis. Neuroimage *54*, 577-593.

Visser, M., Jefferies, E., and Lambon Ralph, M.A. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. J Cogn Neurosci *22*, 1083-1094.

Vittengl, J.R., and Holt, C.S. (2000). Getting acquainted: the relations of self-disclosure and social attraction to positive affect. Journal of Social and Personal Relationships *17*, 53-66.

Waugh, C.E., and Fredrickson, B.L. (2006). Nice to know you: Positive emotions, self-other overlap, and complex understanding in the formation of a new relationship. J Posit Psychol *1*, 93-106.

Weber, A., Grice, M., and Crocker, M.W. (2006). The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. Cognition *99*, B63-72.

Wells, B., and Macfarlane, S. (1998). Prosody as an interactional resource: turn-projection and overlap. Lang Speech *41 (Pt 3-4)*, 265-294.

Whitehouse, A.J., Watt, H.J., Line, E.A., and Bishop, D.V. (2009). Adult psychosocial outcomes of children with specific language impairment, pragmatic language impairment and autism. Int J Lang Commun Disord *44*, 511-528.

Whitney, C., Huber, W., Klann, J., Weis, S., Krach, S., and Kircher, T. (2009). Neural correlates of narrative shifts during auditory story comprehension. Neuroimage *47*, 360-366.

Wiemann, J.M., and Knapp, M.L. (1975). Turn-taking in conversations. Journal of Communication *25*, 75-92.

Wilkes-Gibbs, D. (1995). Coherence in collaboration: some examples from conversation. In Coherence in Spontaneous Text, M.A. Gernsbacher, and T. Givon, eds. (John Benjamins Publishing).

Willems, R.M., de Boer, M., de Ruiter, J.P., Noordzij, M.L., Hagoort, P., and Toni, I. (2010). A dissociation between linguistic and communicative abilities in the human brain. Psychol Sci *21*, 8-14.

Williams, J.H., Waiter, G.D., Perra, O., Perrett, D.I., and Whiten, A. (2005). An fMRI study of joint attention experience. Neuroimage *25*, 133-140.

Wilson, M., and Wilson, T.P. (2005). An oscillator model of the timing of turn-taking. Psychon Bull Rev *12*, 957-968.

Winner, E., Brownell, H., Happe, F., Blum, A., and Pincus, D. (1998). Distinguishing lies from jokes: theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. Brain Lang *62*, 89-106.

Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., and Frackowiak, R. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. Brain *114 (Pt 4)*, 1803-1817.

Wittfoth, M., Buck, D., Fahle, M., and Herrmann, M. (2006). Comparison of two Simon tasks: neuronal correlates of conflict resolution based on coherent motion perception. Neuroimage *32*, 921-929.

Wright, C.I., Williams, D., Feczko, E., Barrett, L.F., Dickerson, B.C., Schwartz, C.E., and Wedig, M.M. (2006). Neuroanatomical correlates of extraversion and neuroticism. Cereb Cortex *16*, 1809-1819.

Xu, J., Kemeny, S., Park, G., Frattali, C., and Braun, A. (2005). Language in context: emergent features of word, sentence, and narrative comprehension. Neuroimage *25*, 1002-1015.

Xu, J., Potenza, M.N., and Calhoun, V.D. (2013). Spatial ICA reveals functional activity hidden from traditional fMRI GLM-based analyses. Front Neurosci *7*, 154.

Yamasaki, H., LaBar, K.S., and McCarthy, G. (2002). Dissociable prefrontal brain systems for attention and emotion. Proc Natl Acad Sci U S A *99*, 11447-11451.

Yarkoni, T., Speer, N.K., and Zacks, J.M. (2008). Neural substrates of narrative comprehension and memory. Neuroimage *41*, 1408-1425.

Yngve, V. (1970). On getting a word in edgewise. Papers from the Sixth Regional Meeting of the Chicago Linguistic Society *6*, 567-577.

Yun, K., Watanabe, K., and Shimojo, S. (2012). Interpersonal body and neural synchronization as a marker of implicit social interaction. Sci Rep *2*, 959.

Zhang, S., and Li, C.S. (2012). Functional connectivity mapping of the human precuneus by resting state fMRI. Neuroimage *59*, 3548-3562.