12-2016

Lateralization of Simulated Sources and Echoes Differing in Frequency Based on Interaural Temporal Differences

Raymond H. Dye Jr.
Loyola University Chicago, rdye@luc.edu

Joseph Boomer
University at Buffalo, The State University of New York

Joleen Frankel

Jacquelyn P. Hill
Loyola University Chicago

Alycia N. Peloquin

Recommended Citation

Lateralization of simulated sources and echoes differing in frequency based on interaural temporal differences[a]

Raymond H. Dye, Jr.,[b] Joseph Boomer,[c] Joleen Frankel, Jacquelyn P. Hill, and Alycia N. Peloquin
Department of Psychology, Loyola University of Chicago, Chicago, Illinois 60660, USA

This study examined listeners' ability to process interaural temporal differences (ITDs) in one of two sequential sounds when the two differed in spectral content. A correlational analysis assessed weights given to ITDs of simulated source and echo pulses for echo delays of 8–128 ms for conditions in which responses were based on the source or echo, a 3000-Hz Gaussian (target) pulse. The other (distractor) pulse was spectrally centered at 1500, 2000, 3000, 4000, or 5000 Hz. Also measured were proportion correct and proportion of responses predicted from the weights. Regardless of whether the echo or source pulse served as the target, target weight, and proportion correct increased with increasing distractor frequency, consistent with low-frequency dominance [Divenyi, J. Acoust. Soc. Am. **91**, 1078–1084 (1992)]. Effects of distractor frequency were observed at echo delays out to 128 ms when the source served as the target, but only out to 64 ms when the echo served as the target. At echo delays beyond 8 ms, recency effects were exhibited with higher proportions correct obtained for judgments based on the echo pulse than the source pulse. © 2016 Acoustical Society of America. [http://dx.doi.org/10.1121/1.4967839]

I. INTRODUCTION

The precedence effect refers to the dominance of earlier arriving sounds over later-arriving ones in determining the position in the horizontal plane for sounds presented in the free field or intracranial position for sounds presented over headphones (e.g., Wallach *et al.*, 1949). It has been demonstrated that later arriving sounds need not be perfect replicas of earlier ones for the precedence effect to occur. For instance, Zurek and Saberi (2003) showed that the extent to which leading pulses dominated later pulses was not affected by their coherence—independent bursts of noise led to as large a precedence effect as perfectly coherent bursts.

Several studies have demonstrated that the precedence effect occurs for conditions in which leading and lagging bursts differ in spectral content. Blauert and Divenyi (1988) and Divenyi (1992) presented stimuli via headphones and measured threshold interaural temporal differences (ITDs) carried by echo pulses as a function of the frequency content of the source pulse and the delay between the simulated source and echo, holding the frequency content of the echo pulse constant. More specifically, Divenyi (1992) fixed the frequency of the echo at 2000 Hz (a narrow band noise having a Gaussian temporal envelope of 5-ms total duration) and varied the center frequency of the simulated source pulse ("conditioner," in Divenyi's terminology) between 500 and 3000 Hz. The conditioner was presented diotically, and

threshold echo ITDs (ΔITD_{echo}'s) were measured as a function of the frequency of the conditioner. The onset-to-onset delays between the conditioner and echo were 2, 5, 10, or 20 ms for conditioner frequencies from 500 to 2000 Hz (up to the echo frequency) and 2 and 5 ms when the conditioner was higher in frequency than the echo. Strong precedence was inferred when there was substantial elevation in ΔITD_{echo} relative to the threshold obtained with a single 2000-Hz pulse presented in isolation. Divenyi found smaller precedence effects for delays of 10 and 20 ms—two of the three listeners yielded ΔITD_{echo}'s that were comparable to those obtained for a single 2000-Hz pulse regardless of the frequency of the conditioner. A third listener produced elevated thresholds at all delays as long as the conditioner was lower in frequency than the echo target. At shorter echo delays, all three listeners showed elevations in ΔITD_{echo} caused by the presence of the diotic conditioner. Furthermore, there was a large effect of conditioner frequency at these shorter echo delays. When the conditioner contained energy at frequencies lower than the echo, ΔITD_{echo} was found to be approximately ten times greater than the threshold obtained for a 2000-Hz pulse alone. Furthermore, the elevations in threshold were greatest when the conditioner was 500-Hz (the lowest conditioner frequency). One might suspect that the findings were due to an upward spread of masking, because it is well known that low frequencies are more effective maskers of high-frequency signals than vice-versa. An argument against this, however, was the finding that a 2000-Hz conditioner did not elevate ΔITD_{echo} to the same extent that a lower frequency conditioner did. Because energetic masking is always greatest for signals and maskers of the same frequency, the elevation in ΔITD_{echo} caused by lower-frequency sources must have a cause other than spectral overlap. One

suggestion is that the dominance exerted by lower frequencies is due to their greater lateralizability (the "localization strength" hypothesis; Divenyi, 1992). It is well known that sensitivity to ITDs at lower frequencies is superior to that at higher frequencies (Divenyi, 1992, verified that threshold ITDs were lower for the pulses creating the greatest elevations in ΔITD_{echo} when used as conditioners).

Shinn-Cunningham et al. (1995) measured the relative weight given to sources and echoes that differed in frequency, confirming low-frequency dominance. Their signals were brief (3-ms Hamming windowed) bursts of narrowband (nominal bandwidth equal to 300 Hz) of noise centered at 450 and 1250 Hz. Because their technique required fused intracranial images, echo delays of 1 ms were used. They found strongest precedence when the low-frequency burst was followed by a high-frequency burst and the weakest for a high-frequency burst followed by a low-frequency burst (most participants actually gave more weight to the lagging event), demonstrating low-frequency dominance.

Yang and Grantham (1997) carried out a similar study in the free field (as opposed to headphone presentation), measuring threshold echo delays for discrimination of echo presentations at 45° vs 55° when the leading pulse was presented at −45°. An adaptive procedure was used, estimating the echo delay required for 70.7% correct discrimination. Stimuli were 5-ms bursts of Gaussian noise that were one-octave wide with center frequencies of 500, 2000, and 3000 Hz. They found that the greatest threshold elevations occurred when the source and echo were the same frequency, e.g., for a 2000-Hz echo pulse, the greatest elevation in threshold was observed when the source pulse was also centered at 2000 Hz (a minimal elevation in threshold was observed when the source pulse was centered at 500 or 3000 Hz). Their results were consistent with a "spectral overlap" hypothesis. Simply stated, this is the hypothesis that greater precedence effects are found when the source and echo overlap in frequency. The reasons for the differences in outcomes for the Divenyi (1992) and Yang and Grantham (1997) studies have remained somewhat mysterious. Yang and Grantham (1997) have argued that the presence of interaural level differences (ILDs) for free-field stimuli might have led to different results.

One goal of the current study was to examine the time interval between the source and echo over which low-frequency dominance is evident. If low-frequency dominance can be demonstrated for temporal separations between source and echo pulses that are beyond the intervals over which temporal masking can be demonstrated, this would be strong support for the "localization strength" hypothesis. At the very least, it would require modifications to the spectral overlap hypothesis such that the intervals of time over which overlap occurred would need to be extended beyond the range of forward and backward masking (typically 20–40 ms, for impulsive stimuli, Raab, 1961).

Our laboratory has reported deleterious effects of later arriving sounds on the ability to extract ITDs from earlier arriving probes (Stellmack et al., 1999; Dye et al., 2006). It seemed important to determine the extent to which pulses of different frequency interact under conditions that produce so-called recency effects (Dye et al., 2006; Stecker and Hafter, 2009). By recency effects, we refer to a superiority of performance when judgments are based on the last events in a sequence. Recency effects in lateralization tasks are frequently accompanied by a greater reliance (i.e., weighting) of ITDs or ILDs carried by the end of a sequence of pulses (Stecker and Hafter, 2009). Thus, another goal was to determine whether or not the low-frequency dominance observed when one attempts to lateralize simulated echo pulses (conditions demonstrating precedence effects) would also be found for conditions in which lateralization is based on the source pulse (conditions demonstrating recency effects). As in Stellmack et al. (1999), the procedure required participants to lateralize the first or second pulse, ignoring the other. Source and echo weights were derived using point-biserial correlations between source and echo ITDs and responses (Lutfi, 1995) for conditions in which listeners were instructed to use information from the first pulse as well as for conditions in which the second pulse served at the target. The focus of the current study is on echo delays that are from 8 to 128 ms. Because conditions were included in which lateralization was to be based on the first pulse, potential frequency effects were examined for conditions that lead to recency effects as well as precedence effects in lateralization.

It should be noted that the shortest echo delay at which data were collected (8 ms) is sufficiently long so that effects are unlikely to be cochlear in origin. First, amplitudes of auditory brainstem responses evoked by sequences of two impulses showed responses to the second to be unaffected by the first once the temporal separation exceeded 4–5 ms (Damaschke et al., 2005; Bianchi et al., 2013). This 5-ms window is consistent with computational modeling work of Xia et al. (2010): simulated medial superior olive responses to lagging pulses at a cell's best ITD are unaffected by preceding pulses for echo delays longer than 5 ms. Because of these findings, there will be no effort in the current paper to model effects by considering cochlear interactions between leading and lagging pulses à la Hartung and Trahiotis (2001). Xia et al. (2010) found simulated inferior colliculus responses to lagging pulses, to be influenced over much larger temporal windows (out to 20 ms), as were actual responses as recorded by Litovsky and Yin (1998). Thus future modeling efforts and neurophysiological investigations of the phenomena described in this paper should likely focus on the inferior colliculus and beyond.

II. METHODS

All procedures including recruitment, consenting, and behavioral testing, were approved by Loyola University's Lakeside Institutional Review Board for the Protection of Human Subjects. Each trial consisted of two intervals. During the first, a 3000-Hz diotic click was presented that marked the intracranial midline and the pitch of the target for the listeners. After 350 ms of silence, the observation interval, consisting of a dichotic sequence of two clicks, was presented. The interaural delays of the lead and lag clicks were independently selected from Gaussian distributions

with means of $0 \mu s$ and standard deviations of $100 \mu s$. The echo delay is defined as the interval of time between the onset of the leading source click and the onset of the leading echo click (as was the case in Divenyi, 1992). Data were gathered at echo delays of 8, 16, 32, 64, and 128 ms for conditions in which the first pulse, the simulated source, served as the target. Data were collected at these echo delays plus 72 ms for conditions in which the second pulse, the simulated echo, served as the target.

Pulses were generated by applying a Gaussian window of 4 ms duration to a sinusoid using the Gauspuls function of MATLAB. The pulses were truncated once the amplitude of the envelope was -60 dB relative to the peak. The center frequency of the pulses was 1500, 2000, 3000, 4000, or 5000 Hz. The bandwidth of the pulses in the frequency domain was 60% of the center frequency at the 6-dB down points. For example, a 3000-Hz pulse had a bandwidth of 1800 Hz, with 6-dB down points at 2100 and 3900 Hz. Stimuli were played at a sampling rate of 100 000 Hz through 24-bit Echo Gina 3 G soundcards under the control of Dell PCs that were used for stimulus generation and experimental control. Interaural delays were presented at integer multiples of the sampling rate, leading to 10-μs resolution. Although the means of both ITD_{source} and ITD_{echo} were $0.0 \mu s$, minimum magnitudes of at least $10 \mu s$ were presented so that there would always be a "correct" response. Stimuli were presented through Sennheiser HD 280 Pro headphones, with the subjects seated in a double-walled sound-attenuating booth. Intensities were set by presenting a continuous train (100/s) of equal-amplitude pulses through the headphones, with the voltage (measured by a Fluke PM 2525 digital root mean square voltmeter) set to produce an overall level of 55 dBA sound pressure level as measured by a type 1 sound-level meter using the slow setting.

Source target: For conditions in which the source pulse was to be lateralized, listeners were instructed to use the information in the first click, ignoring the echo click, and to press one button on a keyboard if the source appeared to the left of midline and another key if the source appeared to the right, using the 3000-Hz diotic cue presented in the first interval as a reference.

Echo target: For conditions in which the echo pulse served as the target, listeners were instructed to use the information carried by the final event, ignoring the first click; again, responses were entered on a keyboard that indicated the laterality of the echo pulse.

For comparison to the conditions in which the source or echo served as the target, baseline conditions were run in which a single click was presented alone in the observation interval (without an echo/source click). Feedback was provided on a trial-by-trial basis for all conditions. For the single-click conditions, the ITDs were also chosen randomly from trial-to-trial from a Gaussian distribution with a mean of $0 \mu s$ and a standard deviation of $100 \mu s$, as they were for the two-pulse conditions.

Prior to each block of trials, listeners were allowed to listen to as many practice trials as they desired. These trials had the same echo delay and source and echo frequencies as the test trials with interaural delays varying from trial to trial. Listeners were instructed to adjust the headphones during the practice trials so that the diotic click during the first interval sounded intracranially centered. When ready, listeners initiated a block of 50 test trials by pressing a particular key on the computer keyboard. Listeners were run in 1.0-h sessions during which 500 to 750 left-right judgments were made. All trials of a particular condition (echo delay/ source frequency/echo frequency) were completed before moving on to a new condition. The order in which conditions were run was random for each individual subject except for the 72 ms echo delay condition, which was run last for all four listeners. The 72-ms condition was added after finding rather large changes in performance between the 64-ms and 128-ms conditions. The plan was to collect additional data at several echo delays between 64 and 128 ms. Because an echo delay of 72 ms yielded echo weights and proportion corrects that were substantially larger than those measured at 64 ms, no other echo delays were added.

Five listeners participated in the conditions in which the source served as the target, and four participated in the conditions in which the echo served as the target. One (S1) ran in both sets of conditions, for a total of eight participants. The one listener (a 21-year old male) who participated in both sets of conditions ran with the source as the target first. Three of the eight listeners had extensive experience in other psychoacoustic tasks, but none had performed in a lateralization experiments prior to the current study. All eight listeners reported normal hearing in both ears. Three of four who lateralized echo pulses were male. Two of five who judged sources were male. Ages of the eight participants ranged from 20 to 25 years. All were naive to the hypotheses under investigation and all were given at least 10 000 trials of training before data collection began. Listeners were paid an hourly wage for their participation. Participants began by making lateralization judgments for a single 3000-Hz Gaussian pulse. It was required that proportion correct be at least 0.80 for single clicks (computed over 200 trials) before moving on to lateralization training sessions with two clicks. Once this criterion was achieved for single 3000-Hz pulses, additional training blocks of 200 trials were run at each of the echo delays for conditions in which the non-target pulse was 1500, 2000, 3000, 4000, and 5000 Hz. The target pulse was fixed at 3000 Hz, whether it was the source or echo. Once the practice trials were completed, data collection commenced. One thousand judgments were made for each condition.

Performance was assessed in three ways. First, simple proportion correct [P(C)] was computed. This is the proportion of trials yielding responses consistent with the interaural delay of the target click. Second, the point-biserial correlation between the response made by the listener and the interaural delay of the "source" was computed, along with the point-biserial correlation between the response and the interaural delay of the "echo" (over 1000 trials). From these two correlations, normalized source and echo weights, W_{source} and W_{echo}, were computed as a measure of relative influence of the source and echo on judgments,

$$W_{Source} = \frac{R_{resp,ITD_{Source}}}{|R_{resp,ITD_{Source}}| + |R_{resp,ITD_{Echo}}|};$$

$$W_{Echo} = \frac{R_{resp,ITD_{Echo}}}{|R_{resp,ITD_{Source}}| + |R_{resp,ITD_{Echo}}|}, \qquad (1)$$

where $R_{resp,ITD_{Source}}$ and $R_{resp,ITD_{Echo}}$ are the point-biserial correlations between the left-right responses and the interaural delays of the source and echo clicks, respectively. The magnitudes of the two relative weights must sum to 1.0. Optimal performance would be obtained when all weight is given to the target ITD and none to the other ITD. For instance, when responses were to be based on the interaural delay of the source, optimal performance would be represented by $W_{source} = 1.0$ and $W_{echo} = 0.0$. Third, we predicted subjects' responses (retroactively) with a simple detection-theoretic model. From the two measured weights that were estimated over 1000 trials, a decision variable (D) was computed for each trial as a linear combination of source and echo ITDs:

$$D = W_{Source}ITD_{Source} + W_{Echo}ITD_{Echo} + \varepsilon, \qquad (2)$$

where ε represents prediction error. We then attempted to predict the responses on each individual trial based on the magnitude of D such that a listener was predicted to respond "left" if $D < 0.0$, "right" if $D > 0.0$, and randomly if $D = 0.0$, because there was no evidence for left-right response bias. Note that negative ITDs lead at the left ear while positive lead at the right. The success of this simple model was assessed by computing the proportion of subjects' responses that were correctly predicted by the decision variable in Eq. (2). This was the third measure of performance, designated as P(Responses Predicted) in figures. High values of P(Response Predicted) demonstrate that weight given to the non-target pulse is behaviorally relevant and serve to validate that proposition that responses are based on a weighted average of the binaural cues of the source and echo pulses.

III. RESULTS

Echo target: First, the results of conditions in which the echo served as the target pulse will be presented. Figure 1(A) shows echo weight as a joint function of echo delay and source frequency, averaged across the four participants. Echo weights increased as the frequency of the source was increased, particularly at the shorter echo delays. At echo delays of 8, 16, and 32, more weight was actually given to the source when it was lower in frequency than the 3000-Hz echo (indicated by echo weights below 0.5). For echo delays of 128 ms, average echo weights exceeded 0.85 for all source frequencies.

Figure 1(B) shows proportion correct plotted as joint function of echo delay and source frequency, also averaged across the four listeners. As was the case for echo weight, proportion correct increased with source frequency, with the effect greatest at shorter echo delays and rather small at longer echo delays. The similarity of these two functions is not mere coincidence—as one gives greater weight to the target pulse, proportion correct is expected to increase as long as the prediction error term (ε) stays constant. For comparison,

FIG. 1. (Color online) Plots of echo weight (A), proportion correct (B), and proportion of responses predicted from source and echo weights (C) are plotted as a function of the source frequency. Separate functions are presented for the different echo delays. Error bars in each panel show ± 1 standard error of the mean.

the proportion correct is shown as a horizontal dashed line for a single 3000-Hz pulse with an ITD picked randomly from a Gaussian distribution with a mean of 0.0 and standard deviation of $100\,\mu$s. Proportion correct approached this value (0.8935) when the echo delay reached 128 ms and the source frequency was 5000 Hz, falling just short.

Figure 1(C) shows the average proportion of responses predicted from source and echo weights [Eq. (2)] as a joint function of echo delay and source frequency. Although there was a trend for the proportion of response predicted correctly to increase with source frequency, this effect did not reach statistical significance as revealed by a repeated measures analysis of variance (5 source frequencies × 6 echo delays) carried out on arcsine transformed versions of the proportion of responses correctly predicted. This transform was used in order to stabilize the variances (Collett, 1991) because this dependent variable is limited to a range of from 0 to 1.0. For source frequency, $F(4,12) = 2.58$, $p = 0.091$ and for echo

delay, $F(5,15) = 1.48$. $p = 0.25$, both failing to reach significance at $\alpha = 0.05$. There was no indication of an interaction between source frequency and echo delay, $F(20,60) = 0.525$, $p = 0.94$.

In the interest of brevity, only means were presented for all three dependent variables (echo weight, proportion correct, and proportion of responses predicted from weights). Individual weights, proportions correct, and proportion of responses predicted from the weights are presented in Tables I, II, and III of the Appendix. The participants for the conditions in which judgments were based on echo pulses are identified as S1, S2, S3, and S4. Individual differences were generally rather small, particularly in terms of qualitative patterns in the data. One of the four participants yielded a slightly different pattern. In particular, he showed nearly equivalent echo weights and proportions correct for echo delays of 72 and 128 ms (whereas the other three showed higher values at 128 than 72 ms). This same participant (S3 in Tables I, II, and III of the Appendix) also showed higher echo weights and proportions correct when the echo was 1500 Hz rather than 2000 Hz for echo delays of 8 to 64 ms. The other three all had lower values at 1500 Hz than 2000 Hz at these echo delays.

Repeated measures analyses of variance (5 source frequencies × 6 echo delays) were also carried out to examine the effects of echo delay and source frequency on the echo weight and proportion correct. The analysis of variance was carried out on arcsine transformed versions of the dependent variables because these dependent variables are also bounded by 0.0 and 1.0.

Echo weight: There were significant main effects of echo delay [$F(5,15) = 27.66$, $p < 0.0001$, $\eta^2 = 0.452$ (η^2 is the proportion of total variance accounted for by the independent variable)] and source frequency [$F(4,12) = 18.50$, $p < 0.0001$, $\eta^2 = 0.207$], as well as a significant interaction between source frequency and echo delay [$F(20, 60) = 7.74$, $p < 0.0001$, $\eta^2 = 0.094$]. It was generally the case that larger echo weights were observed at longer echo delays and higher source frequencies, with source frequency having a greater effect on echo weight when the echo delays were shorter. The significant interaction was further investigated with tests of simple main effects of source frequency. Significant effects of source frequency were found when the echo delay was 8 ms [$F(4,12) = 39.68$, $p < 0.0001$], 16 ms [$F(4,12) = 11.69$, $p < 0.0001$], 32 ms [$F(4,12) = 20.80$, $p < 0.0001$], 64 ms [$F(4,12) = 6.87$, $p = 0.0041$], but not for echo delays of 72 and 128 ms [$F(4,12) = 0.966$, $p = 0.4612$ and $F(4,12) = 1.534$, $p = 0.2543$, respectively]. Clearly, the elevation of echo weight with source frequency was larger at shorter echo delays.

Proportion correct: There were significant main effects of echo delay [$F(5,15) = 24.69$, $p < 0.0001$, $\eta^2 = 0.311$ and source frequency [$F(4,12) = 17.85$, $p < 0.0001$, $\eta^2 = 0.273$], as well as a significant interaction between source frequency and echo delay [$F(20, 60) = 9.90$, $p < 0.0001$, $\eta^2 = 0.090$]. It was generally the case that greater proportions correct were observed at longer echo delays and higher source frequencies, with source frequency having a greater effect on proportion correct when the echo delays were shorter. Again,

the significant interaction was investigated further with tests of simple main effects of source frequency. Significant effects of source frequency were found when the echo delay was 8 ms [$F(4,12) = 62.41$, $p < 0.0001$], 16 ms [$F(4,12) = 21.82$, $p < 0.0001$], 32 ms [$F(4,12) = 12.74$, $p = 0.0003$], 64 ms [$F(4,12) = 5.316$, $p = 0.011$], but not for echo delays of 72 and 128 ms [$F(4,12) = 2.146$, $p = 0.1375$ and $F(4,12) = 1.92$, $p = 0.1725$, respectively]. As was the case for echo weight, the elevation of proportion correct with source frequency was larger at shorter echo delays.

Source target: Fig. 2(A) shows source weight as a joint function of echo delay and echo frequency, averaged across five listeners. Source weights increased as the frequency of the echo was increased. At echo delays of 8 to 64, more weight was given to the echo when it was below the frequency of the 3000-Hz source.

Figure 2(B) shows average proportion correct plotted as joint function of echo delay and echo frequency. As was the case for source weight, proportion correct increased with echo frequency. Proportion correct was highest for echo
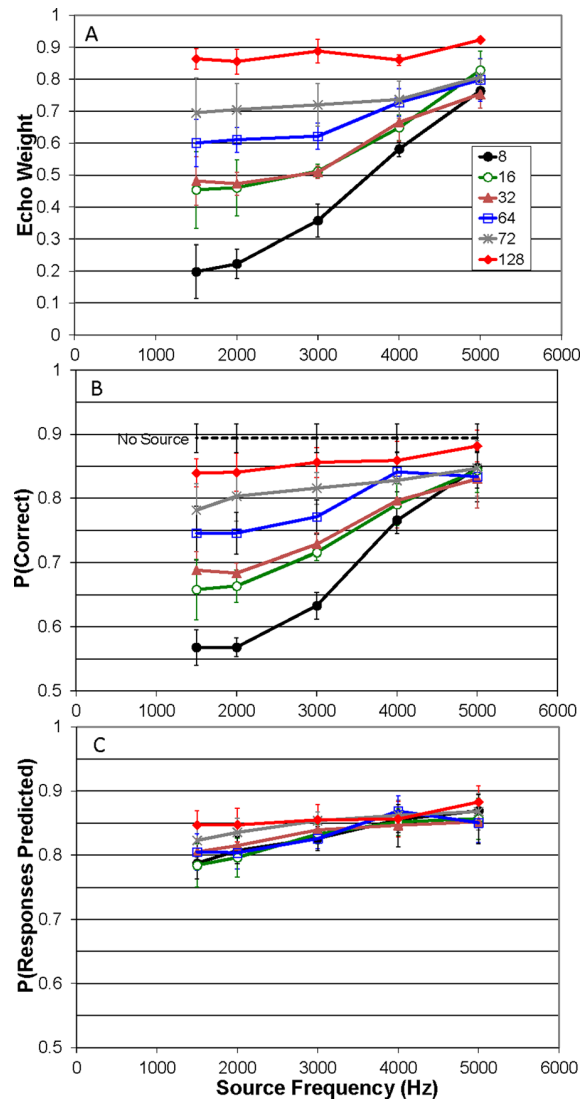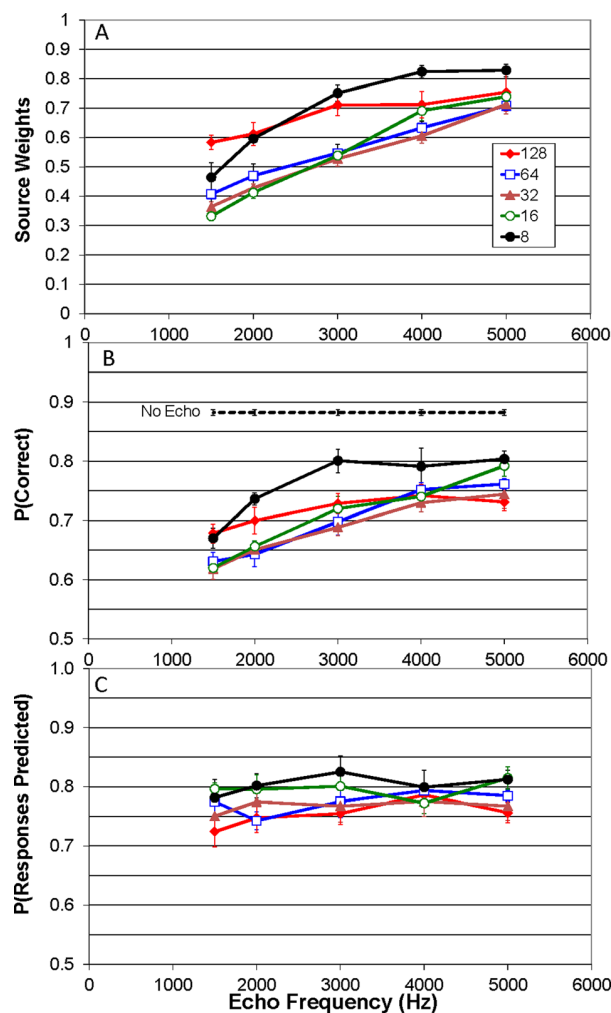

FIG. 2. (Color online) Plots of source weight (A), proportion correct (B), and proportion of responses predicted from source and echo weights (C) are plotted as a function of echo frequency. Separate functions are presented for the different echo delays. Error bars in each panel show ±1 standard error of the mean.

delays of 8 ms, never reaching the value (0.874) obtained for single 3000-Hz pulses (the dashed horizontal line in the panel).

Figure 2(C) shows the proportion of responses predicted from source and echo weights [Eq. (2)] as a joint function of echo delay and echo frequency, averaged across five participants. There was a trend for the proportion of response predicted correctly to increase as the echo delay was made shorter, but this effect did not reach statistical significance as revealed by a repeated measures analysis of variance (5 echo frequencies × 5 echo delays) carried out on arcsine transformed versions of the proportion of responses correctly predicted. For echo delay, $F_{(4,16)} = 2.58$, $p = 0.0774$ and for echo frequency, $F_{(4,16)} = 2.81$, $p = 0.0606$. The interaction between echo frequency and echo delay also did not reach significance, $F_{(16,64)} = 1.358$, $p = 0.192$.

For conditions in which judgments were based on ITDs carried by the source, mean data for source weight and proportion correct reflect the findings for each of the five participants. Individual source weights, proportions correct, and proportions of responses predicted from the weights are presented in Tables IV, V, and VI of the Appendix. Participants in these conditions are identified as S1, S5, S6, S7, and S8. Two participants (S1 and S5), yielded lower proportions of responses predicted from the weights for echo delays of 128 ms than at shorter values. The lower levels of P(Responses Predicted) for these two listeners were found for all echo frequencies.

Repeated measures analyses of variance (5 source frequencies × 5 echo delays) were also carried out to examine the effects of echo delay and source frequency on arcsine transformed source weights and proportions correct.

Source weight: There were significant main effects of echo delay [$F_{(4,16)} = 14.87$, $p < 0.0002$, $\eta^2 = 0.210$ and echo frequency [$F_{(4,16)} = 86.94$, $p < 0.0001$, $\eta^2 = 0.562$], as well as a significant interaction between echo delay and echo frequency [$F_{(16, 64)} = 3.26$, $p = 0.0004$, $\eta^2 = 0.051$]. It was generally the case that larger source weights were observed at echo delays of 8 and 128 ms and for higher echo frequencies. The significant interaction was further investigated with tests of simple main effects of echo frequency. In spite of the interaction, significant effects of echo frequency were found at all echo delays; 8 ms: $F_{(4,16)} = 27.70$, $p < 0.0001$; 16 ms: $F_{(4,16)} = 62.69$, $p < 0.0001$; 32 ms: $F_{(4,16)} = 36.57$, $p < 0.0001$; 64 ms: $F_{(4,16)} = 39.49$, $p < 0.0001$, 128 ms: $F_{(4,16)} = 5.43$, $p = 0.0059$. It may seem paradoxical that greatest source weight was given at the shortest and longest echo delays—at an echo delay of 8 ms. If the precedence effect were still in effect, it might result greater weight given to the source pulse. On the other hand, the two pulses are likely to be perceptually segregated at 128 ms, so the first pulse could be given higher weight appropriate with the demands of the task.

Proportion correct: There were significant main effects of echo delay [$F_{(4,16)} = 7.78$, $p = 0.0011$, $\eta^2 = 0.160$ and echo frequency [$F_{(4,16)} = 44.02$, $p < 0.0001$, $\eta^2 = 0.489$], as well as a significant interaction between source frequency and echo delay [$F_{(16, 64)} = 2.63$, $p = 0.0032$, $\eta^2 = 0.068$]. It was generally the case that greater proportions correct were observed at echo delays of 8 and 128 ms and higher echo

frequencies, with echo frequency having a greater effect on proportion correct at echo delays that were between 16 and 64 ms. Again, the significant interaction was investigated further with tests of simple main effects of echo frequency. Significant effects of echo frequency were found at all echo delays; 8 ms: $F_{(4,16)} = 16.23$, $p < 0.0001$; 16 ms: $F_{(4,16)} = 26.75$, $p < 0.0001$; 32 ms: $F_{(4,16)} = 14.11$, $p < 0.0001$; 64 ms: $F_{(4,16)} = 17.58$, $p < 0.0001$, 128 ms: $F_{(4,16)} = 3.075$, $p = 0.0459$.

Figures 3 and 4 present scatter plots for two typical participants so that the reader may gain a better appreciation of the consequences of differential weighting of source and echo clicks on responses. Different panels in each figure present responses for different distractor frequencies. Each panel presents the three dependent variables: target weight, proportion correct, and the proportion of responses predicted from the weights. Figure 3 shows data for a condition in which responses were to be based on the 3000-Hz second (echo) pulse and Fig. 4 shows responses for a condition in which responses were to be based on the 3000-Hz first (source) pulse. Figure 3 shows data that were collected at an echo delay of 8 ms from S4, and Fig. 4 shows data collected at an echo delay of 32 ms for S1. These conditions were chosen because they were associated with relatively large changes in target weight with distractor frequency. In both figures, left and right responses are segregated by a line with a slope equal to $-W_{Target}/W_{Distractor}$, where target and distactor refer to the click participants attempt to attend to and ignore, respectively. The proportion of responses predicted from the weights is graphically represented by the proportion of right responses above the boundary and left responses below it. Proportion correct is represented by the proportion of left response (filled diamonds) to the left of a vertical line running through 0,0 and right responses (open squares) to the right of a vertical line through the origin.

The response profiles shown in Fig. 3 reveal that, at an echo delay of 8 ms, the responses were nearly completely dependent on the interaural temporal difference carried by the source when it was lower in frequency that the target echo. As such, the boundary between left and right responses was nearly flat, reflecting the fact that ITD_{echo} was given little weight. When both the source and echo were centered at 3000 Hz, the listener gave less weight to the target echo than the source, reflected by the slope being shallower than -1.0. Panels for source frequencies greater than the 3000 Hz echo show boundaries that were steeper than -1.0, reflecting the fact that listeners were giving greater weight to ITD_{echo} than to ITD_{source}. The response profiles of this individual participant are quite representative, although this listener gave even less weight to the echo when the source was lower in frequency than was typical.

Response profiles depicted in Fig. 4 (for an echo delay of 32 ms) show a more gradual change in target (source) weight with echo frequency. As echo frequency increased from 1500 to 3000 Hz, target weight only went from 0.38 to 0.473 for this particular listener (S4). Once the echo frequency was increased to be greater than the target (source) frequency, target weight increased to 0.578 then to 0.652 as echo frequency was 4000 then 5000 Hz.

J. Acoust. Soc. Am. **140** (6), December 2016

Dye, Jr. *et al.* 4477

FIG. 3. (Color online) Scatter plots of left-right responses are shown for one listener (S4) that were obtained at an echo delay of 8 ms for conditions in which the echo served as the to-be-judged target. Each panel depicts data obtained for different source frequencies, from 1500 Hz on the upper left to 5000 Hz on the lower right. Open boxes represent "right" responses and filled diamonds represent "left" responses. The line represents the best-fitting linear boundary between left and right responses that passes through the origin.

IV. DISCUSSION

A. Source vs echo performance

In order to compare conditions in which lateralization was based on the source and echo pulse, results will initially be discussed in terms of the target and distractor. In spite of the rather different visual impression produced by the plots of target weights and proportions correct for the two sets of data (compare 1a with 2a and 1b with 2b), the statistical results were nearly identical in terms of main
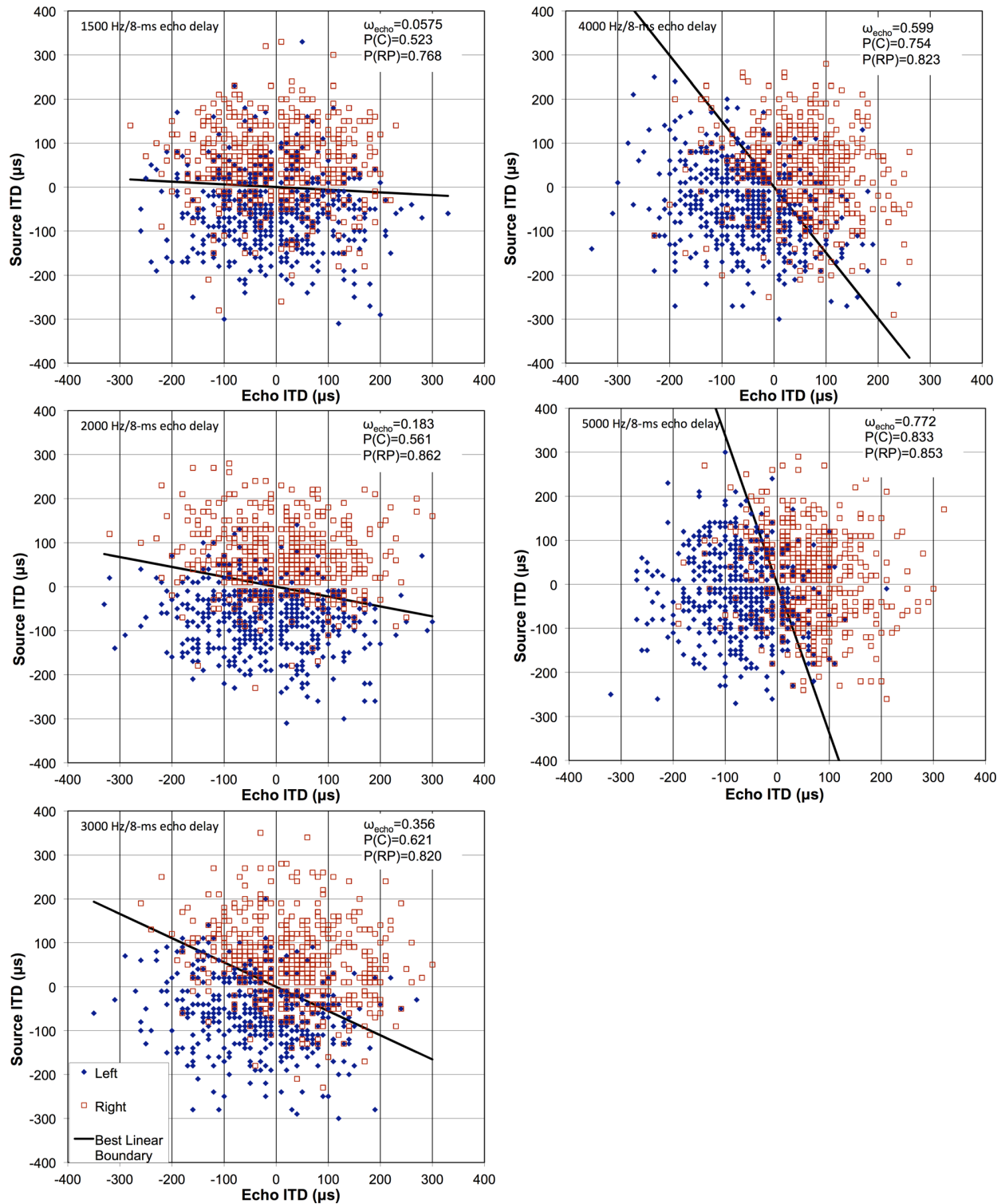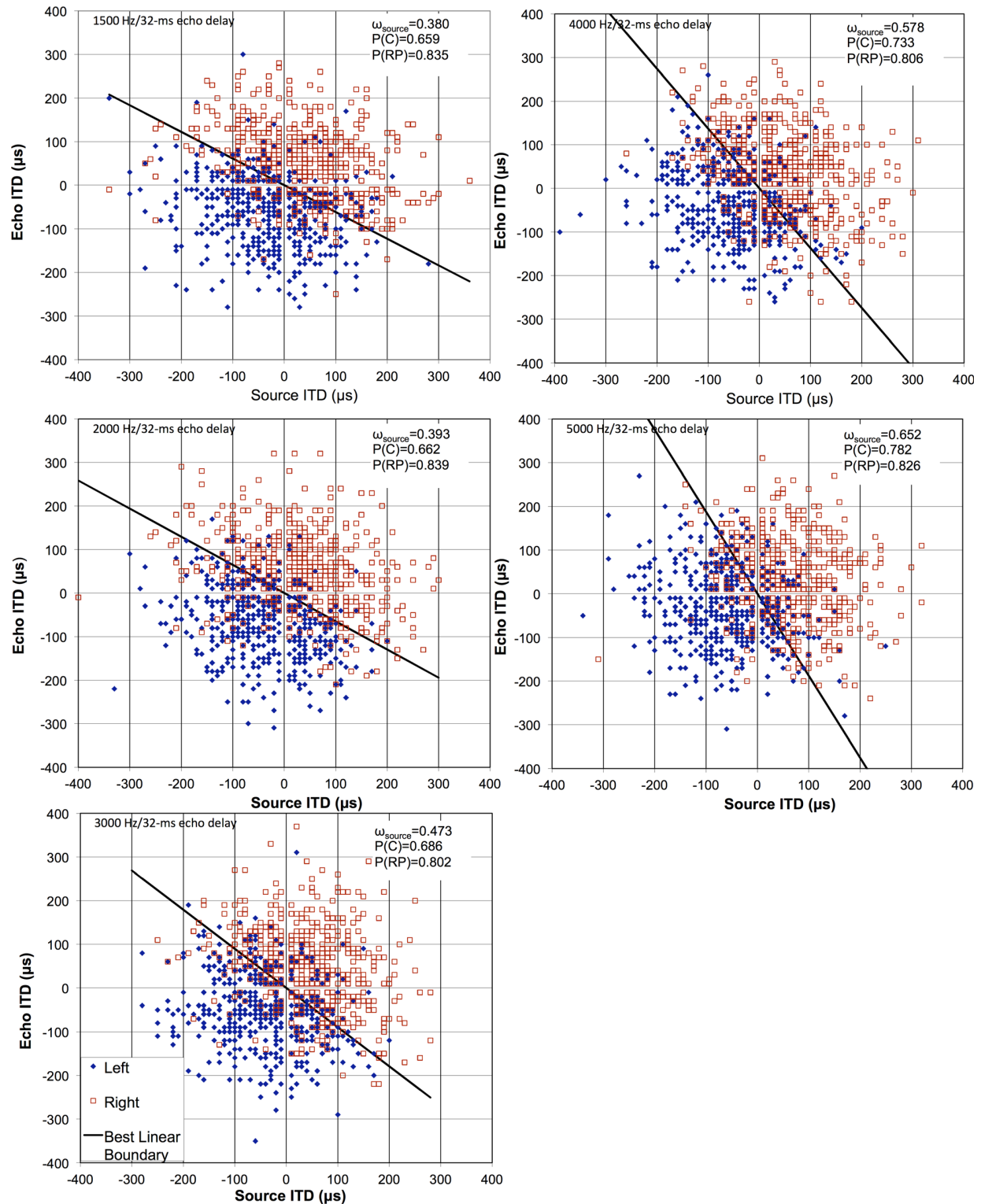
FIG. 4. (Color online) Scatter plots of left-right responses are shown for one listener (S1) that were obtained at an echo delay of 32 ms for conditions in which the source served as the to-be-judged target. Each panel shows scatter plots of data obtained at different echo frequencies, from 1500 Hz (upper left) to 5000 Hz (lower right). As in Fig. 3, open boxes represent "right" responses, filled diamonds represent "left" responses, and the line represents the best-fitting linear boundary between left and right responses that passes through the origin.

effects and interactions. This is not to say, that the results themselves were identical; a quick glance at panels 1b and 2b shows that performance is quite good at an echo delay of 8 ms when the source served as the target (2b) and quite poor when the echo served as the target (1b), particularly

when the distractor frequency was lower than the 3000-Hz target frequency.

In addition, the statistical results were identical regardless of whether the target weight or proportion correct was used as the dependent variable (main effects of distractor

J. Acoust. Soc. Am. **140** (6), December 2016

Dye, Jr. *et al.* 4479

frequency and echo delay, the interactions *and* the analysis of simple effects). As such, effects on target weight and proportion correct will be described together. As was pointed out earlier, this is to be expected—if a listener gives more weight to the to-be-judged target, proportion correct should be higher.

With regard to the effect of distractor frequency, both target conditions (source and echo) yielded higher target weights and proportions correct when the 3000-Hz target was lower in frequency than the distractor. Both target conditions exhibit a dominance of the lower frequency. The only difference between the conditions in which the to-be-judged pulse was the source vs the echo was revealed in the analysis of simple effects that explored the nature of the frequency-by-echo delay interactions. For cases in which the source served as the target, frequency produced significant effects at all echo delays out to 128 ms. When the echo served as the target, significant effects of frequency were found at echo delays out to 64 ms but not at 72 and 128 ms.

Both target conditions also yielded significant main effects of echo delay, although the effects were somewhat different for the two conditions. When the echo served as the target, greatest target weight and proportion correct were found at the longest echo delays. When the source served as the target, highest target weights and proportions correct were produced at the shortest (8 ms) and longest (128 ms) echo delays.

The current study found no effects of echo delay (nor distractor frequency) on the proportion of responses predicted from the weights, P(Responses Predicted). This is typically used as a measure of the extent to which responses can be explained by the ITDs of the source or echo, with lower values associated greater randomness of responses [larger values of ε in Eq. (2)]. In response profiles like those depicted in Figs. 3 and 4, lower values of P(Responses Predicted) lead to a greater number of misclassified response ("left" above and "right" below the best-fitting linear boundary). Stellmack *et al.* (1999) found effects for both target conditions, with smallest P(Responses Predicted) for short echo delays when the echo served as the target and intermediated

echo delays (16–64 ms) when the source served as the target. It should be noted, however, that Stellmack *et al.* varied echo delay from 1 to 256 ms, while the current study only collected data at echo delays between 8 and 128 ms. As such, the current study excluded echo delays producing variation in P(Responses Predicted), so it is hardly surprising that the effect of echo delay did not reach statistical significance.

To facilitate a direct comparison of performance based on sources vs echoes, the proportions correct from Figs. 1(B) and 2(B) are re-plotted in Fig. 5 with echo delay on the abscissa and the distractor frequency as the parameter within the figure. The labels "Judge Echo" (on the left) and "Judge Source" (on the right) indicate the pulse participants attempted to judge. First examine the functions for conditions in which both pulses were spectrally centered at 3000 Hz (the circles). At an echo delay of 8 ms, approximately 80% correct was achieved when the source was the target (right panel), while only 63% correct was obtained when judgments were based on the echo (left panel). At an echo delay of 16 ms, performance was at 72% correct regardless of whether the source or echo served as the target. For longer echo delays, superior performance was obtained for judgments based on the echo, with the difference growing larger with increasing echo delay. Stellmack *et al.* (1999), Litovsky and Shinn-Cunningham (2001), and Goupell *et al.* (2012) have observed this same trend. All found superior performance based on the leading source at delays shorter than 10 ms and better performance based on the lagging echo at longer echo delays [out to 256 ms in Stellmack *et al.* (1999)]. Goupell *et al.* (2012) extended this finding to three-pulse stimuli in which judgments were to be based on the first, second, or third. They reported best performance based on the first when the interval between the first and second pulse was less that 10 ms (as long as the lag between the second and third pulses was also less than 10 ms) and best performance based on the third when the interval between the second and third pulse exceeded 10 ms (regardless of the lag between the first and second pulse).

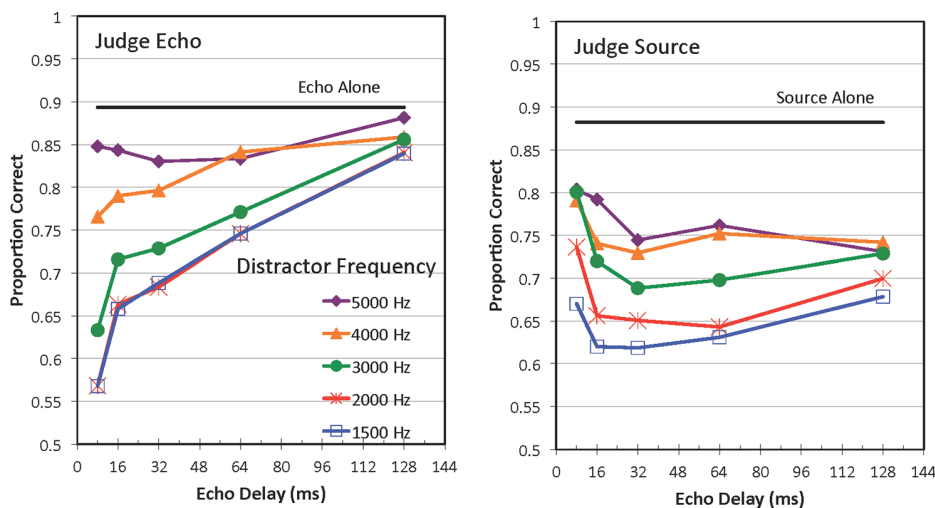FIG. 5. (Color online) Proportions correct [from Figs. 1(B) and 2(B)] plotted as a function of echo delay. The left panel shows conditions in which the echo served as the target; the right panel shows conditions in which the source served as the target. Distractor frequency is the parameter within each panel. For reference, performance for echo alone and source alone are shown as horizontal solid lines.

The current data extend the finding to conditions in which the two pulses differed in spectral content. A comparison of the proportions correct for the two target conditions revealed better performance when the echo pulse served as the target once echo delay exceeded 16 ms. For example, for an echo delay of 64 ms, performance ranged from 75% to 84% as the source (distractor) frequency was varied from 1500 to 5000 Hz when the echo served as the target (left panel). When the source pulse served as the target (right panel), percent correct ranged from 63% to 76% for an echo delay of 64 ms, nearly 9% to 10% worse. Although listeners in the two experiments were not the same, performance was only slightly worse for the listeners judging the 3000-Hz source (88.23%) than those judging the echo (89.35%) when single pulses were presented. At shorter echo delays (16 and 32 ms), the differences in proportion correct between the two conditions were smaller, while they were larger at the largest echo delays (128 ms), showing the same trends as when the source and echo were both 3000 Hz. Only at an echo delay of 8 ms—a delay at which onset dominance is likely to be in effect (Buell and Hafter, 1991)—was performance better when the source served as the target (except when the distractor frequency was 5000 Hz). It appears that once one gets beyond echo delays that produce a precedence effect, performance is actually better when judging the echo pulse than the source pulse. This superior performance when judgments are based on information carried by last acoustic event has been termed a binaural "recency effect" (Dye *et al.*, 2006; Stecker and Hafter, 2009; Goupell *et al.*, 2012).

B. Potential mechanisms for recency

To some extent, the poorer performance when the information bearing pulse is the first one could be due to temporal order confusion (e.g., Hirsh, 1959; Hirsh and Sherrick, 1961). It has been suggested that listeners fail to resolve the temporal order associated with the two intracranial images, one associated with the source pulse and one associated with the echo pulse (Stellmack *et al.*, 1999; Litovsky and Godar, 2010). Resolving the spatial location of the two intracranial images is not sufficient for performing the task. Perhaps listeners in Stellmack *et al.* (1999) and Goupell *et al.* (2012) formed separate intracranial images for the source and echo clicks when echo delays exceeded 8 ms but were confused as to which arose first. Indeed, Stellmack *et al.* (1999) found that performance based on source pulses substantially improved when the ITD carried by the echo was fixed at 0 μs so that temporal order no longer needed to be resolved. In these conditions, listeners simply responded according to the laterality of any intracranial image not at the midline. It should be emphasized that merely being confused about temporal order would not, necessarily, lead to superior performance based on the echo. There must also be a tendency to report the characteristics of the last perceptual event! Massaro *et al.* (1976) demonstrated this propensity in localization experiments. In their study, listeners were to lateralize a 20-ms 1000 Hz tone on the basis of an interaural level difference (ILD) that was either −4 dB or +4 dB. This target

tone was followed on most trials by a "backward interference" tone that was also 1000 Hz and 20 ms in duration. The interval between the target tone and the backward interferer was varied between 0 and 250 ms. The ILD of the interferer was −8, 0, or +8 dB, where negative ILDs were more intense at the left ear and positive ILDs were more intense at the right ear. Performance was poorer when interfering tones followed the target out to 180 ms, particularly when the subsequent interferer was contralateral to the target tone. When the interferer and target carried ILDs that favored opposite ears, responses were generally consistent with the side of the backward interferer at lags of 20 and 40 ms, leading to proportions correct that were below 50% because listeners were responding according the laterality of the last event.

Besides temporal order confusion, a second potential mechanism for recency effects is "backward recognition masking." Massaro (1970) proposed that later arriving sounds disrupt the read-out from a short-term temporary buffer holding information about the first event. In the current context, the representation of the first event (simulated source) is degraded by the latter event (simulated echo). If this were the mechanism for recency effects, then one should find evidence of superiority of judgments based on the second of two pulses even when temporal order confusion could not be a factor. As described earlier, Stellmack *et al.* (1999) argued that temporal order confusion effects could be mitigated by simply fixing the non-target click at 0 interaural delay. When the source click served as the target and the echo was always diotic, proportion correct still showed a minimum at intermediate echo delays, although it was only 6% to 7% lower (relative to what was obtained at the longest echo delays) instead of 18% to 20% lower when the interaural delay of the echo was varied across trials. The time interval over which sensitivity to ITDs conveyed by echoes is superior to information conveyed by sources is generally consistent with the 20–180 ms reported in Massaro *et al.* (1976). It may well be the case that both temporal order confusion and backward recognition masking contribute to recency effects that are observed at echo delays between 16 and 128 ms (or longer).

It should be emphasized that the conditions under which other authors have described recency effects are quite different from those used in the current study. The paradigm used in the current study required listeners to either respond according to the laterality of the first or second pulse. Recency effects are inferred when performance is superior when based on the last of two events. Stecker and his colleagues (Stecker and Hafter, 2002, 2009; Stecker *et al.*, 2013; and Stecker, 2014), on the other hand, required participants to respond according to the laterality or the location of an entire train of clicks when individual pulses were presented with interaural cues there were "perturbed" (randomly chosen from a distribution or range of locations). They refer to recency effects (or "upweighting") as greater weight (derived from multiple linear regression) given to the last event (the Nth) vs the second to N − 1. Both phenomena can reasonably be called recency effects, but the underlying mechanisms may be quite different. Stecker and

J. Acoust. Soc. Am. **140** (6), December 2016

Dye, Jr. *et al.*   4481

Hafter (2009), for instance, model recency effects with the "leaky integrator" model of temporal integration.

Note that the simple backward recognition-masking hypothesis makes no assertions about frequency effects, because the disruption of processing is proposed to be independent of the similarity between signals and backward interferers (Massaro, 1970). Nonetheless, similarity effects have occasionally been found, with greater backward recognition masking observed when the target and backward interferer are more similar (at least in frequency discrimination paradigms, Kallman and Massaro, 1979). The current study finds greater interference exerted by the echo pulse to the extent that it is lower than the target source—when it is more dissimilar. Only when the echo was higher in frequency than the source did proportion correct improve with difference in frequency between the target source and the echo, as was the case for target echoes following distractor sources.

C. Low-frequency dominance

In the current experiment, placing energy into different spectral regions for the source and echo pulse afforded listeners another basis for determining which pulse served as the target—pitch (except when both pulses were centered at 3000 Hz). The hope was that providing a pitch difference between the source and echo could reduce the consequences of temporal order confusion, because the target pulse differed in both intracranial position and pitch from the distractor pulse. Interestingly, making the frequencies of the target and distractor pulse different only elevated target weight and proportion correct when the distractor frequency was higher than the 3000-Hz target. When the distractor pulse was lower in frequency than the target, target weights, and proportions correct were actually lower than when both were 3000 Hz. The data are consistent with "localization strength" hypothesis of Divenyi (1992)—lower frequency components dominate because humans are more sensitive to ITDs at low frequencies. The current study found evidence of low-frequency dominance for echo delays out to 64 ms when the echo served as the target and out to 128 ms when the source served as the target. Divenyi found evidence of stronger precedence effects when the leading pulse was lower in frequency than a 2000-Hz echo pulse. Similarly, Shinn-Cunningham et al. (1995) found largest source weights when a 450-Hz band of noise (300 Hz wide) was followed by a 1250-Hz band of noise (also 300 Hz wide). Both studies examined the suppression of interaural time information carried by the echo pulse for relatively short echo delays (1 ms in the case of Shinn-Cunningham et al., and less than 20 ms in the case of Divenyi, although data were not collected for echo delays greater than 5 ms for conditions in which the source frequency was greater than the echo frequency). A frequency difference *per se* between source and echo pulses did not lead to better target performance (higher target weight/higher proportion correct). This failure except when the distractor was higher in frequency than the target extends out to (at least) 64 ms, and is quite damaging to the

spectral overlap hypothesis of Yang and Grantham (1997). While one could easily extend the spectral overlap hypothesis to durations over which pulses show energetic masking [20–40 ms, according to Raab (1961)], the fact that performance was generally better when both source and echo pulses were 3000 Hz than when the distractor pulse was lower in frequency (1500 or 2000 Hz) renders the spectral overlap hypothesis untenable over the range of echo delays at which across-frequency interactions were observed.

Hartung and Trahiotis (2001) have successfully accounted for low-frequency dominance when echo delays are 4 ms or less, particularly when the spectra are restricted to frequencies below 1000 Hz. Their argument is that source and echo pulses interact within single auditory filters as long as the filters are responding to the first pulse at the time that the second pulse arrives. As long as the impulse responses to the two successive pulses overlap, the instantaneous ITDs and ILDs are derived by comparing the left and right filters will often be quite different than the binaural cues provided by the source and echo pulse. Hartung and Trahiotis (2001) computed the correlogram for binaural clicks after passing the lead and lag pulses through a bank of auditory filters spanning a range of 250 to 1700 Hz, with the outputs processed by the Meddis "hair-cell" model (Meddis, 1986, 1988). While they were able to predict the outcome of the experiment of Shinn-Cunningham et al. (1995) with high accuracy, it should be remembered that the stimuli were brief bursts of noise centered at 450 or 1250 Hz, separated by a lag of 1 ms. Auditory filters at these frequencies are considerably narrower that those at the higher frequencies used in the current study, leading to longer impulse response functions (and greater temporal overlap of the two pulses). It is unlikely that an explanation based on peripheral interactions could account for the low-frequency dominance exhibited in the current study, given the higher frequencies/broader tuning and much longer echo delays that were used here. It is much more likely to be of cortical origin, where interactions between responses to impulses have been observed to occur over several hundred milliseconds (Wehr and Zador, 2005).

Relation to binaural interference: One other binaural phenomenon that shows low-frequency dominance is binaural interference. Studies have shown that the presence of low-spectral frequencies reduces sensitivity to ITDs presented at higher frequencies to a much larger extent than high frequencies affect sensitivity to ITDs at presented at lower frequencies. Yang and Grantham (1997) were the first to note this similarity. For instance, Heller and Trahiotis (1995) measured sensitivity to ITDs carried by 100% SAM carriers at 2 kHz in the presence of diotic interferers that were centered at 500 or 4000 Hz. All carriers were modulated at 250 Hz. Although both diotic SAM tones elevated thresholds at 2 kHz, the 500-Hz carrier had a much larger impact, elevating thresholds by about a factor of 4. The 4-kHz carrier, on the other hand, elevated threshold by less than a factor 2. Explanations of binaural interference typically involve some sort of obligatory non-optimal weighting of non-informative non-target components (Dye, 1990;

Buell and Hafter, 1991; Woods and Colburn, 1992; Heller and Trahiotis, 1995). By the same token, we have argued that precedence and (especially) recency effects might be due to inappropriate weight given to the distractor pulse, at least for tasks in which responses are to be based on either the first or second pulse. Although the term "binaural interference" is typically used to describe the deleterious effects of non-target components that are contemporaneous with target, the spectral dominance exerted by simulated sources and echoes on processing of subsequent or preceding ITDs is so similar that a closer look is warranted. Repeatedly it has been shown that binaural interference is much larger when the low-frequency interferer and the high-frequency target are gated on and off simultaneously (Bernstein and Trahiotis, 1995; Croghan and Grantham, 2010) compared to when the lower frequency stimulus is presented continuously.

Although binaural interference usually occurs for simultaneous stimuli, Kopčo et al. (2007) examined the impact of a distractor click (a 2-ms rectangular pulse) on the perceived location of a subsequent target click (also a 2-ms rectangular pulse). Apparent location was indicated by listeners with a Polhemus electromagnetic sensor attached at the end of a stick. The distractor was presented on 83% of the trials from a speaker at 0° or 90° (constant within a block of trials), with the target click emanating from one of seven loudspeakers between 0° and 90°. Data were collected in a small classroom or in an anechoic chamber. Lags between 25 and 400 ms were presented between the distractor and subsequent target clicks, randomized from trial to trial. For blocks with a distractor at 0°, localization responses were biased away from 0° towards the side for stimuli at all seven locations and *even on trials on which no distractor was presented*. This occurred at all lags from 25 to 400 ms, both for judgments made in the classroom and in the anechoic chamber. For blocks with a distractor at 90°, judgments of frontal targets were biased towards the side, while lateral targets were biased back toward the front. As such, most judgments were biased towards the middle of the range between 0° and 90°. For a distractor at 90°, the biasing effects were most prominent at shorter lags, with the effect decreasing as the lag was increased 400 ms. Judgments of lateral targets were biased towards 90°, an effect found at all lags between 25 and 400 ms. Results for distractors at 90° differed for the two listening environments, with a weaker bias found for frontal targets in the anechoic room, but comparable biases for lateral targets towards the 90° distractor. Again, these effects were also exhibited for trials on which no distractor was actually presented. At least for a distractor at 90° in a classroom, interference could have resulted from a constriction of perceptual space compared to the actual locations of the speakers. Another factor could have been the greater variability of responses that was observed for judgments made in the presence of distractors, particularly for judgments made in a classroom (with reverberations). Regardless of the listening environment, response variability decreased with increasing lag between the source and echo clicks. While interference *per se* was not measured in Kopčo et al. (2007), the greater variability in the judgments induced by distractor

pulse would likely result in larger threshold ITDs and/or ILDs. Furthermore, the smaller range of apparent localizations when the interferer was at 90° in a reverberant classroom might also produce elevations in threshold binaural cues.

Another study that measured interference for non-simultaneous targets and distractors was Stellmack (1994), who found evidence of binaural interference when there were brief temporal notches in diotic distractors. He measured elevations in ITD thresholds at 753 Hz when flanking diotic components were (453, 553, 653, 853, 953, and 1053 Hz) were turned off for anywhere from 10 to 200 ms. Total stimulus duration was 500 ms, with the temporal notch in the diotic components occurring 200 ms into the presentation of the seven-component complex, turning off all but the target component. He found that there was significant interference for brief notches, with thresholds for notch durations of 25 ms generally 1.5 to 2 times those obtained for the target component in isolation. As notch duration increased from 25 to 200 ms, threshold ITDs approached those obtained for the target presented alone. For interferers gated on and off with the 753-Hz target (without a notch), thresholds were 5–10 times larger than those obtained for a target in isolation.

In order to compare the current data with the results of studies of binaural interference, data collected at echo delays of 8, 16, 32, and 64 ms were converted to d's, using "left"|left leading to define hits and "left"|right leading to define false alarms. This was carried out separately for conditions in which the echo and source pulse served as the target. Keep in mind that the ITDs were drawn from Gaussian distributions with a mean of 0 and a standard deviation of 100 μs. Because this was true of all echo delays and distractor frequencies, it allows one to compare the effect of the frequency of the irrelevant distractor pulse on performance and to examine the impact of whether it precedes or follows the 3000-Hz target pulse.

Because most studies of binaural interference have reported ratios by which ITD thresholds are elevated by the presence of irrelevant components (relative to presentation of the target by itself), we have opted to plot and report ratios of $d'_{3000\text{-Hz alone}}/d'_{3000\text{-Hz+distractor pulse}}$. This is referred to as the "interference index" in the current paper. Because of the inverse relation between threshold and d', higher ratios indicate a greater degree of interference; a ratio of 1.0 indicates that absence of any effect of the distractor pulse. Figure 6 shows plots of the interference index for conditions in which judgments were based on the echo. Figure 7 shows data for conditions in which judgments were based on the source. Individual panels present data for different listeners, with the final panel in Figs. 6 and 7 showing averaged interference indices. Note the changes in scale of the ordinate across the two figures. This was necessitated by the greater interference obtained when judgments were based on the echo, particularly at echo delays of 8 ms. For an echo delay of 8 ms, the average interference index was greater than 12 when the judgments were based on the echo (Fig. 6), but only a little more than 2 for judgments based on the source (Fig. 7). When judging the echo, the
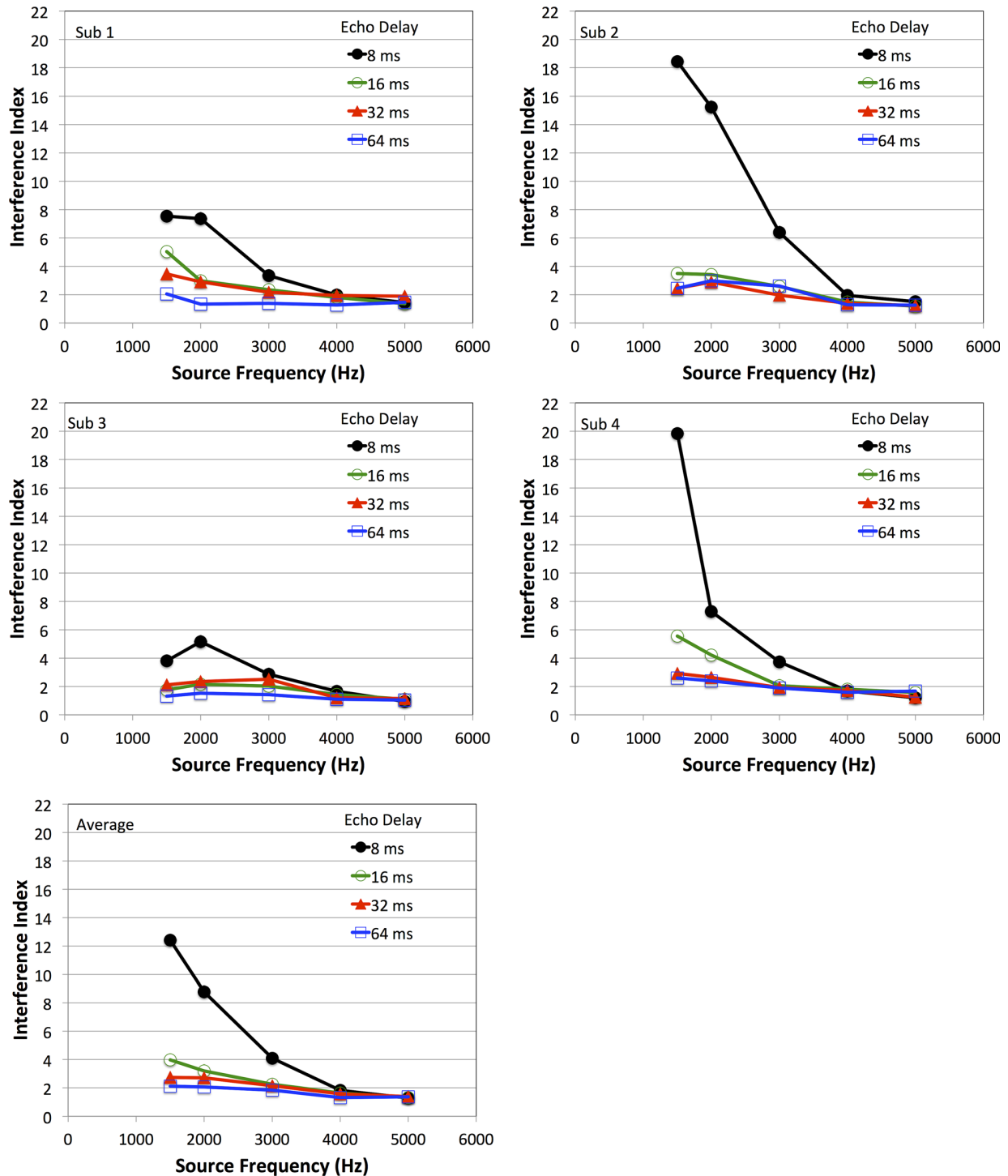
FIG. 6. (Color online) Plots of interference index as a function of source frequency and echo delay for cases in which judgments were based on the ITD presented in the echo pulse. Interference index is defined as the ratio of d′$_{3000\text{-Hz alone}}$/d′$_{3000\text{-Hz + distractor pulse}}$. Data for the four listeners are shown in separate panels, with the final panel depicting data averaged across them.

largest interference indices were produced at echo delays of 8 ms, while they were smallest at 8 ms when judging the source. Remember that high source weights and proportions correct were yielded when the echo delay was 8 ms and judgments were based on the source [refer to Fig. 2, panels (A) and (B)]. For echo delays longer than 8 ms, the average interference indices were less than 4, falling with

increased echo delay for judgments based on the echo (final panel of Fig. 6) and nearly independent of echo delay for judgments based on the source (final panel of Fig. 7). In both figures, the subject numbers correspond to the identifiers the tables in the Appendix. For both cases, it is evident that the interference ratios were higher when the distractor pulse was lower in frequency than the 3000-Hz target pulse.
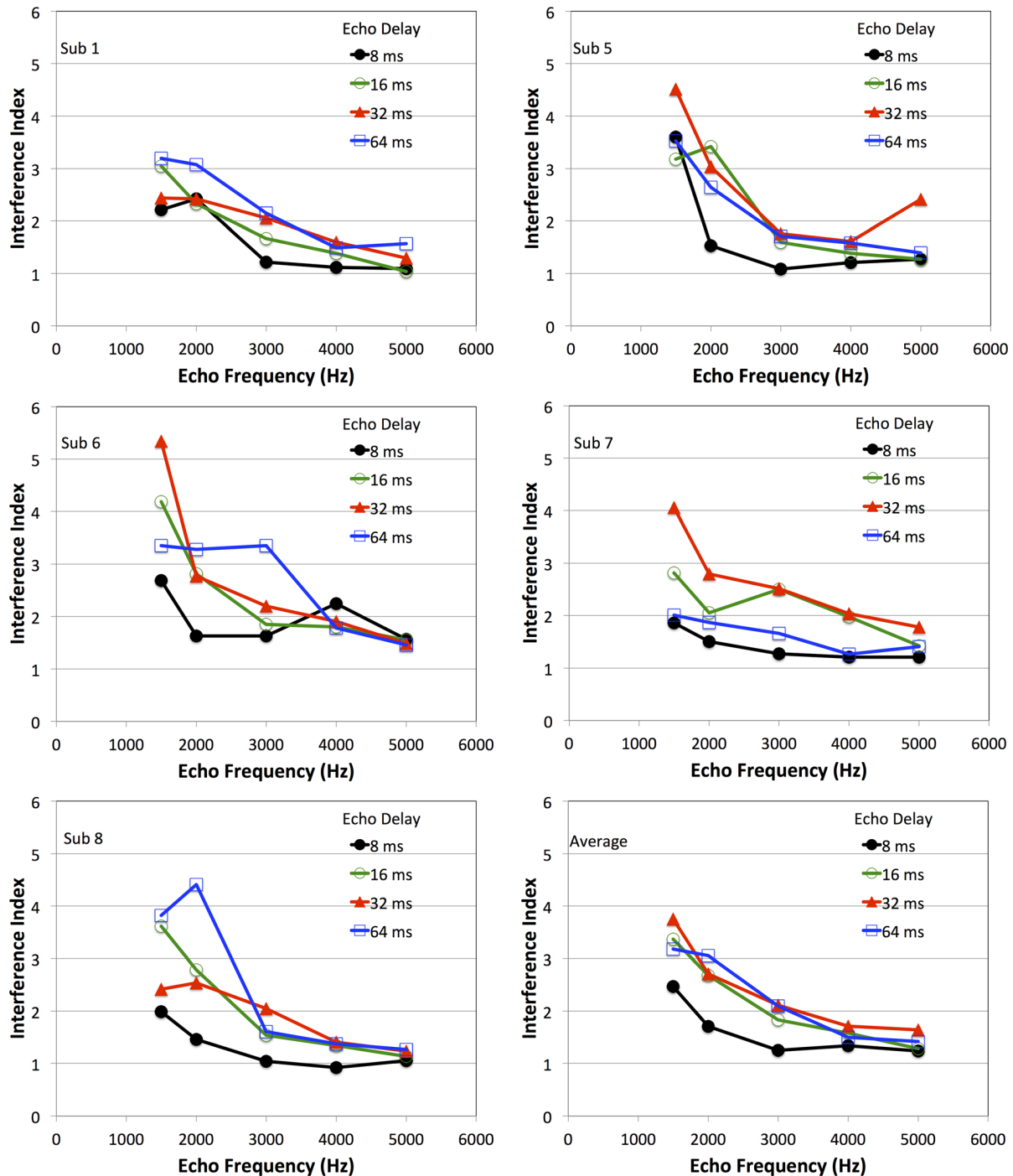
FIG. 7. (Color online) Plots of interference index as a function of echo frequency and echo delay for cases in which judgments were based on the ITD presented in the source pulse. Data for the five listeners are shown in separate panels, with the final panel showing interference indices averaged across them.

In fact, the interference indices were between 1.0 and 2.0 when the distractor frequency was higher than the target frequency, demonstrating a relatively small amount of interference. This was true regardless of which pulse served as the target.

Separate two-way repeated measures analyses of variance were carried out on interference indices to examine the effects of distractor frequency and echo delay. For the conditions in which judgments were based on the echo ITD, both source frequency and echo delay had significant effects on the interference index [source frequency: $F(4,12) = 9.70$, $p < 0.001$; echo delay: $F(3,9) = 8.90$, $p < 0.01$]. Larger interference indices were found at shorter echo delays and lower source frequencies. The interaction between source frequency and echo delay, also proved to be significant [$F(12, 36) = 5.80$, $p < 0.001$], with much larger effects of source frequency found at shorter echo delays. The nature of the interaction was further explored with an analysis of simple

J. Acoust. Soc. Am. **140** (6), December 2016

Dye, Jr. *et al.* 4485

effects, which revealed significant effects of source frequency at all four echo delays (8, 16, 32, and 64 ms). For conditions in which judgments were based on the source ITD, both echo frequency and echo delay again produced significant effects [echo frequency: $F(4,16) = 32.56$, $p < 0.0001$; echo delay: $F(3, 12) = 7.84$; $p < 0.01$], with greater interference indices occurring at lower frequencies and the longer three echo delays. The interaction between echo frequency and echo delay did not reach significance; $F(12, 48) = 1.556$, $p = 0.137$. *Post hoc* Tukey's HSDs holding family-wise α at 0.05 revealed that the mean interference indices were significantly lower at 8 ms than at other echo delays, which were statistically equivalent to one another. Interference indices at echo frequencies of 3000 to 5000 Hz were all statistically equivalent, while those obtained at 1500 and 2000 Hz were significantly higher than those obtained at the three higher echo frequencies. Interference indices obtained at 1500 and 2000 Hz did not significantly differ from one another.

At echo delays of 16, 32, and 64 ms, the amount of interference observed for distractors below the frequency of the target in this study were quite similar to those reported by Heller and Trahiotis (1995) for SAM tones at different spectral locations. In both cases, interference effects are somewhere between a factor of 2 to 4, with greater interference found as the distractors were decreased in frequency in spite of the fact that the difference between the target and distractor was increased in magnitude. This was true for judgments based on the source and on the echo. Interference indices obtained at an echo delay of 8 ms were much larger for judgments based on the echo, reaching a value greater than 12 at a distractor frequency of 1500 Hz vs below 2.5 for conditions in which judgments were based on the source.

The much greater interference ratio at 8 ms when judgments were based on the echo leads one to wonder whether the decrease in d′ should be treated as "interference," since it is likely that there is also discrimination suppression (Litovsky and Shinn-Cunningham, 2001) taking place at 8 ms. Buell and Hafter (1988) found that intervals between pulses needed to be at least 10 ms (on the average) in order for listeners to optimally integrate ITD information from successive events, so it may be reasonable to treat the 8-ms data somewhat differently than those collected at longer echo delays, particularly when judgments are based on the echo and the localization suppression is still a factor. Furthermore, Tollin and Henning (1998) found that most of their listeners still exhibited evidence of precedence effects for intervals of 9.6 ms between pulses.

One of the similarities between binaural interference tasks and precedence/recency studies in which judgments are to be based on one of two pulses (both varying in ITD) is the need to ignore information from irrelevant frequency components (in interference) or the irrelevant pulse (in precedence/recency studies). This non-optimal weighting may be obligatory, but it may reflect differences in auditory selective attention. Note the presence of individual differences, which we believe indicate difference in auditory selective attention. S2 and S4 show much greater

interference when judging the echo (Fig. 6) when the source was 1500 Hz, yielding interference indices of more than 18. S3, on the other hand, has a peak interference index of slightly greater than 5. When judgments were based on the source, S7 appeared to show considerably less interference when the echo delay was lengthened to 64 ms, while others showed as much or greater than was obtained with an echo delay of 32 ms. Substantial individual differences have been reported in studies of binaural interference, generally on the order of a factor of two (see Heller and Trahiotis, 1995; and Bernstein and Trahiotis, 1995). If one ignores the data obtained with an echo delay of 8 ms in the current study, we find approximately the same amount of interference and the same extent of individual differences.

APPENDIX

Individual echo weights (Table I), proportions correct for judgments based on the echo (Table II), proportion of responses predicted correctly from weights for echo target trials (Table III), source weights (Table IV), proportions correct for judgments based on the source (Table V), and proportion of responses predicted correctly from weights for source target trials (Table VI) are presented in the following tables.

TABLE I. Echo weights.

Source frequency (Hz)	W_{Echo} Echo delay					
	8 ms	16 ms	32 ms	64 ms	72 ms	128 ms
S1						
5000	0.716	0.707	0.629	0.651	0.671	0.903
4000	0.512	0.592	0.552	0.685	0.640	0.852
3000	0.358	0.451	0.474	0.636	0.619	0.831
2000	0.165	0.352	0.399	0.634	0.637	0.893
1500	0.178	0.237	0.317	0.509	0.593	0.793
S2						
5000	0.730	0.930	0.811	0.912	0.879	0.948
4000	0.599	0.762	0.778	0.752	0.838	0.894
3000	0.233	0.518	0.562	0.546	0.728	0.995
2000	0.183	0.469	0.528	0.591	0.745	0.933
1500	0.118	0.608	0.617	0.602	0.708	0.925
S3						
5000	0.836	0.932	0.810	0.905	0.913	0.917
4000	0.617	0.652	0.754	0.836	0.837	0.877
3000	0.483	0.542	0.484	0.729	0.905	0.849
2000	0.359	0.705	0.543	0.702	0.910	0.843
1500	0.438	0.709	0.605	0.811	0.989	0.914
S4						
5000	0.772	0.742	0.765	0.723	0.760	0.922
4000	0.599	0.588	0.577	0.639	0.633	0.815
3000	0.356	0.538	0.516	0.576	0.628	0.875
2000	0.183	0.316	0.422	0.515	0.528	0.751
1500	0.058	0.261	0.387	0.480	0.492	0.822

TABLE II. P(C) for judgments based on echoes.

| | P(C) Echo delay | | | | | |
Source frequency (Hz)	8 ms	16 ms	32 ms	64 ms	72 ms	128 ms
S1						
5000	0.782	0.794	0.704	0.778	0.755	0.825
4000	0.713	0.729	0.698	0.811	0.772	0.796
3000	0.631	0.685	0.696	0.789	0.774	0.830
2000	0.561	0.646	0.644	0.796	0.762	0.802
1500	0.561	0.587	0.624	0.705	0.750	0.817
One pulse	0.842					
S2						
5000	0.842	0.908	0.894	0.886	0.928	0.934
4000	0.788	0.852	0.864	0.884	0.886	0.928
3000	0.590	0.720	0.778	0.730	0.822	0.902
2000	0.540	0.678	0.706	0.692	0.818	0.858
1500	0.538	0.674	0.732	0.742	0.766	0.842
One pulse	0.938					
S3						
5000	0.934	0.898	0.896	0.916	0.914	0.908
4000	0.808	0.840	0.872	0.900	0.886	0.890
3000	0.690	0.746	0.714	0.838	0.884	0.888
2000	0.610	0.724	0.712	0.808	0.906	0.916
1500	0.648	0.786	0.742	0.864	0.900	0.900
One pulse	0.922					
S4						
5000	0.833	0.774	0.827	0.755	0.790	0.859
4000	0.754	0.740	0.751	0.770	0.767	0.822
3000	0.621	0.712	0.726	0.727	0.783	0.804
2000	0.561	0.606	0.672	0.687	0.727	0.786
1500	0.523	0.585	0.655	0.672	0.710	0.799
One Pulse	0.872					

TABLE III. P(Responses Predicted) for judgments based on echoes.

| | P(Responses Predicted) Echo delay | | | | | |
Source frequency (Hz)	8 ms	16 ms	32 ms	64 ms	72 ms	128 ms
S1						
5000	0.811	0.838	0.760	0.811	0.814	0.823
4000	0.825	0.816	0.764	0.867	0.821	0.801
3000	0.778	0.799	0.810	0.835	0.830	0.828
2000	0.805	0.885	0.815	0.855	0.808	0.804
1500	0.794	0.857	0.839	0.839	0.817	0.822
S2						
5000	0.88	0.912	0.898	0.882	0.926	0.934
4000	0.864	0.89	0.898	0.896	0.914	0.912
3000	0.87	0.854	0.874	0.808	0.876	0.898
2000	0.76	0.76	0.774	0.746	0.822	0.852
1500	0.85	0.698	0.768	0.754	0.792	0.844
S3						
5000	0.932	0.902	0.908	0.920	0.906	0.912
4000	0.908	0.886	0.902	0.910	0.892	0.896
3000	0.836	0.862	0.834	0.866	0.880	0.894
2000	0.802	0.760	0.828	0.832	0.902	0.920
1500	0.736	0.808	0.808	0.868	0.900	0.910

TABLE III. (*Continued*)

| | P(Responses Predicted) Echo delay | | | | | |
Source frequency (Hz)	8 ms	16 ms	32 ms	64 ms	72 ms	128 ms
S4						
5000	0.853	0.775	0.841	0.788	0.825	0.862
4000	0.823	0.814	0.821	0.801	0.821	0.816
3000	0.82	0.816	0.839	0.793	0.828	0.799
2000	0.862	0.779	0.841	0.781	0.809	0.812
1500	0.768	0.773	0.805	0.759	0.782	0.812

TABLE IV. Source weights.

| | W_{Source} Echo delay | | | | |
Echo frequency (Hz)	8 ms	16 ms	32 ms	64 ms	128 ms
S1					
5000	0.810	0.777	0.652	0.654	0.621
4000	0.775	0.676	0.578	0.595	0.655
3000	0.675	0.532	0.473	0.488	0.659
2000	0.526	0.420	0.393	0.409	0.573
1500	0.459	0.340	0.380	0.416	0.588
S5					
5000	0.904	0.674	0.697	0.700	0.620
4000	0.839	0.650	0.568	0.629	0.589
3000	0.706	0.529	0.542	0.563	0.730
2000	0.587	0.337	0.416	0.503	0.602
1500	0.297	0.327	0.317	0.359	0.616
S6					
5000	0.834	0.772	0.786	0.699	0.732
4000	0.829	0.730	0.677	0.628	0.708
3000	0.779	0.616	0.494	0.510	0.610
2000	0.659	0.437	0.477	0.442	0.508
1500	0.603	0.326	0.330	0.379	0.503
S7					
5000	0.818	0.710	0.640	0.745	0.924
4000	0.782	0.602	0.548	0.716	0.759
3000	0.773	0.478	0.614	0.655	0.816
2000	0.633	0.452	0.429	0.612	0.748
1500	0.497	0.377	0.382	0.576	0.649
S8					
5000	0.774	0.764	0.789	0.743	0.875
4000	0.893	0.797	0.656	0.598	0.850
3000	0.823	0.538	0.512	0.518	0.738
2000	0.575	0.414	0.425	0.381	0.630
1500	0.463	0.287	0.410	0.303	0.560

TABLE V. P(C) for judgments based on sources.

| | P(C) Echo delay | | | | |
Echo frequency (Hz)	8 ms	16 ms	32 ms	64 ms	128 ms
S1					
5000	0.821	0.829	0.782	0.734	0.676
4000	0.814	0.764	0.732	0.745	0.711

J. Acoust. Soc. Am. **140** (6), December 2016

Dye, Jr. *et al.* 4487

TABLE V. (*Continued*)

Echo frequency (Hz)	P(C) Echo delay				
	8 ms	16 ms	32 ms	64 ms	128 ms
3000	0.793	0.723	0.685	0.679	0.664
2000	0.716	0.668	0.662	0.627	0.659
1500	0.684	0.627	0.659	0.624	0.671
One pulse	0.880				
S5					
5000	0.808	0.804	0.674	0.785	0.735
4000	0.800	0.756	0.748	0.760	0.697
3000	0.848	0.762	0.740	0.738	0.737
2000	0.752	0.634	0.644	0.686	0.686
1500	0.618	0.626	0.596	0.610	0.699
One pulse	0.866				
S6					
5000	0.762	0.752	0.758	0.762	0.744
4000	0.680	0.716	0.712	0.708	0.793
3000	0.740	0.708	0.680	0.618	0.747
2000	0.712	0.652	0.650	0.620	0.653
1500	0.642	0.602	0.576	0.626	0.619
One pulse	0.889				
S7					
5000	0.785	0.749	0.705	0.748	0.742
4000	0.791	0.689	0.685	0.771	0.780
3000	0.782	0.648	0.653	0.720	0.765
2000	0.739	0.685	0.637	0.696	0.735
1500	0.706	0.631	0.599	0.692	0.706
One pulse	0.881				
S8					
5000	0.842	0.826	0.802	0.778	0.757
4000	0.870	0.776	0.770	0.776	0.728
3000	0.840	0.758	0.684	0.734	0.730
2000	0.762	0.642	0.660	0.586	0.766
1500	0.700	0.614	0.662	0.602	0.697
One pulse	0.895				

TABLE VI. P(Responses Predicted) for judgment based on sources.

Echo frequency (Hz)	P(Responses Predicted) Echo delay				
	8 ms	16 ms	32 ms	64 ms	128 ms
S1					
5000	0.815	0.841	0.826	0.765	0.710
4000	0.822	0.800	0.806	0.801	0.742
3000	0.826	0.798	0.802	0.782	0.689
2000	0.837	0.812	0.839	0.753	0.717
1500	0.796	0.813	0.835	0.711	0.679
S5					
5000	0.812	0.852	0.702	0.812	0.775
4000	0.812	0.822	0.816	0.823	0.749
3000	0.890	0.858	0.818	0.816	0.745
2000	0.834	0.836	0.742	0.784	0.728
1500	0.834	0.792	0.762	0.800	0.723

TABLE VI. (*Continued*)

Echo frequency (Hz)	P(Responses Predicted) Echo delay				
	8 ms	16 ms	32 ms	64 ms	128 ms
S6					
5000	0.772	0.774	0.776	0.768	0.744
4000	0.694	0.748	0.706	0.752	0.820
3000	0.738	0.782	0.748	0.678	0.787
2000	0.726	0.704	0.734	0.746	0.707
1500	0.664	0.778	0.678	0.762	0.667
S7					
5000	0.794	0.764	0.722	0.776	0.742
4000	0.799	0.725	0.722	0.779	0.796
3000	0.801	0.749	0.670	0.763	0.770
2000	0.789	0.787	0.738	0.738	0.742
1500	0.792	0.774	0.685	0.743	0.738
S8					
5000	0.868	0.842	0.810	0.802	0.810
4000	0.868	0.766	0.826	0.812	0.823
3000	0.870	0.816	0.798	0.836	0.781
2000	0.822	0.838	0.818	0.692	0.843
1500	0.822	0.824	0.792	0.858	0.814

Bernstein, L. R., and Trahiotis, C. (**1995**). "Binaural interference effects measured with masking-level difference and with ITD- and IID-discrimination paradigms," J. Acoust. Soc. Am. **98**, 155–163.

Bianchi, F., Verhulst, S., and Dau, T. (**2013**). "Experimental evidence for a cochlear source of the precedence effect," J. Assoc. Res. Otolaryngol. **14**, 767–779.

Blauert, J., and Diveny, P. L. (**1988**). "Spectral selectivity in binaural contra-lateral inhibition," Acustica **66**, 267–274.

Boomer, J. B., Dye, R. H., Jr., Sheft, S., and Yost, W. A. (**2006**). "The effect of spectral differences on the ability to judge the laterality of a simulated source followed by a simulated echo," J. Acoust. Soc. Am. **119**, 3296.

Buell, T. N., and Hafter, E. R. (**1988**). "Discrimination of interaural differences of time in the envelopes of high-frequency signals: Integration times," J. Acoust. Soc. Am. **84**, 2063–2066.

Collett, D. (**1991**). *Modelling Binary Data*, 2nd ed. (Chapman and Hall, London).

Croghan, N. B. H., and Grantham, D. W. (**2010**). "Binaural interference in the free field," J. Acoust. Soc. Am. **127**, 3085–3091.

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

Divenyi, P. L. (**1992**). "Binaural suppression of non-echoes," J. Acoust. Soc. Am. **91**, 1078–1084.

Dye, R. H., Jr. (**1990**). "The combination of interaural information across frequencies: Lateralization of the basis of interaural delay," J. Acoust. Soc. Am. **88**, 2159–2170.

Dye, R. H., Jr., Brown, C. A., Gallegos, J. A., Yost, W. A., and Stellmack, M. A. (**2006**). "The influence of later-arriving sounds on the ability of listeners to judge the lateral position of a source," J. Acoust. Soc. Am. **120**, 3946–3956.

Goupell, M. J., Yu, G., and Litovsky, R. Y. (**2012**). "The effect of an additional reflection in a precedence effect experiment," J. Acoust. Soc. Am. **131**, 2958–2967.

Hartung, K., and Trahiotis, C. (**2001**). "Peripheral auditory processing and investigations of the 'precedence effect' which utilize successive transient stimuli," J. Acoust. Soc. Am. **110**, 1505–1513.

Heller, L. M., and Trahiotis, C. (**1995**). "Interference in detection of interaural delay in a sinusoidally amplitude-modulated tone produced by a second spectrally remote sinusoidally amplidute-modulated tone," J. Acoust. Soc. Am. **97**, 1808–1816.

Hirsh, I. J. (**1959**). "Auditory perception of temporal order," J. Acoust. Soc. Am. **31**, 759–767.

Hirsh, I. J., and Sherrick, C. E., Jr. (**1961**). "Perceived order in different sense modalities," J. Exp. Psychol. **62**, 423–432.

Kallman, H. J., and Massaro, D. W. (**1979**). "Similarity effects in backward recognition masking," J. Exp. Psych.: Human Percept. Perform. **5**, 110–128.

Kopčo, N., Best, V., and Shinn-Cunningham, B. G. (**2007**). "Sound localization with a preceding distractor," J. Acoust. Soc. Am. **121**, 420–432.

Litovsky, R. Y., and Godar, S. P. (**2010**). "Difference in precedence effect between children and adults signifies development of sound localization abilities in complex listening tasks," J. Acoust. Soc. Am. **128**, 1979–1992.

Litovsky, R. Y., and Shinn-Cunningham, B. G. (**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, R. Y., and Yin, T. C. (**1998**). "Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics," J. Neurophysiol. **80**, 1285–1301.

Lutfi, R. A. (**1995**). "Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks," J. Acoust. Soc. Am. **97**, 1333–1334.

Massaro, D. W. (**1970**). "Preperceptual auditory images," J. Exp. Psychol. **85**, 411–417.

Massaro, D. W., Cohen, M. M., and Idson, W. L. (**1976**). "Recognition masking of auditory localization and pitch judgments," J. Acoust. Soc. Am. **59**, 434–441.

Meddis, R. (**1986**). "Simulation of mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. **79**, 702–711.

Meddis, R. (**1988**). "Simulation of auditory-neural transduction: Further studies," J. Acoust. Soc. Am. **83**, 1056–1063.

Raab, D. H. (**1961**). "Forward and backward masking between acoustic clicks," J. Acoust. Soc. Am. **33**, 137–139.

Shinn-Cunningham, B. G., Zurek, P. M., Durlach, N. I., and Clifton, R. K. (**1995**). "Cross-frequency interactions in the precedence effect," J. Acoust. Soc. Am. **98**, 164–171.

Stecker, G. C. (**2014**). "Temporal weighting functions for interaural time and level differences. IV. Effects of carrier frequency," J. Acoust. Soc. Am. **136**, 3221–3232.

Stecker, G. C., and Hafter, E. R. (**2002**). "Temporal weighting in sound localization," J. Acoust. Soc. Am. **112**, 1046–1057.

Stecker, G. C., and Hafter, E. R. (**2009**). "A recency effect in sound localization?," J. Acoust. Soc. Am. **125**, 3914–3924.

Stecker, G. C., Ostreicher, J. D., and Brown, A. D. (**2013**). "Temporal weighting functions for interaural time and level differences. III. Temporal weighting for lateral position judgments," J. Acoust. Soc. Am. **134**, 1242–1252.

Stellmack, M. A. (**1994**). "The reduction of binaural interference by the temporal nonoverlap of components," J. Acoust. Soc. Am. **96**, 1465–1470.

Stellmack, M. A., Dye, R. H., Jr., and Guzman, S. J. (**1999**). "Observer weighting of interaural delays in source and echo clicks," J. Acoust. Soc. Am. **105**, 377–387.

Tollin, D. J., and Henning, G. B. (**1998**). "Some aspects of the lateralization of echoed sound in man. I. The classical interaural-delay based precedence effect," J. Acoust. Soc. Am. **104**, 3030–3038.

Wallach, H., Newman, E. B., and Rosenzweig, M. R. (**1949**). "The precedence effect in sound localization," Am. J. Psychol. **62**, 315–336.

Wehr, M., and Zador, A. M. (**2005**). "Synaptic mechanisms of forward suppression in rat auditory cortex," Neuron **47**, 437–445.

Woods, W. S., and Colburn, H. S. (**1992**). "Test of a model of auditory object formation using intensity and interaural time difference discrimination," J. Acoust. Soc. Am. **91**, 2894–2902.

Xia, J., Brughera, A., Colburn, H. S., and Shinn-Cunningham, B. (**2010**). "Physiological and psychophysical modeling of the precedence effect," J. Assoc. Res. Otolaryngol. **11**, 495–513.

Yang, X., and Grantham, D. W. (**1997**). "Cross-spectral and temporal factors in the precedence effect: Discrimination suppression of the lag sound in free-field," J. Acoust. Soc. Am. **102**, 2973–2983.

Zurek, P. M., and Saberi, K. (**2003**). "Lateralization of two-transient stimuli," Percept. Psychophys. **65**, 95–106.

J. Acoust. Soc. Am. **140** (6), December 2016

Dye, Jr. *et al.*   4489