Listeners' expectations about echoes can raise or lower echo threshold

Rachel K. Clifton
Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01003

Richard L. Freyman
Department of Communication Disorders, University of Massachusetts, Amherst, Massachusetts 01003

Ruth Y. Litovsky and Daniel McCall
Department of Psychology, University of Massachusetts, Amherst, Massachusetts 01003

(Received 3 June 1993; revised 20 October 1993; accepted 29 November 1993)

Echo threshold increases with exposure to redundant trains of stimuli. Three experiments were conducted to test the hypothesis that a change in the ongoing train would affect listeners' perception of the echo, but only if it signified an unusual change in room acoustics. The stimulus train was composed of 4-ms narrow-band noise bursts, with the leading sound from a loudspeaker placed 45° left of midline and the lagging sound or simulated echo from 45° right, delivered in an anechoic chamber. The lagging sound in the test noise, which followed the train after a 750-ms pause, came randomly from loudspeakers at 35° or 55° right, and the listener's task was to choose which position the echo came from on each trial. In experiment 1 the delay between onsets of the leading and lagging bursts was varied between train and test bursts, which simulated a sudden movement of the reflecting surface either toward the listener (if the delay of the test burst was shorter than the train) or away (if the delay was longer). In both cases listeners detected the echo's direction more easily, compared to trials when there was no change between train and test burst delays. In order to check whether *any* change between train and test bursts would increase echo discriminability, experiment 2 varied frequency and experiment 3 varied intensity. These variations were not expected to affect the echo's detectability because such changes signify that the original sound changed in these characteristics and the echo reflected these changes. These events are highly probable in the listener's everyday experience because sound sources (and their reflections) typically vary in frequency and intensity content from moment to moment. As predicted, echo detectability in experiments 2 and 3 was not affected by whether the test noise bursts' frequency or intensity was the same as the train's or was varied. The results from all three experiments were interpreted in terms of listeners' expectations about echoes. It is proposed that echoes provide information about room acoustics, which the listener picks up during the ongoing sound and uses to form expectations about what will be heard. When expectations are violated by changes in the echo, this disruption can be seen in a lowering of echo threshold, relative to the "built-up" threshold when expectations are fulfilled.

PACS numbers: 43.66.Pn, 43.66.Qp, 43.66.Mk[HSC]

INTRODUCTION

One of the basic problems in psychological acoustics is understanding how listeners are capable of separating complex acoustic waves arriving at the two ears into discrete auditory events. In the natural listening environment several sound sources and their reflections may propagate complex waveforms, resulting in an intricate array of signals arriving at the ears. The auditory system's job is to untangle these complex waveforms into simpler components, assigning each to their respective auditory events. Those components assigned to the same auditory event, such as a clock ticking or a dog barking, are localized together in the same place, while locations of the associated reflected sounds are ignored. The general problem of how extremely ambiguous signals arriving at the ears get organized into a stable, sensible "auditory scene" has been described at length in two recent books, Bregman's *Auditory Scene Analysis* (1990) and Handel's *Listening* (1989; see especially Chaps. 3 and 7). Our research concerns a particular aspect of this problem, namely how the reflected sounds come to be treated as part of the original signal rather than as separate acoustic events.

Sound produced in an enclosed space inevitably (unless the room is anechoic) produces reflections off surrounding surfaces such as walls, ceilings, floors, and nearby objects. These reflections or echoes color the original sound and enhance its loudness, but they are not identified as separate sounds from new sources unless the delay between original sound and echo is quite long. This phenomenon has been variously called the precedence effect, the Haas effect, and Law of the First Wave Front, to emphasize the greater weight given to the directional information in the first wave that strikes the ears (Gardner, 1968). The pre-

0001-4966/94/95(3)/1525/9/$6.00

cedence effect has often been described as an echo suppression mechanism (Green, 1976; Mills, 1972; Moore, 1989) because of its apparent utility in enhancing the location of the preceding original sound at the expense of localizing the delayed echoes. The strength of the lead's temporal advantage is surprisingly strong; even when the echo is produced at the same intensity as the leading sound, the listener will localize the sound at the leading site (Wallach *et al.*, 1949). It should be emphasized that echoes do contribute to localization by pulling the apparent sound source in the direction of the echo (Hartmann, 1983). Perrott *et al.* (1989) found that listeners could discriminate azimuthal shifts in the position of the echo because of its influence on the fused sound image. Thus it is incorrect to think of echo suppression as the elimination of the echo's influence; rather, we use the term to refer to the listener's state when a delayed sound is below echo threshold. Echo threshold is defined as the shortest delay at which the echo is perceived as a separate sound, coming from a different location, and no longer fused with the original sound (Blauert, 1983, pp. 224–225).

One obvious hypothesis about how the brain might accomplish echo suppression can be rejected. It seems intuitively reasonable that the brain might run a cross correlation to check on whether the delayed sound is an exact copy of the original sound coming from a different location. Detection of a mismatch in temporal and spectral qualities would lead to rejection of the delayed sound as an echo because it would be highly unlikely that two identical sounds would originate from two different sources in close temporal relationship. Nonidentical lead and lag sounds would indicate that the delayed sound should be treated as a new source. However, the physics of how sound behaves in enclosed spaces preclude such an easy solution. Rarely are echoes exact copies of the original sound. Sound reflected from walls and ceilings is not "mirrored" in the way light reflected off a mirror represents the original image in accurate detail. A more likely analogy is an image reflected from a wavy glass pane or a dark metal surface because reflecting surfaces in rooms distort the wave and absorb low and high frequencies differentially. Echoes that are filtered and distorted versions of the original sound still get suppressed. What possible commonalities between the original signal and the echo must be present in order for the system to designate a delayed sound to be an echo rather than a different sound source? Remarkably few experiments have been done on this question. Zurek (1980) discovered that brief bursts of *uncorrelated* noise would show echo suppression. Echo suppression is possible even when spectra of lead and lag signals do not overlap. Blauert and Divenyi (1988) and Divenyi (1992) reported that low-frequency signals suppressed high-frequency echoes better than the reverse. Clifton *et al.* (1989) found no asymmetry in echo suppression when stimuli were balanced in SPL for low-frequency domination over high frequencies in determining localization. However, all of these studies agree that the echo does not have to be an exact copy of the original sound; on the contrary, lag can differ spectrally from lead in dramatic ways and still be suppressed.

In a series of experiments we have investigated a phenomenon we refer to as "buildup" in echo suppression. When a train of clicks is presented from two loudspeakers with one onset lagging the other by a few ms, the listener hears both clicks initially, each localized at their respective locations. As the train continues the delayed click fades and only the leading click is perceived at its location (Clifton, 1987; Clifton and Freyman, 1989). We have conceptualized this process as a "buildup" in suppression that takes place over time as the ongoing click train supplies increasing information about the leading and lagging sounds. The result of this changing suppression is to raise echo threshold at the end of the click train several ms above what it would be for a single click pair (Freyman *et al.*, 1991). Many factors contribute to this process. The delay between loudspeaker onsets is critical because the buildup only occurs in the region of the echo threshold. Very short delays (2–3 ms) produce immediate suppression and longer delays (> 10–15 ms, depending on the subject and stimulus) produce no suppression. In the latter case the echo is heard initially and continues to be heard, indicating the delayed sound is above echo threshold (Clifton and Freyman, 1989). Number of clicks in the train is a crucial parameter, but not rate at which clicks are delivered (Freyman *et al.*, 1991, experiment 1). The presence of the echo during the click train is necessary to produce the buildup; preceding a lead–lag click pair by a train of clicks from a single loudspeaker does not produce suppression of the echo click. The train must be composed of lead and lag clicks before buildup is seen (Freyman *et al.*, 1991, experiment 3). When the lagging click is below echo threshold, switching the location of lead and lag clicks in the middle of the train appears to "reset" the system, so that the echo is heard immediately after the switch (Clifton, 1987; Clifton and Freyman, 1989). All of these findings suggest that "new" information, in the form of either the sudden introduction of an echo where none had been previously or a sudden switch in spatial location of the echo, may be the critical feature in lowering echo threshold back to the unadapted level. Conversely, redundant information in an ongoing click train increases echo threshold.

The above description of the buildup process suggests listeners have expectations about what reasonable echoes might be. These expectations are most likely based on the listener's accumulated experience in highly variable acoustic environments as well as the transitory auditory information present at the moment and specific to the listener's current acoustic environment. The acoustic characteristics of the delayed sounds inform the listener about the reflecting surfaces in the room; this is a rapid, automatic, and unconscious process. If neither the listener nor the objects in the room are moving, the listener would expect a stable acoustic environment with predictable echoes. A reasonable hypothesis about these expectations is that changes in echoes from the reflecting surfaces that are apt to be experienced in everyday life will not disrupt echo suppression, but changes that are improbable will disrupt the process. In the present experiments our procedure was to present a train of noise pairs that preceded a single test noise pair that varied from the train in lead–lag delay, frequency, or intensity. If echo suppression was disrupted, listeners were expected to respond to the test noise as though there had been no preceding train. That is, discriminability of the echo's location would be similar to a control condition in which an isolated test noise pair was presented. If echo suppression held despite the difference between train and test noise, echo discrimination was expected to be similar to the condition where train and test noises were the same.

I. EXPERIMENT 1. CHANGING THE DELAY BETWEEN NOISE TRAIN AND TEST NOISE PAIR

If the test noise has a different delay between leading and lagging sounds compared to the preceding noise train, this simulates a quick movement of the reflecting surface. This manipulation was expected to violate the listener's expectations because such a movement would be highly unlikely. If a click train began with a certain delay between leading and lagging sounds, then the delay suddenly changed, this would signify that the reflecting surface either moved abruptly toward the listener (if the delay was shortened) or away (if the delay was lengthened). In either case, because the shift in delay would indicate a highly improbable change in the echo, the listener is likely to conclude that a new sound source is present rather than the same echo continuing. This violation in expectation would result in a lowering of echo threshold for the click pair with the aberrant delay.

A. Method

1. *Stimuli and apparatus*

The stimuli presented during both the conditioning train and test burst were 4-ms segments of computer-generated white noise shaped with 2-ms linear rise/fall times. Each token in the train, as well as the test burst, was randomly selected from a longer (400 ms) sample of noise. Each burst was presented from two channels of a 16-bit D/A converter (TTES QDA1) with a specified delay to the right channel. The outputs of the two signal channels were low-pass filtered at 8500 Hz (TTE J1390), attenuated (TTES PAT1), amplified (NAD 2100), multiplexed (TTE AMUX1) and connected to a set of matched loudspeakers (Realistic Minimus 7) situated in a 4.9 m×4.1 m×3.12 m anechoic chamber. The floor, ceiling, and walls of the chamber were lined with 0.72-m foam wedges. Subjects sat near the center of the room with a total of four loudspeakers situated at 45° left and 35°, 45°, and 55° right of midline at a distance of 1.9 m. The center of the loudspeakers was 1.04 m above the wire mesh floor of the anechoic chamber, the approximate height of the typical subject's ears while seated in the chair. The stimulus level was measured by presenting trains of noise bursts at a rate of four bursts/s. With the microphone placed at the position of the center of the listener's head, and the meter response of a B & K 2204 SLM set on the "fast" meter response, the measured level was approximately 50 dBC, although there were slight variations from token to token.

2. *Procedures*

On each trial the test burst was presented from the left (lead) loudspeaker, and a delayed copy presented from either the 35° or the 55° lag loudspeaker in the right hemifield. The subjects' task was to report, by pressing the appropriate key on a response box held on the lap, which of these two lag loudspeakers presented the delayed sound. Performance on this discrimination task has been shown to be highly correlated with subjective echo thresholds (Freyman *et al.*, 1991, experiment 3). Correct-answer feedback was provided on every trial by illuminating the appropriate light on the button panel. Subjects were instructed to face directly ahead, but were not physically restrained in any way. In all but one condition the test burst was preceded by a train of nine bursts presented at four bursts/s. During the train the lead sound was presented from the loudspeaker at 45° left and the lag sound was from the middle (45°) lag loudspeaker on the right. The train was followed by a brief silent interval of 750 ms, and then the test burst was presented with the lag sound shifting left or right.

The delay of the lag sound during the test burst was determined individually for each subject from preliminary testing with an isolated test burst that had no preceding train (the "no conditioner" or NC condition). The goal of the preliminary testing was to find a delay that produced reasonably good, though not perfect, performance on the NC condition. Good performance on this condition would allow us to evaluate the degree to which the task was made more difficult when the test burst was preceded by the conditioning trains. For most subjects, the appropriate delay for the NC condition was found using an adaptive three-down one-up procedure which estimated 79.4% correct on the psychometric function (Levitt, 1971), corresponding to a d' of approximately 1.63. The correct lag loudspeaker was either at 35° or 55°, each with 50% probability. During the adaptive run the step size on the delay of test burst was 3 ms. The adaptive tracking progressed through a total of 12 reversals, the last eight of which were averaged to estimate threshold. Final thresholds were taken as the mean of three adaptive runs.

During the main part of the experiment the test burst delay was fixed for all conditions at a value 1 ms higher than the adaptive threshold (rounded to the nearest ms) to ensure that performance would be sufficiently high. For one subject (RLF), for whom adaptive thresholds were not obtained, the desired test burst delay was estimated from the results of a previous experiment in which discrimination performance had been evaluated using fixed trial blocks at various delays. All subjects also had adaptive runs with conditioning trains of nine bursts preceding the test burst pair. A comparison of these runs with the NC runs indicated that all subjects showed buildup; that is, they had lower thresholds in the NC condition than in the conditioning train procedure.

A blocked procedure was used to evaluate discrimination performance for the test burst as a function of the delay during the conditioning train. The conditioning train delays included the test burst delay and at least two shorter and three longer delays. For most subjects these delays

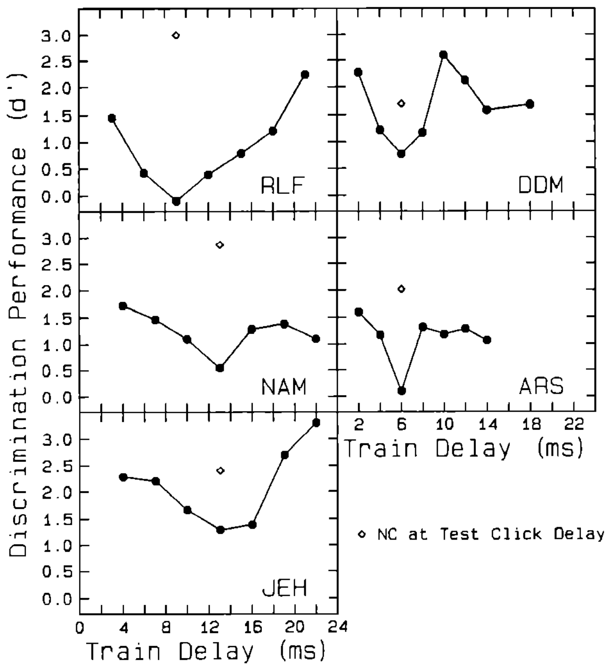

FIG. 1. Performance of five subjects on a task where delay of the test noise burst was held constant and delay of the preceding train of noise bursts was varied. Discriminability of the echo when not preceded by a train is indicated by the isolated open diamond. The delay for this condition varied among subjects because it was based on each individual's echo threshold. Immediately below each diamond is the condition in which a train of the same delay as the test burst preceded it. Surrounding this point are the conditions when the train delay was either shorter or longer than the test burst delay.

were at intervals of 3 ms, but for two subjects DDM and ARS, intervals were 2 ms because of their short test delay (6 ms). The conditioning train delay was fixed within a block of 20 trials. The lag test burst originated from 35° during ten randomly selected trials, and from 55° during the other ten. Subjects heard all conditions three times in a random order, with the constraint that all conditions were presented before any were repeated. Thus, measurement of discrimination performance for each subject was based on 60 total trials for each condition. Finally, three 20-trial blocks of the NC condition were run after the other testing was completed for comparison with the test burst train conditions and to ensure that performance on the test burst delay presented in isolation had not changed.

3. Subjects

Five normal-hearing listeners participated. All listeners had pure-tone air conduction thresholds less than or equal to 15 dB HL (re: ANSI, 1969) at 0.25, 0.5, 1.0, 2.0, 3.0, 4.0, 6.0, and 8.0 kHz, and had no more than a 10-dB difference between the ears at any test frequency. Three of the five subjects (RLF, ARS, JEH) had recently participated in previous research involving similar tasks. The other two listeners were given a minimum of 3-h practice before any data were collected.

B. Results

Data for the five subjects are displayed in Fig. 1, where discrimination performance is plotted as a function of the lag-burst delay during the conditioning train. Data for the NC (test burst in isolation) condition are shown as diamonds and indicate the test burst delay used for individual subjects. The difference between each diamond and the circle immediately below it shows that performance decreased when the test burst was preceded by a conditioning train having the same delay as the test burst. This demonstrates what has been termed the "buildup" of echo suppression (Clifton and Freyman, 1989). For all subjects the poorest discrimination performance occurred when conditioning and test burst delays coincided. The task became easier when the conditioning train delay was moved either above or below the test burst delay, although there were some differences in the shapes of the functions, particularly at the longer delays. Improved performance at the higher and lower delays, when compared with the control condition of the same delay for train and test bursts, suggests that the buildup of echo suppression was reduced or absent when conditioning and test bursts had different delays.

An increase or decrease in the train burst delay relative to test delay simulated a significant movement of the reflecting surface. For example, subject RLF's test burst delay was always 9 ms. If the preceding train delay was 3 ms, this shift simulated a jump of the reflecting surface from about 1 m away from the subject to about 2 m away. Likewise, this subject would experience a surface that "advanced" toward him from an original position of 3.2 m (train delay of 15 ms) to 2 m (test burst at 9 ms). We hypothesized that either of these changes presented the listener with an improbable perceptual event involving rapidly moving surfaces, which would lead to an alternative conclusion that the test burst echo was produced by a new sound source from a new location. In other words, the test burst echo would not be perceived as a reflection of the original sound and, therefore, would not be suppressed by the buildup process of the preceding train.

II. EXPERIMENT 2. CHANGES IN FREQUENCY BETWEEN TRAIN AND TEST NOISE BURSTS

An alternative explanation of the above findings is that *any* change between the train and test bursts would produce a disruption in buildup of echo suppression. Our hypothesis is that only changes that are improbable will disrupt the process. Variations in frequency and intensity should not disrupt echo suppression because such changes would simulate a change in the output of the original sound source, with both source and echo remaining in the same locations. This circumstance is encountered every day; in fact, sounds that exactly repeat are fairly infrequent in natural circumstances. A shift in frequency or intensity between train and test noise would signal that the original sound had simply changed in frequency or intensity, and the echo reflected this change. Echo threshold for the test noise should not be different from a train-test sequence that maintained the same frequency and intensity. In experiment 2 we compared echo threshold for a test noise presented in isolation versus a test noise preceded by a train that was either the same frequency or was different from the test noise. A test noise delay was selected for each subject such that the echo was easily discriminable when presented in isolation. We predicted that echo discriminability at this test noise delay would always be more difficult when preceded by a train, and that variations in frequency between train and test stimuli would not disrupt the suppression that resulted in an increased threshold.

A. Method

1. Stimuli and apparatus

The stimuli were narrow-band noises with bandwidths of 300 Hz and center frequencies of either 450 or 950 Hz. These frequencies were chosen because they are both in the low range (under 1500 Hz), close enough in frequency to have similar base line echo thresholds but far enough apart to be easily discriminable. The noises were digitized from the output of a sine-random generator (B & K type 1024) using a 16-bit A/D converter (TTES QAD1) running at 20 kHz. Approximately 415 ms of digitized noise was stored for each frequency. As in experiment 1, 4-ms segments of the noise were extracted randomly from longer noise segments, then were shaped with 2-ms linear rise/fall times for presentation during the experiment. The stimulus delivery apparatus was the same as in experiment 1. Stimulus level for both noise frequencies was 53 dBC on the "fast" meter response of a B & K sound level meter at the position of the listener's head.

2. Procedures

Using the identical forced choice procedure as in experiment 1, six experimental conditions were tested, three in which the test burst was the low-frequency noise and three in which the test burst was the high-frequency noise. The test burst was presented either in isolation or was preceded by a train consisting of nine noise bursts presented at four bursts/s. The noise bursts during the train were either the low- or high-frequency noise. Thus for the low-frequency test burst, the conditions were (1) no preceding train (NC LOW), (2) test burst preceded by a low-frequency train (LOW–LOW), and (3) test burst preceded by a high-frequency train (HIGH–LOW). The three parallel conditions for the high-frequency test burst were NC HIGH, HIGH–HIGH, and LOW–HIGH. In all cases lead and lag sounds were always identical in this experiment; shifts in frequency occurred between train and test bursts.

Because the purpose of the experiment was to determine whether the conditioning trains would increase the difficulty of localizing the lag signal relative to the same signal in isolation, it was again necessary to find delays which produced at least satisfactory discrimination performance for the NC condition. In this study, the delays were selected individually for each subject and each noise frequency by running preliminary tests on isolated noise bursts (NC condition) in blocks of 30 trials at a fixed delay. The adaptive procedure used for this purpose in experiment 1 was not used again because it was found to be time consuming and difficult for some listeners. Two new normal-hearing subjects and two subjects from the experiment 1 participated in this study. One additional new subject began the preliminary testing, but was excluded because she could not discriminate between the two test loudspeaker positions at long delays well above her threshold for subjectively reporting hearing an echo. The new subjects (DAN, ROB) were first tested on a fixed-delay block at a long delay (18–20 ms) at which the echo was clearly audible in order to familiarize them with the task. This practice continued until greater than 90% correct performance was achieved. For subjects who had run in previous experiments (RLF, NAM), the first delay tested was closer to assumed threshold. For all subjects, the delays were increased or decreased as necessary until performance on 90 trials (three blocks) at a single delay yielded a d' in the range of 1.5 to 2.0. For each subject, the delays arrived at through this procedure were 1–3 ms larger for the high-frequency noise than the low-frequency noise (see Table I, top panel). However, during the main part of the experiment, the delay in the conditioning train always matched the delay of the test burst, even if the frequencies were different to prevent a confounding of frequency and delay effects when going from train to test bursts.

TABLE I. Echo thresholds for the isolated test noise (NC).

	Low-frequency threshold (ms)	High-frequency threshold (ms)
Experiment 2—frequency		
RLF	7	10
ROB	12	15
NAM	14	15
DAN	6	9
Mean	9.75	12.25
	Low-intensity threshold (ms)	High-intensity threshold (ms)
Experiment 3—intensity		
LRA	7	6
KAC	7	8
RKC	6	6
DAN	5	5
RLF	8	8
GRE	12	12
Mean	7.5	7.5

Following the search for the appropriate test burst delay, subjects were screened to exclude any who did not show buildup of echo suppression under presumably optimum conditions, i.e., HIGH–HIGH and LOW–LOW. Single 30-trial blocks were run for both of these conditions and the results compared to the relevant NC data obtained above. One subject (DDM), who participated in experiment 1, was excluded because the results with the conditioning trains were not different from the NC condition for either high- or low-frequency noises, indicating no buildup.

In the main part of the experiment, three 30-trial blocks were run for each of the six conditions, for a total of 18 blocks. Blocks for all six conditions were run in a random order once before being repeated twice more with new random orders. The 18 blocks typically required three experimental sessions. At the beginning of each session, sub-

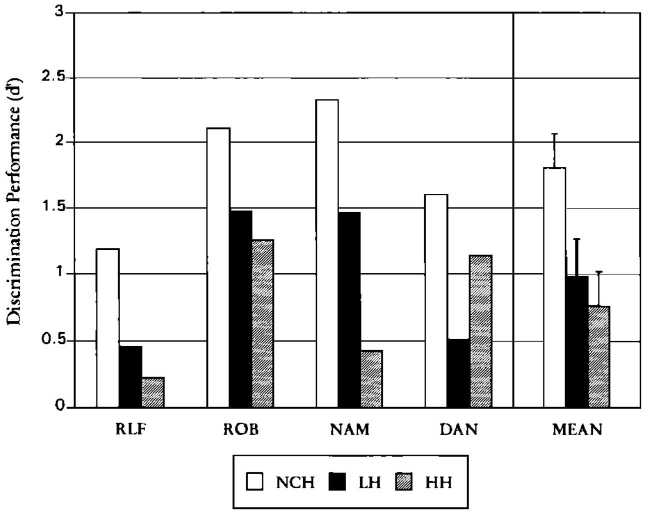

FIG. 2. Discrimination performance of subjects for the high-frequency test burst when the preceding train of noise bursts was lower (LH) or was the same (HH), compared to when the test burst was presented in isolation (NCH). Four subjects' data are plotted, along with the averaged data in the far right column with error bars.

jects listened to 20 practice trials, ten each of NC HIGH and NC LOW.

B. Results

Discrimination performance (d') for each subject is plotted in Fig. 2 for the high test burst frequency and in Fig. 3 for the low test frequency. Average data are shown at the right side of the figures. The location of the echo was easier to discriminate when the test burst was presented in isolation (NCH and NCL), regardless of whether the train was the same frequency or a different frequency. The data were analyzed in a 3 (condition)×2 (frequency) analysis of variance. A main effect of condition [$F(2,6)=23.42$, $p<0.001$] was the only significant source of variance. A follow-up ANOVA compared the NC conditions (NC HIGH and NC LOW) with the train-test sequences that featured a frequency change (HIGH–LOW and LOW–HIGH). This comparison was highly significant [$F(1,3)=93.26$, $p<0.002$], indicating the buildup in echo suppression was maintained across the shift in frequency. A second follow-up ANOVA found no difference between conditions when train and test bursts had the same frequency (HIGH–HIGH and LOW–LOW) versus different frequencies (HIGH–LOW and LOW–HIGH). This was further evidence that changes in frequency do not disrupt echo suppression. (Note: we did not test the remaining orthogonal comparison of NC conditions versus train-test sequences with the same frequency because subjects were screened to ensure that they had a difference in echo threshold for these conditions—see Sec. II A.)

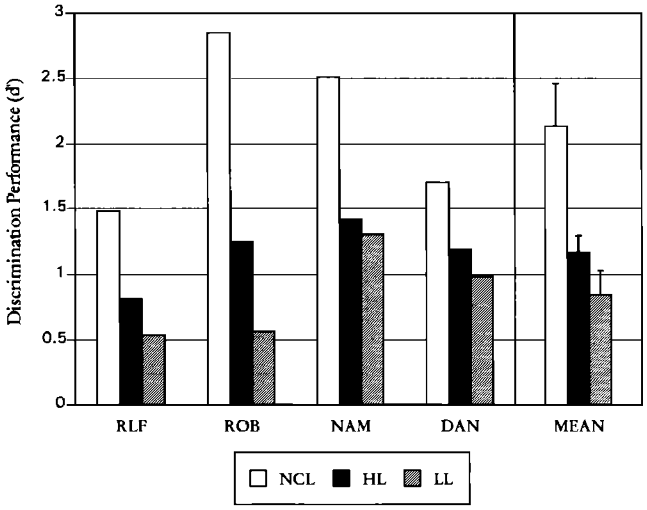

FIG. 3. Discrimination performance of subjects for the low-frequency test bursts when the preceding train of noise bursts was higher (HL) or was the same (LL), compared to when the test burst was presented in isolation (NCL). Four subjects' data are plotted, along with the averaged data in the far right column with error bars.

III. EXPERIMENT 3. CHANGE IN INTENSITY BETWEEN TRAIN AND TEST NOISE BURSTS

As predicted, variation in frequency had no effect on the buildup in echo suppression produced by ongoing stimulation. In experiment 3 we varied intensity in a similar manner and again predicted this would not affect the buildup process because variations in intensity signal nothing new or improbable about the acoustic situation in the room. Such variations would be interpreted as intensity changes in the original sound source. In experiment 3 a test noise delay was selected for each subject that allowed easy discrimination of the echo. Again, we predicted that echo discriminability at this test noise delay would always be more difficult when preceded by a train, and that variations in intensity between train and test stimuli would have no effect on performance.

A. Method

As in experiment 1, the stimuli were 4-ms segments of white noise shaped with linear 2-ms rise/fall durations. Presentation level was either at 50 dBC (HIGH) or 40 dBC (LOW) as measured by equipment and procedures described previously. The design of the experiment, number of blocks and trials, and screening and practice procedures were identical to those of experiment 2. The sole difference was that in the current experiment, HIGH and LOW refer to intensities rather than frequencies. When subjects were screened with the NC condition to determine their test noise delay, there were no differences in echo threshold for high- and low-intensity sounds ($M=7.5$ ms for both conditions; see Table I).

Six subjects participated who met the criteria for normal-hearing described previously. Two additional subjects were screened; one was dropped because of no buildup and one because of failure to meet normal-hearing requirements. Of the six participants, RLF had participated in both experiments 1 and 2. DAN had been a subject in experiment 2. RKC and KAC participated for the first time in this set of studies, but had previous experience with similar stimuli and tasks. LRA and GRE were new subjects.

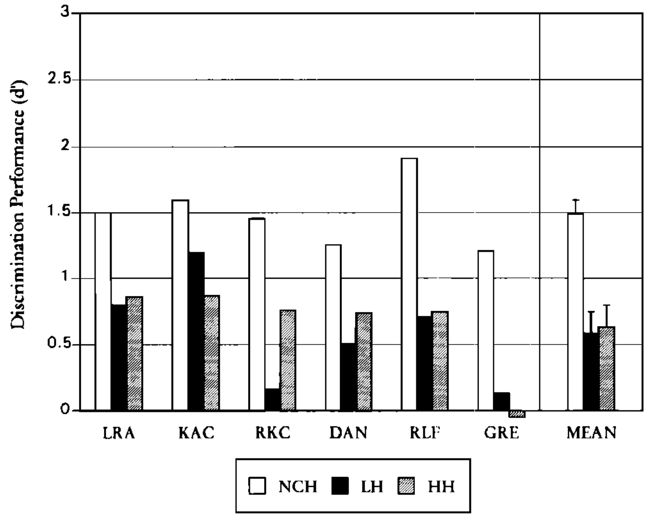

FIG. 4. Discrimination performance of subjects for the high-intensity test bursts when the preceding train of noise bursts was lower (LH) or was the same (HH), compared to when the test burst was presented in isolation (NCH). Six subjects' data are plotted with the averaged data in the far right column.

B. Results

Figures 4 and 5 display subjects's discrimination performance for the high- and low-intensity test noises, respectively. Across the three conditions, the intensity data show a similar pattern to that obtained for frequency variations in experiment 2. As before, performance was always best when there was no conditioning train. A 3 (condition)×2 (intensity) analysis of variance (ANOVA) revealed a main effect of condition [$F(2,10) = 25.34$, $p < 0.001$]. A follow-up ANOVA comparing the NC conditions with the train-test conditions differing in intensity (HIGH–LOW and LOW–HIGH) indicated that echo direction was discriminated significantly better in the NC condition [$F(1,5) = 23.07$, $p < 0.005$]. The train-test sequences with different intensities (HIGH–LOW and LOW–HIGH) were not different from those with the same intensities (HIGH–HIGH and LOW–LOW) [$F(1,5) < 1.0$]. No effect of test burst intensity was found. These results suggest that the buildup of echo suppression during the conditioning train is not broken down by a sudden intensity change between the train and test burst.

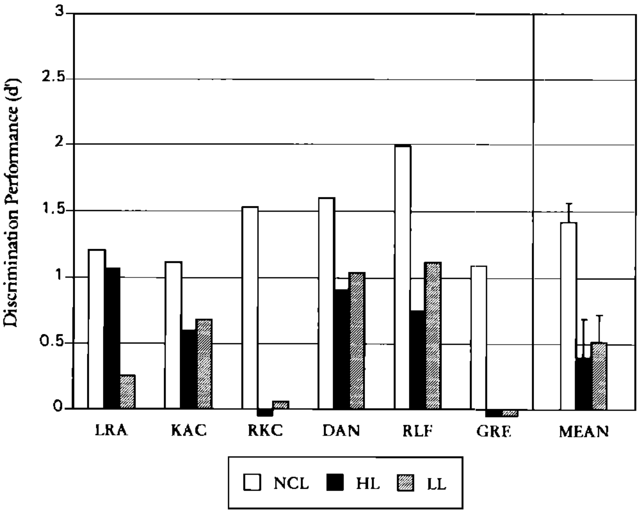

FIG. 5. Discrimination performance of subjects for the low-intensity test bursts when the preceding train of noise bursts was higher (HL) or was the same (LL), compared to when the test burst was presented in isolation (NCL). Six subjects' data are plotted with the averaged data in the far right column.

IV. GENERAL DISCUSSION

Three experiments tested the hypothesis that improbable echoes would reset echo threshold to an unadapted level whereas "reasonable" or expected echoes would not. In the present experiments probable and improbable echoes were created by first presenting the listener with a train of lead–lag noise pairs that were immediately followed by a single instance of a test pair that differed in some way from the train. We hypothesized that expectations built-up during the repetitious train were either violated or confirmed on the test burst, with echo discriminability improved in the first case and not changed in the second case.

An improbable echo was considered to be one that would occur rarely in natural surroundings, a circumstance implemented in experiment 1 by a delay between lead and lag that suddenly changed, specifying that the reflecting surface had moved either toward or away from the listener. Such an abrupt movement would be unusual in the listener's experience and was predicted to lower echo threshold relative to an expected echo. All subjects exhibited better discrimination of the echo on trials when the test burst delay was either shorter or longer than the train burst delay, compared to trials when the test and train burst delays were the same. Experiments 2 and 3 featured changes in the test burst that were not expected to violate the listener's expectations built-up during the train. Frequency (experiment 2) and intensity (experiment 3) changes were introduced between train and test bursts; these changes inform the listener that the sound source itself changed with regard to these acoustic properties. Such changes are not unusual and should be unsurprising to any listener. The key feature is not a test burst change *per se*, but a change that informs about the acoustic properties of the room.

The broader theoretical context for these findings grows out of Bregman's (1990) approach of trying to understand how we construct reasonable auditory events to make up the "auditory scene." The traditional view of the precedence effect as an inhibitory process based on binaural cross correlation (e.g., Lindemann's model, 1986) appears to be too limited. Rather, the precedence effect should be viewed as a complex decision-making process whose end product is the assignment of sounds to their apparent locations and objects. We are not the first to propose a role for cognitive processes in the precedence effect. Hafter *et al.* (1988), in discussing their findings on binaural adaptation, proposed that the precedence effect is "a case of sensory rivalry" (p. 670). Echoes present the listener with conflicting sensory information about a sound's direction, a conflict that is usually resolved in favor of the first arriving wave front. They made an analogy between precedence and ventriloquism, with echoes being "captured" at the leading sound's location just as the voice gets captured by vision in ventriloquism. A second example of how the listener's expectations about the world influence decisions comes from Rakerd and Hartmann (1985). These authors developed a "plausibility hypothesis" to explain results from time-intensity trade experiments. Their subjects appeared to weight interaural time cues less when they were extreme values, that is beyond the range that would be plausible for their head size. In these cases listeners weighted interaural intensity cues more heavily in their localization judgements, more or less discounting the "implausible" interaural time differences (ITDs).

The hallmark that distinguishes Hafter *et al.* (1988; Hafter, 1984) and Rakerd and Hartmann (1985) from traditional views of the precedence effect is the proposal that localization of sound is a dynamic, interactive process that has input from higher cognitive levels. Our present data extend this view to include the assumption that listener's expectations are influenced by information about the acoustic properties of the room carried in the echoes. These properties include the volume of the room, the placement and distance of reflecting surfaces and absorbency of surfaces. We assume that expectations are built-up quickly and automatically by hearing sound produced in the room. Several hypotheses emerge from this assumption. (1) Echo threshold should be a dynamic process, subject to change during an ongoing sound as information is evaluated. In support of this hypothesis Clifton and Freyman (1989) found that subjects who were instructed to report continually on whether they heard an echo had systematic variations in threshold during a click train lasting several seconds. (2) Echo threshold should be lower at the beginning of a train than at the end (Clifton and Freyman, 1989; Freyman *et al.*, 1991). (3) Echo threshold should be lower when changes in the ongoing echo signify either new properties of reflective surfaces in the room or unusual movements of these surfaces. The third hypothesis is the richest vein for experimental exploration because a wide variety of stimulus manipulations can be devised to test it. Current support for this hypothesis comes from a number of studies. Clifton (1987) reported that an abrupt switch in the lead and lag positions in an ongoing click train would disrupt echo suppression so that the listener heard clicks from both lead and lag sides for a few clicks, then localization of the lag click "faded out," leaving only the leading click. Clifton and Freyman (1989) quantified this effect by varying delay between lead and lag clicks and click rate. The breakdown in echo suppression produced by switching lead and lag positions can be predicted in terms of violating expectations. Before the switch the listener localizes the click train only on the leading side because the train is presented with a delay below echo threshold. Although it is not localized as a separate auditory event, the echo contributes to the perception of the lead click by influencing its timbre and apparent position. The abrupt switch in lead and lag positions surprises the listener by informing him or her that the reflective surface producing the echo has suddenly moved from one hemifield to the other. Blauert and Col (1992) found that if the switch in lead and lag positions was repeated several times, subjects appeared to adjust to the switch and echo threshold was no longer lowered. The effect of repetition should lead subjects to expect this odd, but predictable movement of the echo. In this procedure the sound source also switched sides, but this may or may not be critical. We are accustomed to sound sources coming from a variety of locations in a room because sound can originate from any direction and is often produced by moving objects (people, animals, bouncing balls, etc.). However, most of the large reflective surfaces in a room (walls, ceiling, furniture, floor) are apt to remain stable. Further research is needed to determine the extent to which movement of the lead sound versus movement of the lag sound affects echo threshold, but at a minimum our expectation hypothesis predicts the lowering of echo threshold when both lead and lag are switched.

Freyman *et al.* (1991, experiment 3) offered additional confirmation of the hypothesis that only changes in the echo that signify new properties of reflective surfaces would lower echo threshold. Click trains were presented to subjects in three conditions: (1) brief noise bursts from the leading side only; (2) noise bursts from the lagging side only, and (3) bursts from both leading and lagging sides. The same test burst from lead and lag followed each type of train to assess the train's effect. Discrimination of the echo was most difficult when the echo was present during the train. When the train was composed of either lead only or lag only, echo threshold for the test bursts was 7–8 ms lower than when preceded by trains having both lead and lag present. During a train of single source sounds composed of either lead only or lag only, the expectation would be built-up that no reflecting surface was present (the subjects were run in an anechoic chamber). The test burst was a sudden disconfirmation in which an echo appeared where none had been heard before. We hypothesize that this disconfirmation produced a lower echo threshold.

The results from the present three experiments provide a final example of how expectations influence our perception of echoes. Our general prediction was that any difference between train and test bursts that signified an unexpected change in reflective surfaces should affect echo threshold but that other changes would have little or no effect. This was confirmed in all three experiments. In addition to the manipulations attempted here, many others can be suggested. For example, a rapid lateral movement of the lag sound should affect echo threshold, as did the simulated movement in depth in experiment 1. Another interesting change would be an increase or decrease in intensity of the lag sound (suggested by Les Bernstein, 1992). Unlike the intensity difference in experiment 3 that specified a change in the sound source's intensity which did not affect echo threshold, a change in only the lag sound's intensity would simulate a change in absorbency of the reflecting surface and should affect echo threshold. A final consideration is that greater changes in frequency or intensity than were used in the present experiments might produce different results, although there is no *a priori* reason to expect this.

In summary, our conceptualization of the precedence effect as a decision-making process that is influenced by ongoing auditory stimulation in the room is a departure from previous descriptions of the phenomenon as a static echo suppression mechanism. It emphasizes the role of experience in modulating the listener's perception, and fits generally with Bregman's (1990) and Handel's (1989) analysis of how listeners resolve ambiguous signals. One interesting developmental implication from our formulation is that very young organisms should not have echo suppression until they have accrued sufficient experience in spatial hearing. Neither human newborns (Clifton *et al.*, 1981; Clifton *et al.*, 1984) nor newborn puppies (Ashmead *et al.*, 1986) orient their heads toward the leading side of a precedence effect sound, although human infants do orient toward a single source sound at birth and puppies do so around 2 weeks of age. Orientation toward the leading side of a precedence effect sound is seen around 16 weeks of age for infants and at sometime past 6 weeks for puppies. Once the infant has shown a basic ability to weight the leading sound stronger than the delayed sound, continuing changes in echo perception can be observed throughout infancy and the preschool years (Morrongiello *et al.*, 1984; Litovsky, 1991). While there are many possible reasons for these developmental changes in the precedence effect, at least the pattern shown during development does not contradict the role of experience proposed here.

ACKNOWLEDGMENTS

This research was supported by a grant from National Institute on Deafness and Other Communication Disorders, DC 01625, to RLF and RKC, and by a Research Scientist Award from the National Institute of Mental Health (MH00332) to RKC. Ruth Litovsky's current address is Department of Neurophysiology, University of Wisconsin Medical Center, Madison, WI 53706.

ANSI (**1969**). ANSI S3.6-1969, "Specifications for Audiometers" (American National Standards Institute, New York).

Ashmead, D., Clifton, R., and Reese, E. (**1986**). "Development of auditory localization in dogs: Single source and precedence effect sounds," Dev. Psychobiol. **19**, 91–104.

Bernstein, L. (**1992**). Personal communication.

Blauert, J. (**1983**). *Spatial Hearing* (MIT, Cambridge, MA).

Blauert, J., and Col, J. P. (**1992**). "Irregularities in the precedence effect," in *Auditory Physiology and Perception*, edited by Y. Cazals, K. Horner, and L. Demany (Pergamon, Oxford), pp. 531–538.

Blauert, J., and Divenyi, P. L. (**1988**). "Spectral selectivity in binaural contralateral inhibition," Acustica **66**, 267–274.

Bregman, A. S. (**1990**). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT, Cambridge, MA).

Clifton, R. K. (**1987**). "Breakdown of echo suppression in the precedence effect," J. Acoust. Soc. Am. **82**, 1834–1835.

Clifton, R. K., and Freyman, R. L. (**1989**). "Effect of click rate and delay on breakdown of the precedence effect," Percept. Psychophys. **46**, 139–145.

Clifton, R. K., Morrongiello, B. A., and Dowd, J. M. (**1984**). "A developmental look at an auditory illusion: The precedence effect," Dev. Psychobiol. **17**, 519–536.

Clifton, R. K., Morrongiello, B. A., Kulig, J., and Dowd, J. M. (**1981**). "Newborns' orientation toward sound: Possible implications for cortical development," Child Dev. **52**, 833–838.

Clifton, R. K., Shinn-Cunningham, B., Zurek, P., and Durlach, N. (**1989**). "Cross-frequency interactions in the precedence effect," J. Acoust. Soc. Am. Suppl. 1 **85**, S83.

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

Freyman, R. L., Clifton, R. K., and Litovsky, R. Y. (**1991**). "Dynamic processes in the precedence effect," J. Acoust. Soc. Am. **90**, 874–884.

Gardner, M. D. (**1968**). "Historical background of the Haas and/or precedence effect," J. Acoust. Soc. Am. **43**, 1243–1248.

Green, D. (**1976**). *Introduction to Hearing* (Erlbaum, Hillsdale, NJ).

Hafter, E. R. (**1984**). "Spatial hearing and the duplex theory: How viable is the model?" in *Aspects of Neocortical Function*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 425–448.

Hafter, E.R., Buell, T. N., and Richards, V. M. (**1988**). "Onset-coding in lateralization: Its form, site, and function," in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 647–676.

Handel, S. (**1989**). *Listening: An Introduction to the Perception of Auditory Events* (MIT, Cambridge, MA).

Hartmann, W. M. (**1983**). "Localization of sound in rooms," J. Acoust. Soc. Am. **74**, 1380–1391.

Levitt, H. (**1971**). "Transformed up–down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lindemann, W. (**1986**). "Extension of a binaural cross-correlation model by contralateral inhibition II: The law of the first wavefront," J. Acoust. Soc. Am. **80**, 1623–1630.

Litovsky, R. Y. (**1991**). "Developmental changes in sound localization precision under conditions of the precedence effect," Ph.D. thesis, University of Massachusetts-Amherst.

Mills, A. (**1972**). "Auditory localization," in *Foundations of Modern Auditory Theory, Vol. 2*, edited by J. V. Tobias (Academic, New York).

Moore, B. C. J. (**1989**). *Introduction to the Psychology of Hearing*, 3rd ed. (Academic, London).

Morrongiello, B. A., Kulig, J., and Clifton, R. K. (**1984**). "Developmental changes in auditory temporal perception," Child Dev. **55**, 461–471.

Perrott, D. R., Marlborough, K., Merrill, P., and Strybel, T. A. (**1989**). "Minimum audible angle thresholds obtained in conditions in which the precedence effect is assumed to operate," J. Acoust. Soc. Am. **85**, 282–288.

Rakerd, B., and Hartmann, W. M. (**1985**). "Localization of sound in rooms, II: The effects of a single reflecting surface," J. Acoust. Soc. Am. **78**, 524–533.

Wallach, H., Newmann, E. B., and Rosenzweig, M. R. (**1949**). "The precedence effect in sound localization," Am. J. Psychol. **52**, 315–336.

Zurek, P. M. (**1980**). "The precedence effect and its possible role in the avoidance of interaural ambiguities," J. Acoust. Soc. Am. **67**, 952–964.