

Dissecting corollary discharge dysfunction in schizophrenia

JUDITH M. FORD,^{a,b} MAX GRAY,^c WILLIAM O. FAUSTMAN,^{c,d} BRIAN J. ROACH,^{a,b} AND DANIEL H. MATHALON^{a,b}

^aPsychiatry Service, Veterans Affairs Connecticut Healthcare System, West Haven, Connecticut, USA

^bDepartment of Psychiatry, Yale University School of Medicine, New Haven, Connecticut, USA

^cPsychiatry Service, Veterans Affairs Palo Alto Health Care System, Palo Alto, California, USA

^dDepartment of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California, USA

Abstract

During talking, a corollary discharge prepares cortex for self-generated sounds, minimizing responsiveness and providing a way to recognize sounds as self-generated. When we talk, we are the agent producing the sound and know what sound to expect. The auditory ERP N1 is normally suppressed during talking, but less so in schizophrenia, perhaps due to deficits in agency and expectancy inherent to talking. N1 was assessed in 27 patients (23 schizophrenia, 4 schizoaffective) and 26 controls. During talking, subjects said “ah” every 1–2 s. During agency, subjects pressed a button to deliver “ah” every 1–2 s. During expectancy, “ah” followed a visual warning. Talking yielded greatest N1 suppression in controls and greatest suppression failure in patients. Agency and expectancy had modest suppression effects on N1 and only in controls. Group differences in expectancy and agency could not account for failed corollary discharge during talking in patients.

Descriptors: Schizophrenia, Corollary discharge, Talking, Expectancy, Agency, ERP

Self-monitoring is a fundamental element of normal cognitive and motor functioning. It allows us to modify our actions online, as they are being planned and executed. Helmholtz (1925) first described the need for a mechanism that would allow us to discriminate between moving objects and movements on the retina resulting from eye movements. Von Holst and Mittelstaedt (1950) and Sperry (1950) later suggested that a motor action is accompanied by an efference copy of the action that produces a corollary discharge in sensory cortex. In the visual system, a corollary discharge may serve to stabilize the visual image during eye movements, maintaining visuo-spatial constancy. In the somatosensory system, it may explain why we cannot tickle ourselves (Blakemore, Wolpert, & Frith, 1998). It has been suggested that an efference copy of a planned action is sent through a corollary discharge mechanism to the appropriate sensory cortex, preparing it for the arrival of the impending sensation. In its simplest form, the forward model works to suppress or dampen

sensation resulting from a self-generated action. Thus, in addition to serving as a mechanism for learning and fine-tuning our actions, the forward model may allow an automatic distinction between internally and externally generated sensory percepts.

During talking, forward models from speech production and vocalization regions in the frontal lobes may prepare the auditory cortex for imminent self-generated speech sounds, minimizing the auditory cortical response to these sounds and providing a mechanism for recognizing these sounds as self-generated. Support for this mechanism comes from studies in which recordings were made during a presurgical planning procedure from the exposed surface of the right and left temporal cortices while patients talked and listened to others talking (Creutzfeldt, Ojeman, & Lettich, 1989). During overt talking, suppression of ongoing activity in approximately one-third of the middle temporal gyrus neurons was observed. This is consistent with noninvasive, human electrophysiological studies employing electroencephalographic (EEG) or magnetoencephalographic (MEG) potentials synchronized to the onset of vocalization (Curio, Neuloh, Numminen, Jousmaki, & Hari, 2000; Ford, Mathalon, Heinks, Kalba, & Roth, 2001; Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Houde, Nagarajan, Sekihara, & Merzenich, 2002) or to sounds presented during vocalization (Ford, Mathalon, Kalba, et al., 2001). In these studies, the N1 of the EEG-based event-related potential (ERP), or the M100 of the MEG-based response to the sound, was smaller during speaking than during passive listening. This suppression was not seen to the same extent when the voice was pitch-shifted down during talking

This work was supported by the VA Schizophrenia Biological Research Center and grants from National Institute of Mental Health (MH40052, MH 58262, MH067967) and the National Alliance for Research in Schizophrenia and Affective Disorders (NARSAD). We thank Drs. Alison Adcock and Sophia Vinogradov for referring patients to this study.

Address reprint requests to: Judith M. Ford, Ph.D., Department of Psychiatry, Yale University School of Medicine, 950 Campbell Ave., VA CT Healthcare System, 116A, West Haven, CT 06517, USA. E-mail: Judith.ford@yale.edu

(Heinks-Maldonado et al., 2005; Houde et al., 2002) or when the spoken sound was replaced by another sound during talking (Heinks-Maldonado et al., 2005; Houde et al., 2002).

Patients with schizophrenia show less reduction in N1 during talking, suggesting that this mechanism of auditory cortical dampening is dysfunctional in schizophrenia (Ford, Mathalon, Heinks, et al., 2001; Ford, Mathalon, Kalba, et al., 2001). These findings are consistent with suggestions that failures of the corollary discharge mechanism characterize schizophrenia, perhaps contributing to self-monitoring deficits and auditory verbal hallucinations (Feinberg, 1978). Specifically, if a forward model of a thought, memory, or other internally generated percept does not produce a corollary discharge representation in the auditory cortex, these internal events may be experienced as having an external source. However, although N1 suppression failures are sometimes related to auditory hallucinations (Heinks-Maldonado, Mathalon, Houde, Gray, & Ford, 2007), it is not always true (Ford, Mathalon, Heinks, et al., 2001).

There is a small research literature indicating that advance warning that a particular stimulus is about to be delivered is enough to dampen sensory cortical response to that stimulus, even though the stimulus does not result *directly* from a motor act, as it does during talking. In fact, in a classic self-stimulation paradigm, when a sound results from *indirect* motor acts of pushing a button, it elicits a smaller N1 than when that sound is passively heard (Martikainen, Kaneko, & Hari, 2005; McCarthy & Donchin, 1976; Schafer & Marcus, 1973). This type of finding supports a forward model that transmits information about one's agency in bringing about sensory stimulation. Thus, there may be a forward model signal that prepares sensory cortex for sensations that result from self-initiated acts even when the causal chain between the act and its sensory consequence is indirect and dependent on external technologies such as button boxes, recorded sounds, and headphones. In self-stimulation or agency paradigms, simple expectancy could also contribute to sensory suppression; however, to the extent that expectancies heighten attention, they may actually enhance, rather than suppress, sensation (Hillyard, Hink, Schwent, & Picton, 1973). Perhaps because of these possible countervailing forces, expectancies generated by presentation of warning stimuli signaling the impending delivery of a sound appear to have little or no effect on

N1 to the sound (Scaife, Groves, Langley, Bradshaw, & Szabadi, 2006).

To explore whether predictive information resulting from expectancy (i.e., warning) and from agency (i.e., self-stimulation) modulates auditory cortex in the same way and to the same degree as talking and whether these modulatory effects hypothesized to operate in healthy controls are deficient in patients with schizophrenia, we performed three experiments. First, we recorded N1 to the onset of a speech sound as it was being spoken, as we have done before (Ford, Mathalon, Heinks, et al., 2001; Heinks-Maldonado et al., 2007). Second, we recorded N1 to the onset of the subject's own previously recorded speech sound when it was delivered via a self-paced button press. Third, we recorded N1 to the onset of the subject's prerecorded speech sound when its impending delivery was heralded by a count-down visual warning.

Methods and Materials

Participants

EEG data were acquired from patients with schizophrenia ($n = 23$) or schizoaffective disorder ($n = 4$) and age-matched healthy comparison subjects ($n = 26$). All gave written informed consent after procedures had been fully described. Demographic and clinical data are summarized in Table 1 for the subjects participating in each of the three experiments. This study was approved by the Human Subjects Committees at the Palo Alto VA Healthcare System and Stanford University.

Patients were recruited from community mental health centers as well as from inpatient and outpatient services of the Veterans Affairs Healthcare System in Palo Alto and San Francisco, California. All patients were on stable, therapeutic doses of antipsychotic medications and met *DSM-IV* (American Psychiatric Association, 1994) criteria for schizophrenia or schizoaffective disorder based on a Structured Clinical Interview for *DSM-IV* (SCID; First et al., 1995) conducted by a clinical psychologist ($n = 24$) or a SCID conducted by a clinically trained research assistant followed by a clinical interview with a clinical psychologist ($n = 2$). Patients were excluded if they met *DSM-IV* criteria for alcohol or drug abuse within 30 days prior to study. In

Table 1. Demographics of Populations Studied

Variable	Normal control subjects ($n = 26$)				Patients with schizophrenia ($n = 27$)			
	Mean	<i>SD</i>	Min	Max	Mean	<i>SD</i>	Min	Max
Age (years) ^a	42.23	10.57	25	72	43.00	10.38	21.0	67.0
Education (years) ^b	15.42	1.63	12.0	19.0	14.04	2.01	11.0	18.0
Parental socioeconomic status ^c	37.80	15.82	9.5	77.0	37.28	10.74	16.5	58.0
Total BPRS ^d					34.75	7.43	20.00	47.50
Handedness	26 right-handed				25 right-handed, 2 left-handed			
Gender	19 men, 7 women				23 men, 4 women			
Diagnosis					12 undifferentiated schizophrenia 9 paranoid schizophrenia 4 schizo-affective 2 residual			
Medication type					22 atypical, 3 typical, 2 both			
Medication dose in chlorpromazine equivalents ^e					665.5 mg/day			

^aControls versus patients, $p = .79$.

^bControls versus patients, $p = .008$.

^cControls versus patients, $p = .89$.

^dScoring on BPRS: 1 = *not present*, 7 = *extremely severe*.

^eCalculations were based on Woods (2003).

addition, patient and control participants were excluded for significant head injury, neurological disorders, or other medical illnesses compromising the central nervous system. Symptom ratings were derived from the mean of two independent raters attending the same rating session, using the 18-item Brief Psychiatric Rating Scale (BPRS; Faustman & Overall, 1999). We have reported N1 data to the onset of speech during talking using two other paradigms, in which 4 (Ford, Mathalon, Heinks, et al., 2001) and 21 (Ford, Roach, Faustman, & Mathalon, 2007) of these patients were included. Talking was self-paced in the former and cued in the latter.

Comparison subjects were recruited by newspaper advertisements and word of mouth, screened by telephone using questions from the SCID (First et al., 1995) nonpatient screening module, and excluded for any history of Axis I psychiatric illness.

Experiments

Subjects participated in three experiments, presented in a random order.

Talking experiment (say “ah,” hear “ah”). In the talking condition, subjects were told to vocalize [a:] about every 1–2 s and were stopped after 60 trials. Depending on the condition, sounds subjects heard over headphones could be unaltered in pitch or pitch-shifted downward, in a blocked design. Only ERPs to the unaltered sounds are presented here. During listening, recorded sounds from the talking condition were played back, and subjects were instructed simply to listen. Loudness was the same in the talking and listening conditions based on equilibration of the headphone audio output across conditions as measured by a dB meter. It is worth noting that although the same sequence of sounds was heard during talking and listening, each individual sound during both is unique.

During talking, the speech signal was picked up by a microphone and sent through a preamplifier to a computer equipped with sound processing software and hardware. It was split so it could both generate a trigger pulse and be amplified and heard in real time via headphones. The trigger pulse was generated on the rising edge of the rectified, low-pass filtered signal and inserted in the EEG data collection system.

The median speech signal duration was longer in patients (250 ms) than in controls (215 ms; $p < .05$). There were no group differences in median speech signal intensity ($p = .83$).

Agency experiment (press a button, hear an “ah”). To assess the effect of agency on N1 suppression, subjects caused speech sounds to be delivered by pressing a button. One sample of speech recorded at the beginning of the session was used as the auditory stimulus (see below for description of how this sound was generated). The speech sound was heard instantaneously following a self-paced button press (zero delay). The temporal sequence of sounds generated by the subject was preserved for playback, as done by others (Martikainen et al., 2005). Although the same “ah” (recorded before the start of the experiment) was heard on each trial, the interval between each was unique. In addition, we asked subjects to press a button at approximately the same pace, and no sound was delivered; the brain response associated with simple button pressing was subtracted from the self-stimulation condition, as done by others (Martikainen et al., 2005). In two other blocks in this experiment not reported here, self-stimulation resulted in the delivery of the sound with a 50-ms or 100-ms delay. The order of the three blocks was randomized across subjects.

Expectancy experiment (see a warning, hear an “ah”). To assess the effect of expectancy on N1 suppression, a visual warning prepared subjects for the impending speech sound. The same sample of speech used in the agency experiment was used as the auditory stimulus. During the warned condition where expectancies were generated, subjects were warned that a sound would occur by the appearance of three dots on a screen, which changed to two dots and then one dot. The speech sound was heard at the moment the single (last) dot appeared. There were 500 ms between the three- and two-dot displays, but the delay between the two-dot and one-dot display varied by about 30 ms, depending on system performance. This difference was not noticeable. During the unwarned condition, the sounds were played back without visual warning to the subjects in the exact sequence as the warned sounds. To remove the brain response associated with the last dot from the response to the warned “ah,” subjects saw visual warnings alone. Because a trigger pulse occurred with each dot, we were able to remove the ERP to the last dot during visual warnings alone from the ERP to the “ah” during the warned condition. The sound-to-sound interval was jittered between 2.5 and 3 s. Although subjects were told to fixate on the visual warning cues, there was no control for fixation.

In summary, ERPs to “ah” were recorded in these three similar experiments. The brain responses associated with pressing the button during the agency experiment and with the visual warning during the expectancy experiment were subtracted from responses associated with hearing the “ah.” There were no parallel brain responses to subtract in the talking experiment, as it was impossible to know what internal cue subjects used to initiate speaking. It is worth mentioning that the three experiments are not exactly parallel: In the talking experiment, a unique “ah” was heard on each trial according to each subject’s unique articulation of each “ah”; in the agency and expectancy experiments, the same prerecorded “ah” was heard throughout both; in the agency experiment, subjects individually determined the time between “ah”s; in the expectancy experiment, inter-“ah” intervals were predetermined.

Acoustic Calibration and Standard Stimulus Generation

Subjects were trained to say short (<300 ms duration), sharp “ah”s, which were heard through headphones, while the experimenter gave feedback on performance. Subjects saying “ah”s louder than 85 dB SPL (sound pressure level) were encouraged to speak more softly. After training, the average “ah” intensity was measured with a Quest Electronics, model 215 Sound Level Meter, set on C scale and fast mode, and held ~6 cm in front of subject’s mouth.

A series of “ah”s was recorded. Using CoolEditPro, a sample “ah” (250–350 ms) was selected and edited to eliminate background noise before and after the “ah” and to ensure a clean sharp onset. It was saved as a Windows PCM .WAV file. This was used as the “ah” for both the expectancy and agency experiments. The intensity of the sample “ah” did not differ in controls and patients ($p = .16$), 75 dB SPL ($SD = 5.9$) and 77 dB SPL ($SD = 5.3$), respectively.

For all three experiments, the sound was amplified 15 dB SPL during both conditions of each experiment, to mask bone conduction that occurs during talking.

Data Acquisition and Preprocessing

EEG data were acquired with Neuroscan SynAmps (0.05–100 Hz band pass filter, 1000 Hz analog-to-digital conversion rate,

12 dB/octave roll-off) from 42 sites referenced to the nose. During preprocessing, data were re-referenced to the mastoid electrodes (TP9 and TP10), to minimize talking artifacts from the nose. Additional electrodes were placed on the outer canthi of both eyes and above and below the right eye to measure eye movements and blinks (vertical and horizontal electrooculogram; VEOG, HEOG). EEG data were separated into 500-ms epochs time-locked to onset of the speech sound, with 100 ms before speech onset and 400 ms after. The 100-ms period preceding speech onset was used for baseline correction after VEOG and HEOG data were used to correct EEG for eye movements and blinks in a regression-based algorithm (Gratton, Coles, & Donchin, 1983). Trials containing artifacts (voltages exceeding $\pm 50 \mu\text{V}$) were rejected. Subjects with fewer than 20 trials in an average were excluded from the analysis. Final numbers of subjects remaining in each experiment are as follows: talking experiment, controls = 24, patients = 22; agency experiment, controls = 25, patients = 21; expectancy experiment, controls = 26, patients = 25.

ERP Analysis

Prior to identification of N1, data were bandpass filtered between 0.5 and 15 Hz to optimize measurement of N1. N1 was identified as the most negative point between 50 and 175 ms after “ah” onset. The voltage at that point was measured relative to a pre-“ah” baseline (-100 to 0 ms). N1 was measured in the subtracted waveforms for the agency (Martikainen et al., 2005) and expectancy experiments, as described above.

Statistical Analysis

Analyses of variance. For each experiment, N1 peak amplitudes and latencies from 35 sites were assessed in a four-way ANOVA for the between-subjects factor of Group and the within-subjects factors of Laterality (far, far left [F7, FT7, T3, TP7, T5], far left [F5, FC5, C5, CP5, P5], left [F3, FC3, C3, CP3, P3], midline [Fz, FCz, Cz, CPz, Pz], right [F4, FC4, C4, CP4, P4], far right [F6, FC6, C6, CP6, P6], far, far right [F8, FT8, T4, TP8, T6]), Anterior-Posterior (A-P) (frontal [F7, F5, F3, Fz, F4, F6, F8], frontal-central [FT7, FC5, FC3, FCz, FC4, FC6, FT8], central [T3, C5, C3, Cz, C4, C6, T4], central-parietal [TP7, CP5, CP3, CPz, CP4, CP6, TP8], parietal [T5, P5, P3, Pz, P4, P6, T6]), and Condition (talk vs. listen for the talking experiment; self-stimulation vs. playback for the agency experiment; warning vs. no warning for the expectancy experiment). Greenhouse–Geisser corrections were used for nonsphericity.

Bivariate and multivariate analyses. For these analyses, N1 suppression effects for each of the three experimental conditions were calculated. To calculate the N1 suppression effect due to talking, we subtracted N1 at Cz during talking from the N1 at Cz during listening. To calculate N1 suppression due to expectancy, we subtracted N1 at Cz with the warning from N1 at Cz without the warning. To calculate N1 suppression due to agency, we subtracted N1 at Cz with self-stimulation from N1 at Cz during playback.

The correlations of N1 suppression from each of the three experiments with the BPRS auditory hallucinations item were assessed in the patients using Pearson product–moment correlation coefficients with alpha set to $p < .05$, two-tailed (because of our a priori hypotheses about these correlations). The other 17 items on the BPRS were also correlated with N1 suppression measures from each of the three experiments, but because these

analyses were exploratory, we used a more conservative Bonferroni-corrected alpha level of $p < .05/51$, or $p < .001$.

Results

Only effects involving Condition and Group are discussed. Exact probability levels are given, except when $p = .0000$, in which case it is listed as $p < .00001$.

Talking Experiment

There was a main effect of Condition, $F(1,44) = 15.46$, $p < .0001$, with N1 amplitude being reduced during talking compared to listening. There was a Condition \times Group interaction, $F(1,44) = 10.41$, $p = .002$, which we parsed in two ways. First, we tested the talk/listen effect separately in controls and patients, and found it was significant in controls, $F(1,23) = 35.71$, $p < .0001$, but not in patients, $F(1,21) = 0.19$, $p = .67$. Second, we tested the Group effect separately for talk and listen, and found that patients had a larger N1 during talking than did healthy controls, $F(1,44) = 9.69$, $p = .003$, with no Group difference in N1 during listening, $F(1,44) = 0.42$, $p = .52$. Figure 1 shows ERPs at midline sites, overlaid for talking and listening.

There was a Condition \times Laterality \times Group interaction, $F(6,264) = 2.46$, $p = .05$, which was parsed two ways. In one, we inspected the Group \times Condition interaction at each lateral strip. It was significant at the five central lateral strips. The Condition effect at each of these five strips was then inspected separately for each group. N1 during talking was smaller than during listening for controls at these five strips of electrodes (far left: $F[1,23] = 17.22$, $p < .0001$; left: $F[1,23] = 48.68$, $p < .0001$; midline: $F[1,23] = 56.06$, $p < .0001$; right: $F[1,23] = 37.56$, $p < .00001$; far right: $F[1,23] = 19.26$, $p < .0001$), but not for patients (far left: $F[1,21] = 0.02$, $p = .88$; left: $F[1,21] = 0.14$, $p = .72$; midline: $F[1,21] = 1.48$, $p = .24$; right: $F[1,21] = 1.75$, $p = .20$; far right: $F[1,21] = 0.17$, $p = .68$). In the other parsing, we inspected the Condition \times Laterality interaction separately in the controls and patients. As can be seen in Figure 2, this interaction was not significant in patients, $F(6,126) = 1.71$, $p = .18$, but was in controls, $F(6,138) = 13.28$, $p = .0001$. The Condition effect in controls was significant at all lateral strips of electrodes, but was stronger closer to midline sites (effect sizes ranging from .71 to .20).

Talking significantly delayed N1 by about 10–15 ms, $F(1,44) = 52.10$, $p < .0001$. Otherwise, N1 latency was not affected by Group, $F(1,44) = 3.03$, $p = .09$, or Condition \times Group, $F(1,44) = .63$, $p = .43$.

Agency Experiment

Although neither the main effect of Condition, $F(1,44) = 2.38$, $p = .13$, nor the interaction of Group \times Condition, $F(1,44) = 1.92$, $p = .17$, were significant, there was a Group \times Condition \times AP interaction, $F(4,176) = 5.97$, $p = .01$. The Condition \times AP interaction was not significant for either group. Because the literature and the data shown in Figure 2 suggest suppression of N1 with self-stimulation at the midline strip in controls, we assessed it at midline sites for each group separately. In controls, the effects of agency at each site (Fz, FCz, Cz, CPz, and Pz), respectively, were $F(1,24) = 0.30$, $p = .59$; $F(1,24) = 1.19$, $p = .29$; $F(1,24) = 2.54$, $p = .12$; $F(1,24) = 3.07$, $p = .09$; $F(1,24) = 3.29$, $p = .08$. With a one-tailed test, N1 reduction would be a significant for controls ($p = .045$) at CPz. In patients, the effects of agency were $F(1,20) = 0.92$, $p = .35$; $F(1,20) = 0.59$, $p = .45$; $F(1,20) = 0.83$,

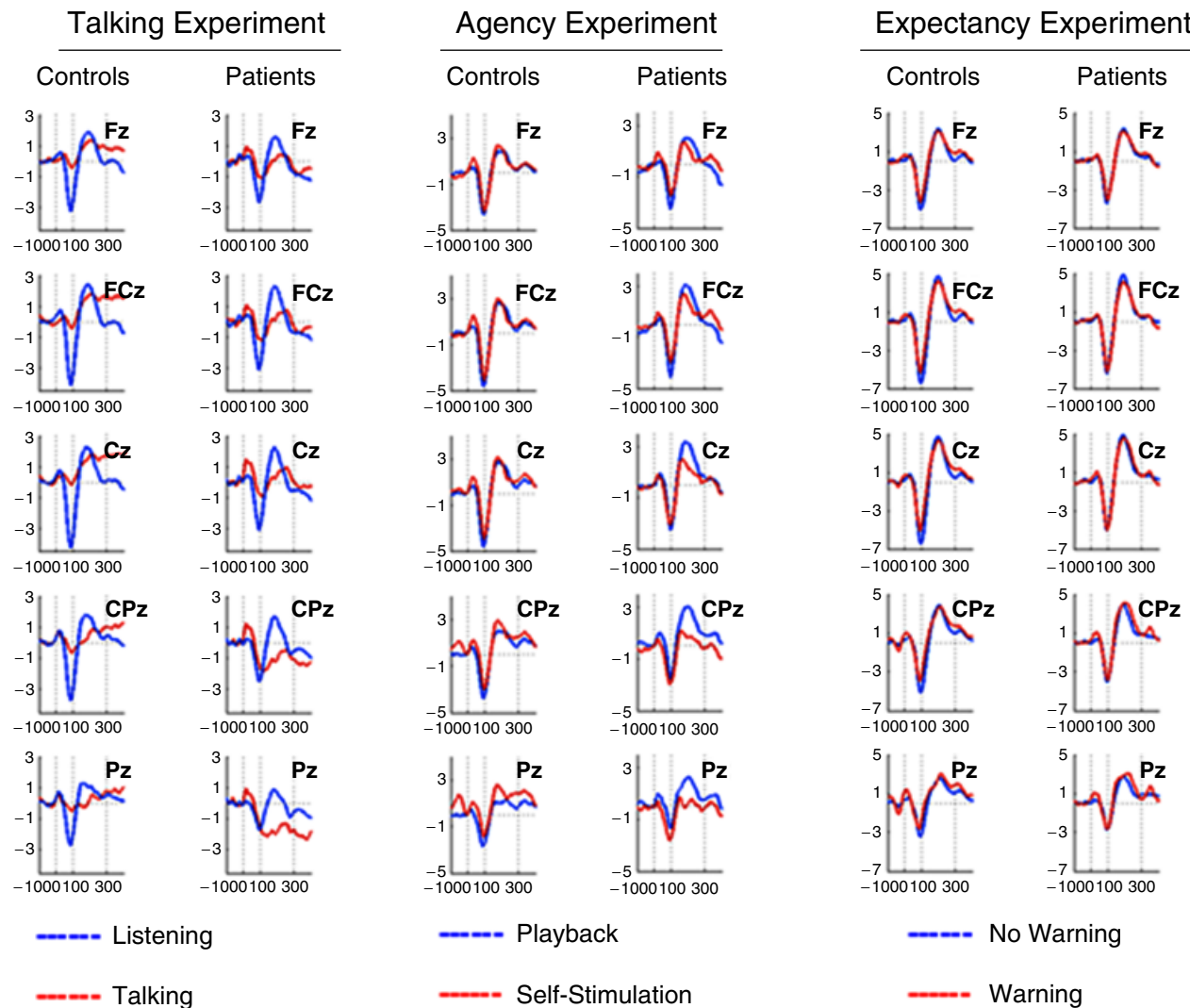


Figure 1. ERPs from Fz, FCz, Cz, CPz, and Pz, locked to speech sound onset (0 ms) are shown for each of the three experiments: talking, expectancy, and agency. ERPs elicited during the two conditions for each experiment are overlaid. In each panel, data from controls appear on the left and those from patients on the right. Amplitude (in microvolts) is on the y-axis and time (in milliseconds) is on the x-axis. Negativity is plotted down.

$p = .37$; $F(1,20) = 3.91$, $p = .06$; $F(1,20) = 4.99$, $p = .037$. Although significant or nearly significantly, the effects of agency CPz and Pz were the opposite of those seen in controls.

N1 latency was not affected by Group, $F(1,44) = 0.31$, $p = .58$, or Condition, $F(1,44) = 2.62$, $p = .11$, or Group \times Condition, $F(1,44) = 0.79$, $p = .38$.

Expectancy Experiment

The main effect of Condition was not significant, $F(1,49) = 3.32$, $p = .07$, but there was a trend for N1 to be suppressed with warning. There was a Group \times Condition \times Laterality \times AP interaction, $F(24,1176) = 22.39$, $p = .03$. This was parsed by looking for the Condition \times Laterality \times AP for each Group separately, but it was not significant for controls, $F(24,600) = 1.31$, $p = .22$, or patients, $F(24,576) = 2.19$, $p = .07$. Nevertheless, because warning appeared to suppress N1 at the midline sites in controls (Figure 2), we assessed it for each group separately. In controls, the effects of expectancy at each midline site (Fz, FCz, Cz, CPz, and Pz), respectively, were $F(1,26) = 5.72$, $p = .024$; $F(1,26) = 7.93$, $p = .009$; $F(1,26) = 13.14$, $p = .001$; $F(1,26) = 7.60$, $p = .01$;

$F(1,26) = 1.22$, $p = .28$. In patients, the effects of expectancy were $F(1,23) = 0.03$, $p = .87$; $F(1,23) = 0.00$, $p = .99$; $F(1,23) = 0.02$, $p = .89$; $F(1,23) = 0.05$, $p = .82$; $F(1,23) = 0.11$, $p = .78$.

N1 latency was not affected by Group, $F(1,48) = 0.67$, $p = .42$, or Condition, $F(1,48) = 0.63$, $p = .43$, or Group \times Condition, $F(1,48) = 0.85$, $p = .36$.

Contributions of Expectancy and Agency to N1 Suppression during Talking

Although the experiments are not perfectly parallel, we did compare them statistically. To assess contributions of expectancy and agency to suppression effects produced by talking, we first assessed the bivariate correlations between them. Next, to examine the independent effects of expectancy and agency on the talking effect, the talking effect was regressed on both expectancy and agency effects using a multiple regression analysis. The weakness of the agency and expectancy effects, both within and between groups, does not preclude the possibility that some subjects showed the effect more strongly than others and that these individual differences in agency and expectancy suppression effects

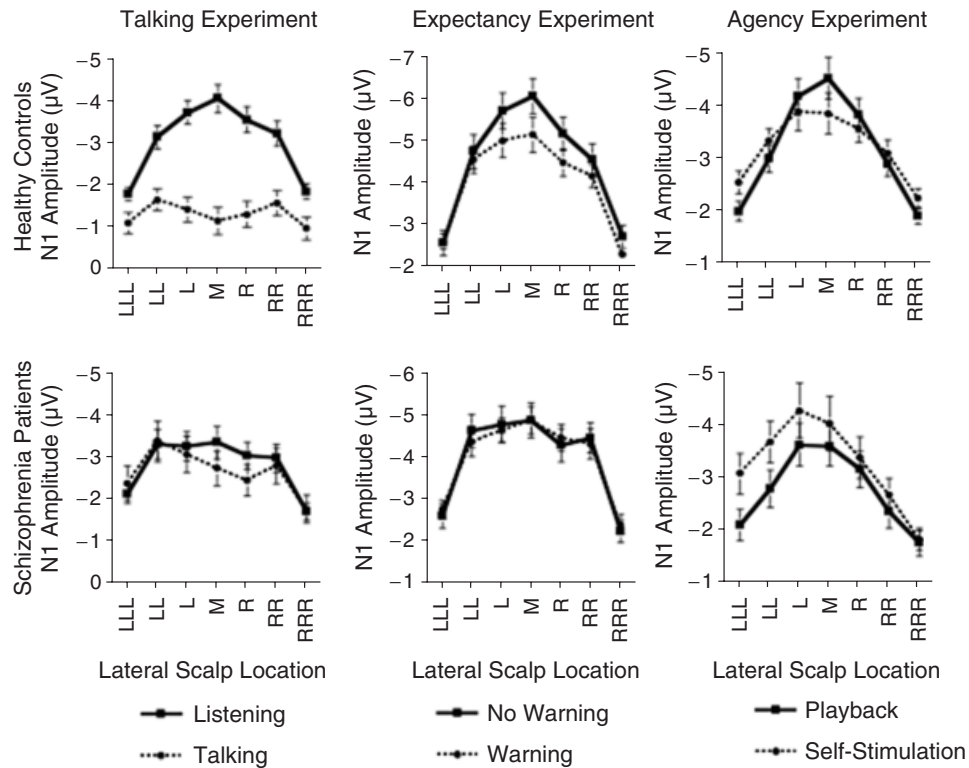


Figure 2. Mean N1 amplitudes collapsed across each lateral strip of electrodes are plotted for each of the three experiments: talking, expectancy, and agency. Data from both conditions of each experiment are overlaid. Data from healthy controls appear at the top and those from patients at the bottom. LLL: far, far left (F7, FT7, T3, TP7, T5); LL: far left (F5, FC5, C5, CP5, P5); L: left (F3, FC3, C3, CP3, P3); M: midline (Fz, FCz, Cz, CPz, Pz); R: right (F4, FC4, C4, CP4, P4); RR: far right (F6, FC6, C6, CP6, P6); RRR: far, far right (F8, FT8, T4, TP8, T6).

could significantly correlate with the individual differences in the talk–listen N1 suppression effects.

For the 23 controls who had data from all experiments, N1 suppression due to talking was significantly related to N1 suppression due to expectancy (visual warning; $r = .57$, $p = .004$). Examination of the regression equation for this correlation shows that when N1 suppression due to expectancy is 0, the predicted N1 suppression effect due to talking (i.e., the intercept) is $-2.94 \mu\text{V}$. Thus, the mean N1 suppression effect due to talking of $-3.79 \mu\text{V}$ was reduced by only $0.85 \mu\text{V}$ after taking into account its correlation with the N1 suppression due to expectancy. N1 suppression due to talking was not related to N1 suppression due to agency (self stimulation; $r = .199$, $p = .36$). Finally, the multiple regression analysis showed that N1 suppression due to expectancy was significantly related to the N1 suppression effect due to talking (beta = $.66$, $p = .005$), independent of the effect of agency (beta = $.14$, $p = .37$).

For the 19 patients who had data from all experiments, N1 suppression due to talking was not significantly related to N1 suppression due to agency (self-stimulation; $r = .34$, $p = .15$) or expectancy ($r = .09$, $p = .71$).

Contributions of Expectancy and Agency Effects to Group Difference in N1 Suppression Due to Talking

Group differences in N1 suppression effect due to talking were assessed in ANCOVA models covarying for N1 suppression effects of expectancy and agency. The Group \times Covariate interactions were deleted from the model after demonstrating that

they were not significant, satisfying the assumption of homogeneous slopes for the ANCOVA. Expectancy ($p = .03$) showed significant relationships with N1 suppression due to talking, and agency ($p = .10$) showed trend level relationships. Most importantly, the Group effect remained significant in both the expectancy ANCOVA ($p = .02$) and the agency ANCOVA ($p = .008$), even when these relationships were taken into account. Thus, group differences in expectancy and agency could not account for patients' reduced suppression of N1 during talking.

Correlations with Clinical Symptoms

Correlations between auditory hallucinations and N1 suppression were not significant for the talking experiment ($r = -.16$, $p = .49$), expectancy experiment ($r = .18$, $p = .38$), or agency experiment ($r = -.08$, $p = .72$). No correlations between N1 suppression and the other 17 BPRS items were significant at the Bonferroni corrected level of $p < .001$.

Discussion

The goal in this series of experiments was to confirm earlier reports of N1 suppression during talking in a larger group of subjects (Ford, Mathalon, Heinks, et al., 2001) and to explore the contributions of expectancy and agency to this suppression. In brief, the earlier work was confirmed and demonstrated that talking produces more suppression of cortical responses to speech sounds than either simple expectancy effects based on visual warning, or simple agency effects based on self-delivery, of those sounds.

As in the prior studies, subjects uttered “ah” every 1–2 s and found suppression of N1 to speech sounds during talking compared to playback (Ford, Mathalon, Heinks, et al., 2001). As before, there was significantly less N1 suppression in patients than in controls (Ford, Mathalon, Heinks, et al., 2001). Also, with this *larger* sample, patients had significantly *larger* N1 amplitudes during talking than did the controls. This is counter to a growing literature showing smaller N1 amplitudes in patients listening to sequences of tones presented at these interstimulus intervals (e.g., Boutros et al., 1997; Ford et al., 1994; Roth, Horvath, Pfefferbaum, & Kopell, 1980; Shelley, Silipo, & Javitt, 1999).

To explore the effects of agency, subjects pressed a button to deliver the “ah.” Others (Martikainen et al., 2005; McCarthy & Donchin, 1976; Schafer & Marcus, 1973) have shown that this suppresses N1 to sounds in healthy controls, but this effect was only significant with a one-tailed test, and only at CPz, in the current study. Clearly, the effects of this manipulation, in the current study, were weak in every way: It did not show strong effects in either group, nor did it show strong interactions with group. There were several small, but perhaps important, differences between the other studies and this one. In this study, subjects were asked to press the button every 1–2 s, a somewhat faster rate than used by others. Because the subjects in the current study were age matched to chronic schizophrenia patients, they were older than subjects used in the other studies. Also, in the other studies, the button press produced short duration, pure tones, whereas the button press in the current study produced a longer duration, personal speech sound.

To explore expectancy, subjects were warned about the impending arrival of the “ah” using a visual count-down cue. With this foreknowledge, N1 was modestly reduced at midline sites in controls, but not in patients. The need to switch attention from the visual to the auditory modality in this paradigm may have attenuated the effect of warning on N1 amplitude in controls, having even greater impact in the patients (Mannuzza, 1980). Although an attempt was made to remove the response to the visual cue from the response to the “ah,” the multisensory nature of the stimulus is likely to be more than the sum of its parts (Foxy et al., 2002; Molholm et al., 2002; Senkowski et al., 2007). Although it is difficult to estimate the differential effects in the two groups, deficits in multisensory integration are well documented in schizophrenia (de Gelder, Vroomen, Annen, Masthof, & Hodiament, 2003).

Why does talking suppress auditory cortex responsiveness more than self-stimulation or warning? Is it due to knowledge of content, to knowledge of timing, or to the natural hard-wired connections between the act and the resulting sensation? There were subtle but perhaps important differences in the content of the sound across the three experiments. Although the “ah” in both the expectancy (visual warning) and the agency (self-stimulation) experiments was constant, it was not produced in the moment, and, therefore, knowledge about the impending “ah” was not exact. The individual “ah” produced during the talking experiment did vary on each trial, but because it was produced in the moment, advance knowledge about its exact sound characteristics should be close to perfect. Indeed, when the pitch of what you say is shifted in pitch or when an alien sound is substituted for what you say, suppression during talking is greatly reduced (Heinks-Maldonado et al., 2005), suggesting that perfect knowledge of content will affect N1 suppression. Similarly, although timing information is perfect in all three experiments, it could be argued that a person has greater timing control when talking, as

the sound being spoken can be slowed, stopped, and started up again, midcourse. Thus, although expectancy and agency effects may contribute to the N1 suppression effects seen during talking, it cannot be the whole story. In controls, although the effects due to talking were related to the effects due to expectancy, controlling for the effects of the visual warning only slightly reduced the size of the talking effect. Although superior knowledge about timing and content might produce stronger suppression of auditory cortical responsiveness to the spoken sound, in reality, few of us know exactly what we are going to say or how the vocal apparatus will cooperate, until after we have spoken. (See the writings of Libet regarding the timing of conscious experience; Libet, Gleason, Wright, & Pearl, 1983.) Accordingly, auditory cortical suppression effects during talking may be primarily a reflection of the specific ways motor systems are wired to send a corollary discharge of the planned actions to the specific sensory cortical areas that will experience the sensory consequences of these actions.

Deficits in either expectancy- or agency-related N1 suppression effects in the patients did not contribute to their diminished suppression of N1 during talking. Any group differences in expectancy and agency could not account for patients’ reduced suppression of N1 during talking; controls showed larger talking-related suppression effects than patients even after the contributions of expectancy and agency were taken into account statistically. Thus, the mechanisms underlying the deficits during talking are more compromised by schizophrenia than those underlying either expectancy or self-stimulation deficits.

N1 suppression during talking was not related to auditory hallucination severity. There has been mixed success relating N1 suppression to hallucinations; in a small sample of 8 patients, there was no relationship, but in a larger, recent sample, they were related (Heinks-Maldonado et al., 2007). Relating symptom severity to measures of neural synchrony has been more successful than to measures of neural power. In one experiment, low coherence in the EEG signal between frontal and temporal sites during talking was modestly related to hallucination severity (Ford, Mathalon, Whitfield, Faustman, & Roth, 2002). More recently, trial-to-trial phase coherence of the neural signal preceding talking was posited to be a signature of the forward model: It is correlated with subsequent suppression of N1 during talking, it is reduced in patients, and it is most reduced in patients with severe hallucinations (Ford et al., 2007). It is worth noting that a more exacting assessment of the nature of voices (e.g., voices coming from inside the head vs. voices coming from outside the head) might improve the success of relating symptoms to neurobiology.

Among the limitations of this study is the fact that all patients were medicated, leaving open the possibility that these effects are due to antipsychotic medications and not to the illness itself. Because of the difficulty in taking patients off their medications or finding patients who took themselves off their medications, future studies are planned involving subjects who are taking the same antipsychotics as the schizophrenia patients but who do not have schizophrenia. Bipolar patients are such a group.

Instead of approaching brain imaging only from the sensory side, in which subjects react to input in a bottom-up fashion, this approaches it from the motor side, in which subjects initiate actions and interact with the environment, in a top-down manner. The temporal resolution of EEG and ERPs allows investigators to study how the brain plans, executes, and responds to its own self-produced input.

REFERENCES

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*, 635–640.
- Boutros, N., Nasrallah, H., Leighty, R., Torello, M., Tueting, P., & Olson, S. (1997). Auditory evoked potentials, clinical vs. research applications. *Psychiatry Research*, *69*, 183–195.
- 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.
- de Gelder, B., Vroomen, J., Annen, L., Masthof, E., & Hodiomont, P. (2003). Audio-visual integration in schizophrenia. *Schizophrenia Research*, *59*, 211–218.
- Faustman, W. O., & Overall, J. E. (1999). The Brief Psychiatric Rating Scale. In M. Maruish (Ed.), *The use of psychological testing for treatment, planning and outcome assessment* (2nd ed., pp. 791–830). Hillsdale, NJ: Erlbaum.
- Feinberg, I. (1978). Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin*, *4*, 636–640.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995). *Structured Clinical Interview for DSM-IV Axis I Disorders*. New York: Biometrics Research Department, New York State Psychiatric Institute.
- Ford, J. M., Mathalon, D. H., Heinks, T., Kalba, S., & Roth, W. T. (2001). 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. (2001). 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*, *21*, 485–492.
- Ford, J. M., Roach, B. J., Faustman, W. O., & Mathalon, D. H. (2007). Synch before you speak: Auditory hallucinations in schizophrenia. *American Journal of Psychiatry*, *164*, 458–466.
- Ford, J. M., White, P. M., Csernansky, J., Faustman, W. O., Roth, W. T., & Pfefferbaum, A. (1994). ERPs in schizophrenia: Effects of antipsychotic medication. *Biological Psychiatry*, *36*, 153–171.
- Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., et al. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: An fMRI study. *Journal of Neurophysiology*, *88*, 540–543.
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography & Clinical Neurophysiology*, *55*, 468–484.
- Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*, 180–190.
- Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, E., & Ford, J. M. (2007). Imprecise corollary discharge in schizophrenia is related to auditory hallucinations. *Archives of General Psychiatry*, *64*, 286–296.
- Helmholtz, H. (1925). *Treatise on physiological optics* (pp. 44–51). Rochester, NY: Optical Society of America.
- Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.
- Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience*, *14*, 1125–1138.
- Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*, *106*, 623–642.
- Mannuzza, S. (1980). Cross-modal reaction time and schizophrenic attentional deficit: A critical review. *Schizophrenia Bulletin*, *6*, 654–675.
- Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex*, *15*, 299–302.
- McCarthy, G., & Donchin, E. (1976). The effects of temporal and event uncertainty in determining the waveforms of the auditory event related potential (ERP). *Psychophysiology*, *13*, 581–590.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory-visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Brain Research Cognitive Brain Research*, *14*, 115–128.
- Roth, W. T., Horvath, T. B., Pfefferbaum, A., & Kopell, B. S. (1980). Event related potentials in schizophrenics. *Electroencephalography and Clinical Neurophysiology*, *48*, 127–139.
- Scaife, J. C., Groves, J., Langley, R. W., Bradshaw, C. M., & Szabadi, E. (2006). Sensitivity of late-latency auditory and somatosensory evoked potentials to threat of electric shock and the sedative drugs diazepam and diphenhydramine in human volunteers. *Journal of Psychopharmacology*, *20*, 485–495.
- Schafer, E. W., & Marcus, M. M. (1973). Self-stimulation alters human sensory brain responses. *Science*, *181*, 175–177.
- Senkowski, D., Gomez-Ramirez, M., Lakatos, P., Wylie, G. R., Molholm, S., Schroeder, C. E., et al. (2007). Multisensory processing and oscillatory activity: Analyzing non-linear electrophysiological measures in humans and simians. *Experimental Brain Research*, *177*, 184–195.
- Shelley, A., Silipo, G., & Javitt, D. (1999). Diminished responsiveness of ERPs in schizophrenic subjects to changes in auditory stimulation parameters: Implications for theories of cortical dysfunction. *Schizophrenia Research*, *37*, 65–79.
- 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.
- Woods, SW. (2003). Chlorpromazine equivalent doses for the newer atypical antipsychotics. *Journal of Clinical Psychiatry*, *64*, 663–667.

(RECEIVED January 21, 2007; ACCEPTED March 15, 2007)