

The bilateral symmetry observed in the production network challenges the current model of a production system that is strongly lateralized to the left hemisphere (Hickok and Poeppel, 2007; Poeppel et al., 2012). This model assumes a dorsal linguistic stream associated with speech production that includes left-lateralized structures in the posterior frontal lobe and the posterior dorsal-most aspect of the temporal lobe and parietal operculum (Hickok and Poeppel, 2007). The model is supported by a number of lesion and neuroimaging studies also advocating for left-lateralized participation in speech production (see (Indefrey, 2011) for a review). A major difference between this study and all others investigating language processing is the production task. Here the speakers produced a long, unconstrained, real-life narrative, which is likely to recruit a larger network of brain areas relative to prior studies that have focused mainly on the production of short and unrelated utterances (Pickering and Garrod, 2004; Indefrey, 2011).

Some recent publications support the lack of lateralized responses, however, and report bilaterality in speech production (Zatorre et al., 2002; Giraud et al., 2007; Federmeier et al., 2008; Indefrey, 2011; Poeppel et al., 2012). Most notably, Braun et al (2001) used PET imaging to study spontaneous speech production in both English and American Sign Language speakers and found robust common activation in brain areas beyond the classical language areas and including extrasylvian regions in both right and left hemispheres (Braun et al., 2001). The growing imaging and physiological reports of right hemisphere homologues for many of the left hemisphere areas involved in speech production introduce an interesting puzzle; they are in apparent conflict with lesion data that persistently point to the lateralization of speech to the dominant hemisphere.

Resolving this puzzle is beyond the scope of this study, as it was not designed to discern the nature of the right hemisphere contribution. Moreover, against the background of predominantly bilateral production-related neural activity, we did note some laterality in the production system. For example, the anterior portion of the left IFG exhibited stronger production-related reliability than in the right hemisphere. In addition, we observed greater production responses in the left MTG than the right.

Activity during speech production was found in areas traditionally thought to be involved in speech comprehension, such as the STG, MTG, and angular gyrus (Hickok and Poeppel, 2007) (Dronkers et al., 1994; Caplan et al., 1996; Binder et al., 1997; Dronkers et al., 2000; Price, 2000; Bookheimer, 2002; Damasio and Damasio, 2002; Friederici, 2002; Bates et al., 2003; Dronkers et al., 2004; Vigneau et al., 2006; Caplan et al., 2007; Hickok and Poeppel, 2007; Ferstl et al., 2008; Tyler and Marslen-Wilson, 2008; Binder et al., 2009; Price, 2010). Although it is initially surprising to observe production activity in traditionally comprehension-based areas, these results are consistent with recent findings. The left MTG, for example, seems to be crucial in the retrieval of lexical items, a process that may be necessary for the production of words (Peelle et al., 2010). The left pSTG also appears to play a role in language production according to PET scans during the spontaneous speech production of sign language (Braun et al., 2001). (For a more in depth review of these areas see Chapter 3 Discussion).

Real-world speech production also activated a collection of extra-linguistic areas known to be involved in the processing of semantic and social aspects of the story (Xu et al., 2005), including the precuneus, anterior cingulate, and medial prefrontal cortex.

These results are in contrast to current models of speech production that are confined to structures in the posterior frontal lobe and the posterior dorsal-most aspect of the temporal lobe and parietal operculum (Hickok and Poeppel, 2007). The recruitment of such extra-linguistic areas during real-world speech production follows from the complexity of the stimuli used in this study, which requires the use of extra-linguistic functions. For example, a recent meta-analysis of the precuneus suggests a role for the region in episodic memory retrieval, self-processing, and consciousness (Cavanna and Trimble, 2006), all functions relied upon in the production of a personally relevant, real-world narrative. Further, it is acknowledged that the mPFC mediates social event knowledge (Krueger and Grafman, 2008), and has been shown to encode the stable representation of predictive event sequences (Koechlin et al., 2000). That brain areas mediating such functions would be activated during real-world speech production is not surprising, as the complexity and emotional saliency of the task is far more complicated than the production of isolated units of speech.

This new map of the production system active in a speaker's brain is significant in that it is both reliable and selective across many retellings of a highly complex narrative. Previous studies have isolated brain areas active during specific and constrained speech. Here we show that the entire network of shared neural activity during speech production of real-world storytelling is as extensive and as reliable as the comprehension system during listening, further explored below.

II) Speech Comprehension Network and its overlap with the Speech Production Network

Introduction:

Recent findings suggest that a large portion of the cortex evokes reliable and selective responses to natural stimuli (e.g. listening to a story), which are shared across all subjects (Hasson et al., 2004; Golland et al., 2007; Hasson et al., 2008a; Wilson et al., 2008). These studies use the inter-subject correlation method to characterize the similarity of cortical responses across individuals during natural viewing conditions (for a recent review see (Hasson et al., 2009)). Despite the unrestrained task of freely listening to a story and the complexity of the stimuli, the similarity in response patterns across all subjects listening to the same story suggests that incoming verbal information is processed in similar ways. Recent findings further expose the complexity of the shared brain responses among listeners by characterizing the time scale of processing at different stages of the auditory system (Lerner et al., 2011). In particular, the reliability of neural responses evoked throughout the cortex was measured while participants listened to parametrically scrambled versions of a real-life auditory story. The story was scrambled at the word, sentence, paragraph, and full story level, allowing for the detection of each brain area's sensitivity to information arriving at differing points in the past. Results show that each brain area shared across listeners reliably integrates information over its preferred time scale. Specifically, early auditory areas responded reliably to all conditions regardless of their temporal structure, while the reliability of responses decreased gradually for temporally unstructured segments moving rostrally and caudally from A1+ to higher-order areas (Lerner et al., 2011). This demonstrates that the brain relies on a

distributed hierarchical network of brain areas to accumulate information over time that is shared among all listeners during comprehension.

In this study, we first recorded the brain responses of listeners listening to the story told by the speaker in Chapter 1 in order to map the entire network of areas reliably activated during the comprehension of an unconstrained, real-world story. The agreement with previous work is far from assured: the story here was both personal and spontaneous, and was recorded in the noisy environment of the scanner. A similarity in the response patterns across all listeners would therefore underscore a strong tendency to process incoming verbal information in similar ways.

The full spatial overlap between the production and comprehension systems is not known for two main reasons. First, the extent of the production system remains uncertain (see Chapter 1). Second, few neurolinguistic studies measure production and comprehension of speech using the same speech materials (Wilson et al., 2004; Okada and Hickok, 2006). Some studies have attempted to map the extent of overlap between brain areas dedicated to the production and comprehension of speech (Indefrey et al., 2004; Awad et al., 2007; Heim, 2008; Price, 2010; Indefrey, 2011; Menenti et al., 2011; Price, 2012; Segaert et al., 2012). Spatial overlaps between the two systems have been seen in left inferior frontal gyrus (IFG), left medial temporal gyrus (MTG), left superior temporal sulcus (STG) and left Sylvian fissure at the parietal-temporal boundary (SPt). This map of the overlap between production and comprehension is incomplete, however, because it is unknown whether such overlap changes in the context of real-life communication.

Having mapped the production network in Study 1, here I measure the extent of the overlap between the production and comprehension systems by mapping the network active during the comprehension of the same unconstrained real-life narrative. Such a comprehensive map of the spatial overlap between production and comprehension processes will shed light on the degree to which the two processes share common ground.

Materials and Methods:

Subject Population: Three speakers and eleven listeners, ages 21-40, participated in the experiment. All participants were right-handed native English speakers. Procedures were in compliance with the safety guidelines for MRI research and approved by the University Committee on Activities Involving Human Subjects at Princeton University. All participants provided written informed consent.

Production procedure: We used the same production procedure and data from Chapter 1 (see above).

Comprehension procedure: Next, we measured the listeners' brain responses during audio playback of the original recorded story (see Chapter 1). We synchronized the functional MRI signals to the speaker's vocalization using the scanner's TTL pulse, which precedes each volume acquisition. Eleven listeners listened to the recording of the story. Our experimental design thus captured both the production and comprehension sides of the simulated communication. Following the fMRI scan, each listener was asked to freely recall the content of the story. Six independent raters scored each of these

listener records according to a 115–point true/false questionnaire, and the resulting score provided a quantitative measure of each listener's understanding.

MRI Acquisition: Subjects were scanned in the same 3T head-only MRI scanner (Allegra; Siemens, Erlangen, Germany) with the same custom radio frequency coil for the structural scans (NM-011 transmit head coil; Nova Medical, Wakefield, MA) and parameters as in Chapter 1. High fidelity MRI-compatible headphones (MR Confon; Magdeburg, Germany) were fitted to present the audio stimuli to the subjects while attenuating scanner noise.

Data Preprocessing: fMRI data were preprocessed in the same manner as in Chapter 1.

Inter-Subject Correlation Analysis: To measure the reliability of the response timecourses between the listeners' brains, we used the same inter-subject correlation (inter-SC) analysis as in Chapter 1 (Figure 4B, and see for eg. Figure 5B-C), but compared the BOLD response timecourses across different listeners listening to the same story as opposed to different speakers producing the same story.

Direct statistical testing between production and comprehension maps: To identify areas that show increase in response reliability for speech production (using the map generated in Study 1) over speech comprehension, a t-test ($\alpha = 0.05$) was performed within each voxel that exceeded the threshold in both of the conditions (i.e. voxels that seems to participate in both the comprehension and production of the speech signal). Thus, the t-test was performed by comparing the correlation values of the speaker's retellings during production $\{r_j, r_{j+1} \dots r_n\}$ to the correlation values of the listeners during comprehension

III) Coupling between speech production and speech comprehension

Introduction:

Verbal communication is a joint activity by which interlocutors share information (Pickering and Garrod, 2004). As discussed previously, however, little is known about the neural mechanisms underlying the transfer of linguistic information across brains, and thus the ongoing interaction between the two systems during everyday communication remains largely unknown. Communication between brains may be facilitated by a shared neural system dedicated to both the production and the perception–comprehension of speech (Liberman and Mattingly, 1985; Levelt, 1989; Branigan et al., 2000; Pickering and Garrod, 2004; Wilson and Knoblich, 2005; Chang et al., 2006; Hari and Kujala, 2009). Mapping the extent of spatial overlap between comprehension and production of real-life speech (Chapter 2) is necessary, but does not address the nature of the relationship between the two processes. Just because an area is involved in both speech production and speech comprehension does not necessarily mean it is performing the same function across the two tasks. If similar functions are being recruited during the production and comprehension of speech, then the brain responses of a speaker during production will be coupled over time to the brain responses of a listener during comprehension. A better characterization of the temporal dependencies between production-based and comprehension-based neural processes requires a paradigm that allows for such a direct comparison of the neural response timecourses across both functions. In this study I directly address the neural basis of the interaction between speech production and comprehension by analyzing the spatial and temporal coupling

between production and comprehension across brains during natural verbal communication.

The premise underlying such brain-to-brain coupling is that the perceptual system of one brain can be coupled to the motor system of another (Hasson et al., 2012). The mechanism works by directly inducing similar brain patterns in a listener by a speaker in the absence of any stimulation other than speech. Under the assumption that a given thought is represented by a given neuronal state in the speaker's brain, a successful communication of the thought will result in an inducement of a similar neuronal state in the listener's brain. Recent evidence suggests that people become implicitly coupled at motor, perceptual, and cognitive levels (Sebanz et al., 2003; Keller, 2008; Lindenberger et al., 2009; Böckler et al., 2011b, a; van der Wel et al., 2011). For example, two people on rocking chairs synchronize their rocking as if they were mechanically coupled despite differences in inherent rocking chair frequencies (Richardson et al., 2007). During dialogue, interlocutors will align on spatial reference frames (if one speaker refers to objects egocentrically, the other tends to use an egocentric perspective (Watson et al., 2004)), on grammar (speakers repeat the syntactic structure used by their interlocutors for cards describing events (Branigan et al., 2000)), and even on accent and speech rate (Giles et al., 1991). Even two people standing back-to-back will assume similar postures simply by talking (Shockley et al., 2003; Shockley et al., 2007).

Coupled systems can generate complex behaviors that cannot be performed in isolation (Hasson et al., 2012). Examples of this range from herding behaviors to coupled oscillatory networks in the brain (Buzsáki and Draguhn, 2004; Couzin, 2009). The emergence of language is a particularly strong example. In human infant communication,

- van Schie HT, van Waterschoot BM, Bekkering H (2008) Understanding action beyond imitation: reversed compatibility effects of action observation in imitation and joint action. *J Exp Psychol Hum Percept Perform* 34:1493-1500.
- Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, Mazoyer B, Tzourio-Mazoyer N (2006) Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *Neuroimage* 30:1414-1432.
- Vigneau M, Beaucousin V, Hervé PY, Jobard G, Petit L, Crivello F, Mellet E, Zago L, Mazoyer B, 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.
- Vissers CT, Chwilla DJ, Kolk HH (2006) Monitoring in language perception: The effect of misspellings of words in highly constrained sentences. *Brain Res* 1106:150-163.
- Vygotski LS, Cole M (1978) *Mind in Society: The Development of Higher Psychological Processes*. In: Harvard University Press.
- Warren RM (1970) Perceptual restoration of missing speech sounds. *Science* 167:392-393.
- Watkins KE, Strafella AP, Paus T (2003) Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989-994.
- Watson ME, Pickering MJ, Branigan HP (2004) Alignment of reference frames in dialogue. In: *Proceedings of the 26th Annual Conference of the Cognitive Science Society*.
- Weber A, Cutler A (2006) First-language phonotactics in second-language listening. *J Acoust Soc Am* 119:597-607.
- Weber A, Grice M, Crocker MW (2006) The role of prosody in the interpretation of structural ambiguities: a study of anticipatory eye movements. *Cognition* 99:B63-72.
- Wernicke C (1969) The symptom complex of aphasia. In, pp 34-97. *Proceedings of the Boston Colloquium for the Philosophy of Science 1966/1968*.
- Whitney C, Grossman M, Kircher TT (2009) The influence of multiple primes on bottom-up and top-down regulation during meaning retrieval: evidence for 2 distinct neural networks. *Cereb Cortex* 19:2548-2560.
- Wicha NY, Moreno EM, Kutas M (2003) Expecting gender: an event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex* 39:483-508.
- Wilson M, Knoblich G (2005) The case for motor involvement in perceiving conspecifics. *Psychological Bulletin* 131:460-473.
- Wilson SM, Molnar-Szakacs I, Iacoboni M (2008) Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. *Cereb Cortex* 18:230-242.
- Wilson SM, Saygin AP, Sereno MI, Iacoboni M (2004) Listening to speech activates motor areas involved in speech production. *Nat Neurosci* 7:701-702.
- Wise SP, Fried I, Olivier A, Paus T, Rizzolatti G, Zilles KJ (1996) Workshop on the anatomic definition and boundaries of the supplementary sensorimotor area. *Supplementary Sensorimotor Area* 70:489-495.