

Published in final edited form as:

J Cogn Neurosci. 2011 October ; 23(10): 2892–2904. doi:10.1162/jocn.2010.21589.

The corollary discharge in humans is related to synchronous neural oscillations

Chi-Ming A. Chen^{1,2}, Daniel H. Mathalon^{2,3}, Brian J. Roach^{2,3}, Idil Cavus², Dennis D. Spencer⁴, and Judith M. Ford^{2,3}

¹Department of Psychiatry, Columbia University, 1051 Riverside Drive, Unit 21, New York, NY 10032

²Department of Psychiatry, Yale University, 300 George St., New Haven, CT 06511

³Department of Psychiatry, University of California, San Francisco, SFVA, 4150 Clement Street, Building 8, Room 9B-30 116D, San Francisco, CA 94121

⁴Department of Neurosurgery, Yale University, 789 Howard Ave., New Haven, CT 06519

Abstract

How do animals distinguish between sensations coming from external sources and those resulting from their own actions? A corollary discharge system has evolved that involves the transmission of a copy of motor commands to sensory cortex, where the expected sensation is generated. Through this mechanism, sensations are tagged as coming from self, and responsiveness to them is minimized. The present study investigated whether neural phase synchrony between motor command and auditory cortical areas is related to the suppression of the auditory cortical response. We recorded electrocorticograms from the human brain during a vocalizing/listening task. Neural phase synchrony between Broca's area and auditory cortex in the gamma band (35 Hz to ~50 Hz) in the 50 ms time window preceding speech onset was greater during vocalizing than listening to a playback of the same spoken sounds. Because pre-speech neural synchrony was correlated ($r = -0.83$, $p = 0.006$) with the subsequent suppression of the auditory cortical response to the spoken sound, we hypothesize that phase synchrony in the gamma band between Broca's area and auditory cortex is the neural instantiation of the transmission of a copy of motor commands. We suggest that neural phase synchrony of gamma frequencies may contribute to transmission of corollary discharges in humans.

INTRODUCTION

There is rapidly growing evidence that synchrony among large populations of neurons is essential for cognitive functions (Engel, Fries, & Singer, 2001; Von Stein, Chiang, & Konig, 2000), and that gamma band activity correlates with various cognitive processes (Pesaran, Pezaris, Sahani, Mitra, & Andersen, 2002; Fries, Reynolds, Rorie, & Desimone, 2001; Varela, Lachaux, Rodriguez, & Martinerie, 2001; Singer, 1999a). It has been hypothesized that cognitive functions require flexibility in the routing of neuronal signals through

anatomically connected structures; specifically, communication between two neuronal groups (either within a region or between regions) mechanistically depends on coherence (or phase-locking or synchrony) of neural activity between them (Fries, 2005). Synchronous oscillations in gamma-frequency band (e.g., 30 Hz to 80 Hz) have been implicated in the mechanisms that integrate spatially distributed processing and signal the relatedness of neurons coding for different features of the same object (Varela et al., 2001; Singer, 1999a). Long-range neural integration may be mediated by neuronal groups oscillating together, precisely phase-locking neural activity in the gamma-frequency bands over short periods of time (Womelsdorf et al., 2007; Fell et al., 2001; Lachaux, Rodriguez, Martinerie, & Varela, 1999; Miltner, Braun, Arnold, Witte, & Taub, 1999; Rodriguez et al., 1999). This mechanism provides an explanation for integration of cortical information processing between primary and association cortical areas (e.g., primary visual and middle temporal cortical areas (Chen et al., 2007)) and different sensory cortical areas (e.g., somatosensory and auditory cortical areas (Lakatos, Chen, O'Connell, Mills, & Schroeder, 2007)). Thus, neuronal communication between cortical areas via gamma-band coherence may subserve the integration of cognitive processes (Fries, 2005; Varela et al., 2001; Singer, 1999b).

The experimental evidence mentioned above supports the proposal that synchronous neuronal oscillation in gamma-frequency band is the “functional building blocks” of information processing (Basar-Eroglu, Struber, Schurmann, Stadler, & Basar, 1996), supporting “binding” of brain functions across cortical networks (Phillips & Singer, 1997). Recent studies show that gamma-band synchrony is used for long-range communication across distant brain areas in monkeys, e.g., between frontal eye field and area V4 (Gregoriou, Gotts, Zhou, & Desmone, 2009); between lateral intraparietal area and medial temporal area (Saalmann, Pigarev, & Vidyasagar, 2007). One neural mechanism that may rely on cortical communication via gamma-band phase synchrony is the “forward model” (Webb, 2004), involving the transmission of a copy of motor commands from cortical areas involved in planning and executing a motor act, to sensory areas involved in processing the resulting sensations. Because the forward model mechanism involves long-range coordinated motor-sensory communication, enhancement of neural synchrony might be evident prior to the execution of motor acts. In fact, local field potential recordings from somatosensory barrel cortex of rats show neural synchrony preceding exploratory whisking in the 25 to 45 Hz band, perhaps triggered by the transfer of a copy of motor commands to somatosensory cortex (Hamada, Miyashita, & Tanaka, 1999). Additional support for a forward model utilizing gamma-band phase synchrony comes from data showing coherent 25 to 35 Hz oscillations in the motor and somatosensory cortices of awake behaving rhesus monkeys (Murthy & Fetz, 1992) and gamma-band coherence between frontal and temporal electrodes in humans (Ford, Gray, Faustman, Heinks, & Mathalon, 2005).

An increasingly recognized forward model is the “corollary discharge.” It has been suggested that the corollary discharge mechanism suppresses sensations that match anticipated sensory consequences of self-generated motor acts (Crapse & Sommer, 2008; Poulet & Hedwig, 2006; Sperry, 1950; Von Holst & Mittelstaedt, 1950; Helmholtz, 1925; Helmholtz, 1867). The theoretical concept of a forward model has been shared and studied by multiple disciplines in a broad variety of scientific fields such as physiology (Poulet & Hedwig, 2002; Sommer & Wurtz, 2002; Bell, 1981; Davis, Siegler, & Mpitsos, 1973),

psychology (Grusser, 1986), psychiatry (Shergill, Samson, Bays, Frith, & Wolpert, 2005; Ford, Mathalon, Heinks, Kalba, & Roth, 2001a; Feinberg, 1978), and cybernetics (Guenther, Hampson, & Johnson, 1998; Wolpert, Ghahramani, & Jordan, 1995; Jordan & Rumelhart, 1992). Taking the singing male cricket (Poulet & Hedwig, 2006) as an example, crickets hear through their ears located on their forelegs very close to where 100 dB sound-pressure level chirps are generated. However, singing crickets respond to other crickets' songs; therefore, the cricket's central nervous system must have a way to deal with self-generated intensive sounds during chirping. To avoid being overwhelmed by its own chirps, there is tight coordination between wing motor and sensory systems (Poulet & Hedwig, 2006) utilizing mechanisms of corollary discharge (Crapse & Sommer, 2008). Additionally, these concepts apply to cybernetics, as autonomous robots need to distinguish between the consequences of their own actions and the actions of others moving toward them (McKinstry, Seth, Edelman, & Krichmar, 2008).

A forward model of motor-sensory modulation has been proposed to explain how people tell the difference between their own inner verbal experiences, or thoughts, and "voices", or auditory verbal hallucinations, that sound as if they are coming from external sources (Ford, Roach, Faustman, & Mathalon, 2007). This model proposes that a copy of motor commands, transferred from speech production regions in the frontal lobes to the auditory cortex, produces an expectancy of consequences in the auditory cortex to prepare it for an imminent self-generated speech sound. This minimizes the auditory cortical response to these self-generated sounds and provides a mechanism for recognizing these sounds as their own. Support for this mechanism in the human auditory system comes from both intra-cortical and non-invasive studies. Intra-cortically, single unit recordings were made during a pre-surgical planning procedure from the exposed surface of the right and left temporal cortex; while patients talked, auditory cortical responses were reduced compared to when the patients listened to others talking (Creutzfeldt, Ojemann, & Lettich, 1989). Findings from non-invasive, human electrophysiological studies, employing electroencephalography (EEG) and magnetoencephalography (MEG) synchronized to the onset of vocalization, show similar auditory cortical suppression (i.e., N100 or M100 suppression) during vocalizing (Ford et al., 2007; Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Houde, Nagarajan, Sekihara, & Merzenich, 2002; Ford, Mathalon, Heinks, Kalba, & Roth, 2001b; Curio, Neuloh, Numminen, Jousmaki, & Hari, 2000). In these studies, the N100 of the EEG-based event-related potential (ERP), or the M100 of the MEG-based response, reveals that activity in auditory cortex, 100 ms after speech onset, is reduced during vocalizing compared to when the same speech sound is recorded and played back. The corollary discharge also works during self-stimulation paradigms; when a tone burst (Martikainen, Kaneko, & Hari, 2005; Schafer & Marcus, 1973) or speech sound (Ford, Gray, Faustman, Roach, & Mathalon, 2007) results from indirect motor acts of pushing a button, it elicits a smaller N100 than when that exact sound is passively heard. Data from humans are consistent with data from non-human primates (Eliades & Wang, 2003; Muller-Preuss & Ploog, 1981) and crickets (Poulet & Hedwig, 2006; Poulet & Hedwig, 2002). Eliades and Wang (Eliades & Wang, 2003) recorded from single units in primary auditory cortex of marmoset monkeys during vocalizing. Like Creutzfeldt and colleagues, they reported vocalization-induced suppression beginning before vocalization, with both inhibition and

excitation of units beginning after vocal onset (Eliades & Wang, 2008; Eliades & Wang, 2005; Eliades & Wang, 2003). They suggested that the origin of the inhibition was from the speech production areas and excitation from the re-afference of the self-produced vocalization.

A recent diffusion tensor imaging study (Frey, Campbell, Pike, & Petrides, 2008) suggested that Brodmann's Area 45 (Broca's area) connects with the superior temporal gyrus via the extreme capsule fiber system, providing a plausible route by which a copy of motor commands from Broca's area travels to auditory cortex. In the present study, we studied forward model mechanisms linking Broca's area and auditory cortex in patients implanted with cortical electrodes for neurosurgical evaluation while they spoke (Vocalizing condition) and then listened to a playback of these recorded sounds (Listening condition). We used the amplitude of the N100 component of the auditory ERP to index the cortical response to the self-produced speech sounds during both Vocalizing and Listening. The N100 has a complex set of subcomponents with both tonotopic and non-tonotopic generators (Woods, 1995). While one of its subcomponents may emanate from primary auditory cortex (Naatanen & Picton, 1987), it peaks too late to reflect the initial cortical volley which is seen on the medial aspect of Heschl's gyrus (Liegeois-Chauvel, Musolino, & Chauvel, 1991; Celesia, Broughton, Rasmussen, & Branch, 1968) and occurs within the first 15 ms after stimulus onset. Determined by clinical necessity, our intracranial electrode grid coverage of the temporal lobe did not allow us to record these early components, but only the activity in the more lateral aspects of the supratemporal plane. The high signal-to-noise ratio of these intracranial recordings allowed us to investigate the mechanisms of long-range neural oscillations in a wide frequency range as well as to localize where signals reflecting the copy of motor commands start and where the electrophysiological effects of its expected sensory consequence are seen. While miniature saccades have been shown to affect the gamma-band response recorded with scalp electrodes during the time period from ~200 ms to 300 ms following stimulus onset (Yuval-Greenberg, Tomer, Keren, Nelken, & Deouell, 2008), intracranial recordings provide a sensitive measure of cortically-generated gamma activity. Based on our previous findings (Ford et al., 2007; Ford et al., 2007; Ford et al., 2005), we hypothesized that there would be greater neuronal phase synchrony in the gamma frequency range (Ford et al., 2005) between Broca's area and auditory cortex during vocalizing than during listening. We further hypothesized that this neural phase synchrony during vocalizing would reflect the copy of motor commands, as evidenced by a correlation between this synchrony and auditory cortical suppression of responsiveness during vocalizing (Ford et al., 2007; Ford et al., 2007).

METHODS

Participants

Participants were three male patients with treatment resistant epilepsy, being evaluated for surgical resection of seizure foci. The experiments were undertaken with the understanding and written consent of each participant. The use of human participants is in conformity with the "Guiding Principles for Research Involving Animals and Human Beings" and data are HIPAA compliant. Protocols were approved in advance by the Yale Human Investigation

Committee. These patients underwent a phased evaluation that included magnetic resonance imaging, interictal single photon emission computed tomography, interictal positron emission tomography, neuropsychological assessment, video/EEG with ictal and interictal scalp recording in phase I, and intracerebral sodium amobarbital testing in phase II. Only patients in whom the epileptic area(s) were not determined and/or in whom there was significant discordant data were offered phase III evaluation (i.e., intracranial EEG monitoring). This monitoring was done in order to localize 1) the seizure focus for later resection, and 2) critical language and motor areas to be avoided during surgery.

Patient A was a 45-year-old right handed Caucasian man with medication-resistant complex partial seizures. Magnetic resonance imaging (MRI) showed bilateral hippocampal atrophy and increased signal change on T2-weighted images. He had resection of right anterior temporal pole and anteromedial hippocampectomy. Pathology showed diffuse glial proliferation of white matter in the right temporal lobe and neuronal loss and gliosis of the hippocampus consistent with hippocampal sclerosis. Patient B was a 37-year-old right handed Caucasian man with medication-resistant complex partial seizures. MRI showed small T2 change in subcortical/cortical lesion in the left frontal lobe, suggesting cortical dysplasia, and some T2 signal changes in the right parietal lobe, which were nonspecific. He had resections of posterior superior temporal gyrus and middle temporal gyrus with multiple subpial transections of inferior frontal and parietal regions. Pathology showed fragments of superficial neocortex with artifactually altered neurons. Patient C was a 52-year-old right handed Caucasian man with medication-resistant complex partial seizures. MRI showed bilateral hippocampal atrophy and the amygdala was larger on the right than left. He had trans-sylvian amygdalohippocampectomy with preservation of lateral temporal cortex. Pathology showed left anterior hippocampus had partial loss of neurons and gliosis consistent with hippocampal sclerosis.

Grid placement was determined entirely by clinical necessity. None of the reported electrode sites in this study was identified as epileptogenic or within the resected area at surgery. We were limited to the patient population presenting for evaluation of their seizure focus for treatment of medically intractable seizures and did not have any influence over the selection and demographic makeup of the patient population. There was no bias in the surgical program regarding patients' gender, socioeconomic status, or ability to pay for clinical care.

Electrode Localization

For each patient, preoperative MRI scans were acquired, and following surgical installation of the electrodes, both postoperative anatomical MRI scans (with a slice thickness of 1.6 mm) and computed tomography (CT) images (with a slice thickness of 1.1 mm) were also acquired. The electrodes were interactively identified in the CT images and then were localized by using the bioimage suite software package developed at Yale University (www.bioimagesuite.org). To determine the location of electrodes included in the analyses, individual electrode locations were marked on postoperative CT images. Briefly, the metal contacts in each electrode are extremely bright in the postoperative CT images. They can be visualized using simple thresholding which yields clusters of bright voxels. Then interactively these clusters are labeled by a trained technologist and the centroid of these

clusters is used to define the electrode position. Marked CT images were registered to postoperative MRI scans using a 6-parameter rigid transformation, and postoperative MRI scans were then registered with preoperative MRI scans using a nonlinear grid-based transformation (Papademetris, Jackowski, Schultz, Staib, & Duncan, 2004) to account for the distortion of the brain that occurs as a result of the craniotomy. Prior to the later registration the skulls are stripped from both MRI scans to enhance convergence.

Experimental Paradigm and Instrumentation

Participants were tested in a vocalizing/listening experiment using Presentation software (Neurobehavioral Systems Inc., Albany, CA). The experimental paradigm consisted of blocks where the word “Ready” was first presented on the screen for one second, followed by an exclamation point for one second, and then a plus sign for seven seconds. Participants were instructed to prepare to say “ah” when the “Ready” message was presented, and then, when the plus sign appeared (and remained on the screen for 7 seconds), they said “ah” roughly every one to two seconds until the plus sign disappeared. Finally, the word “Rest” would appear on the screen for five seconds, indicating that participants could stop vocalizing and prepare for the next trial. This vocalizing condition block was repeated 15 times, with an extended break in the middle of the whole recording. During listening condition blocks, recorded sounds from the vocalizing condition were played back and participants were instructed simply to listen. The participants were trained to produce uniform, brisk utterances with minimal tongue, jaw, and throat movements. Before data acquisition, the participants uttered “ah” several times to facilitate sound system calibration and acclimation to the environment.

The Presentation software was used to run both conditions from a laptop running Windows XP and a Sound Blaster Audigy 2 ZS PCMCIA sound card (Creative Technology Ltd., Jurong East, Singapore). The sound card was used to monitor and record the participant’s vocalization and to send trigger codes mark the beginnings of the recordings to the EEG data collection system. Voice was recorded at 44.1 KHz and no explicit filters were applied to the recorded data. The same recorded sounds were played back using Presentation software at an intensity level set based on offline tests to calibrate the vocalizing and listening recording sessions. This was achieved by introducing a single output of 114 dB at 1000 Hz through the microphone using a Quest QC Sound Calibrator (Quest Technologies, Inc., Oconomowoc, WI) during the vocalizing condition, measuring its output on a Quest Sound Level Meter (C-scale weighting) through the insert earphones, and adjusting playback volume to match the sound level through the same earphones. We did not use an online algorithm to insert “ah” onset trigger markers during recording. The onset of each “ah” was determined offline using a detection algorithm (Ford, Roach, & Mathalon, 2010) on the recorded data of the vocalizing condition to avoid any jitter in timing between voice onset and the trigger pulses. Another issue related to the reliability and timing of trigger makers is the difference between the onset of a trigger marker to an EEG recording system and the onset of sound recording or playback. This difference using Windows XP and the Audigy 2 ZS PCMCIA Notebook sound card in studies have been rigorously tested by our lab and others (please see detailed test results of the Audigy 2 ZS Notebook sound card latency in the webpage of Presentation’s Hardware Corner about Sound Cards by Neurobehavioral Systems here:

http://www.neurobs.com/presentation/hardware_corner/sound_cards/test_results#nbs36).

The difference between these two onset signals is less than 2 ms, and the variability in this difference is under half of a millisecond. It is worth noting that while the same sequence of sounds is heard during vocalizing and listening, each individual sound is unique. All patients heard speech sounds binaurally through Etymotic ear insert headphones (Etymotic Research, Inc., Elk Grove Village, IL).

Data Collection and Event-related Potential Processing

Electrocorticography (ECoG) data were recorded from a selection of 128 subdural strip and grid electrodes implanted in the neurosurgery patients. The impedance of each electrode was between 5 KOhm to 10 KOhm. Signals were amplified using alternating current amplifiers (SA Instrumentation Co., San Diego, CA) prior to analog band-pass filtering from 0.1 Hz to 100 Hz. The software for collecting ECoG data is homemade (Yale New Haven Hospital, New Haven, CT). Signals were digitized at 14-bit resolution (DAP 4200a, Microstar Laboratories, Bellevue, WA) and sampled at either 1000 Hz (patients A and B) or 2000 Hz (patient C). ECoG signals were referenced to a peg electrode. The peg electrode was designed so that the contact lay within the diploic space. The location of the peg electrode was individualized in that the surgeon tried to place it as far as possible from the intracranial electrodes used for the patient. The ground electrode is a four-contact strip placed facing the dura. For offline processing, the ECoG data were epoched to onsets of participant-generated speech sounds and were baseline corrected. Then, a 60 Hz second-order infinite impulse response notch filter was used.

Statistical and Data Analysis

Our analysis methods for Figures 1 to 3 are described in detail here. Supplementary Material Figure S1a is a schematic illustration of sequences of permutation. The classical Broca's area (i.e., yellow-circled electrode in Fig. 1A; blue-circled electrodes in Figs. 3A, 3B, and 3C) was determined by intraoperative language mapping for each patient by interruption of speech production (complete speech arrest) with direct electrical stimulation at each electrode. Auditory cortex (i.e., red-circled electrode in Fig. 1A; yellow-circled electrodes in Figs. 3A, 3B, and 3C) was not determined by functional mapping, but instead by gyral anatomy and auditory N/M100 component (Reite et al., 1994). Since the inter-electrode spacing was ~1 cm and primary auditory areas extended for an average of 24 mm (Artacho-Perula et al., 2004), we decided to select three electrodes for representing auditory cortex in each patient. None of the recorded electrode sites was identified as epileptogenic or within the resected area at surgery. No data from tissue involved in seizure activity are reported here.

Measuring phase synchrony between electrodes: Phase locking value (PLV)—

Intracranial electrode placements of interest of patient A are shown in Figure 1A. The ECoG data were epoched to onsets of participant-generated speech sounds, and the averaged event-related potentials to speech onset of electrodes A12 and G8 are shown in Figures 1B and 1C, respectively. The red and blue traces are the ERPs for the vocalizing and listening conditions, respectively. The most N100 suppression of patient A during Vocalizing compared to Listening was observed and illustrated in Figure 1C at the electrode contact G8,

suggesting G8 is in or adjacent to primary auditory cortex (Eliades & Wang, 2008; Eliades & Wang, 2005). Eliades and Wang (Eliades & Wang, 2005) suggested that the origin of the suppression was from the speech production areas, so we selected the electrode contact A12 in Figure 1A located in Broca's area.

The neuroelectric signals at the two circled electrode contacts (i.e., A12 and G8) were used to measure phase synchrony between those two contacts by phase locking values (PLVs) for Vocalizing and Listening, Figures 1D and 1E, respectively. These PLVs represent the intertrial variability of the phase differences between two electrode contacts (Fell et al., 2001; Lachaux et al., 1999; Rodriguez et al., 1999). If the phase difference between two electrode contacts varies little across trials, PLV is close to 1 (perfect phase synchrony). Phase synchrony of neuronal oscillation in the gamma-band frequencies has been proposed and shown as an integrative mechanism that brings long-range sets of neurons together into coherently oscillating neuronal groups that can interact effectively.

To generate Figures 1D and 1E, first, continuous recordings were epoched offline from -1000 ms to 1000 ms to avoid edge effects of the wavelet transformation for time window -300 ms to 300 ms. 76 epoched trials were then spectrally decomposed by a Morlet wavelet transform in the gamma frequency range from 25 Hz to 50 Hz in 1 Hz step using `freganalysis_wltconvol.m` from FieldTrip toolbox (Radboud University, Nijmegen, the Netherlands) in MATLAB (The MathWorks, Inc., Natick, MA). The wavelet family we used is defined by a constant ratio, $f_0/\sigma_f = 7$, where f_0 is the center frequency and $\sigma_f = 1 / (2\pi\sigma_t)$. At 25 Hz, this leads to a wavelet duration ($2\sigma_t$) of 89.1 msec and to a spectral bandwidth ($2\sigma_f$) of 7.1 Hz and, at 40 Hz, to a duration of 55.7 msec and a bandwidth of 11.4 Hz. Second, for each of the pair of electrodes, instantaneous phase was extracted by the wavelet decomposition for each time point and frequency on the data from each trial. The result of the wavelet decomposition is a matrix with complex elements. Using the function, `angle.m`, in MATLAB, we extracted the phase information from the complex matrix. Third, for each time point of each trial, a phase difference ($\Delta\theta_{(\text{time}, \text{trial})}$) between two selected electrode contacts (e.g., A12-and-G8-paired electrodes in Fig. 1A) was determined. Using phase differences of N number of trials, PLV_t (see below formula), ranging from 0 (no synchronization) to 1 (perfect synchronization), was calculated at each time point (Lachaux et al., 1999). Lastly, the third step was repeated for all selected frequencies.

$$PLV_t = \frac{1}{N} \left| \sum_{n=1}^N e^{i\Delta\theta_{(t,n)}} \right|$$

Permutation procedure for testing differences between two conditions—To demonstrate how to differentiate significant PLVs from background fluctuations in each of different experimental conditions, one could implement certain techniques of randomization as described by Lachaux et al (1999). However, because the results of those randomization techniques would only reveal the significant PLVs within a condition, they would not be able to determine whether the differences of PLVs between two conditions (e.g., Fig. 1f) are significant. Based on the principles of permutation tests (Maris, Schoffelen, & Fries, 2007;

Wilcox, 2004), we developed a new permutation procedure specifically designed for testing the significances of the differences of PLVs between two conditions (e.g., the differences of PLVs between vocalizing and listening conditions in Fig. 1F). The details of applied permutation procedure were described and illustrated in Supplementary Material Figure S1a. The significant differences of PLVs between two conditions in Figure 1F are shown in Supplementary Material Figure S1b.

Percentage changes of significant differences of PLVs between conditions—

Percentage changes of significant differences of PLVs between conditions are shown in Figure 2 and the formula is (t is the time point in ms and f is the frequency in Hz):

$$\text{Percentage Change}_{(t, f) \text{ Significant } t} = \frac{(\text{PLV}_{(t, f) \text{ Vocalizing}} - \text{PLV}_{(t, f) \text{ Listening}})_{\text{Significant } t}}{\text{PLV}_{(t, f) \text{ Listening}}} \times 100\%$$

Since it was based only on the “significant” differences of PLVs between vocalizing and listening conditions (i.e., $[\text{PLV}_{\text{Vocalizing}} - \text{PLV}_{\text{Listening}}]_{\text{significant}}$, Suppl. Material Fig. S1b), red regions would represent significant enhanced changes of PLVs during vocalizing compared to the listening condition, blue regions would represent significant suppressed changes of PLVs during vocalizing compared to the listening condition, and green regions would represent non-significant changes of PLVs between two conditions. One of the advantages of percentage changes (or modulation) is that enhanced/suppressed phase synchronies across different patients can be compared.

Relationship between N100 suppression and PLVs of gamma frequencies

PLV between Broca’s area and auditory cortex: One Broca’s area electrode and three auditory cortex electrodes were selected for each patient, making three pairs of electrodes for each patient (i.e., A12-G8, A12-G9, A12-G10 for patient A in Fig. 3A; G3-G50, G3-G42, G3-G34 for patient B in Fig. 3B; G6-G45, G6-G37, G6-G29 for patient C in Fig. 3C).

For each pair of the electrodes, we first calculated the percent change between conditions for all PLVs that were significantly different during vocalizing and listening (i.e., the method for generating Fig. 2). Next, we averaged these values between -50 ms to 0 ms time window, avoiding possible event-related changes in coherence that might cause significant differences of PLVs. These mean values represented the degree of synchrony for each electrode pair for each participant and higher gamma phase synchrony is reflected in larger mean values in the x-axis of Figure 3D. These mean values were used to assess the degree of association between inter-region synchrony during vocalizing and subsequent suppression of auditory cortical activity to the spoken sound.

Relating N100 suppression to gamma phase synchrony during vocalizing

(i) **N100 suppression:** The averaged ERPs to speech onset of auditory cortex electrodes (three for each patient) are yellow-circled and shown in Supplementary Material Figures S2a, S2b, and S2c. The averaged ERPs to speech onset of green-rectangle-selected electrodes are shown in bottom-left corner of Figures 3A, 3B, and 3C (magenta trace is

Vocalizing and blue trace is Listening; also shown in the Supplementary Material Figs). From the averaged ERP, the N100 was measured as the most negative peak between 50 ms and 150 ms, area ± 25 ms around the peak were calculated. N100 suppression was calculated by taking the area difference between N100s of listening and vocalizing conditions (i.e., $N100Area_{Listening} - N100Area_{Vocalizing}$). Greater N100 suppression is seen in more negative numbers in the y-axis of Figure 3D.

N100 was seen best at electrodes over primary auditory cortex (i.e., the green rectangle within the yellow circle shown in Figs. 3A and 3C) in patients A and C. However, in patient B, N100 was best seen at an electrode well above the electrodes placed over primary auditory cortex (i.e., the green rectangle shown in Fig. 3B and Supplementary Material Fig. S2b). Because each participant's anatomy is different, and because electrodes are placed for clinical necessity not for scientific convenience, it is likely that the electrode placement over the temporal lobe in patient B was not ideal for recording N100. We suggest that in patients A and C, N100 was generated in superior temporal gyrus, directly under the electrodes shown in the green rectangle, with the dipole generating N100 being perpendicular to the electrode. In patient B, we suggest that N100 was generated deep in the superior temporal sulcus, at an orientation tangential to the electrodes on the gyrus, making it electrically invisible at those sites. Instead, the dipole generating N100 was likely perpendicular to electrodes well above the generator, but visible to them. Depending on the orientation of the dipole plane generating the auditory evoked potentials, one might expect to observe evoked potentials generated in auditory cortex volume conducted to a distant site.

(ii) Predicting N100 Suppression from phase synchrony between Broca's area and auditory cortex: N100s from the yellow-circled electrodes were used for the first correlation analysis and were from the same three auditory cortex electrodes selected for calculating PLV (i.e., G8, G9, G10 for patient A in Fig. 3A; G50, G42, G34 for patient B in Fig. 3B; G45, G37, G29 for patient C in Fig. 3C). N100s from the green-circled electrode were used for the second correlation analysis and were based on where N100s were the largest during Listening. The selection for the best N100 electrode for each patient was shown by plotting vocalizing/listening ERPs for the G electrode grid in the Supplementary Material Figures S2a, S2b, and S2c for patients A, B, and C, respectively.

For each participant, there were three mean values for depicting the averaged percentage change of significant differences of PLVs between conditions. Since we had three patients, there were nine mean values used for our first correlation analysis. We calculated a two-tailed Pearson's correlation coefficient (r) between the yellow-circled N100 suppression and its corresponding averaged percentage change of significant differences of PLVs (i.e., the hollow circles in Fig. 3D; both were normally distributed; Shapiro-Wilk normality tests, $p > 0.05$). Additionally, a mixed model analysis of covariance (ANCOVA) was applied to variables in Figure 3D. Since the patients are assumed to be randomly selected from an infinite population of possible levels, the "patient" variable was classified as a random factor in the general linear model. The mixed model ANCOVA design for categorical (i.e., "patient" variable) and continuous predictor variables (i.e., "averaged percentage change of significant differences of PLVs" variable) is only appropriate when the categorical and continuous predictors do not interact in influencing responses on the outcome, the dependent

variable (i.e., “N100 suppression” variable). The homogeneity of slopes test was applied to determine whether the continuous and categorical predictors interact in influencing responses. As none of these interactions was significant ($p > .05$), all interaction terms were dropped from the ANCOVA model, thereby imposing the assumption that slopes of the regression lines were homogeneous across the three patients.

In a second correlation analysis, we used the best N100 suppression (green-circled electrode) and one grand mean of the averaged percentage changes of significant differences of PLVs for representing an overall gamma phase synchrony for each patient. We calculated a two-tailed Pearson's r between the best N100 suppression and the grand mean (i.e., the solid circles in Fig. 3D; both were also normally distributed; Shapiro-Wilk normality tests, $p > 0.05$).

RESULTS

Phase Locking Values (PLVs) between Broca's and Auditory Cortical Areas

The PLVs between electrodes over Broca's area and auditory cortex are greater during Vocalizing (Fig. 1D) than Listening (Fig. 1E), which suggests there is more phase synchrony between those two electrode sites during Vocalizing than Listening. Figure 1F shows the difference between PLVs during Vocalizing and Listening. We found a similar vocalizing-related correlation in the scalp recorded EEG over Broca's and Wernicke's areas in the gamma band (Ford et al., 2005), without the spatial precision afforded by this study.

Based on the results of our permutation test (see details in METHODS and Suppl. Material Fig. S1a and b), Figure 2 shows the percent changes in the significant differences of PLVs between vocalizing and listening conditions with listening condition as the baseline. Between -50 ms and 100 ms, phase synchronies between Broca's and auditory cortical areas during Vocalizing were significantly greater (warm colors in Fig. 2) than those during Listening for two frequency ranges within the 25 Hz to 50 Hz frequency window (i.e., 25 Hz to ~ 33 Hz and 35 Hz to ~ 50 Hz). In the time period of -50 ms to 0 ms, prior to the onset of speech, there are more significant gamma band phase synchronies between Broca's and auditory cortical areas during Vocalizing in the frequency range from 35 Hz to ~ 50 Hz. In the time period of 40 ms to 120 ms, after the onset speech, there are more significant phase synchronies between those two areas during Vocalizing in frequency ranges from 25 Hz to ~ 33 Hz, and also, from ~ 38 Hz to 50 Hz.

Cortical Interaction and N100 Suppression

As can be seen in Figure 3D, higher gamma phase synchrony is reflected in larger values on the x-axis, and more N100 suppression is seen in smaller numbers on the y-axis. A significant negative correlation (indicated by the dashed line in Fig. 3D; $r = -0.825$, $p = 0.006$) means that higher gamma phase synchronies between Broca's and auditory cortical areas correlate with greater N100 suppression in auditory cortices. To avoid problems associated with non-independent data points (three pairs of electrodes from each participant), we also calculated the relationship between N100 suppression and gamma phase synchrony using a mixed model ANCOVA with the “patient” variable included as a

random factor, gamma phase synchrony as a covariate, and N100 suppression as a dependent variable. Patient did not affect N100 suppression, but phase synchrony did ($F_{(1,5)} = 9.058$, $p = 0.03$), consistent with the Pearson's correlation analysis.

Although there were only three points (one for each patient) in our second correlational analysis, the correlation coefficient showed a significant trend that higher gamma phase synchronies between Broca's and auditory cortical areas correlated with greater N100 suppression (i.e., the solid line in Fig. 3D; $r = -0.998$, $p = 0.044$).

As described previously in our methods, all the ECoG data that were used for calculating Figure 3D preceded the onset of vocalization (i.e., between -50 ms to 0 ms time window; 0 ms is the onset of vocalization). Due to the fact that movement artifacts are much less prominent in ECoG than scalp recorded EEG, and that data during vocalization were excluded from the analyses for Figure 3D, it is unlikely that this finding could result from movement in the vocalizing condition. Lastly, we illustrated that the percentage changes of significant differences of PLVs were not widespread and not due to artifacts recorded from the same electrode grid in the Supplementary Material Figures S3a and S3b.

DISCUSSION

Helmholtz (1867) described a forward model mechanism that allows us to discriminate between real moving objects and movements on the retina resulting from eye movements. This can be demonstrated by comparing the visual image during a saccade to the visual image experienced when the eye is externally moved by tapping the corner of the eye. In the former case, but not in the latter, visual perception is suppressed, a phenomenon known as saccadic suppression (Thiele, Henning, Kubischik, & Hoffmann, 2002). After Helmholtz (Helmholtz, 1925; Helmholtz, 1867), others suggested that this suppression was accomplished through the action of an efference copy (Von Holst & Mittelstaedt, 1950) or a corollary discharge (Sperry, 1950). This mechanism is responsible for our critical ability to unconsciously recognize and disregard sensations resulting from our own actions (Von Holst & Mittelstaedt, 1950).

While the action of this mechanism is undeniable and essential to negotiating and sensing our complex environment, until recently its neurophysiological basis was unknown. It is now being studied in vertebrates and invertebrates with electrophysiological methods similar to those used with humans, making findings cross-translational. The concept of a forward model is increasingly recognized as an important modulator of motor control, sensory processing, and cognition (Crapse & Sommer, 2008). It is argued that the forward model is an internal loop that uses the copy of motor commands to predict the expected sensation and modulate its sensory processing. To this end, there should be a specific output stream to sensory input areas from areas generating this output. This kind of stream or pathway has been identified in animals with "simple" nervous systems, such as crickets, by analyses of synaptic connectivity (Poulet & Hedwig, 2006).

In the present study, we investigated neuronal communication between Broca's area and auditory cortex. Our data suggest that neural phase synchrony of gamma frequencies (e.g.,

35 Hz to 50 Hz) between Broca's area and auditory cortex may at least partially explain the neural mechanism of corollary discharge in humans. Gamma-band phase synchrony in the 50 ms period immediately preceding speech onset may be the instantiation of a copy of motor commands being sent from Broca's area to auditory cortex. Although we cannot argue causality from correlation, that, it is correlated with the suppression of auditory cortical responses, 100 ms after speech onset, supports this interpretation. Furthermore, the gamma-band phase synchrony *precedes* speech onset; therefore, these synchronies were not due to the auditory responses of the uttered speech.

Although the functional meaning of specific frequencies is still unclear, synchronous oscillations at specific frequencies could identify even distant neuronal populations as belonging to the same functional network of spatially distributed neuronal assemblies (Bastiaansen & Hagoort, 2006). Possibly, only a portion of the gamma band is used for modulating sensory processing in the output stream from Broca's area to auditory cortex in our proposed forward model, and different gamma frequencies may participate in distinct neurophysiological mechanisms and reflect different aspects of cortical activation during information processing (Roopun et al., 2008). However, the range of frequencies above 80 Hz remains largely unexplored and the function of high gamma oscillation is unknown.

To rule out the simple effects of agency and predictability on our findings, we studied the effects of agency (subjects pressed a button to deliver an "ah") and predictability (a visual warning stimulus gave precise information about when an "ah" would be presented); in neither case was cortical suppression of the sensation as great as when subjects said "ah" (Ford et al., 2007). It is also worth mentioning that when we shifted the pitch of the "ah" during talking, there is still suppression, but less (Heinks-Maldonado et al., 2005).

We are becoming increasingly aware that locally specialized functions in the brain must be coordinated with each other. In our vocalizing/listening paradigm, the data show that participants with higher gamma phase synchrony between Broca's area and auditory cortex (prior to the onset of speech) have greater N100 suppression to the spoken sound. We speculate that (1) the transmission of a copy of motor commands to appropriate sensory cortex may be an emergent property of a self-organizing system, (2) communication among neuronal populations requires precise matching of the relative phase of distinct frequencies, and (3) integration of those mechanisms is responsible for our critical capability to unconsciously recognize and disregard sensations from our own actions. This type of ongoing motor-sensory interaction illustrates how important the neural system's constant-changing "context" is in processing new external sensory "content" (Arieli, Sterkin, Grinvald, & Aertsen, 1996).

The pioneering work by Creutzfeldt et al (1989), the only other such study using intra-cortical recordings from the human auditory cortex during speech, showed that while patients talked, auditory cortical responses were reduced compared to when the patients listened to others talking. However, this work is not directly comparable to ours, because they used tape-recorded words and sentences of different lengths, spoken by different speakers and not by the patients themselves. Also, they used single unit recordings and their recordings were restricted to regions to be resected which excluded recordings from

language areas. Nevertheless, their findings also suggested possible mechanisms for a forward model that could explain how people tell the difference between external and self-generated sounds.

Our results are consistent with the findings in auditory-visual (AV) speech studies (Kauramaki et al., 2010; Skipper, van Wassenhove, Nusbaum, & Small, 2007; van Wassenhove, Grant, & Poeppel, 2005). Van Wassenhove et al. (2005) suggested the first computational stage of AV speech integration is a feature-extraction stage where visual information enables the prediction of the auditory input, and this early visual processing can be used in a forward model, i.e. “analysis-by-synthesis” model, for computing the residual error between the following auditory input and the internal prediction. Later, Skipper et al. (2007) illustrated this predictive mechanism using functional magnetic resonance imaging, and showed AV and visual speech perception yielded more robust activities in the motor system associated with speech production, supporting their hypothesis that motor commands can be used to predict the sensory consequences of those corresponding movements through the transmission of a copy of those motor commands. Kauramaki et al. (2010) suggest the origin of the top-down inputs that cause N100 or M100 suppression during lip-reading is from the speech-production system that acts like mirror neurons (Rizzolatti & Craighero, 2004). Specifically, it was hypothesized that during lip-reading the observers do not vocalize themselves, but their speech-production cortical areas are activated through “mirroring” of observed vocal actions; a copy of this mirroring response could then be sent to auditory cortex and enhance auditory processing (Kauramaki et al., 2010). In the present study, our findings provide new information for explaining the possible underlying mechanisms of a forward model or corollary discharge/efference copy in AV speech research. Additionally, our results support the consensus that interacting cortical areas are associated with the production and coordination of speech, illustrated and investigated by a comprehensive speech acquisition and production model, the DIVA (Directions Into Velocities of Articulators) model (Guenther, Ghosh, & Tourville, 2006; Guenther et al., 1998). According to the DIVA model, the expected auditory consequences of self-generated speech are encoded in the forward model projections from the “Speech Sound Map” (hypothesized to lie in the left posterior inferior frontal gyrus and ventral premotor cortex) to the “Auditory Target Map” (hypothesized to lie in bilateral planum temporale and superior temporal gyrus); if the incoming auditory signals are within the predicted range, there is an inhibition that cancels the excitatory effects of those incoming auditory signals. Converging with the hypothesis of auditory expectations projected from Speech Sound Map to Auditory Target Map in the DIVA model, our findings may suggest how multiple cortical areas interact with one another. Furthermore, the present study may contribute potential mechanisms to be used in EEG-based brain-machine interfaces for providing communication capabilities to profoundly paralyzed patients (Guenther et al., 2010).

In addition to dampening irrelevant sensations resulting from our own actions, the forward model provides a mechanism for automatic distinction between internally and externally generated percepts across sensory modalities and may even operate in the realm of covert thoughts, which have been viewed as our most complex motor act (Feinberg, 1978). An extension of the forward model to the process of *thinking* would be that the covert rehearsal

of a word or phrase requires activation of the auditory cortical areas involved in speech perception, and in this case, the generation of corollary discharge is essential to distinguishing between self- and externally-generated sensations (Shergill et al., 2002). The translational potential for the present study and other studies of the corollary discharge mechanism is illustrated by Feinberg (1978) who suggested that patients with schizophrenia may have a deficit in this basic neural mechanism, making it difficult for them to distinguish their own inner thoughts and memories from external voices, resulting in auditory hallucinations. Specifically, if a copy of a memory, a thought, or other inner experience is not transmitted, and if its expected sensory consequence is not produced, internal experiences may be experienced as external and perceived as auditory verbal hallucinations. Indeed, previous work from our lab using scalp recorded electrophysiological responses (Ford et al., 2007) suggest that a copy of the speech plan precedes speech onset and may work to dampen the auditory cortical response to that speech in healthy participants. However, the copy of the speech is reduced in patients with schizophrenia, especially in those with severe auditory hallucinations (Ford et al., 2007). Because our method of measuring gamma band phase synchronies between Broca's and auditory cortical areas during Vocalizing does not provide information about the directionality of the transmission of the copy of the vocalization, additional evidence will be needed to elucidate the directional specifics of the forward model.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Arieli A, Sterkin A, Grinvald A, Aertsen A. Dynamics of ongoing activity: Explanation of the large variability in evoked cortical responses. *Science*. 1996; 273:1868–1871. [PubMed: 8791593]
- Artacho-Perula E, Arbizu J, Arroyo-Jimenez MM, Marcos P, Martinez-Marcos A, Blazot X, et al. Quantitative estimation of the primary auditory cortex in human brains. *Brain Res*. 2004; 1008:20–28. [PubMed: 15081378]
- Basar-Eroglu C, Struber D, Schurmann M, Stadler M, Basar E. Gamma-band responses in the brain: a short review of psychophysiological correlates and functional significance. *Int J Psychophysiol*. 1996; 24:101–112. [PubMed: 8978437]
- Bastiaansen, MCM.; Hagoort, P. Oscillatory brain dynamics during language comprehension. In: Klimesch, W.; Neuper, C., editors. *Event-related dynamics of brain oscillations*. Amsterdam: Elsevier; 2006.
- Bell CC. An efference copy which is modified by reafferent input. *Science*. 1981; 214:450–453. [PubMed: 7291985]
- Celesia GG, Broughton RJ, Rasmussen T, Branch C. Auditory evoked responses from the exposed human cortex. *Electroencephalogr Clin Neurophysiol*. 1968; 24:458–465. [PubMed: 4172362]
- Chen CM, Lakatos P, Shah AS, Mehta AD, Givre SJ, Javitt DC, et al. Functional anatomy and interaction of fast and slow visual pathways in macaque monkeys. *Cereb Cortex*. 2007; 17:1561–1569. [PubMed: 16950866]
- Crapse TB, Sommer MA. Corollary discharge across the animal kingdom. *Nat Rev Neurosci*. 2008; 9:587–600. [PubMed: 18641666]
- Creutzfeldt O, Ojemann G, Lettich E. Neuronal activity in the human lateral temporal lobe. II. Responses to the subjects own voice. *Exp Brain Res*. 1989; 77:476–489. [PubMed: 2806442]
- Curio G, Neuloh G, Numminen J, Jousmaki V, Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. *Hum Brain Mapp*. 2000; 9:183–191. [PubMed: 10770228]

- Davis WJ, Siegler MVS, Mpitsos GJ. Distributed neuronal oscillators and efference copy in the feeding system of Pleurobranchaea. *J Neurophysiol.* 1973; 36:258–274. [PubMed: 4350359]
- Eliades SJ, Wang X. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J Neurophysiol.* 2003; 89:2194–2207. [PubMed: 12612021]
- Eliades SJ, Wang X. Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cereb Cortex.* 2005; 15:1510–1523. [PubMed: 15689521]
- Eliades SJ, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature.* 2008; 453:1102–1106. [PubMed: 18454135]
- Engel AK, Fries P, Singer W. Dynamic predictions: Oscillations and synchrony in top-down processing. *Nat Rev Neurosci.* 2001; 2:704–716. [PubMed: 11584308]
- Feinberg I. Efference copy and corollary discharge: implications for thinking and its disorders. *Schizophr Bull.* 1978; 4:636–640. [PubMed: 734369]
- Fell J, Klaver P, Lehnertz K, Grunwald T, Schaller C, Elger CE, et al. Human memory formation is accompanied by rhinalhippocampal coupling and decoupling. *Nat Neurosci.* 2001; 4:1259–1264. [PubMed: 11694886]
- Ford JM, Gray M, Faustman WO, Heinks T, Mathalon DH. Reduced gamma-band coherence to distorted feedback during speech when what you say is not what you hear. *Inter J Psychophysiol.* 2005; 57:143–150.
- Ford JM, Gray M, Faustman WO, Roach BJ, Mathalon DH. Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology.* 2007; 44:522–529. [PubMed: 17565658]
- Ford JM, Mathalon DH, Heinks T, Kalba S, Roth WT. Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *American Journal of Psychiatry.* 2001a; 158:2069–2071. [PubMed: 11729029]
- Ford JM, Mathalon DH, Heinks T, Kalba S, Roth WT. Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *American Journal of Psychiatry.* 2001b; 158:2069–2071. [PubMed: 11729029]
- Ford JM, Roach BJ, Faustman WO, Mathalon DH. Synch before you speak: Auditory hallucinations in schizophrenia. *American Journal of Psychiatry.* 2007; 164:458–466. [PubMed: 17329471]
- Ford JM, Roach BJ, Mathalon DH. How to assess the corollary discharge in humans using non-invasive neurophysiological methods. *Nature Protocols.* 2010 in press.
- Frey S, Campbell JSW, Pike GB, Petrides M. Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *J Neurosci.* 2008; 28:11435–11444. [PubMed: 18987180]
- Fries P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci.* 2005; 9:474–480. [PubMed: 16150631]
- Fries P, Reynolds JH, Rorie AE, Desimone R. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science.* 2001; 291:1560–1563. [PubMed: 11222864]
- Gregoriou GG, Gotts SJ, Zhou H, Desimone R. High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science.* 2009; 324:1207–1210. [PubMed: 19478185]
- Grusser OJ. Interaction of efferent and afferent signals in visual perception a history of ideas and experimental paradigms. *Acta Psychologica.* 1986; 63:3–21. [PubMed: 3591440]
- Guenther FH, Brumberg JS, Wright EJ, Nieto-Castanon A, Tourville JA, Panko M, et al. A wireless brain-machine interface for real-time speech synthesis. *PLoS One.* 2010; 4:e8218. [PubMed: 20011034]
- Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* 2006; 96:280–301. [PubMed: 16040108]
- Guenther FH, Hampson M, Johnson D. A theoretical investigation of reference frames for the planning of speech movements. *Psychol Rev.* 1998; 105:611–633. [PubMed: 9830375]
- Hamada Y, Miyashita E, Tanaka H. Gamma-band oscillations in the “barrel cortex” precede rat’s exploratory whisking. *Neuroscience.* 1999; 88:667–671. [PubMed: 10363807]
- Heinks-Maldonado TH, Mathalon DH, Gray M, Ford JM. Fine-tuning of auditory cortex during speech production. *Psychophysiology.* 2005; 42:180–190. [PubMed: 15787855]

- Helmholtz, H. Handbuch der physiologischen optik. In: Karsten, G., editor. Allgemeine Encyklopadie der Physik. Leipzig: Voss; 1867.
- Helmholtz, H. Helmholtz's treaties on physiological optics. In: Southall, JPC., editor. Optical Society of America. Rochester, NY: 1925. p. 44-51.
- Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: An MEG study. *J Cogn Neurosci*. 2002; 14:1125–1138. [PubMed: 12495520]
- Jordan MI, Rumelhart DE. Forward models: supervised learning with a distal teacher. *Cognitive Science*. 1992; 16:307–354.
- Kauramaki J, Jaaskelainen IP, Hari R, Mottonen R, Rauschecker JP, Sams M. Lipreading and covert speech production similarly modulate human auditory-cortex responses to pure tones. *J Neurosci*. 2010; 30:1314–1321. [PubMed: 20107058]
- Lachaux JP, Rodriguez E, Martinerie J, Varela F. Measuring phase synchrony in brain signals. *Hum Brain Mapp*. 1999; 8:194–208. [PubMed: 10619414]
- Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. Neuronal mechanisms and functional significance of multisensory integration in primary auditory cortex. *Neuron*. 2007; 53:279–292. [PubMed: 17224408]
- Liegeois-Chauvel C, Musolino A, Chauvel P. Localization of the primary auditory area in man. *Brain*. 1991; 114:139–153. [PubMed: 1900211]
- Maris E, Schoffelen JM, Fries P. Nonparametric statistical testing of coherence differences. *J Neurosci Methods*. 2007; 163:161–175. [PubMed: 17395267]
- Martikainen MH, Kaneko K, Hari R. Suppressed responses to self-triggered sounds in the human auditory cortex. *Cereb Cortex*. 2005; 15:299–302. [PubMed: 15238430]
- McKinstry JL, Seth AK, Edelman GM, Krichmar JL. Embodied models of delayed neural responses: spatiotemporal categorization and predictive motor control in brain based devices. *Neural Networks*. 2008; 21:553–561. [PubMed: 18495424]
- Miltner WHR, Braun C, Arnold M, Witte H, Taub E. Coherence of gamma-band EEG activity as a basis for associative learning. *Nature*. 1999; 397:434–436. [PubMed: 9989409]
- Muller-Preuss P, Ploog D. Inhibition of auditory cortical neurons during phonation. *Brain Research*. 1981; 215:61–76. [PubMed: 7260601]
- Murthy VN, Fetz EE. Coherent 25- to 35-Hz oscillations in the sensorimotor cortex of awake behaving monkeys. *Proc Natl Acad Sci USA*. 1992; 89:5670–5674. [PubMed: 1608977]
- Naatunen R, Picton T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiol*. 1987; 24:375–425.
- Papademetris, X.; Jackowski, AP.; Schultz, RT.; Staib, LH.; Duncan, JS. Integrated intensity and point-feature non-rigid registration. In: Barillot, C.; Haynor, D.; Hellier, P., editors. Medical image computing and computer-assisted intervention. Saint Malo, France: Springer; 2004. p. 763-770.
- Pesaran B, Pezaris JS, Sahani M, Mitra PP, Andersen RA. Temporal structure in neuronal activity during working memory in macaque parietal cortex. *Nat Neurosci*. 2002; 5:805–811. [PubMed: 12134152]
- Phillips WA, Singer W. In search of common foundations for cortical computation. *Behav Brain Sci*. 1997; 20:657–722. [PubMed: 10097008]
- Poulet JF, Hedwig B. A corollary discharge maintains auditory sensitivity during sound production. *Nature*. 2002; 418:872–876. [PubMed: 12192409]
- Poulet JF, Hedwig B. The cellular basis of a corollary discharge. *Science*. 2006; 311:518–522. [PubMed: 16439660]
- Reite M, Adams M, Simon J, Teale P, Sheeder J, Richardson D, et al. Auditory M100 component 1: relationship to Heschl's gyri. *Cogn Brain Res*. 1994; 2:13–20.
- Rizzolatti G, Craighero L. The mirror-neuron system. *Annu Rev Neurosci*. 2004; 27:169–192. [PubMed: 15217330]
- Rodriguez E, George N, Lachaux JP, Martinerie J, Renault B, Varela F. Perception's shadow: Long-distance synchronization of human brain activity. *Nature*. 1999; 397:430–433. [PubMed: 9989408]

- Roopun AK, Kramer MA, Carracedo LM, Kaiser M, Davies CH, Traub RD, et al. Temporal interactions between cortical rhythms. *Frontiers in Neuroscience*. 2008; 2:145–154. [PubMed: 19225587]
- Saalmann YB, Pigarev IN, Vidyasagar TR. Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. *Science*. 2007; 316:1612–1615. [PubMed: 17569863]
- Schafer EWP, Marcus MM. Self-stimulation alters human sensory brain responses. *Science*. 1973; 181:175–177. [PubMed: 4711735]
- Shergill SS, Brammer MJ, Fukuda R, Bullmore E, Amaro E Jr, Murray RM, et al. Modulation of activity in temporal cortex during generation of inner speech. *Hum Brain Mapp*. 2002; 16:219–227. [PubMed: 12112764]
- Shergill SS, Samson G, Bays PM, Frith CD, Wolpert DM. Evidence for sensory prediction deficits in schizophrenia. *Am J Psychiatry*. 2005; 162:2384–2386. [PubMed: 16330607]
- Singer W. Neuronal synchrony: a versatile code for the definition of relations? *Neuron*. 1999a; 24:49–65. [PubMed: 10677026]
- Singer W. Striving for coherence. *Nature*. 1999b; 397:391–393. [PubMed: 9989402]
- Skipper JI, van Wassenhove V, Nusbaum HC, Small SL. Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cereb Cortex*. 2007; 17:2387–2399. [PubMed: 17218482]
- Sommer MA, Wurtz RH. A pathway in primate brain for internal monitoring. *Science*. 2002; 296:1480–1482. [PubMed: 12029137]
- Sperry RW. Neural basis of the spontaneous optokinetic response produced by visual inversion. *J Comp Physiol Psych*. 1950; 43:482–489.
- Thiele A, Henning P, Kubischik M, Hoffmann KP. Neural mechanisms of saccadic suppression. *Science*. 2002; 295:2460–2462. [PubMed: 11923539]
- van Wassenhove V, Grant KW, Poeppel D. Visual speech speeds up the neural processing of auditory speech. *Proc Natl Acad Sci USA*. 2005; 102:1181–1186. [PubMed: 15647358]
- Varela F, Lachaux JP, Rodriguez E, Martinerie J. The brainweb: Phase synchronization and large-scale integration. *Nat Rev Neurosci*. 2001; 2:229–239. [PubMed: 11283746]
- Von Holst E, Mittelstaedt H. Das Reafferenzprinzip. *Naturwissenschaften*. 1950; 37:464–476.
- Von Stein A, Chiang C, Konig P. Top-down processing mediated by interareal synchronization. *Proc Natl Acad Sci USA*. 2000; 97:14748–14753. [PubMed: 11121074]
- Webb B. Neural mechanisms for prediction: do insects have forward models? *Trends Neurosci*. 2004; 27:278–282. [PubMed: 15111010]
- Wilcox, RR. Introduction to robust estimation and hypothesis testing. Second. New York: Academic Press; 2004.
- Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. *Science*. 1995; 269:1880–1882. [PubMed: 7569931]
- Womelsdorf T, Schoffelen JM, Oostenveld R, Singer W, Desimone R, Engel AK, et al. Modulation of neuronal interactions through neuronal synchronization. *Science*. 2007; 316:1609–1612. [PubMed: 17569862]
- Woods DL. The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalogr Clin Neurophysiol Suppl*. 1995; 44:102–109. [PubMed: 7649012]
- Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron*. 2008; 58:429–441. [PubMed: 18466752]

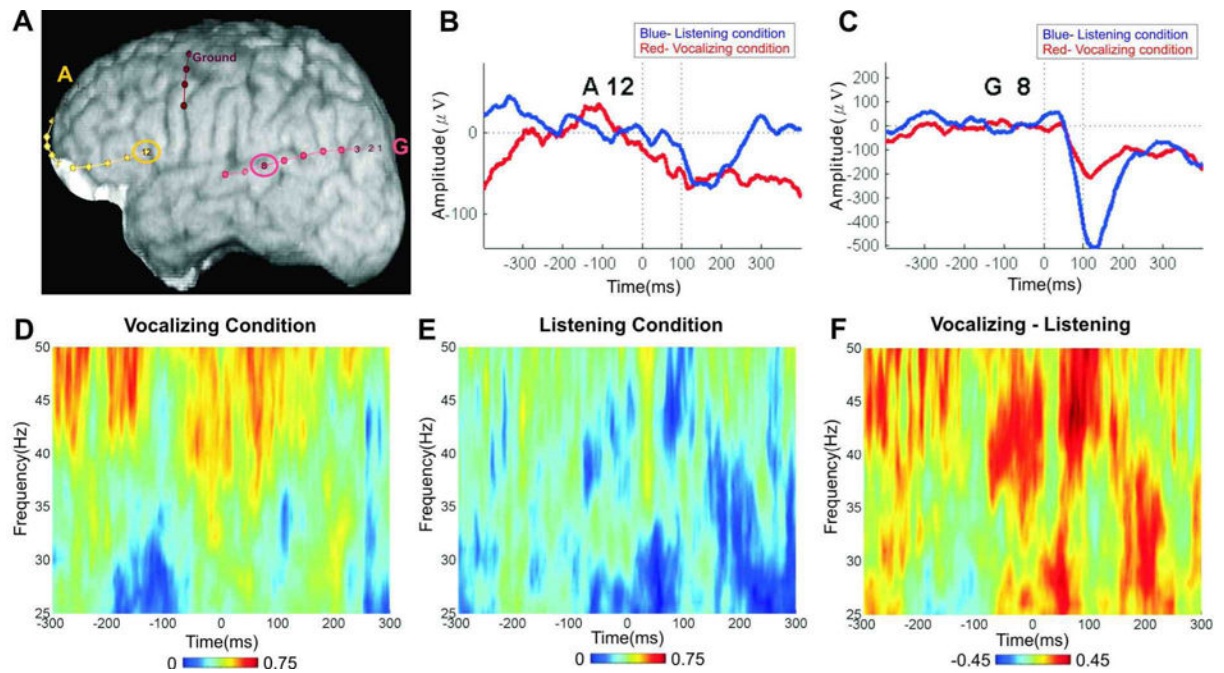


Figure 1.

Phase locking value (PLV) between Broca's area and auditory cortex. (A) Intracranial electrode contact placements of patient A are shown and the neuroelectric signals of two circled electrode contacts (i.e., A12 circled in yellow and G8 circled in red) were used to calculate PLVs for both of vocalizing and listening conditions. Electrode contacts A12 and G8 are located in Broca's area and auditory cortex, respectively. (B) electrocorticogram data were epoched to onsets of participant-generated speech sounds. The averaged event-related potential of electrode contact A12 is shown (listening condition is in blue trace and vocalizing condition is in red trace). (C) As in (B), for electrode contact G8 and N100 suppression in vocalizing condition (red trace) is observed. (D) In vocalizing condition, PLVs between electrode contacts A12 and G8 from -300 ms to 300 ms in gamma frequencies, ranging from 25 Hz to 50 Hz are shown. PLV ranges from 0 (no phase synchrony) to 1 (perfect phase synchrony). (E) As in (D), for listening condition. (F) The difference plot of Figures 1D and 1E (i.e., PLVs of vocalizing condition minus PLVs of listening condition) shows enhancement of PLVs in warm colors and suppression of PLVs in cold colors during Vocalizing compared to Listening.

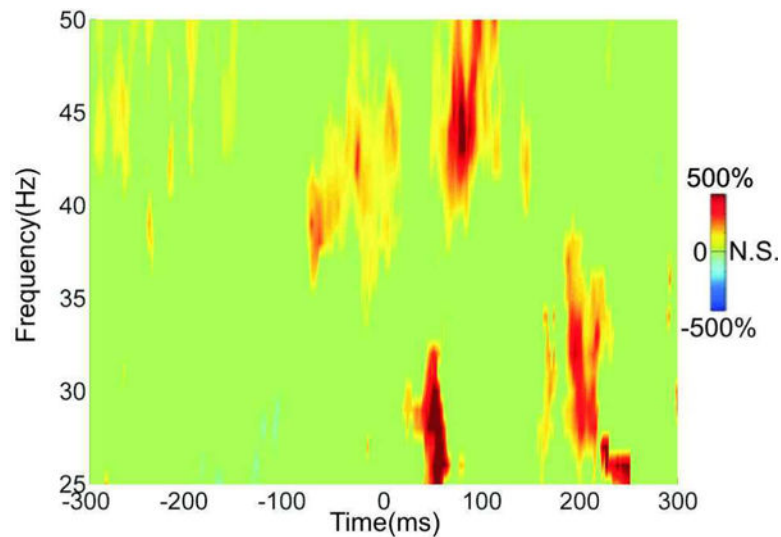


Figure 2.

Percentage changes of significant differences of phase locking values (PLVs) between vocalizing and listening conditions. PLV represents the phase synchrony between Broca's area and auditory cortex at the given frequency and time point. This figure shows the percent changes in the significant differences of PLVs between Vocalizing and Listening with Listening as the baseline ($\{[PLV_{Vocalizing} - PLV_{Listening}]_{significant} / PLV_{Listening}\} \times 100\%$, see details in "METHODS"). Since it is based on significant differences of PLVs between two conditions, red regions would represent significant enhanced changes of PLVs from the baseline, blue regions would represent significant suppressed changes of PLVs from the baseline, and green regions would represent non-significant changes of PLVs between two conditions.

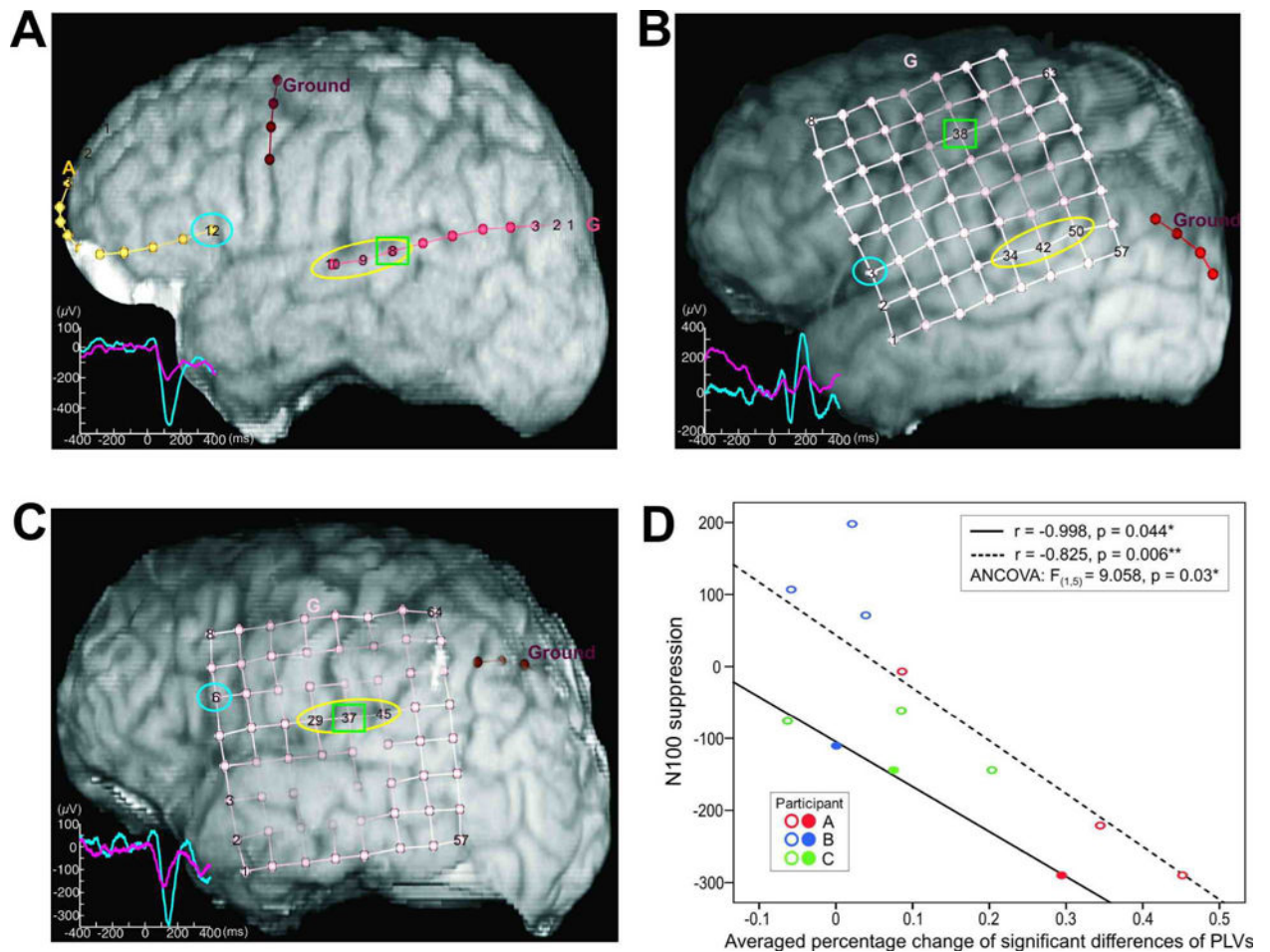


Figure 3.

Relationship between N100 suppression and averaged percent change of significant differences of PLVs. Intracranial electrode placements of patients A, B, and C are shown, in (A), (B) and (C), respectively. The neuroelectric signals of circled electrode contacts (i.e., Broca's area circled in blue and auditory cortex circled in yellow) were used for calculating phase locking values (PLVs) and the averaged percent change of significant differences of PLVs (see details in "METHODS"). The averaged event-related potentials to speech onset of green-rectangle-selected electrodes are shown in bottom-left corner of (A), (B), and (C) (magenta trace is Vocalizing and blue trace is Listening). (D) The averaged percent change of significant differences of PLVs is on the x-axis and N100 suppression is shown on the y-axis (see details in "METHODS"). Two-tailed Pearson's correlation coefficients (r) and ANCOVA results are shown in the upper-right corner. A two-tailed Pearson's correlation coefficient (r) between the yellow-circled N100 suppression and its corresponding averaged percent change of significant differences of PLVs was calculated (i.e., the hollow circles and the dashed line; $r = -0.825$, $p = 0.006$). Additionally, a mixed model analysis of covariance (ANCOVA) was applied to those hollow-circled values ($F_{(1,5)} = 9.058$, $p = 0.03$). The other two-tailed Pearson's r between the best N100 suppression (from the green-rectangle-selected electrode) and the grand mean of averaged percent changes of

significant differences of PLVs was computed (i.e., the solid circles and line; $r = -0.998$, $p = 0.044$).