

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

International Journal of Psychophysiology xx (2005) xxx – xxx

INTERNATIONAL  
JOURNAL OF  
PSYCHOPHYSIOLOGY[www.elsevier.com/locate/ijpsycho](http://www.elsevier.com/locate/ijpsycho)

## Reduced gamma-band coherence to distorted feedback during speech when what you say is not what you hear

Judith M. Ford<sup>a,b,\*</sup>, Max Gray<sup>b</sup>, William O. Faustman<sup>b</sup>, Theda H. Heinks<sup>b</sup>, Daniel H. Mathalon<sup>a</sup><sup>a</sup>*Department of Psychiatry, Yale University School of Medicine, New Haven, CT Psychiatry Service, Veterans Affairs Connecticut Health Care System, West Haven, CT, United States*<sup>b</sup>*Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA Psychiatry Service, Veterans Affairs Palo Alto Health Care System, Palo Alto, CA, United States*

Received 15 October 2004; received in revised form 1 March 2005; accepted 15 March 2005

### Abstract

**Background:** Communication between the frontal lobes, where speech is generated, and the temporal lobes, where it is perceived, may occur through the action of an efference copy/corollary discharge mechanism that prepares the temporal lobes for the expected sound. We suggest that coherence of EEG in gamma-band between frontal and temporal lobes may reflect the successful action of such a mechanism. We tested the hypothesis that there would be a disruption of gamma-band coherence when the expected auditory consequence of speech does not match the auditory experience.

**Method:** EEG was recorded from 21 healthy adult subjects as they uttered the sound [a:] (Talking condition) and as they heard these recorded sounds played back (Listening condition). As they spoke, subjects heard real-time feedback of the sounds that were: (1) pitch-shifted down one semi-tone, (2) pitch-shifted down one-half of a semi-tone, or (3) not pitch-shifted (veridical), each in separate runs. Event-related gamma coherence between frontal and temporal sites was calculated relative to the onset of the sound, *as it was being spoken* during Talking, and as it was being played back during Listening.

**Results:** Frontal–temporal gamma-band coherence spanning 33–43 Hz was greater during Talking than Listening and greater when speech was veridical than when it was distorted a whole semi-tone.

**Conclusions:** Gamma-band fronto-temporal synchrony may reflect a “binding of expectation with experience.” Disruption of this synchrony may provide feedback to the frontal lobes, particularly regions subserving vocalization, to implement sensorimotor adaptations to either adjust motor programs for speech production in the short run, or to reorganize expectations in the long run.

© 2005 Elsevier B.V. All rights reserved.

**Keywords:** Self-monitoring of distorted speech; EEG coherence

### 1. Introduction

Traditionally, neuroscientists have thought that sensory information processing was revealed by changes in firing patterns of “smart” neurons in a bottom-up fashion (Singer, 1999). Single-cell recordings were made from neurons, or populations of neurons, in one region of the brain at a time,

and firing rates revealed that neuronal units were tuned to respond to a specific sound frequency (in auditory cortex) or a specific line orientation (in visual cortex). However, more recently, evidence is emerging to suggest that sensory processing is accomplished by a dynamic matching of anticipatory self-generated activity with activity generated by physical input, in a combined top-down/bottom-up fashion (Engel et al., 2001). Studying this type of interactive dynamic system is more complicated and requires simultaneous recording of activity from multiple sites. The quality of perceptual experience may depend on the degree of temporal synchronization between the firing rates of these

\* Corresponding author. Department of Psychiatry, Yale University School of Medicine, New Haven, CT, United States. Tel.: +1 203 932 5711x4642; fax: +1 230 937 3886.

E-mail address: [judith.ford@yale.edu](mailto:judith.ford@yale.edu) (J.M. Ford).

distributed neuronal units. For example, columns of cells that prefer lines of a certain orientation all oscillate in phase, at an a priori rate (Gray and Singer, 1989). If the neuronal systems that subserve perception are essentially “synchrony detectors”, then the exquisite temporal sensitivity afforded by such a mechanism could account for the capacity of assemblies of neurons to respond instantaneously to sensation, allowing rapid modifications in behavior if necessary.

While this type of synchrony can be studied at the single-cell level, it can also be studied from large groups of neurons in animals, and from the scalps of human volunteers, using electroencephalography (EEG). Although coherences between directly recorded local fields in animals are rarely observed at distances observed in humans with recordings from the scalp, evidence is accumulating that gamma range (20–50 Hz) oscillations in neurons and groups of neurons, and synchrony of the oscillations between neurons and neuronal groups, are a fundamental mechanism by which information is coded and integrated in the mammalian brain (Fernandez et al., 2002; Singer, 1999). For example, attention is associated with an increase in synchrony, not a change in firing rates (Fries et al., 2001), and when expected events are experienced, there is an increase in synchrony. Such synchronous oscillations may reflect higher-order perceptual and cognitive functions involved in making sense of a chaotic environment, binding disparate information into a single percept, and forming new memories (Singer, 1999). In particular, synchronous, or correlated, neuronal firing in the gamma range may signal communication across distant brain regions.

Of interest to us is how the brain communicates motor plans to sensory cortex where the sensations resulting from those actions are received. Such communication provides a mechanism for anticipation of the sensory consequences of self-generated actions. Detection of discrepancies between anticipated and actual sensory consequences of self-generated acts has been proposed as a fundamental mechanism for rapidly recognizing and responding to sensations that result from external stimuli and for distinguishing them from those sensations that result from self-generated actions (Angel, 1976; Jeannerod, 2003; Wolpert and Kawato, 1988). Across sensory modalities, the operation of motor-sensory “forward model” systems has been proposed to account for the capacity of organisms to rapidly detect such discrepancies. These models posit the transmission of “efference copies” of motor commands to sensory regions where they produce “corollary discharges” representing the anticipated sensory consequences of the impending motor act (Von Holst and Mittelstaedt, 1950; Angel, 1976, Jeannerod, 2003, Wolpert and Kawato, 1988). A subtractive comparison of the corollary discharge with the actual sensory feedback (i.e., “sensory reafference”) is thought to provide the specific mechanism by which discrepancies are detected. When the corollary discharge signal representing the

anticipated sensory consequence of the impending motor act is well-matched by the actual sensory reafference produced by this act, the net subtraction of signals results in little discrepancy and the sensory activation is minimal. However, when the sensory reafference is not matched by a corollary discharge signal, the subtraction does not “cancel” the sensory reafference, and the activation of the sensory cortex is enhanced and attributed to external sources. This allows rapid orienting to the sensory input and rapid responding to potentially important stimuli in the local environment. Moreover, repeated discrepancies between the corollary discharge and sensory reafference signals have been shown to invoke sensorimotor adaptations to bring the anticipated and actual sensory consequences of self-generated actions back into alignment (see Houde and Jordan, 1998).

Physiological studies have documented a corollary discharge mechanism during self-generated sounds across organisms from crickets (Poulet and Hedwig, 2002) to primates (Muller-Preuss and Ploog, 1981). In subhuman primates, about 50% of call-responsive neurons are inhibited during phonation compared to when those phonations are recorded and played back (Muller-Preuss and Ploog, 1981). In humans, pre-operative intracranial recordings show a reduction in responsiveness of neurons in the middle temporal gyrus, starting 100 ms before and continuing during the patient’s speech but not occurring when another person speaks to the patient (Creutzfeldt et al., 1989). Magnetoencephalography (MEG) data from intact human volunteers show that auditory cortical responsiveness to self-produced speech is attenuated compared to when those same sounds are played back (Curio et al., 2000; Houde et al., 2002). Using EEG, we showed a similar effect on the auditory N1 component of the event-related brain potential (ERP) (Ford et al., 2001a), with the N1 being smaller to speech sounds during talking than during playback.

These results could be accounted for by the operation of a forward model in which an efference copy of the speech commands and a corollary discharge representing their predicted auditory consequences modulate the responsiveness of the auditory cortex. More recently, we have shown that this forward model operates with relative precision such that cortical dampening is greatest when there is a close match between the spoken (intended) and heard (observed) sounds (Heinks-Maldonado et al., 2005). By altering the auditory feedback during speaking, we found that N1 reduction was greatest when the spoken sound was unaltered, consistent with the notion of a precise forward model modulating cortical responsiveness to self-generated speech. The operation of forward model/corollary discharge systems in the speech-audition network may serve language acquisition during childhood, second-language acquisition in adulthood, and acquisition of regional accents throughout life.

While we assume that this efference copy/corollary discharge mechanism is instantiated in a connection

between frontal lobes (where speech is produced) and temporal lobes (where speech is perceived), none of the studies cited above offer direct anatomical evidence that the frontal lobes were involved. However, using EEG coherence analysis, we addressed the frontal–temporal connection (Ford et al., 2002). We showed that event-related frontal–temporal coherence in all EEG bands to random acoustic probes was greater when subjects were repeating sentences aloud than when they were simply listening to a recording of those sentences.

The current study was intended to extend this finding in the following ways:

- (1) In the previous study, we measured coherence to externally presented acoustic probes during speech, not to speech sounds themselves as they were being produced. Given the prosodic nature of speech, some of the probes may have fallen between words, diluting speaking-related coherence. In the current study, the probes were the actual sounds being produced. We predicted that we would replicate our original finding of greater coherence during talking than listening when the probe was the spoken sound.
- (2) The acoustic probes in the previous study (pre-recorded [ba] and noise burst) were not the same as the sounds being produced by the subject who was repeating sentences aloud. To the extent that the corollary discharge has a tuning curve anchored to the spoken sound, increases in talking-related coherence to these externally generated acoustic probes may have been underestimated. In the current study, we predicted that when the probe sound is the same as the spoken sound, communication between frontal and temporal lobes would be manifest as synchronous, coherent gamma-band activity.
- (3) Because gamma-band activity has been associated with binding of percepts, we focused on gamma coherence as a function of the match between the expected auditory consequences of the spoken sound and the sound that was actually heard. That is, we distorted the auditory reafference feedback the subjects heard as they vocalized the sounds. We predicted that gamma coherence between the frontal and temporal brain regions would be reduced when the planned speech and its anticipated consequences did not match the auditory percept, because these are the regions involved in producing speech, checking the correspondence between the expected and actual sounds, and adjusting speech accordingly.
- (4) Although we analyzed other frequencies within the broader gamma-band (~20–50 Hz), we focused on two bands of activity, 33–38 Hz and 38–43 Hz. These frequencies have been the target of recent research on schizophrenia (Kwon et al., 1999), the focus of our program of research. Data from patients will be reported separately.

To address gamma-band EEG coherence during speaking with distorted feedback, we altered the sounds that subjects heard as they were vocalizing them. Here we report data from an experiment in which the auditory feedback subjects heard as they vocalized was either undistorted, distorted by a half semi-tone, or distorted by a whole semi-tone. These three blocks of speech were recorded and played back to the subject, enabling a comparison of frontal–temporal gamma coherence during Talking and Listening, as well as a comparison of coherence to veridical and distorted speech.

## 2. Methods and materials

### 2.1. Subjects

Data from 21 healthy adults (14 men, 7 women), ranging in age from 25 to 55 years (mean=41.9, S.D.=7.2). All gave written informed consent after the procedures had been fully described. Subjects were recruited by newspaper and internet advertisements, as well as word-of-mouth, screened by telephone using the psychiatric screening questions from the Structured Clinical Interview for DSM-IV (SCID) (First et al., 1995). Subjects were excluded for any significant history of Axis I psychiatric illness, significant head injury (loss of consciousness greater than 30 min or resulting in neurological sequelae) or neurological or other medical illnesses compromising the central nervous system. All were right handed.

### 2.2. Experimental procedure

#### 2.2.1. Pre-testing procedures

Before electrodes were attached, subjects practiced saying [a:] while minimizing movement artifacts. During the practice, the following steps were taken: (1) sound equipment input levels were adjusted to obtain an average speaking level; (2) subjects were familiarized with the different levels of pitch-shifted feedback, veridical, one-half semi-tone, or one full semi-tone; (3) a short series of veridical [a:] was recorded, from which an exemplar was chosen for conversion to Neuroscan .SND format; (4) the exemplar was down-sampled to 22,050 Hz; (5) the exemplar was used to calibrate the sound system to produce a headphone sensation level 15 dB SPL (C scale) above an individual's average speaking level, determined in the first step. No output level exceeded 100 dB SPL.

#### 2.2.2. Talking and Listening conditions

Two speech distortion experiments were conducted in a counter-balanced order, the Predictable Distortion and Unpredictable Distortion experiments. Only data from the Predictable Distortion experiment are presented here. In the Talking condition, subjects were told to vocalize [a:] about every 1–2 s. The sounds subjects heard over headphones as

they vocalized were either their unaltered voice (Veridical), their voice pitch-shifted downward by a half semi-tone (Half Semi-Tone), or their voice pitch-shifted down by a whole semi-tone (One Semi-Tone). The level of distortion was predictable, as each was repeated for about 60 trials during one of three blocks. The order of blocks within the Talking condition was randomized across subjects. In the Listening condition, the recorded feedback sounds from the Talking condition were played back, in the same order, and subjects were instructed to simply listen. Loudness was the same in the Talking and Listening conditions.

### 2.2.3. Instrumentation

To create the different feedback conditions, we used an audio presentation system that allowed us to detect the subject's vocalization and, in real-time, modulate the subject's voice. When the subject vocalized, the speech signal was picked up by a microphone and sent through a pre-amplifier to a personal computer equipped with sound processing software and hardware. The incoming audio signal was used to generate a trigger pulse that initiated a pitch-shift that was amplified and played back to the subject via headphones, essentially in real-time.

A trigger pulse, signaling onset of vocalization, was generated within the software program on the rising edge of the rectified and low-pass filtered channel of the split incoming audio signal. This internal trigger pulse initiated all other software processing including modification of the original incoming audio channel, resetting the trigger production module, and inserting a trigger code in the EEG data collection system. The rectified and filtered signal was also used internally to drive an envelope follower that modulated the processed output signal to match the incoming audio signal in amplitude and duration. The average duration of the subjects' vocalizations was approximately 350 ms.

## 2.3. EEG

### 2.3.1. Recording

EEG data were collected from 40 sites on the scalp, all referenced to the nose. To reduce the complexity of the data, we analyzed EEG coherence between frontal and temporal–parietal electrodes within each hemisphere. We report data from electrodes placed over left (F7) and right (F8) inferior frontal regions in the vicinity of “Broca's area”, and over left (CP3, CP5, P3, P5) and right (CP4, CP6, P4, P6) temporo-parietal regions in the vicinity of “Wernicke's area”, as well as posterior auditory cortex. Vertical electro-oculogram (VEOG) was recorded from electrodes placed above and below the right eye, and horizontal EOG (HEOG) from electrodes placed at the outer canthus of each eye. EEG and EOG were sampled at 1000 Hz continuously and band pass filtered between 0.05 and 50 Hz. EEG data were epoched from –100 ms to 923 ms around the onset of [a:]. Individual epochs were corrected for the effects of eye

blinks and eye movements based on correlations of the VEOG and HEOG with the EEG recorded at each electrode site (Gratton et al., 1983). Epochs were rejected from analysis if voltages exceeded  $\pm 100 \mu\text{V}$ . All remaining samples were subjected to visual screening by the investigators to remove trials with incorrect/inaccurate sound triggers, and trials with gross noise contamination including excessive alpha activity, transient spikes, muscle activity or subject movement. On average, 38 trials (60%) remained in the analysis.

### 2.3.2. Analysis of EEG coherence

To translate data from the time domain to the frequency domain, a Fast Fourier Transform was calculated on all 1024 points.

We report coherence data for gamma-band frequencies between 21 and 48 Hz, calculated in 1 Hz bins, for 5 sub-bands: Gamma 25 (20.51–28.31 Hz), Gamma 30 (28.32–33.19 Hz), Gamma 35 (33.2–38.08 Hz), Gamma 40 (38.09–42.96 Hz), and Gamma 45 (42.97–47.84 Hz). Coherence was calculated as the spectral cross-correlation between two electrodes normalized by their power spectra (NeuroscanLabs, 1999) for four left hemisphere pairs (F7–CP5, F7–CP3, F7–P5, F7–P3) and four right hemisphere pairs (F8–CP6, F8–CP4, F8–P6, F8–P4). The coherence values for each electrode pair were averaged across the single trials. The resulting average event-related coherence values for each electrode pair were the dependent variables in the five analyses of variance (ANOVA), one for each sub-band.

Although subjects were trained to make short utterances every 1–2 s, some produced sounds that were longer and more frequent, with very little “utterance-free” time. Thus, no baseline coherence estimates could be reliably made.

### 2.3.3. Statistical analyses

Four-way repeated measures ANOVAs were performed on the average event-related coherence values for the specific frontal–temporal electrode pairs listed above, for each of the 5 sub-bands. There were 4 factors: Condition (Talk/Listen), Distortion (Veridical, Half Semi-Tone, One Semi-Tone, Hemisphere (Left, Right), and Temporal Site (CP3/4, CP5/6, P3/4, P5/6). Follow-up tests between means were conducted using planned contrast tests using the pooled error term. Huyn–Feldt epsilon corrections for non-sphericity were used to evaluate all  $F$  ratios for effects involving two or more degrees of freedom. In addition, we compared gamma coherence during the Talking and Listen conditions for each level of distortion and each electrode pair using paired  $t$ -tests.

## 3. Results

As can be seen in Fig. 1, gamma-band coherence during Talking was greater than during Listening, and this effect

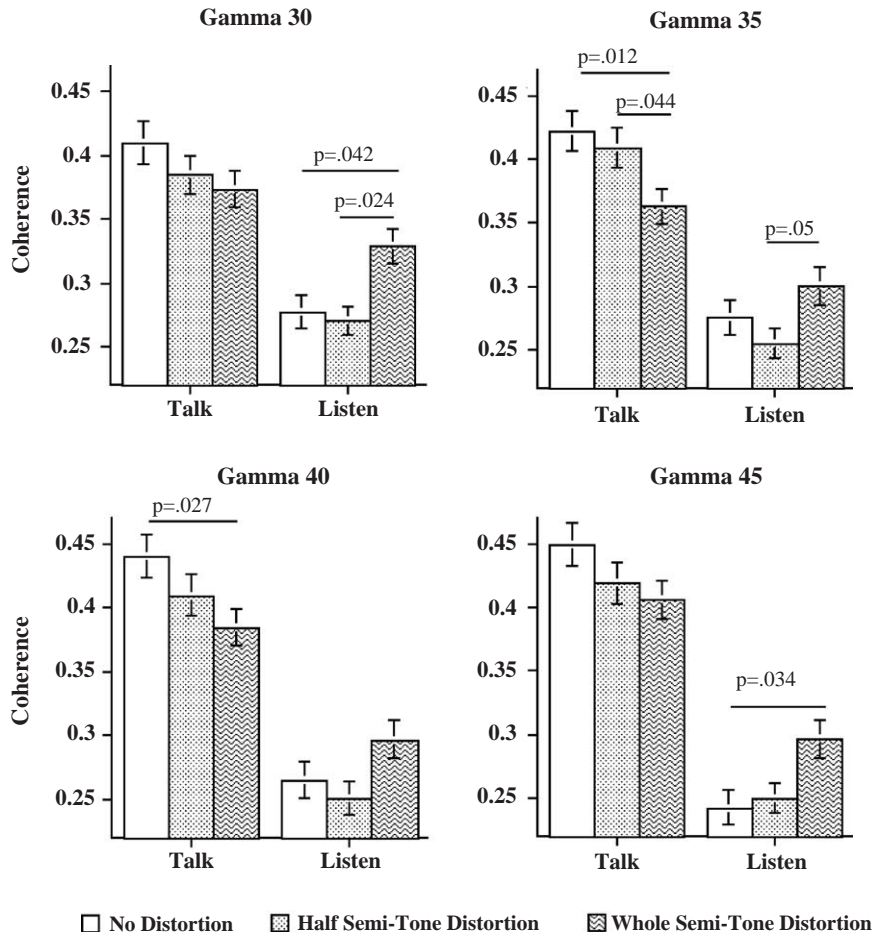


Fig. 1. Gamma-band coherence, averaged across sites and hemispheres, is plotted separately for talking and listening condition for each of the sub-bands having a significant Condition  $\times$  Distortion interaction as seen in Table 1. Significance levels for pair-wise comparisons between different levels of distortion are indicated separately for Talking and Listening.

was reduced when the sound produced during Talking (auditory refference) was distorted for the sub-bands, Gamma 30, Gamma 35, Gamma 40, and Gamma 45. This impression was confirmed by significant interactions

between Talk/Listen and Distortion obtained from the 4-way ANOVA (Table 1) for each of these sub-bands.

There was also a significant main effect of Talk/Listen for each sub-band, with coherence being greater during

Table 1  
Four-way analysis of variance for gamma coherence between frontal (F7, F8) and temporal sites

Source	df	Gamma 25		Gamma 30		Gamma 35		Gamma 40		Gamma 45	
		F-Value	H-F	F-Value	H-F	F-Value	H-F	F-Value	H-F	F-Value	H-F
Distortion (Veridical, Half Semi-Tone, One Semi-Tone)	2,40	0.67	0.52	1.01	0.37	0.64	0.54	0.82	0.45	0.53	0.57
Hemisphere (Left, Right)	1,20	0.04	0.84	0.11	0.75	0.31	0.58	0.44	0.52	0.22	0.64
Temporal Site (CP3/4, CP5/6, P3/4, P5/6)	3,60	51.36	0.0001	28.00	0.0001	23.03	0.0001	25.34	0.0001	26.64	0.0001
Condition (Talk/Listen)	1,20	12.59	0.002	7.36	0.01	8.48	0.01	9.33	0.01	13.51	0.0015
Distortion * Hemisphere	2,30	0.54	0.59	0.24	0.76	0.32	0.73	1.40	0.26	0.06	0.91
Distortion * Temporal Site	6,120	0.41	0.80	0.96	0.44	1.56	0.19	0.66	0.64	0.51	0.77
Hemisphere * Temporal Site	3,60	0.12	0.91	0.37	0.70	0.28	0.77	0.17	0.88	0.19	0.83
<b>Distortion * Condition</b>	<b>2,40</b>	<b>1.30</b>	<b>0.28</b>	<b>3.55</b>	<b>0.04</b>	<b>5.20</b>	<b>0.01</b>	<b>3.68</b>	<b>0.03</b>	<b>3.96</b>	<b>0.03</b>
Hemisphere * Condition	1,20	0.22	0.65	0.28	0.60	0.00	0.98	0.07	0.80	0.10	0.75
Temporal Site * Condition	3,60	7.14	0.0038	4.94	0.01	9.06	0.0006	9.41	0.0002	8.37	0.0004
Distortion * Hemisphere * Temporal Site	6,120	1.17	0.33	0.53	0.62	1.56	0.20	1.25	0.30	0.78	0.54
Distortion * Hemisphere * Condition	2,40	0.59	0.56	0.45	0.64	0.06	0.94	0.13	0.85	0.49	0.59
Distortion * Temporal Site * Condition	6,120	0.46	0.77	0.88	0.47	0.70	0.58	0.32	0.85	0.28	0.89
Hemisphere * Temporal Site * Condition	3,60	0.26	0.79	0.54	0.58	0.38	0.64	1.56	0.21	1.07	0.36
Distortion * Hemisphere * Temporal Site * Condition	6,120	0.19	0.91	0.08	0.97	0.03	1.00	0.45	0.75	1.15	0.34

Talking than Listening. This effect was especially strong for the F7–CP3 pair (Talk/Listen  $\times$  Temporal Site interactions). A main effect of Temporal Site revealed that coherence was greatest for this electrode pair, perhaps because of the small inter-electrode distance. It is important to note that this inter-electrode distance effect on coherence did not affect our primary finding of greater distortion effects during Talking than Listening, as the Talk/Listen  $\times$  Distortion  $\times$  Site interaction was not significant for any of the sub-bands. There was no main effect of Distortion for any sub-band.

Planned comparisons between the different levels of distortion during the Talking condition revealed that frontal–temporal gamma coherence was reduced when speech feedback was distorted by one semi-tone compared to veridical feedback for Gamma 40 and Gamma 35, and by a half semi-tone compared to veridical feedback for Gamma 35. None of the pair-wise comparisons were significant for Gamma 30, or Gamma 45 during Talking. Of note, because of the low-pass filter imposed by the mastoid bone, sounds are more “self-like” or veridical when pitch is slightly shifted down, as with a half semi-tone pitch shift (Shuster and Durrant, 2003). Nevertheless, there were linear trends in the means across the different levels of Distortion during Talking for Gamma 35 ( $p < 0.02$ ), Gamma 40 ( $p = 0.04$ ), and a trend for Gamma 45 ( $p < 0.10$ ).

Planned comparisons between different levels of distortion during Listening revealed that coherence was somewhat greater to the sounds that were distorted by a whole

semi-tone than those distorted by a half semi-tone or that were veridical, as can be seen in Fig. 1.

We compared gamma coherence during Talking to coherence during Listening for each electrode pair at each level of distortion. The resulting  $t$ -values and probability levels are portrayed graphically in Fig. 2, for Gamma 40. Coherence during Talking was greater than during Listening when the sounds were not distorted or only distorted by half a semi-tone in 15 of the 16 electrode pairs tested.

#### 4. Discussion

To the best of our knowledge, this is the first report of event-related gamma-band coherence to sounds *as they are being spoken*. Although earlier, we reported that event-related gamma-band coherence was greater during talking than listening, the events were randomly presented, externally generated acoustic probes ([ba] and white noise burst). Because these probes may have fallen during pauses in speaking, our estimate of talking-related coherence was sub-optimal. However, by calculating gamma-band coherence to sounds *as they are being spoken*, we have maximized our ability to obtain talking-related coherence. We can now conclude that there is greater gamma-band coherence during talking than during listening, indicating greater communication in the form of synchronous gamma oscillations between frontal and temporal lobes during talking than

T-tests Indicating Greater Gamma40 Coherence during Talking than Listening (p values)

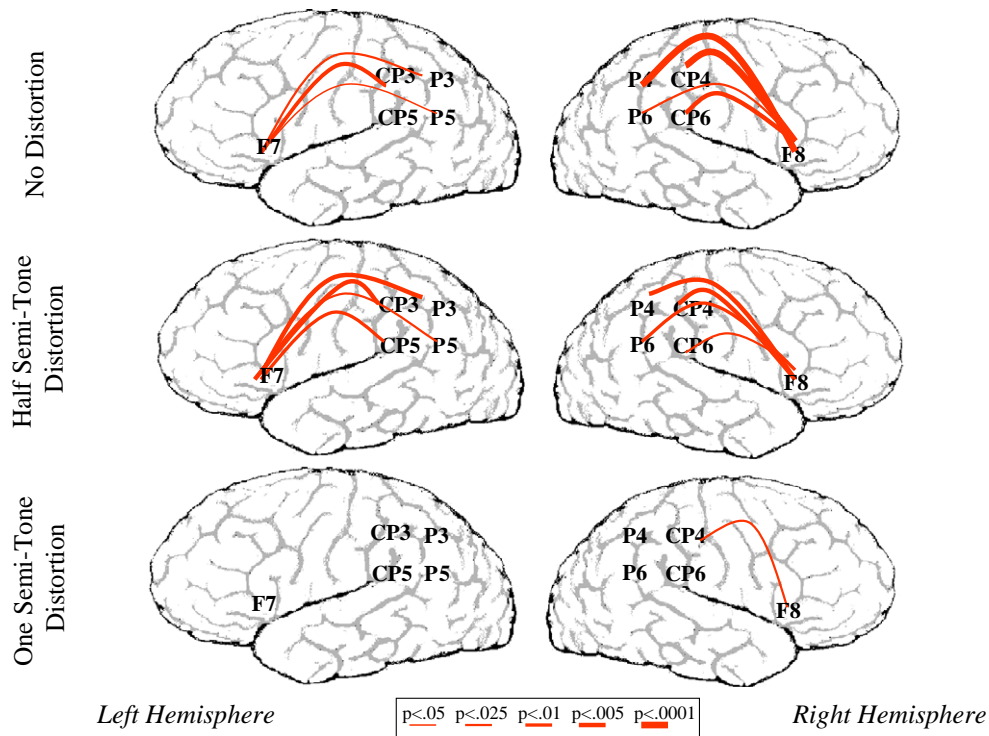


Fig. 2. Probability levels for  $t$ -tests showing greater fronto-temporal Gamma 40 coherence during Talking than during Listening are superimposed on lateral view of right and left hemispheres.

listening. Although not the focus of this paper, we also performed the same analysis on the theta band EEG. As before (Ford et al., 2001b), we found greater coherence during talking than listening ( $p=0.0006$ ) in the theta band, but it was not affected by the level of distortion ( $p=0.28$ ) during talking.

This gamma coherence finding is consistent with the literature on language activation tasks; simple listening of verbal stimuli engages auditory cortex bilaterally with little frontal lobe activation, while talking engages both frontal and temporal lobe structures (Demonet et al., 1992; Warburton et al., 1996). It is worth noting that even on the right, there were significant increases in coherence during talking compared to listening. Although left hemisphere structures are more involved than right in generation of verbal thoughts and speech, the right hemisphere is typically engaged as well (Sperry, 1950; Warburton et al., 1996), especially during repetition of material and listening to repetitive material (Ingvar and Schwartz, 1974; Roland, 1984; Roland et al., 1981). Moreover, lateralization of speech was less likely in the present paradigm because vocalization of the phoneme [a:] does not necessarily depend on semantic language circuitry.

Also, to the best of our knowledge, this is the first report that event-related gamma-band coherence is affected by the match between the *actual* auditory experience and the *expected* experience during talking. gamma-band synchrony may reflect a “binding of expectation with experience” (Singer, 1999). Disruption in this synchrony may signal the frontal lobe regions subserving speech vocalization to implement sensorimotor adaptation routines (Houde and Jordan, 1998) and may facilitate revision of the anticipated auditory consequences of vocalization. In particular, while we did not prolong the distortion blocks beyond 60 trials (about 2 min), if we had, subjects may have made noticeable, but unconscious, adjustments to their speaking to make the experience sound more like the expectation in the short run; in the long run, repeated exposures to distorted auditory feedback may lead to a revision of the anticipated auditory consequences of vocalization, in effect updating the “corollary discharge” signal that prepares the auditory cortex for the impending auditory reafference associated with speaking. Such neuroplastic accommodations to long-term speech alteration may be a natural occurrence during adolescence when vocal cord maturation leads to voice changes. The capacity for the brain to reorganize in this was demonstrated in the visual system where subjects wearing prism goggles learn to adjust to a visually distorted world (Welch, 1978).

In spite of our attempt to inspect each individual trial and exclude trials with muscle artifacts, we must consider the possible contribution of subtle muscle artifacts to our finding of reduced gamma coherence during distorted talking. If both frontal and temporal aspects of the temporalis muscle are similarly involved in making fine articulatory-motor adjustments, synchronous activation of

both parts of the muscle could *increase* coherence. If muscle activity were different at frontal and temporal areas, then its involvement would *decrease* coherence. Because it spans both frontal and temporal areas, we suggest that its involvement might be similar at frontal and temporal sites and its contribution would increase, not decrease, coherence. Nevertheless, muscle artifacts remain a possible explanation for our finding.

Because our sensory world is far richer and more complex than our nervous system can process and appreciate, we need a top-down mechanism to simplify bottom-up sensory processing. This mechanism may incorporate efference copies of the motor commands that predict the sensory consequences (corollary discharge) of the action (Sperry, 1950; Von Holst and Mittelstaedt, 1950). Frontal-temporal gamma synchrony during speech may reflect the successful operation of the efference copy/corollary discharge mechanism that helps us to (1) anticipate the sensory consequences of our own actions, (2) efficiently process what we experience, (3) rapidly detect discrepancies between our experience and our expectations, and (4) implement instantaneous sensorimotor adaptations to bring our experience and expectations into alignment. Thus, a failure in this mechanism could have far-reaching effects including difficulties in language learning and motor awkwardness, and may contribute to symptoms in schizophrenia including auditory hallucinations, delusions, and formal thought disorder (Ford and Mathalon, *in press*).

## Acknowledgements

This work was supported by grants from NARSAD, National Institute of Mental Health (MH40052, MH 58262) and by the Department of Veterans Affairs.

## References

- Angel, R.W., 1976. Efference copy in the control of movement. *Neurology* 26, 1164–1168.
- Creutzfeldt, O., Ojeman, G., Lettich, E., 1989. Neuronal activity in the human lateral temporal lobe: II. Responses to the subject's own voice. *Experimental Brain Research* 77, 476–489.
- Curio, G., Neuloh, G., Numminen, J., Jousmaki, V., Hari, R., 2000. Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping* 9, 183–191.
- Demonet, J.F., Celsis, P., Nespoulous, J.L., Viallard, G., Marc-Vergnes, J.P., Rascol, A., 1992. Cerebral blood flow correlates of word monitoring in sentences: influence of semantic incoherence—a SPECT study in normals. *Neuropsychologia* 30, 1–11.
- Engel, A., Fries, P., Singer, W., 2001. Dynamic predictions: oscillations and synchrony in top-down processing. *Nature Reviews. Neuroscience* 10, 704–716.
- Fernandez, G., Fell, J., Fries, P., 2002. Response: the birth of a memory. *Trends in Neurosciences* 25, 281–282.
- First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 1995. Structured Clinical Interview for DSM-IV Axis I Disorders. Biometrics Research Department, New York State Psychiatric Institute, New York, NY.

- Ford J.M., Mathalon, D.H., in press. Corollary discharge dysfunction in schizophrenia: Can it explain auditory hallucinations? *International Journal of Psychophysiology*.
- Ford, J.M., Mathalon, D.H., Heinks, T., Kalba, S., Roth, W.T., 2001a. Neurophysiological evidence of corollary discharge dysfunction in schizophrenia. *American Journal of Psychiatry* 158, 2069–2071.
- Ford, J.M., Mathalon, D.H., Kalba, S., Whitfield, S., Faustman, W.O., Roth, W.T., 2001b. Cortical responsiveness during talking and listening in schizophrenia: an event-related brain potential study. *Biological Psychiatry* 50, 540–549.
- Ford, J.M., Mathalon, D.H., Whitfield, S., Faustman, W.O., Roth, W.T., 2002. Reduced communication between frontal and temporal lobes during talking in schizophrenia. *Biological Psychiatry* 51 (6), 485–492 (Mar 15).
- Fries, P., Reynolds, J., Rorie, A., Desimone, R., 2001. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291, 1560–1563.
- Gratton, G., Coles, M.G.H., Donchin, E., 1983. A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology* 55, 468–484.
- Gray, C., Singer, W., 1989. Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization with reflects global stimulus properties. *Nature* 338, 334–337.
- Heinks-Maldonado, T.H., Mathalon, D.H., Gray, M., Ford, J.M., 2005. Fine-tuning of auditory cortex during speech production. *Psychophysiology* 42 (2), 180–190 (Mar).
- Houde, J.F., Jordan, M.I., 1998. Sensorimotor adaptation in speech production. *Science* 279, 1213–1216.
- Houde, J.F., Nagarajan, S.S., Sekihara, K., Merzenich, M.M., 2002. Modulation of auditory cortex during speech: an MEG study. *Journal of Cognitive Neuroscience* 14 (8), 1125–1138 (Nov 15).
- Ingvar, D.H., Schwartz, M.S., 1974. Blood flow patterns induced in the dominant hemisphere by speech and reading. *Brain* 97, 273–288.
- Jeannerod, M., 2003. The mechanism of self-recognition in humans. *Behavioural Brain Research* 142, 1–15.
- Kwon, J., O'Donnell, B., Wallenstein, G., Greene, R., Hirayasu, Y., Nestor, P.G., ME, H., GF, P., ME, S., McCarley, R., 1999. Gamma frequency-range abnormalities to auditory stimulation in schizophrenia. *Archives of General Psychiatry* 56, 1001–1005.
- Muller-Preuss, P., Ploog, D., 1981. Inhibition of auditory cortical neurons during phonation. *Brain Research* 215, 61–76.
- NeuroscanLabs, 1999. Edit 4.1 Offline Analysis of Acquired Data: Coherence, SCAN™ User Guide Volume II. Neurosoft, Inc., Sterling, VA, pp. 128–129.
- Poulet, J., Hedwig, B., 2002. A corollary discharge maintains auditory sensitivity during sound production. *Nature* 418, 872–876.
- Roland, P.E., 1984. Organization of motor control by the normal human brain. *Human Neurobiology* 2, 205–216.
- Roland, P.R., Skinhoj, E., Lassen, N.A., 1981. Focal activations of human cerebral cortex during auditory discrimination. *Journal of Neurophysiology* 45, 1139–1151.
- Shuster, L., Durrant, J., 2003. Toward a better understanding of the perception of self-produced speech. *Journal of Communication Disorders* 36, 1–11.
- Singer, W., 1999. Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24, 49–65.
- Sperry, R.W., 1950. Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology* 43, 482–489.
- Von Holst, E., Mittelstaedt, H., 1950. Das reafferenzprinzip. *Naturwissenschaften* 37, 464–476.
- Warburton, E., Wise, R.J.S., Price, C.J., Weiller, C., Hadar, U., Ramsay, S., Frackowiak, R.S.J., 1996. Noun and verb retrieval by normal subjects. studies with PET. *Brain* 119, 159–179.
- Welch, R.B., 1978. *Perceptual Modification: Adapting to Altered Sensory Environments*. Academic Press, New York.
- Wolpert, D., Kawato, M., 1988. Multiple paired forward and inverse models for motor control. *Neural Networks* 11, 1317–1329.