Behavioral/Systems/Cognitive

Neural Time Course of Echo Suppression in Humans

Kristina C. Backer,[1] **Kevin T. Hill,**[1] **Antoine J. Shahin,**[1] **and Lee M. Miller**[1,2]

[1]Center for Mind and Brain and [2]Department of Neurobiology, Physiology, and Behavior, University of California, Davis, California 95616

In reverberant environments, the brain can suppress echoes so that auditory perception is dominated by the primary or leading sounds. Echo suppression comprises at least two distinct phenomena whose neural bases are unknown: spatial translocation of an echo toward the primary sound, and object capture to combine echo and primary sounds into a single event. In an electroencephalography study, we presented subjects with primary-echo (leading–lagging) click pairs in virtual acoustic space, with interclick delay at the individual's 50% suppression threshold. On each trial, subjects reported both click location (one or both hemifields) and the number of clicks they heard (one or two). Thus, the threshold stimulus led to two common percepts: *Suppressed* and *Not Suppressed*. On some trials, a subset of subjects reported an intermediate percept, in which two clicks were perceived in the same hemifield as the leading click, providing a dissociation between spatial translocation and object capture. We conducted time–frequency and event-related potential analyses to examine the time course of the neural mechanisms mediating echo suppression. Enhanced gamma band phase synchronization (peaking at ∼40 Hz) specific to successful echo suppression was evident from 20 to 60 ms after stimulus onset. N1 latency provided a categorical neural marker of spatial translocation, whereas N1 amplitude still reflected the physical presence of a second (lagging) click. These results provide evidence that (1) echo suppression begins early, at the latest when the acoustic signal first reaches cortex, and (2) the brain spatially translocates a perceived echo before the primary sound captures it.

Introduction

Many everyday environments are highly reverberant, yet we often fail to hear echoes. This complex perceptual phenomenon (Wallach et al., 1949), known as the precedence effect or echo suppression, improves our ability to localize and identify sounds of interest (Freyman et al., 1999). It requires at least two mechanisms: spatial translocation and object capture (Blauert, 1999; Litovsky et al., 1999; Litovsky and Shinn-Cunningham, 2001). Spatial translocation, occurring over echo delays of ∼1–9 ms for clicks, moves the perceived location of an echo toward the location of the leading sound (Litovsky et al., 1999). Object capture, occurring at the shorter translocated delays (∼1–5 ms for clicks), combines the leading and lagging sounds into a single, fused auditory object (Litovsky et al., 1999). Interestingly, at delays just above object capture threshold, an echo can be heard but at a location transposed toward the leading sound (Pecka et al., 2007). That is, spatial translocation can occur without object capture, providing a key dissociation between the mechanisms (Yang and Grantham, 1997). As echo delay increases further, both leading and lagging sounds are heard distinctly at their veridical locations.

Although the psychoacoustics associated with echo suppression have been well described, its underlying neural mechanisms are poorly understood. The cochlear nucleus and inferior colliculus (IC) have been suggested by animal studies as the first levels where echo suppression begins (Hafter et al., 1988; Wickesberg and Oertel, 1990; Yin, 1994; Fitzpatrick et al., 1995; Litovsky, 1998; Litovsky and Yin, 1998a,b; Pecka et al., 2007), consistent with a human IC lesion study (Litovsky et al., 2002). In contrast, electroencephalography (EEG) studies in humans have shown that auditory brainstem responses index the physical presence of an echo regardless of perception (Damaschke et al., 2005), whereas late cortical potentials, including mismatch negativity (Damaschke et al., 2005) and object-related negativity (Sanders et al., 2008), may reflect perceptual consequences of echo suppression. Another recent study (Spierer et al., 2009) suggests that right temporoparietal cortical activity is associated with a fused percept. Together, these electrophysiological studies leave a large time window between brainstem firing and late cortical potentials, wherein the precise timing and neural mechanisms underlying echo suppression remain unknown. Furthermore, the neural processes related to spatial translocation and object capture have never been dissociated.

In the present study, we used EEG to characterize the neural time course of echo suppression. Subjects listened to click pairs in virtual acoustic space, with interclick delay calibrated to each subject's threshold, yielding three distinct perceptions of physically identical stimuli: *Suppressed* (translocation and object capture), *Intermediate* (translocation without object capture), and *Not Suppressed* (neither translocation nor capture). Two additional conditions, single click and obvious double click, were perceptually or physically comparable with the threshold conditions, providing key reference points for the EEG analysis. We analyzed the EEG for changes in latency and amplitude of event-related potentials (ERPs) as well as spectrotemporal power and

intertrial phase locking. We hypothesized that only physical stimulus properties would be manifest in the earliest EEG responses, followed by neural evidence of echo suppression mechanisms. Still later, we should observe the perceptual correlates of spatial translocation and, finally, object capture.

Materials and Methods

The protocol described here was approved by the Institutional Review Board at the University of California, Davis. All subjects gave written, informed consent before participating in the study and were paid for their participation.

Participants. All participants were right-handed, reported no neurological disorders, had not used any psychoactive medications within the month before their participation, and had normal pure-tone (200–12,500 Hz) hearing thresholds. Of the 15 subjects completing the study, 13 (five males; mean ± SD age, 22.3 ± 2.8 years) had adequate behavior and EEG signal quality. Two subjects could not be included in the analysis because of too few trials after artifact rejection or poorly balanced behavior (trial numbers) among suppressed, intermediate, and not suppressed percepts.

Subjects were separated into two groups based on stimulus configuration, left or right leading hemifield, as explained below. Five (three males) of the 13 subjects completed the experiment for both stimulus configurations to test for intrasubject differences in echo threshold and electrophysiological results between hemifields. For these five subjects, EEG data for left- and right-leading sounds were collected in separate sessions on different days. Because the EEG analyses disregard overall mean differences across subjects [repeated-measures ANOVA (RMANOVA); see below], each subject's two sessions were considered independently. Therefore, each configuration consisted of nine subjects (four males per group) for a total of 18 complete datasets.

Stimuli. Head-related transfer functions (HRTFs) were created for each subject in a sound-dampened chamber using AuSIM in-the-canal microphones (AuSIM) and a Tannoy (Precision 6; Tannoy Ltd.) loudspeaker placed in the horizontal plane ~1.2 m away. Subjects were seated in a swivel chair, and white noise (3 s duration) was recorded when subjects were positioned with the loudspeaker 45° to the left and 45° to the right of midline. HRTFs were estimated in the frequency domain, by dividing the magnitude and subtracting the phase of the fast Fourier transform (FFT) of the recorded sound in each ear relative to the FFT of the presented sound. From this HRTF, a head-related impulse response (HRIR) function was obtained by inverse FFT, and, for practical purposes, the HRIR was truncated to the first 3000 time samples (i.e., filter order). This HRIR was then convolved with the stimuli.

All stimuli were biphasic clicks, consisting of 12 samples (six positive and six negative) at a sampling rate of 96 kHz (0.125 ms long). The clicks were filtered with each subject's HRTF to place them in virtual acoustic space (for more details, see Wightman and Kistler, 1989). Click pairs contained a leading click at either +45 or −45°, followed by an equally intense lagging click in the opposite location. Subjects were randomly assigned to one of two groups (left-leading or right-leading) before the experiment. For the left-leading group, the leading sound was always presented from −45° (left) and the lagging sound from +45° (right) in virtual acoustic space (and vice versa for the right-leading group). All stimuli were presented at a comfortable listening level, with peak ~98 dBA sound pressure level.

Trial structure. Trials consisted of a series of conditioning click pairs followed by a pause, a test click pair, and a cue for subjects to respond. The conditioning click train was used to prevent any one trial from affecting perception in subsequent trials, because short-term auditory context strongly affects the probability of echo suppression (Clifton, 1987; Clifton and Freyman, 1989; Clifton et al., 1994). Conditioning click trains consisted of 11 click pairs or single clicks (identical to the test pairs) and lasted ~1.8 s. On each trial, subjects listened to the conditioning click train, followed by a 1.6 s pause and a test click or click pair (identical to the one click or click pair in the conditioning train). There was ~0.8 s of silence between the test click and the conditioning train of the subsequent trial. Thus, the total length of each trial was 4.2 s.

Subjects were seated in a comfortable chair in a sound-dampened chamber, held the response keyboard on their lap, and faced a computer screen. Presentation software (NeuroBehavioral Systems) was used to control stimulus timing and to present the auditory and visual (fixation cross) stimuli. Auditory stimuli were delivered through shielded ER-4 earphones (Etymotic Research). Participants visually fixated on a white cross in the center of the computer screen, and the cross flashed green at the start of the conditioning click train in each trial. Subjects were instructed to wait for the green flash before making any button responses for the trial they just heard. This prevented motor cortical activity from contributing to the EEG signals of interest. After seeing the green crosshair, subjects made two button responses in succession. First, they indicated with the right forefinger or middle finger whether they heard a click in one or both hemifields. They were told before the task that they would always hear a click on the left side (for left-leading subjects) or on the right side (for right-leading subjects). Second, they indicated whether they heard either one or two clicks with a button press of their left forefinger or middle finger. They were further instructed that hearing a click on both sides always counted as two clicks, even if they sounded simultaneous. All subjects completed a practice session with feedback on each trial before beginning the behavioral calibration.

Behavioral calibration. In this session, subjects were presented with three conditions (Fig. 1) in pseudorandom trial order: (1) single clicks in the leading location, (2) threshold click pairs (i.e., click pairs presented with a delay equal to the subject's echo threshold), and (3) obvious double clicks (35 ms lag). Each subject completed a 20 min calibration to determine his/her echo threshold, defined as the delay at which the subject reported the lagging sound on ~50% of the threshold click trials. The calibration algorithm used an adaptive one-up-one-down procedure, based on the subject's responses (starting at 5 ms echo lag, consisting of 0.5 ms steps, and comprising 290 trials). Each subject's echo threshold was determined as the most frequently occurring echo delay (on threshold trials) during the second half of the calibration. This threshold from the behavioral session became the starting point for the EEG session abbreviated calibration.

EEG session procedure. The EEG session consisted of a 10 min echo threshold recalibration, placing the cap and electrodes on the subject's scalp, and eight 15 min blocks of experimental recording time. The calibration was done to confirm the echo threshold assessed during the behavioral session for each subject. After the calibration, the echo delay for threshold trials was chosen for the duration of the EEG session. During the EEG recording, we verified that the subjects' suppression rate (or proportion of suppressed to not suppressed trials) after each block was ~50:50. In some cases, a subject's suppression rate deviated substantially (e.g., 20:80) from this optimal ratio after the first block of EEG recording, so we adjusted the threshold to account for this and continued recording. EEG analyses only used data collected with this new threshold.

Data were collected using high-density (128 channel) EEG (BioSemi System; BioSemi) at a sampling rate of 1024 Hz. The EEG experiment lasted a minimum of 2 h, and subjects were given breaks between blocks. Each block consisted of 215 trials, and the task was similar to the previous behavioral calibration tasks. However, instead of three stimulus conditions, there were five during the EEG experiment, also presented pseudorandomly: (1) single clicks (in the leading click location), (2) threshold click pairs, (3) obvious double clicks (35 ms lag), (4) single click catch trials (single click conditioning train but no test click), and (5) threshold pair catch trials (threshold click pair conditioning train but no test pair). Subjects did not respond on catch trials, and the fixation cross did not flash green after these trials. Catch trials were used to ensure that no sustained brain activity from the conditioning train overlapped with the later activity locked to the test click. Other than the addition of catch trials, the task was the same as in the behavioral session.

EEG data analysis. EEG data preprocessing was done in MATLAB (MathWorks) using both EEGLAB software (Delorme and Makeig, 2004) and in-house code. First, each dataset was downsampled to 512 Hz, re-referenced to the average reference (i.e., the average activity of all channels was subtracted from each channel), high-pass filtered at 0.5 Hz [zero-phase finite impulse response (FIR)], and separated into epochs (segments) spanning −1 s to +0.5 s, relative to the first test click on each
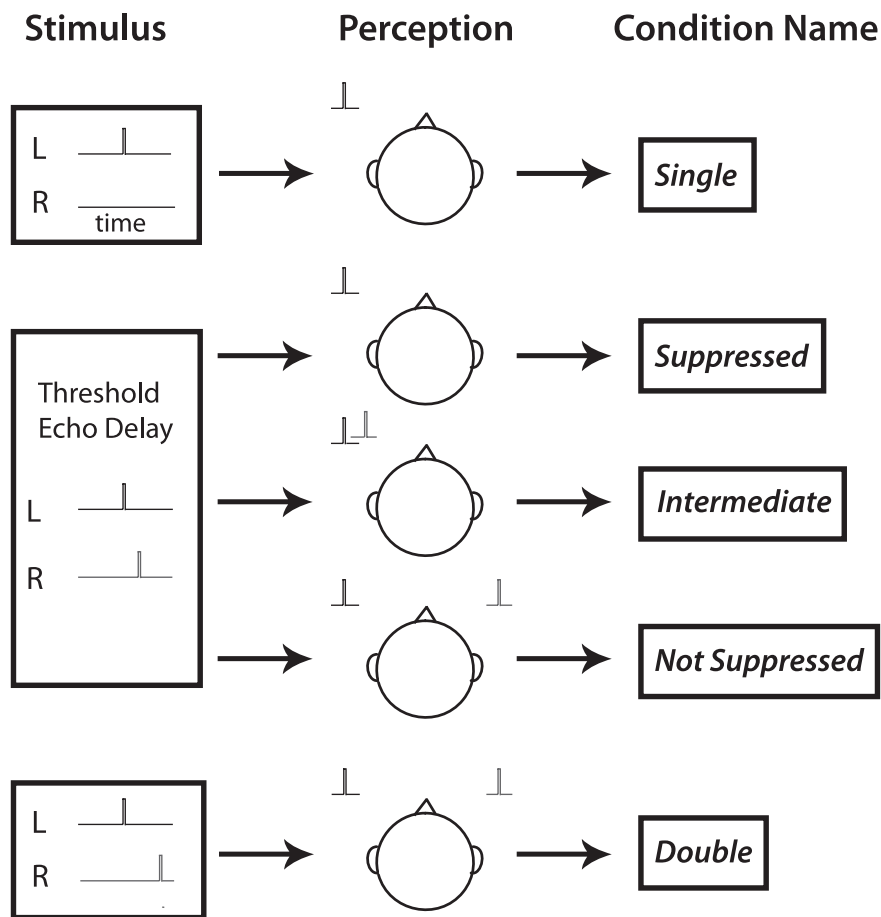
Stimulus

Perception

Condition Name

Figure 1. Experimental design. The left column shows the three stimulus types used in the EEG session for a left-leading subject. Left (L; primary) and right (R; echo) click time courses are shown aligned to emphasize their temporal lag. The possible perception(s) for each stimulus type is demonstrated in the middle column. Here, perceived lead and lag clicks are depicted in space, on a schematic of the subject's head as viewed from above, to emphasize their perceived number and location. The right column shows the condition names, which are based on both the stimulus type and (for threshold trials) the response.

trial. Thus, each epoch was 1.5 s long. There was one epoch for each trial of recorded EEG data, so each subject recorded a maximum of 1720 epochs (trials) before artifact rejection. Then, all epochs were baselined (normalized) to the average value measured during the 100 ms preceding the test click. All trials were then entered into independent components analysis (ICA) [as implemented by EEGLAB (Makeig et al., 1997)], using 50 principal components. Canonical eyeblink and lateral eye-movement independent components were conservatively removed from the data by visual inspection of each component map. After combining all non-artifactual ICA components, all trials were re-baselined (renormalized) to the 100 ms preceding the test click. Artifact rejection removed epochs with data exceeding $\pm 100 \ \mu V$ during the period -300 to $+500$ ms. Rejected epochs were removed from all 128 channels. Trials were then sorted according to stimulus type and the subject's responses. As shown in Figure 1, EEG trials were separated into seven perceptual categories: *Single* (single click), *Suppressed* (full suppression of the threshold echo), *Intermediate* (two clicks in the leading hemifield), *Not Suppressed* (perception of the echo), *Double* (obvious double click pair), *Catch Single* (single click catch trial), and *Catch Threshold* (threshold pair catch trial). This epoched data (i.e., data that has been segmented and separated by condition) then was used in both an ERP analysis and a time–frequency analysis, as detailed below.

Event-related potential analysis. The mean ERP waves of each perceptual condition were low-pass filtered at 20 Hz using a zero-phase FIR filter. These filtered, mean waves for each condition were averaged across all subjects to obtain the group mean waves for each condition.

We examined the amplitude and latency differences among conditions (*Single*, *Suppressed*, *Not Suppressed*, and *Double*) for three auditory ERP

components (P1, N1, and P2). To analyze effects of laterality of the leading click in the P1 and N1, 10 left frontal channels and 10 right frontal channels (Fig. 2B), at which the P1 and N1 reached their maximum and minimum values, respectively, were averaged, and the peak latencies within each condition were identified manually based on the group average ($n = 18$) of these 20 channels. These latencies were then used in an automated windowing algorithm, which recorded the local maximum (P1) or minimum (N1) amplitude and latency within 15 samples (~ 29.3 ms) on either side of the group peak latency for each subject within each condition and each side (ipsilateral or contralateral to the leading click location). Similarly, the peak latency for the P2 was based on the group average of Cz and its adjacent five channels (Fig. 2B). The local maximum amplitude and latency for each subject within each condition was automatically chosen. For ERP effects that revealed differences between *Suppressed* and *Not Suppressed*, we also analyzed the subset of subjects ($n = 8$) who reliably (≥ 50 trials after artifact rejection) reported the *Intermediate* percept.

Time–frequency analyses. Time–frequency analysis was performed using EEGLAB, which generates spectrograms for both intertrial phase coherence (ITPC) and event-related spectral perturbation (ERSP). ITPC shows the extent to which neural oscillations are consistent in phase from one trial to another, within each channel. In other words, the term "phase locking" as used here refers to the trial-to-trial consistency of the phase of a particular frequency band relative to the stimulus onset (first click). Thus, the phrase "enhanced phase locking" indicates that the phase of a certain frequency within the EEG signal had a tighter temporal relationship with stimulus onset for one condition compared with another. ERSP, conversely, reflects a relative change in power over time after the stimulus, regardless of phase. Subjects ($n = 2$) who had fewer than 90 trials in any one condition were not included in the main time–frequency analyses, leaving 16 datasets. First, artifact rejection was repeated with the same threshold ($\pm 100 \ \mu V$) but an earlier starting point of -1 s, because a longer baseline period is necessary for low-frequency ERSP baselining. A sliding Hanning-windowed two-cycle sinusoidal wavelet (short time discrete Fourier transform) of the EEG signal was used, in which the frequency increments were 2 Hz and the average step size was 6.23 ms. The sliding window size was 128 samples (250 ms, 2 cycles) at the lowest frequency (8 Hz), and it decreased linearly, whereas number of cycles increased with frequency, resulting in ~ 12 cycles at the highest frequency (100 Hz).

Time–frequency region of interest analysis. We conducted a time–frequency region of interest (ROI) analysis to determine whether enhanced phase locking during echo suppression was suppression specific or reflected perception. Note that the time–frequency ROI is not a spatial ROI (as in some neuroimaging analyses) but rather an analysis of a specific time–frequency region within an EEG spectrogram. As mentioned above, we analyzed two types of spectrograms: (1) ITPC and (2) ERSP. Using the same ROI for ITPC and ERSP, we can examine whether differences between conditions in intertrial phase locking are also associated with differences in spectral power. ITPC generally returns phase-locking factors (PLFs) at each time–frequency bin; PLF is the dependent measure in an ITPC spectrogram and ranges from 0 to 1, in which 1 represents perfect phase locking and 0 reflects phase desynchronization (Tallon-Baudry et al., 1996). All ITPC values were converted from PLFs to Ray-

leigh's z value using the following equation: $z = n * PLF^2$, where n equals the number of trials for that subject and condition (Fisher, 1993). (Note that Rayleigh's z value should not be confused with a z-score or normalized variate.) We conducted permutation tests (as detailed below) on the z-value data to contrast phase-locking activity during *Suppressed* and *Not Suppressed* trials. From the thresholded ($p < 0.05$) ITPC contrast between *Suppressed* and *Not Suppressed* averaged across all channels, two main activation clusters, in which phase locking during *Suppressed* was greater than during *Not Suppressed* (as depicted in Fig. 5A), were chosen for an ITPC ROI analysis. The boundary for the early gamma ROI included 30–60 Hz from ~0 to 82 ms after stimulus. The alpha/beta ROI included 8–28 Hz from ~50 to 212 ms.

The following steps describe the ITPC ROI analysis. (1) A box was drawn around each cluster to use as the time–frequency ROI mask for each channel. (2) Channels in which the ROI mask had 50 or more significant time–frequency bins were used. (3) Within these channels, each subject's z values at each of the significant group contrast pixels were returned. This step was done separately for each condition but used the same ROI mask. (4) For each condition, each subject's z values were averaged. Steps 3 and 4 were repeated for the ERSP spectrograms to determine whether changes in spectral power occurred at the same time–frequency bins as in the ITPC spectrograms. Note that the ERSP values were not converted into Rayleigh's z value because ERSP does not have a circular distribution.

Statistical analysis. Statistica software (version 8; StatSoft) was used to compute statistics, particularly t tests for behavioral results and RMANOVA and *post hoc* Fisher's least significant difference (LSD) tests for all EEG results. The proportion of suppressed and not suppressed trials remaining after artifact rejection of the EEG data was used to calculate suppression rate on threshold trials. Independent t tests were used to compare the suppression rates and echo thresholds of the left- and right-leading groups. For the ERP analysis, a 4 (condition) × 2 (side) × 2 (group) RMANOVA was done for the latency and amplitude of the P1 and N1. For the P2 amplitude and latency, a 4 (condition) × 2 (group) test was done. Because P2 topography was centralized, unlike P1 and N1, laterality effects were not examined.

For the time–frequency analyses, permutation tests were used for statistical thresholding of the main contrast (*Suppressed − Not Suppressed*) at the group level, as described in detail previously by Chau et al. (2004) and Shahin et al. (2008). Briefly, permutation methods are nonparametric, in that they determine the null distribution by resampling the data rather than assuming an analytic form for the population distribution. This null distribution is derived by randomly relabeling a subject's data with a condition, such as *Suppressed* or *Not Suppressed*, and quantifying the ITPC or ERSP differences between these conditions. Each random assignment of data therefore produces a chance (null) distribution of ERSP or ITPC differences across electrodes and time–frequency bins, based on the prestimulus period. This distribution reflects the null hypothesis that oscillatory activity does not differ between conditions. To handle the problem of multiple comparisons, permutation tests were applied based on the null distributions of the maximum values obtained in repeated resamplings of the data (Holmes et al., 1996). This maximal
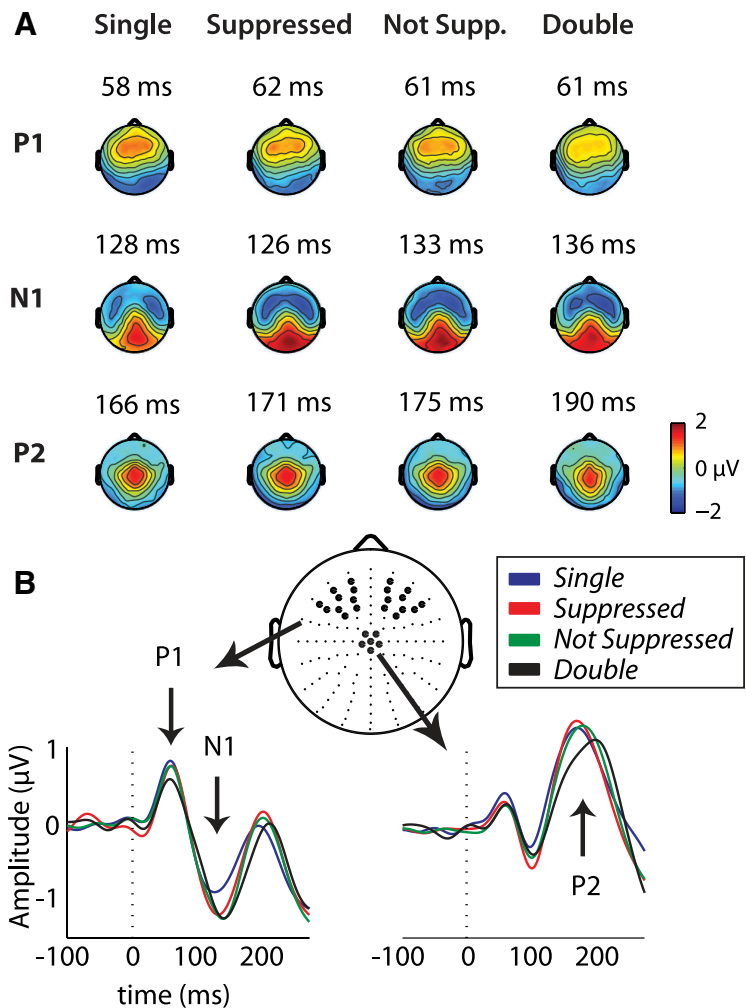

Figure 2. ERP topography and waveforms. *A* shows the ERP topography for each component (P1, N1, and P2) and condition (*Single*, *Suppressed*, *Not Suppressed*, and *Double*) at the group (*n* = 18) peak latencies. The group mean waveforms from 20 frontal channels (P1 and N1) and six central channels (P2) are shown in *B*. Dotted black lines in *B* represent time 0, when the first click occurred.

null distribution was used to determine the threshold for a given p value (here, $p < 0.05$).

RMANOVAs were also done to compare phase locking and spectral power within each ROI. The catch trials acted as a control, to measure any activity from the conditioning train that sustained through the test click. For each time–frequency ROI, the Rayleigh's z values (for ITPC) or spectral power values (decibels, for ERSP) were averaged for the *Catch Threshold* and *Catch Single* conditions; this mean was subtracted from the z values or spectral power of the other four conditions (single, suppressed, not suppressed, and double). For these four conditions (*Single − Catch*, *Suppressed − Catch*, *Not Suppressed − Catch*, and *Double − Catch*), 4 (condition) × 2 (group) RMANOVAs were computed for ITPC z values and spectral power. Because the creation of the ROI clusters were based on differences between phase locking of *Suppressed* and *Not Suppressed* trials, the ITPC RMANOVA is likely significant. However, we were mainly interested in the *post hoc* LSD tests to compare phase locking among all four conditions, not only differences between *Suppressed* and *Not Suppressed*. Conversely, because the time–frequency bins in the spectral power plots examined were not based on preexisting differences in spectral power, neither the RMANOVA nor *post hoc* LSD tests for the ERSP analysis was biased. In both ITPC and ERSP, the *post hoc* LSD tests showed whether an effect was specific to echo suppression or generally related to perception, as discussed below.
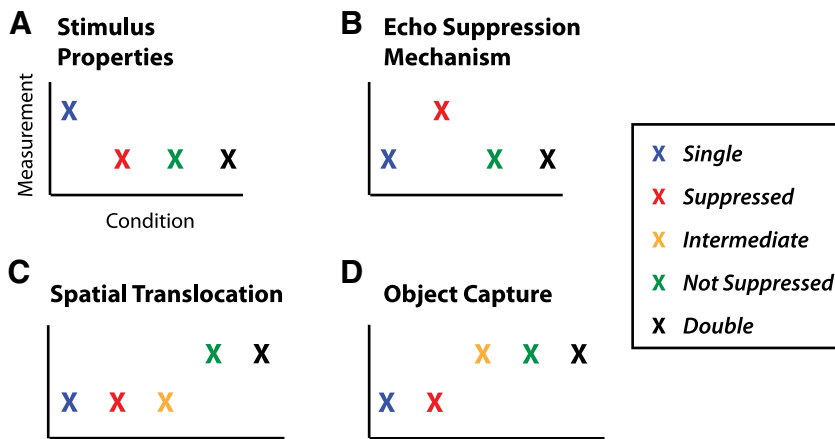
Figure 3. Profiles for hypothesized EEG effects across conditions. **A**, If the single condition differentiates from the other three conditions, the measure reflects the physical presence of an "echo" click, regardless of perception. **B**, If a suppressed value is different from the other conditions, the effect is specific to the echo suppression mechanism. Effect profiles **C** and **D** reflect perceptual consequences of echo suppression, either spatial translocation or object capture, respectively. Without the *Intermediate* percept, these two interpretations cannot be dissociated.

Results

Behavior

Since only eight subjects reliably reported the *Intermediate* percept on threshold trials, these trials were not included in the overall suppression rate. Thus, each subject's overall suppression rate is their proportion of *Suppressed* to *Not Suppressed* responses. There was no significant difference ($p = 0.93$) in overall suppression rate between left-leading (mean $\pm$ SD, 59.0 $\pm$ 13.5%) and right-leading (mean $\pm$ SD, 59.6 $\pm$ 16.1%) groups. Thus, these proportions give a balanced sample of both *Suppressed* and *Not Suppressed* trials for the EEG analyses. Also, the echo thresholds for the two groups were very similar ($p = 0.83$). The mean $\pm$ SD echo threshold of the left-leading group was 6.33 $\pm$ 2.02 ms, whereas the mean $\pm$ SD of the right-leading group was 6.11 $\pm$ 2.33 ms.

Event-related potentials

A standard ERP analysis was done to examine interactions among either amplitude or latency of auditory evoked potentials (P1, N1, and P2), condition (single, suppressed, not suppressed, or double), and laterality of activity relative to the leading click hemifield (for P1 and N1). However, no main effects or interactions were found involving laterality of activity for either component, so all results reported collapse activity across hemispheres (Fig. 2*B*). We therefore used RMANOVAs to test whether the latency or amplitude of each component, averaged within subject (see Materials and Methods), showed a main effect of condition.

We expected a series of neural responses that index different aspects of echo suppression. Four possible relationships among conditions, or profiles, and their corresponding interpretations are presented in Figure 3. (1) Because the *Single* condition was the only condition without a lagging "echo" click, difference in activity between the *Single* and the other conditions indexes the physical presence of the "echo" regardless of perception (Fig. 3*A*). (2) If *Suppressed* activity differs from the other conditions, it reflects the echo suppression mechanism (Fig. 3*B*). Finally, if both *Single* and *Suppressed* activities are similar, both *Not Suppressed* and *Double* are similar, and these two pairs are distinct, this pattern reflects the perceptual consequences of echo suppression. However, only with the *Intermediate* percept can we attribute this effect unambiguously to either spatial translocation or object

capture. Namely, (3) if the *Intermediate* value is similar to *Single* and *Suppressed* values, the effect indexes spatial translocation (Fig. 3*C*). (4) However, if the *Intermediate*, *Not Suppressed*, and *Double* values are similar, the effect reflects object capture (Fig. 3*D*).

We analyzed the P1 amplitude and latency (Fig. 4*A*) to determine whether there were any early effects in the ERP related to echo suppression. P1 amplitude showed a main effect of condition ($F_{(3,48)} = 3.30$, $p = 0.028$), such that the P1 of the *Double* click was smaller than the other three conditions (*post hoc*, $p < 0.05$). Because the lead and lag were separated by 35 ms in this condition, it is probable that the P1 responses to each click destructively interfered with one another, causing a summed component to flatten and broaden. However, the lead and lag of the threshold pair were only separated on average by ~6 ms; thus, the destructive interference would be minimal, resulting in P1 amplitudes similar to that of the single click. There was no main effect of condition on P1 latency ($F_{(3,48)} = 1.75$, $p = 0.17$).

Whereas the P1 did not reveal any early perceptual or echo suppression-specific effects, perception was reflected in the N1. Both N1 amplitude and latency showed a main effect of condition ($F_{(3,48)} = 4.35$, $p = 0.0087$; $F_{(3,48)} = 7.05$, $p < 0.001$, respectively), as shown in Figure 4*B*. *Post hoc* LSD tests showed that the amplitude for the *Single* click condition (-1.2 ± 0.23 μV) was smaller than all other conditions (*Suppressed*, -1.4 ± 0.20 μV; *Not Suppressed*, -1.5 ± 0.23 μV; and *Double* click, -1.4 ± 0.23 μV). For the N1 latency, according to *post hoc* LSD tests, there was no difference between that of the *Single* (128 $\pm$ 5.9 ms) and *Suppressed* (126 $\pm$ 4.4 ms) conditions or the *Not Suppressed* (133 $\pm$ 4.9 ms) and *Double* (136 $\pm$ 4.7 ms) conditions. Furthermore, the N1 occurred earlier for the *Single* and *Suppressed* trials than for *Not Suppressed* and *Double* trials. Thus, N1 amplitude reflects stimulus properties (same effect profile as in Fig. 3*A*), and N1 latency reflects echo perception (one or two clicks heard, on one or both sides).

There was no main effect of condition on P2 amplitude ($F_{(3,48)} = 1.01$, $p = 0.40$). A main effect of condition ($F_{(3,48)} = 14.34$, $p < 0.001$) on P2 latency was found, such that the P2 latency was smallest for *Single* (166 $\pm$ 4.1 ms) and *Suppressed* (171 $\pm$ 3.8 ms) trials and largest for *Double* (190 $\pm$ 4.5 ms) click trials (Fig. 4*C*). Specifically, *post hoc* LSD tests revealed a significant difference between the P2 latency of *Double* clicks and of the other three conditions. Also, P2 latency was significantly reduced for *Single* compared with *Not Suppressed* (175 $\pm$ 3.9 ms) trials. The *Suppressed* trial latency was similar to that of both the *Single* condition ($p = 0.23$) and the *Not Suppressed* P2 latency ($p = 0.24$).

To distinguish between the mechanisms of spatial translocation and object capture (as illustrated in Fig. 3*C*,*D*, respectively), we analyzed the effects that generally reflected perception (N1 and P2 latency) within the eight *Intermediate* percept subjects (Fig. 4*D*,*E*, respectively.) N1 latency showed a main effect of condition ($F_{(4,24)} = 4.38$, $p = 0.0084$). *Post hoc* LSD tests revealed that the N1 latencies of *Suppressed* (125 $\pm$ 7.6 ms) and the *Intermediate* (128 $\pm$ 5.3 ms) percept were both faster than those of *Not Suppressed* (141 $\pm$ 7.3 ms) ($p < 0.01$). Because subjects reported

hearing sounds in only the leading hemi-field for *Suppressed* and *Intermediate* per-cept trials versus both hemifields for not *Suppressed* trials, the decrease in N1 la-tency is related to spatial translocation rather than object capture. Also, for P2 latency, there was a main effect of condi-tion ($F_{(4,24)} = 3.08$, $p = 0.035$). P2 laten-cies for *Suppressed* (163 ± 5.6 ms) and *Not Suppressed* (172 ± 4.6 ms) were signifi-cantly different, but the P2 latency of the *Intermediate* percept (166 ± 6.5 ms) was not different from either *Suppressed* or *Not Suppressed*, providing no clear dis-tinction between translocation and cap-ture mechanisms.

Time–frequency analysis
Because echo suppression is a rapid, time-sensitive phenomenon, a time–frequency approach may complement the tradi-tional ERP analysis, for at least two rea-sons. First, neural processing can manifest as oscillatory activity, which is naturally suited to frequency-domain representa-tion. Second, ERPs tend to measure effects only evident in trial-averaged signals, whereas time–frequency approaches first quantify effects in single trials and then take the average of those effects. If neural timing and amplitude are ever uncoupled, condition differences in timing (phase) and amplitude (power) will be confounded in ERP averaging but will remain distinct in time–frequency analyses. For instance, ITPC shows the degree to which neural activity is phase locked to a particular event, such as sound onset, thus focusing on timing rather than power. We com-pared ITPC between *Suppressed* and *Not Suppressed* conditions, revealing two signif-icant spectrotemporal clusters that were consistent across channels: an early gamma cluster (30–60 Hz from ~0 to 82 ms after stimulus) and a later alpha/beta cluster (8–28 Hz from ~50 to 212 ms). These clusters were then used as time–frequency regions of interest (see Materials and Methods) to determine whether enhanced phase locking was specific to an echo suppression mechanism or reflected perceptual consequences, following the same logic discussed above and shown in Figure 3. To reiterate, an effect (here, enhanced phase locking) that distinguishes *Sup-pressed* from the other conditions reflects the echo suppression mechanism (Fig. 3B). Conversely, if conditions are grouped by what subjects heard (one click for *Single* and *Suppressed* vs both clicks for *Not Suppressed* and *Double*), the activity indexes the perceptual consequences suppressing an echo (Fig. 3C,D). We used RMANOVAs ($n = 16$) to determine significant differences in phase locking among conditions. The same ROI masks were used to examine spectral power within each condition. Results from this analysis are shown in Figure 5, *B* and *C*.

In the early gamma phase-locking ROI, there was a significant main effect of condition ($F_{(3,42)} = 10.97$, $p < 0.0001$). *Post hoc* LSD tests showed that *Suppressed* had higher phase locking than
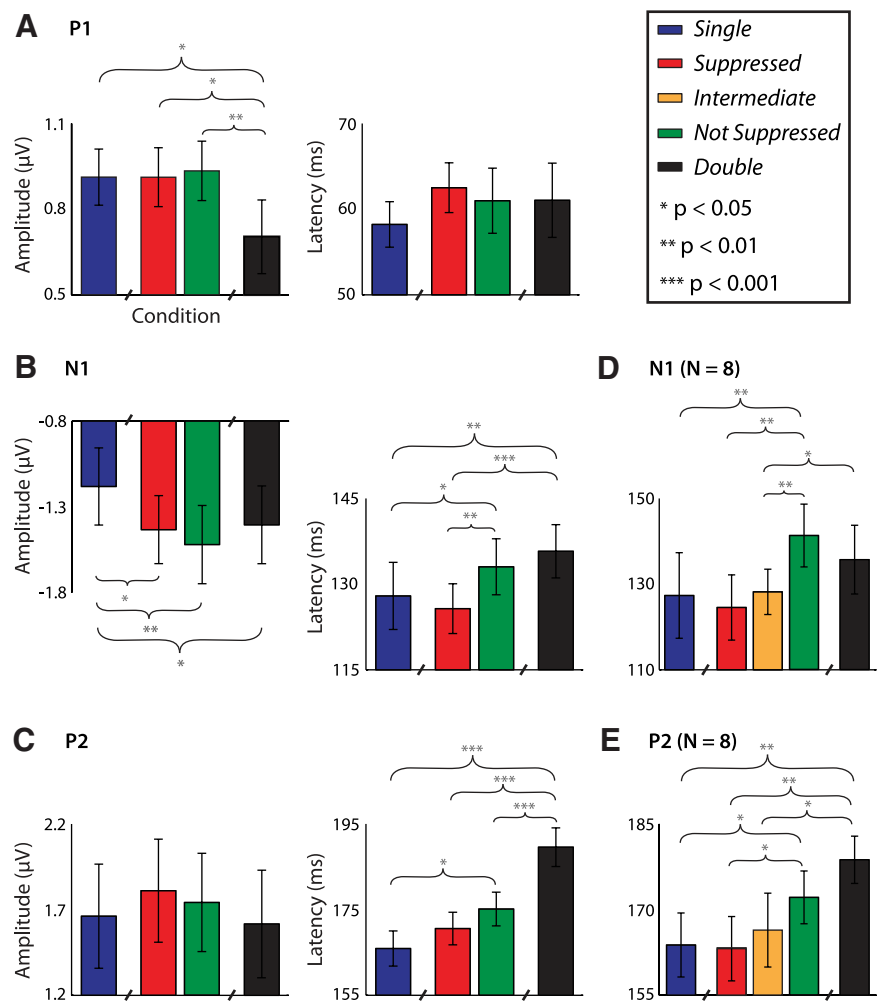

Figure 4. Latency and amplitude of ERP components differentiate stimulus properties, suppression mechanism, and perceptual consequences. *A–C* show group results ($n = 18$). *A*, P1 amplitude distinguishes conditions based on physical stimulus properties (click number or delay) rather than perception. *B*, N1 amplitude reflects physical click number, whereas N1 latency is categorically sensitive to perceptual echo suppression. *C*, P2 latency increases with likelihood of hearing the echo as a separate object but not categorically. *D* and *E* show results from the *Intermediate* percept subgroup ($n = 8$). *D*, N1 latency reflects spatial translocation of the echo rather than object capture. *E*, As in *C*, P2 latency increases noncategorically with the likelihood of hearing the echo as a separate object. Symbols indicate *p* values of *post hoc* LSD tests based on the RMANOVA, whereas error bars in all figures represent the least squares SE of each mean across subjects. Conditions that differ in their stimulus properties are separated by axis ticks.

the other conditions ($p < 0.001$). There was also a main effect of condition within the low-frequency phase-locking ROI ($F_{(3,42)} = 24.77$, $p < 0.0001$). Similar to the first ROI, phase locking during *Suppressed* trials was greater than all other conditions ($p < 0.0001$). In contrast to phase locking, there was no main effect of condition for spectral power within the gamma ROI ($F_{(3,42)} = 1.43$, $p = 0.25$) and only a strong trend in the alpha/beta ROI ($F_{(3,42)} = 2.86$, $p = 0.048$), driven by decreased power for the double relative to the *Suppressed* and *Not Suppressed* conditions ($p < 0.05$). Because phase locking in both the early gamma and low-frequency ROIs was enhanced only for *Suppressed*, the one condition when echo suppression was both required and successful, then this enhanced phase locking likely indexes the mechanism underlying suppression (as represented in Fig. 3B). Furthermore, the enhanced phase locking was not accompanied by differential spectral power. Therefore, the echo suppression mechanism is reflected by increased temporal precision in neural firing both early (20–60 ms, ~40Hz) and for hundreds of milliseconds after a stimulus (~50–212 ms, 8–28 Hz).
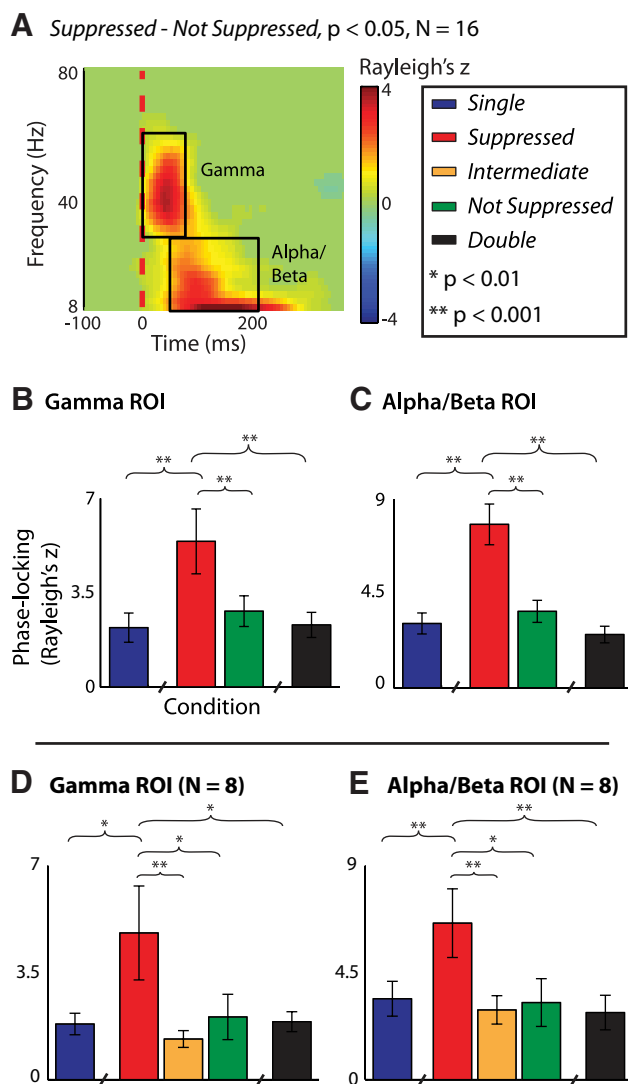
Figure 5. Time–frequency ROI analysis (*n* = 16). *A* shows the ITPC contrast *Suppressed − Not Suppressed* (*p* < 0.05) in the grand mean average spectrogram of all channels. The black boxes bound the time–frequency bins used for the ROI analysis. Bar graphs show the ITPC results for the early gamma ROI (*B*) and later alpha/beta ROI (*C*). *Suppressed* dissociates from the other conditions, indicating that both ROIs reflect the echo suppression mechanism. Bar graphs in *D* and *E* show the ITPC results for the *Intermediate* percept subjects (*n* = 8) for the early gamma ROI and later alpha/beta ROI, respectively. Because the *Intermediate* percept phase locking is similar to *Not Suppressed*, phase-locking enhancement is specific to the mechanism underlying full echo suppression (translocation and object capture). Conditions that differ in their stimulus properties are separated by axis ticks.

Because there were only eight *Intermediate* percept subjects, some of whom had fewer than 90 trials in one or more conditions, there was not sufficient power for an independent time–frequency analysis. However, we conducted a supplementary analysis to determine whether the echo suppression mechanism, evident in ITPC, reflected translocation alone or both translocation and object capture. The time–frequency ROI masks generated from the whole-group analysis *Suppressed − Not Suppressed* contrast were used in this analysis. Despite low trial numbers of some subjects, the results were very similar to the whole-group results. There was a main effect of condition for both the early gamma and low-frequency ROIs ($F_{(4,24)}$ = 4.66, p < 0.01; $F_{(4,24)}$ = 6.10, p < 0.005, respectively). In both time–frequency ROIs, phase locking was enhanced for only the *Suppressed* condition,

whereas the *Intermediate* condition was similar to *Single, Not Suppressed*, and *Double* (*post hoc* LSD, p < 0.01) (Fig. 5 *D*, *E*). Thus, unlike the N1 latency effect, which reflected the perceptual consequences of echo suppression (because *Single, Suppressed*, and *Intermediate* were distinct from *Not Suppressed* and *Double*), this phase-locking enhancement is specific to the echo suppression mechanism. These results suggest that enhanced phase locking occurs only during full echo suppression, that is, when the echo is both translocated and captured by the leading sound.

Discussion

In this study, we characterize the temporal dynamics of echo suppression in humans using EEG. Using a subject-specific threshold echo delay, we compared activity related to three distinct perceptions (*Suppressed, Intermediate*, and *Not Suppressed*) that arose from the same physical stimulus. Single and obvious double clicks provided a controlled measure by which we distinguished EEG responses based on stimulus attributes, suppression mechanism, and perceptual consequences. The *Intermediate* percept was particularly informative in that it dissociated two key percepts, spatial translocation and object capture. We used both ERP and time–frequency analysis methods to determine the following neural timeline for echo suppression: (1) the echo suppression mechanism is evident in enhanced early gamma followed by low-frequency phase locking, (2) perceptual spatial suppression is reflected in shorter N1 latency, whereas N1 amplitude still reflects the physical rather than perceptual presence of an echo, and (3) P2 latency may index the probability of fusing the echo and leading sounds or the interaction between physical stimulus properties and perception.

Evidence for the echo suppression mechanism was apparent very early, peaking from 20 to 60 ms after stimulus. We observed enhanced low gamma (~40 Hz) phase locking specific to successful suppression, peaking at the vertex and frontocentral channels. Based on the timing and topography of this enhancement, this activity is likely related to the auditory middle latency response (AMLR). The AMLR occurs after the auditory brainstem response and consists of three main peaks: Na, Pa, and Nb (Goldstein and Rodman, 1967). The latency of the early phase enhancement matches that of the Na (~20–24 ms), Pa (~31–35 ms), and Nb (~42–50 ms) (Goldstein and Rodman, 1967; Neves et al., 2007). Because the P1 peaked at ~60 ms in this study, the later portions of this enhanced synchrony could partially reflect the P1 as well. However, the Na, Pa, and Nb usually peak at the vertex, matching the topography of this enhancement, whereas the P1 was strongest in frontal channels. Previous findings suggest that the Pa and Nb are generated in Heschl's gyrus and thalamocortical circuits (Picton et al., 1974; Hall, 1992; Liégeois-Chauvel et al., 1994; Yvert et al., 2001). Furthermore, such evoked 40 Hz gamma band responses have been shown to be modulated by top-down or contextual processes, especially auditory expectations (Widmann et al., 2007; Schadow et al., 2009) of the sort that may build up during the conditioning click train. Thus, the echo suppression mechanism likely reflects recent acoustic context (built up during the conditioning click train) and is manifest, at the latest, in either thalamus or Heschl's gyrus.

This echo suppression mechanism, evidenced by enhanced intertrial phase synchrony, continued in the alpha/beta frequency range from ~50 to 200 ms after stimulus. Although the amount of low-frequency intertrial phase synchronization between trials may be proportional to the amplitude of ERP components, this enhanced phase locking was not evident in the N1

or P2 amplitudes. This suggests that the time–frequency analyses used in this study are more sensitive to fine fluctuations in phase. Furthermore, enhanced phase locking during echo suppression indicates that a higher degree of temporal precision in neural firing, and thus more aligned phase, is likely important for the mediation of full echo suppression, including both spatial translocation and object capture mechanisms. Under this interpretation, if the neural response mediating echo suppression does not reach this level of temporal alignment to the stimulus, suppression will likely fail. The enhanced low-frequency phase locking suggests that the echo suppression mechanism continues through the timeframe of the N1 and P2, during which the resulting neural correlates of perception emerge.

N1 amplitude and latency index stimulus attributes and perception, respectively. The single click contained half the acoustic energy compared with the other conditions, resulting in smaller N1 amplitude. The N1 amplitude thus faithfully encodes the acoustic presence rather than perceptual suppression of the echo. In contrast to amplitude, N1 latency reflects perceptual consequences of echo suppression. Specifically *Single*, *Suppressed*, and *Intermediate* conditions had faster N1 latencies than *Not Suppressed* and *Double* trials. Recall that, when subjects reported the *Intermediate* percept, they suppressed most of the directional information of the echo (translocation) because they heard two objects in the hemifield of the leading click while preserving the echo as a separate object (no object capture). Thus, the *Intermediate* percept provides a clear dissociation between spatial translocation and object capture. In contrast, during *Suppressed* trials, both translocation and object capture occurred, resulting in only one perceived object. However, when subjects detected the echo (*Not Suppressed*), both aspects of suppression failed. In light of this, shorter N1 latency is a categorical marker of spatial translocation. The N1 in the present study is likely an overlap of neural activity from the N1b, which peaks at the vertex at ∼100 ms (Vaughan and Ritter, 1970; Näätänen and Picton, 1987), and the N1c, which reaches its maximum amplitude in temporal electrodes from 138 to 155 ms (Perrault and Picton, 1984; Shahin et al., 2003; Bosnyak et al., 2004). An overlap of the N1b and N1c components explains both the late N1 latency reported here as well as the topography, as shown in Figure 2*A*.

Whereas N1 latency categorically represents spatial translocation, the P2 latency may reflect the additional consequences of object capture. P2 latencies increased with echo lag as follows: *Single* ∼ *Suppressed* ≤ *Intermediate* ≤ *Not Suppressed* < *Double*. Because, in general, the probability of object capture decreases with increasing echo lag, the P2 latency could index the ease or likelihood of capture. However, this view is complicated by the fact that *Intermediate* latency is between *Suppressed* and *Not Suppressed* rather than grouped with *Not Suppressed*. Alternately, the P2 latency may reflect an interaction among stimulus properties, translocation (which was already evident in the N1 latency), and object capture. This integration of perceptual outcomes could result in the graded P2 latency for *Suppressed*, *Intermediate*, and *Not Suppressed*. Recall that the echo is still clearly encoded as a second object in the N1 amplitude, suggesting that capture has not yet occurred by the time spatial translocation is evident. Furthermore, a recent EEG study (Sanders et al., 2008) showed that echo detection, and thus hearing two sounds, was indexed by object-related negativity (ORN). The ORN is a negativity associated with perceiving two versus one sound that overlaps with the N1 and P2 (Alain et al., 2001). Together, the neural mechanisms underlying object capture likely occur after the N1 and thus after spatial translocation.

Our results therefore suggest that spatial translocation temporally precedes and may be necessary for object capture. Furthermore, the processes underlying spatial translocation and object capture may be mediated in communicating yet parallel pathways. This mechanistic dissociation recalls the large-scale organization of the auditory cortical system, which, like the visual system (Ungerleider and Haxby, 1994), has dorsal and ventral pathways for processing spatial and nonspatial information, respectively (Rauschecker, 1998). Consistent with our results, two recent studies suggest that spatial processing in the dorsal auditory pathway occurs faster than acoustic feature analysis in the ventral pathway (Ahveninen et al., 2006; Altmann et al., 2007). Ahveninen et al. (2006) demonstrated that the "where" pathway was activated 30 ms before the "what" pathway, whereas Altmann et al. (2007) showed that changes in sound location were processed ∼100 ms faster than changes in acoustic patterns. Because the brain must know whether a sound is an echo before suppressing its directional information, basic acoustic features (e.g., spectral profile) must be evaluated to accurately label an echo. This rudimentary feature information may be passed from the ventral to dorsal stream so that the brain can spatially translocate the correct echo. Alternately, potential echoes could already be "tagged" in subcortical structures such as the inferior colliculus without ventral cortical stream involvement (Pecka et al., 2007). A third possibility is that the dorsal stream might rely on its own lower-fidelity feature representations to identify the echoes. In any case, the spatial translocation system must influence the final object representation, because we cannot perceive a fused auditory object originating from two distinct spatial locations. Thus, if spatial translocation is successful, the processes underlying object capture continue and usually (but not always) fuse the leading and lagging sounds into one object.

Finally, our design, with a buildup click train on every trial, differs from numerous studies that use only one click pair on each trial (for review, see Litovsky et al., 1999). The buildup of echo suppression strongly depends on short-term auditory context, which informs the brain's "model" or expectations about its acoustical environment (Clifton, 1987; Clifton and Freyman, 1989; Clifton et al., 1994). Thus, measuring electrophysiological markers of echo suppression after a single click pair in isolation would likely reflect not only echo suppression per se but also the buildup process during which perception is not yet stabilized. In the present study, we were interested in echo perception once it had stabilized, that is, once the expectations of the brain had been established. Because these two designs are aimed at different neural processes related to echo suppression, they will likely yield different yet complementary electrophysiological results.

The present study sheds light on the neural time course underlying echo suppression. The first neural marker, enhanced gamma intertrial phase synchronization, begins at ∼20 ms after stimulus onset and reflects the neural mechanism that predicts successful echo suppression. This suppression-specific phase coherence enhancement continues in low frequencies from ∼50 to 200 ms, as correlates of spatial suppression and object capture emerge. By ∼130 ms, suppression of the spatial information of the echo is categorically indexed by the N1 latency. However, the acoustic attributes of the stimuli, specifically the presence of a second "echo" click, is still encoded by the N1 amplitude. Finally, P2 latency may reflect the ease of object capture or an interaction between stimulus properties and perception. Thus, cortical mechanisms underlying echo suppression are evident first, followed by correlates of spatial translocation and, finally, object capture.

References

Ahveninen J, Jaaskelainen IP, Raij T, Bonmassar G, Devore S, Hamalainen M, Levanen S, Lin FH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW (2006) Task-modulated "what" and "where" pathways in human auditory cortex. Proc Natl Acad Sci U S A 103: 14608–14613.

Alain C, Arnott SR, Picton TW (2001) Bottom-up and top-down influences on auditory scene analysis: evidence from event-related brain potentials. J Exp Psychol Hum Percept Perform 27:1072–1089.

Altmann CF, Bledowski C, Wibral M, Kaiser J (2007) Processing of location and pattern changes of natural sounds in the human auditory cortex. Neuroimage 35:1192–1200.

Blauert J (1999) Spatial hearing: the psychophysics of human sound localization. Cambridge, MA: Massachusetts Institute of Technology.

Bosnyak DJ, Eaton RA, Roberts LE (2004) Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. Cereb Cortex 14:1088–1099.

Chau W, McIntosh AR, Robinson SE, Schulz M, Pantev C (2004) Improving permutation test power for group analysis of spatially filtered MEG data. Neuroimage 23:983–996.

Clifton RK (1987) Breakdown of echo suppression in the precedence effect. J Acoust Soc Am 82:1834–1835.

Clifton RK, Freyman RL (1989) Effect of click rate and delay on breakdown of the precedence effect. Percept Psychophys 46:139–145.

Clifton RK, Freyman RL, Litovsky RY, McCall D (1994) Listeners' expectations about echoes can raise or lower echo threshold. J Acoust Soc Am 95:1525–1533.

Damaschke J, Riedel H, Kollmeier B (2005) Neural correlates of the precedence effect in auditory evoked potentials. Hear Res 205:157–171.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21.

Fisher NI (1993) Statistical analysis of circular data. Cambridge, UK: Cambridge UP.

Fitzpatrick DC, Kuwada S, Batra R, Trahiotis C (1995) Neural responses to simple simulated echoes in the auditory brain stem of the unanesthetized rabbit. J Neurophysiol 74:2469–2486.

Freyman RL, Helfer KS, McCall DD, Clifton RK (1999) The role of perceived spatial separation in the unmasking of speech. J Acoust Soc Am 106:3578–3588.

Goldstein R, Rodman LB (1967) Early components of averaged evoked responses to rapidly repeated auditory stimuli. J Speech Hear Res 10:697–705.

Hafter ER, Buell TN, Richards VM (1988) Onset-coding in lateralisation: its form, site and function. In: Auditory function: the neurobiological bases for hearing (Edelman GM, Gall WE, Cowan WM, eds). New York: Wiley.

Hall III JW (1992) Handbook of auditory evoked responses. Boston: Allyn and Bacon.

Holmes AP, Blair RC, Watson JD, Ford I (1996) Nonparametric analysis of statistic images from functional mapping experiments. J Cereb Blood Flow Metab 16:7–22.

Liégeois-Chauvel C, Musolino A, Badier JM, Marquis P, Chauvel P (1994) Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalogr Clin Neurophysiol 92:204–214.

Litovsky RY (1998) Physiological studies of the precedence effect in the inferior colliculus of the kitten. J Acoust Soc Am 103:3139–3152.

Litovsky RY, Shinn-Cunningham BG (2001) Investigation of the relationship among three common measures of precedence: fusion, localization dominance, and discrimination suppression. J Acoust Soc Am 109:346–358.

Litovsky RY, Yin TC (1998a) Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics. J Neurophysiol 80:1285–1301.

Litovsky RY, Yin TC (1998b) Physiological studies of the precedence effect in the inferior colliculus of the cat. II. Neural mechanisms. J Neurophysiol 80:1302–1316.

Litovsky RY, Colburn HS, Yost WA, Guzman SJ (1999) The precedence effect. J Acoust Soc Am 106:1633–1654.

Litovsky RY, Fligor BJ, Tramo MJ (2002) Functional role of the human inferior colliculus in binaural hearing. Hear Res 165:177–188.

Makeig S, Jung TP, Bell AJ, Ghahremani D, Sejnowski TJ (1997) Blind separation of auditory event-related brain responses into independent components. Proc Natl Acad Sci U S A 94:10979–10984.

Näätänen R, Picton T (1987) The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology 24:375–425.

Neves IF, Gonçalves IC, Leite RA, Magliaro FCL, Matas CG (2007) Middle latency response study of auditory evoked potentials' amplitudes and latencies in audiologically normal individuals. Braz J Otorhinolaryngol 73:69–74.

Pecka M, Zahn TP, Saunier-Rebori B, Siveke I, Felmy F, Wiegrebe L, Klug A, Pollak GD, Grothe B (2007) Inhibiting the inhibition: a neuronal network for sound localization in reverberant environments. J Neurosci 27:1782–1790.

Perrault N, Picton TW (1984) Event-related potentials recorded from the scalp and nasopharynx. I. N1 and P2. Electroencephalogr Clin Neurophysiol 59:177–194.

Picton TW, Hillyard SA, Krausz HI, Galambos R (1974) Human auditory evoked potentials. I. Evaluation of components. Electroencephalogr Clin Neurophysiol 36:179–190.

Rauschecker JP (1998) Parallel processing in the auditory cortex of primates. Audiol Neurootol 3:86–103.

Sanders LD, Joh AS, Keen RE, Freyman RL (2008) One sound or two? Object-related negativity indexes echo perception. Percept Psychophys 70:1558–1570.

Schadow J, Lenz D, Dettler N, Fründ I, Herrmann CS (2009) Early gamma-band responses reflect anticipatory top-down modulation in the auditory cortex. Neuroimage 47:651–658.

Shahin A, Bosnyak DJ, Trainor LJ, Roberts LE (2003) Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. J Neurosci 23:5545–5552.

Shahin AJ, Roberts LE, Chau W, Trainor LJ, Miller LM (2008) Music training leads to the development of timbre-specific gamma band activity. Neuroimage 41:113–122.

Spierer L, Bourquin NM, Tardif E, Murray MM, Clarke S (2009) Right hemispheric dominance for echo suppression. Neuropsychologia 47:465–472.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 16:4240–4249.

Ungerleider LG, Haxby JV (1994) "What" and "where" in the human brain. Curr Opin Neurobiol 4:157–165.

Vaughan HG Jr, Ritter W (1970) The sources of auditory evoked responses recorded from the human scalp. Electroencephalogr Clin Neurophysiol 28:360–367.

Wallach H, Newman EB, Rosenzweig MR (1949) The precedence effect in sound localization. Am J Psychol 62:315–336.

Wickesberg RE, Oertel D (1990) Delayed, frequency-specific inhibition in the cochlear nuclei of mice: a mechanism for monaural echo suppression. J Neurosci 10:1762–1768.

Widmann A, Gruber T, Kujala T, Tervaniemi M, Schröger E (2007) Binding symbols and sounds: evidence from event-related oscillatory gamma-band activity. Cereb Cortex 17:2696–2702.

Wightman FL, Kistler DJ (1989) Headphone simulation of free-field listening. I. Stimulus synthesis. J Acoust Soc Am 85:858–867.

Yang X, Grantham DW (1997) Echo suppression and discrimination suppression aspects of the precedence effect. Percept Psychophys 59:1108–1117.

Yin TC (1994) Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. J Neurosci 14:5170–5186.

Yvert B, Crouzeix A, Bertrand O, Seither-Preisler A, Pantev C (2001) Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cereb Cortex 11:411–423.