

Release and re-buildup of listeners' models of auditory space

Rachel Keen^{a)}

University of Virginia, Charlottesville, Virginia 22901

Richard L. Freyman^{b)}

Department of Communication Disorders, University of Massachusetts, Amherst, Massachusetts 01003

(Received 4 August 2008; revised 11 February 2009; accepted 13 February 2009)

When listeners hear sound presented repeatedly in a room with reflections, echo threshold rises. The current experiments tested how long this buildup in echo threshold would last when exposure to a different simulated space (designated as room B) intervened before returning to the original space (designated room A). Stimuli were trains of lead-lag click pairs (room A) and trains of clicks with no reflections (room B) in an ABA sequence. After buildup in room A, echo threshold for click pairs in room A decreased in direct relation to amount of intervening exposure to room B. After 11 click pairs of room B, the effect of exposure to room A was gone. A second buildup in echo threshold in room A was not differentially affected by prior exposure to room A or a different simulated room, room C. Listeners appear to form a model when exposed to sound in a particular space, which is lost quickly upon hearing sound in a different space. Storing previous models is inefficient because the processes of buildup and breakdown occur quickly to sound in a new space.

© 2009 Acoustical Society of America. [DOI: 10.1121/1.3097472]

PACS number(s): 43.66.Pn, 43.66.Qp [BCJM]

Pages: 3243–3252

I. INTRODUCTION

Imagine that you are walking through a house while continuously conversing with an acquaintance. As you progress through the house, the acoustics will vary from room to room, depending on room size, height of ceiling, furniture, rugs, curtains, and so on. In addition to sound waves from the original source, reflections from various surfaces in the room will bounce back to create an overall perception of the room's acoustics. Unless the room is large enough to create long delays between the sound source and its reflections, the reflected sounds will be below echo threshold. That is, they will not be heard as separate sounds localized apart from the original source. Rather, the reflections fuse in location with the original sound, exerting their effect by changing the timbre and loudness of the fused sound (Blauert, 1997; Litovsky *et al.*, 1999). The fusion of the original and reflected sounds into a single image that is perceived at the location of the original sound is referred to as the precedence effect and serves to enable the listener to localize the sound source with reasonable accuracy despite the presence of potentially conflicting directional information from reflections.

In addition to aiding sound localization, we have proposed that the precedence effect plays an important role in informing the listener about a room's acoustics (Clifton *et al.*, 1994; Clifton and Freyman, 1997). Reflected sound, although below echo threshold, is analyzed to form a model of the auditory space. Writing about musical acoustics, Benade (1976) (pp. 208–210) proposed such a process as part of the precedence effect, noting that it informed the listener

about objects in the room, distance of walls, ceiling, and so on. Benade (1976) claimed that the listener was very sensitive to frequency and amplitude differences between reflections and the original sound, a claim that has been upheld by later empirical work (Bech, 1995, 1996, 1998). Our research has borne out Benade's (1976) intuitions about this aspect of the precedence effect. Furthermore, we hypothesized that the listener develops expectations about the room's acoustics based on a model formed by ongoing input (Clifton *et al.*, 1994; Clifton and Freyman, 1997). When these expectations are violated by incompatible input, the model collapses, to be replaced by a new model built on the latest input. We refer to this hypothesized process as the "room-acoustics model" in this paper.

This process of buildup and breakdown of the model can be monitored by measuring the listener's echo threshold (for review, see Clifton and Freyman, 1997). It can be raised (called buildup) or lowered (called breakdown), depending on the ongoing sound. A typical trial sequence is to present a train of repeated click pairs followed by a test click pair that is identical to the train (testing buildup) or differs from the train (testing breakdown). The control condition is the test click presented in isolation, without the preceding train. The listener's task is to respond to the test click, indicating whether or not the lagging sound was heard at a location separate from that of the leading sound. While repeating the same click pairs raises echo threshold (Freyman *et al.*, 1991), it can then be lowered by introducing any of several changes. All of the following changes from the repeating train will lower echo threshold: a test click with a different delay between lead and lag from what was heard in the repeating train (Clifton *et al.*, 1994), a lag sound with a different spectrum from what it had in the train while the lead remains unchanged (McCall *et al.*, 1998), and when there is a switch in lead and lag sound locations (Clifton, 1987; Clifton and

^{a)}Formerly Rachel K. Clifton.

^{b)}Author to whom correspondence should be addressed. Electronic mail: rlf@comdis.umass.edu

Freyman, 1989; Blauert and Col, 1992; Yost and Guzman, 1996). Clifton and Freyman (1997) proposed that only those changes that give information about the room's acoustics would affect echo threshold. For example, a different time delay between source and reflection would signal that the reflecting surface had moved, and a different spectrum for the delayed sound echo would signal that some property of the reflecting surface had changed.

A change between train and test click that does not signal a change in room acoustics does not affect echo threshold. For example, when both lead and lag sounds were changed in frequency or in intensity between train and test, echo threshold was unchanged (Clifton *et al.*, 1994). Congruent changes signal that the source changed, and the "echo" reflected this change. As long as congruent changes in both source and reflections are made, echo threshold is not affected (Clifton, *et al.*, 1994; Yost and Guzman, 1996). This conception of how changes in echo threshold are produced is akin to the "plausibility hypothesis" proposed by Rakerd and Hartmann (1985) and Hartmann and Rakerd (1989). They noted that subjects in time-intensity trade experiments tend to discount values of interaural time delay (ITD) that are outside the range that could be plausible given their head size and location within the room. Rakerd and Hartmann (1985) found that when the ITD cue was implausible, listeners responded on the basis of the interaural intensity difference (IID) cue. In their case and ours, it is proposed that listeners are influenced by their implicit knowledge about how sound behaves in a room when making localization judgments.

By measuring echo threshold under varying conditions, we can examine timing parameters for buildup and breakdown of the precedence effect. The time course of these processes can tell us not only about the functioning of the auditory system but also about the nature of the neural processes themselves. Shifts in echo threshold reflect a decision-making process in the brain. Sanders *et al.* (2008) found larger negativity in the event related potentials at anterior-central and medial sites when listeners reported hearing an echo, than when not hearing an echo, for the same click pair at a delay chosen to be at each listener's echo threshold. The mystery of the precedence effect is why the brain sometimes hears the reflected sound as an independent sound source and at other times suppresses its location information while preserving information about room acoustics. There are many unanswered questions about this process. For example, once buildup in echo suppression has occurred, how long will it last? Does it decay naturally with the passage of time? Using the analogy of walking through rooms while talking, if the talkers fell silent for a few moments while standing in the room, would the listener's model of that room's acoustics collapse? Djelani and Blauert (2000) tested the effect of 1, 4, and 9 s of silence between train and test click. They found that 1 s of silence had little effect, but the longer periods led to monotonic decreases in echo threshold. The time course of the model's decay needs to be systematically examined in relation to echo threshold for a click pair with no preceding train that produces buildup.

The room acoustics hypothesis specifies what type of input would disrupt the model, but not how the process unfolds. Our previous research featured a train of identical click pairs followed by a test click pair that differed from those in the train. This procedure tested whether the listener detected a difference between train and test, evidenced by a lower threshold to the changed test click compared to a test click like those in the train. However, this procedure leaves unanswered the question of whether echo threshold for the stimuli in the ongoing train was disrupted by changing the test click. Djelani and Blauert (2001) found that a brief injection of one aberrant click pair (they reversed the location of the lag) at the end of the train had no effect on echo threshold for a test click pair identical to those in the train. After replicating that basic result, Freyman and Keen (2006) (Exp. 3) increased the number of aberrant sounds to five consecutive presentations before representing the original click pair configuration. Echo threshold for the original configuration decreased by several milliseconds. We concluded that breakdown of a room model does not occur instantly, but like buildup it is a graded process that depends on the amount of input. To understand the time course of a model's decay, the full range of input, from completely ineffective to that sufficient to produce complete breakdown, needs to be investigated.

Another critical question is whether there are lingering effects of buildup such that re-exposure to the same acoustic parameters after breakdown would show savings. To use the room analogy again, if after exposure to other rooms the listener came back to a previously visited room, would buildup be faster the second time around? This issue is related to whether models of familiar rooms are stored in the brain. A study by Robart and Rosenblum (2005) suggests that this is possible. They found that listeners could identify in which of several rooms a sound had been recorded in (e.g., a gym, restroom, classroom), suggesting that models of various spaces could be held simultaneously in memory.

The current experiments attempt to answer several questions about the formation and disruption of a listener's model for how sound is reflected in a room. The purpose of experiment 1 was to determine how varying amounts of exposure to a new space or "room" would affect retention of a model of a previous space. Freyman and Keen (2006) found that five exposures to a new space was sufficient to decrease echo threshold for an existing model, but we do not know what the smallest number for causing a disruption is, or if more than five exposures would reduce echo threshold still further. In experiment 2 listeners experienced a second buildup after the initial model had been broken down. After buildup followed by maximum breakdown for one stimulus configuration was accomplished, we tested whether re-exposure to that room would produce a more rapid buildup than that seen initially. Experiment 2 featured manipulations and controls to test whether savings would occur under a variety of conditions.

II. EXPERIMENT 1

The primary purpose of this experiment was to determine the boundaries for minimum and maximum breakdown

TABLE I. List of stimulus conditions for experiment 1.

Segment 1	Segment 2	Segment 3	Name/description
Train	Train Room A5	Test Room A1	Basic buildup A5 A1
Train Room A5 Room A5 Room A5 Room A5 Room A5 Room A5	Train Room B1 Room B3 Room B5 Room B7 Room B9 Room B11	Test Room A1 Room A1 Room A1 Room A1 Room A1 Room A1	Buildup then breakdown A5 B1 A1 A5 B3 A1 A5 B5 A1 A5 B7 A1 A5 B9 A1 A5 B11 A1
Train Room A5 Room A5	Train S5 S11	Test Room A1 Room A1	Buildup then silence A5 S5 A1 A5 S11 A1
Train	Train Room B1 Room B3 Room B5 Room B11	Test Room A1 Room A1 Room A1 Room A1	Breakdown only B1 A1 B3 A1 B5 A1 B11 A1
Train	Train	Test Room A1	No conditioning train NC

of a room acoustics model. The procedure was to first allow the model to be built with a click train of five exposures that featured a left-side leading click with a right-side lagging click to simulate a single reflective surface. Without interrupting the click train, there immediately followed clicks from only the left side, ending with the test click, which was a single click pair identical to the pairs used in the first train. The amount of “new space” input varied from a single click on the left side (not expected to have an effect) to 11 consecutive left-only clicks (expected to lead to complete breakdown of the previous model). A necessary control in experiment 1 was used to evaluate whether the built-up model would last if an equivalent length of silence intervened rather than the new input before the test click.

A second control tested listeners’ response to the test click without prior buildup to that input; that is, a train of clicks from only the left side preceded the test click pair. This condition simulated sound in an anechoic room, followed by presentation of the same sound in the same location but with a simulated reflection present. By using this anechoic condition as the intervening or “new” room in all experimental conditions described below (referred to as room B in Table I), we sought to present a highly contrasting space with the initial buildup condition. In previous research (Freyman *et al.*, 1991), we found that exposure to a delayed sound after hearing single-source sounds in an anechoic condition lowered echo threshold even below the threshold for the isolated test click. In other words, preceding a lead-lag click pair with a train of lead-side-only clicks caused the echo to “pop out,” producing very low thresholds of 6 ms. Interestingly, the pop out does not occur, at least to the same extent, if the preceding click train has a different echo delay

or location from that of the test click (as opposed to having no echo) (Clifton *et al.*, 1994, 2002). Thus, the contrast between an anechoic space versus a space with echoes appears to be greater than the contrast between two spaces having different echoes. Using the anechoic condition as the intervening space before the test click was expected to widen the range of echo thresholds among experimental conditions, compared to the introduction of a subtle change in room acoustics. The control condition of the anechoic stimulus preceding the test click without the preceding buildup was a necessary check on the powerful influence of the anechoic clicks on the test click after the preceding buildup.

A. Methods

1. Stimuli and apparatus

Stimuli were pairs of computer generated 150-ms pulses presented from two channels of a 16-bit digital/analog converter (TTES QDA1). The outputs of the two signal channels were low-pass filtered at 8.5 kHz (TTE J1390), attenuated (TTES PAT1), amplified (CROWN D40), and delivered to a pair of loudspeakers (Realistic Minimus 7). The loudspeakers rested on a semicircular arc constructed of foam-covered wood that was housed in an anechoic chamber measuring $4.9 \times 4.1 \times 3.12$ m. The floor, ceiling, and walls of the chamber were lined with 0.72 m foam wedges. Subjects sat in a chair in the center of the room with the loudspeakers situated at 45° left (-45°) and 45° right ($+45^\circ$) at a distance of 1.9 m. The center of the loudspeakers was at ear height for the typical listener seated in the chair. The stimulus level was measured by presenting the click stimuli through the loudspeakers at a rate of 4 clicks/s. A microphone was lowered to the position of the center of the subject’s head with the subject absent. The microphone output was fed to a sound level meter (B&K 2204) set on the “fast” meter response on the A-scale. Unattenuated outputs through the system were 61 dBA from either loudspeaker. The experiments were run at 43 dBA (with attenuators set to 18 dB).

2. Procedures and conditions

The primary conditions of the experiment employed a train-test method used previously in a number of studies (e.g., Freyman *et al.*, 1991; Clifton *et al.*, 1994; Grantham, 1996; Yost and Guzman, 1996; Yang and Grantham, 1997; Djelani and Blauert, 2001; Freyman and Keen, 2006). On each trial, repeated pairs of clicks (one to each loudspeaker) were delivered at a rate of 4 clicks/s to form a click “train.” Following the train and a pause of 750 ms, a “test click” was presented. In all conditions, the test click from the right loudspeaker was delayed relative to that from the left loudspeaker by 2–14 ms in 2-ms steps. The clicks during the train were either click pairs (one from the left loudspeaker and one from the right), single-source clicks from the left loudspeaker, or a sequential mixture of the two. In all cases in which click pairs were used, the delay used during the train matched that for the test click. The listeners’ task was to report whether

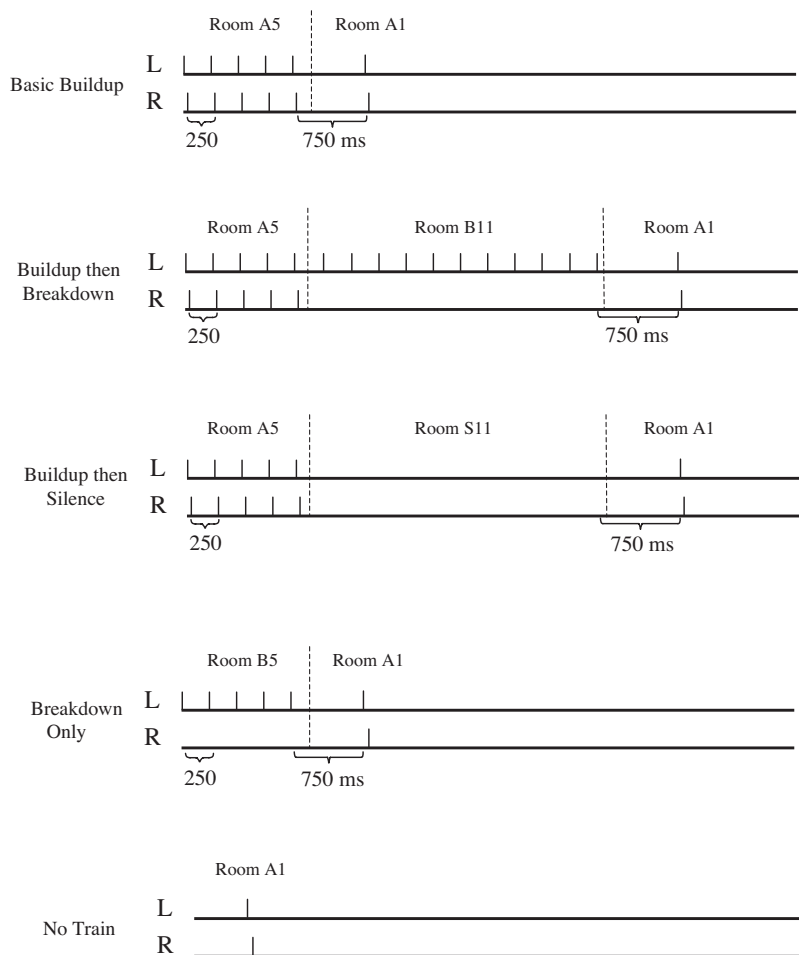


FIG. 1. Schematic illustrations of examples of stimuli used in experiment 1.

they heard a sound in the vicinity of the right (lagging) loudspeaker during the test click. In this design, the test click pair had the same locations (lead left, lag right) in all conditions, so the effects of various conditioning trains on the same test click could be assessed.

Table I lists all conditions tested in this experiment, and Fig. 1 shows example stimuli from the different conditions. The table refers to the lead-lag configurations as representing crude simulations of auditory space referred to as rooms. We call the room with a reflective surface only on the right “room A.” The first row of the table shows the basic buildup condition against which other conditions were compared. The train presents five repetitions of room A (called A5), followed by a pause, and then the test click, which is a single presentation of room A (labeled A1). The next six rows illustrate stimulus manipulations intended to lead to breakdown of the listener’s acoustic model of room A. Room B presentations are single-source clicks from the left loudspeaker only. The source has not moved relative to its location in room A, but the reflection and, thus, the simulated reflective surface have been eliminated. Room B was presented for 1, 3, 5, 7, 9, or 11 clicks. Thus, these conditions consisted of room A5, room B_n, then room A1 (the test click) at the end of the entire train. The next two rows show the configurations of conditions that control the time lapse in room presentations before the test click. After the initial presentations of room A were completed, the buildup effect of room A might be expected simply to dissipate over time. In order to

understand the influence of the room B presentations, it was necessary to determine what the effect would be of presenting no sound during an equivalent period. Two conditions were included, silence for the periods of time equivalent to 5 and 11 clicks (S5 and S11, respectively). Adding the standard pause of 750 ms between train and test, the total silence was 2 s for S5 and 3.5 s for S11. The next four conditions shown in the table were also control conditions. Only the single-source clicks (room B) were presented before the test click. Finally the last condition included no preceding train to determine the response to the test click presented as a solitary sound [no conditioning train (NC)].

The different conditions were run in blocks of 35 trials. The room condition was fixed (e.g., A5|B7|A1) during a block. The seven lag-click delays (2, 4, 6, 8, 10, 12, and 14 ms) were presented five times each within a block in a randomized order. An individual experimental session lasted for approximately 1 h and consisted of a total of 13 blocks, one for each condition. Note that the B11|A1 condition was run later when it was decided that it was necessary for comparison to A5|B11|A1. It was interspersed among the other conditions for experiment 2. The order of blocks was randomized separately for each listener and experimental session. Each listener completed four such sessions so that, across all sessions, a single condition at a specific delay was based on 20 judgments (five repetitions per block × four blocks per condition).

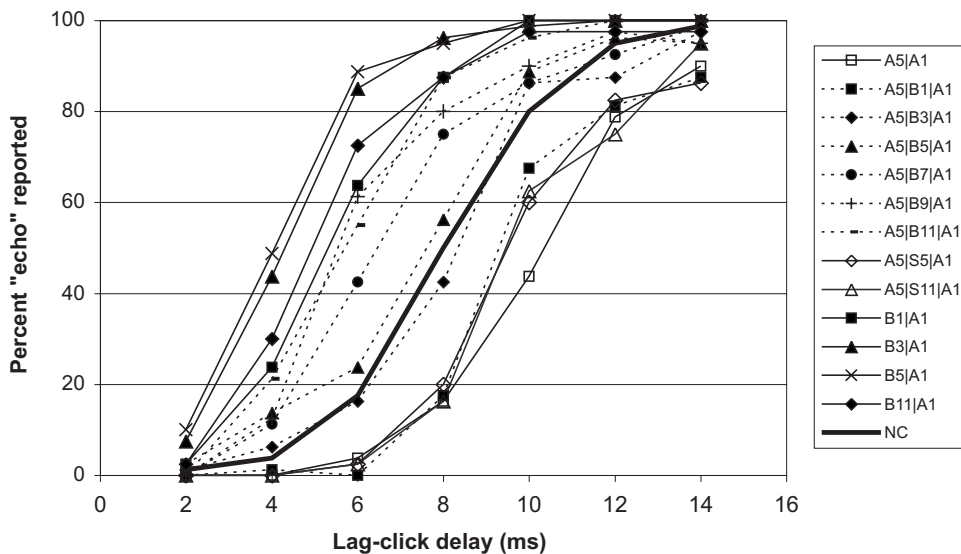


FIG. 2. Mean percentage of trials on which an echo was reported as a function of the delay of the lag click for all conditions of experiment 1. A description of each of the conditions is given in Table I.

3. Listeners

Four graduate students from the University of Massachusetts, all with hearing thresholds ≤ 20 dB HL (ANSI, 1996) from 500 to 4000 Hz, participated in the study.

B. Results

1. Psychometric functions

To get an overall view of the data, the results for each condition and delay were averaged over subjects to form the group psychometric functions shown in Fig. 2. The percentage of trials on which an echo was reported is plotted as a function of the lag-click delay. As expected, the functions were generally monotonic, showing increasing reporting of echoes as delay was increased; however, these functions also display a large effect of the acoustic stimulation that preceded the test click. It should be noted again that exactly the same test click, i.e., the lead-lag click pair of room A, was the stimulus that subjects responded to in every condition. The thick solid line indicates the function obtained for the NC condition, where this test click was presented in isolation. NC crossed the 50% point around 8 ms. At the delay of 8 ms the percentage of echoes reported ranged from 17% for the basic buildup condition (A5|A1) to 95% for the breakdown condition B5|A1, indicating the dramatic influence of context on the listeners' reporting of echoes at the same lead-lag delay. The intermediate values reflect variable amounts of input from the different simulations of acoustic space. Consistent with previous reports (Djelani and Blauert, 2001; Freyman and Keen, 2006), a single instance of room B had little or no effect on buildup, but each additional presentation of room B degraded echo threshold for room A stimuli.

2. Echo thresholds

To quantify the effects observed in Fig. 2, echo thresholds (the delay at which echoes were reported on 50% of the trials) were estimated from the psychometric functions obtained from each individual subject. Each function was fitted with a logistic equation of the form $1/(1+\exp(-(t-m)/s))$, where t is the lag-click delay, m is midpoint of the function,

and s is the slope. The parameter m represents an estimate of the delay at which 50% echoes were reported, the echo threshold in this case. The fits were generally very good, 85% of the r^2 values being above 0.95. As might be expected, the slopes of the individual functions tended to be slightly steeper than those of the mean functions shown in Fig. 2. No formal comparison of the slopes across conditions was undertaken. However a slight tendency for slopes to be steeper for the functions with lower echo thresholds was apparent in both the mean (e.g., compare B5 and A5 in Fig. 2) and individually fitted functions.

The echo thresholds are displayed in Fig. 3, which plots the group mean thresholds for all conditions shown in Table I. Higher echo thresholds indicate more buildup, that is, more suppression of echoes, and conversely low thresholds indicate that more echoes are heard at shorter delays. The abscissa shows the number of clicks in room B that preceded test click A1, with zero along the axis referring to the basic buildup condition with no room B clicks following buildup

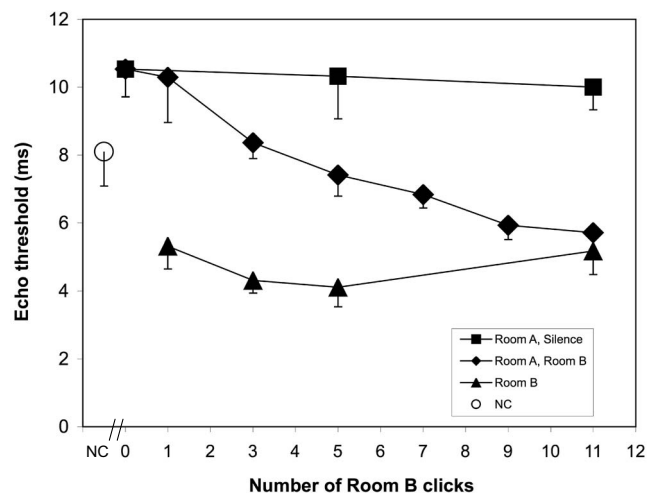


FIG. 3. Mean echo thresholds from experiment 1 as a function of the number of room B clicks presented alone (triangles) or following five room A clicks (diamonds). Squares show results of a control condition using five room A clicks followed by silence equivalent in duration to 0, 5, and 11 room B clicks. Error bars indicate one standard error of the mean.

(A5|A1). Echo threshold was 10.5 ms for buildup in room A with no exposure to room B. The continuation of this line represents the silent control condition (A5|Sn|A1, squares) and shows that, with up to 3.5 s of silence, there was little or no change in built-up threshold. The A5|Bn|A1 condition (diamonds) shows that exposure to room B resulted in a gradual effect, with each additional click reducing the previously built-up threshold. A two-way analysis of variance (ANOVA) comparing the room A conditions (A5, A5|S5, and A5|S11) with the comparable room B conditions (A5|B1, A5|B5, and A5|B11) yielded a main effect of condition [$F(1,3)=13.21$, $p<0.036$]. There was also a main effect of amount of time since the A5 buildup [$F(2,6)=5.67$, $p<0.042$], indicating a decrease in echo threshold over time. The effect of increasing the room B input was analyzed with a one-way ANOVA and trend test on the six levels of room B (A5|Bn). A main effect of exposure level was obtained [$F(1,15)=5.94$, $p<0.003$], with a linear trend that just missed significance [$F(1,3)=8.70$, $p<0.06$].

Another way to assess the effect of the buildup of the model of room A is to compare the effect of exposure to room B with and without prior exposure to room A (A5|Bn versus Bn). The lower thresholds for the latter condition, shown in Fig. 2, indicate that the initial presentation of the five buildup clicks (A5) still had a strong influence on the reporting of echoes to the room A test click when that room was re-introduced after several presentations of room B. A two-way ANOVA on condition (A5|Bn versus Bn) and amount of room B exposure (B1, B5, and B11 for the two conditions) yielded a main effect of condition [$F(1,3)=22.83$, $p<0.017$], confirming the lingering effect of the prior room A exposure. There was also a main effect of amount of room B exposure, with echo threshold for A decreasing significantly as room B presentations progressed [$F(2,6)=5.11$, $p<0.05$].

Reductions in echo threshold began to level off by 9–11 presentations of room B. At this point, the A5|B11|A1 threshold was within 1 ms of the B11|A1 threshold, suggesting that this may be the point where the model of room A was virtually lost. One might assume that during the increasing input of room B the difference between echo thresholds for Bn|A1 and A5|Bn|A1 was due to the influence of prior room A input. This assumption could be wrong because the experience of any other room before the introduction of room B may have had the same elevating effect on echo threshold. To ensure that the response to A1 was affected specifically by prior exposure to room A, a different space (room C) with new locations for lead and lag clicks should replace room A as the initial exposure before room B. In experiment 2, room B was presented in between room C and the test click (A1) to learn how the breakdown created by the single-source clicks progressed under this circumstance. Echo threshold for the test click may be affected by prior exposure to room C because, like room A, room C has a delayed sound, even though it comes from a different direction.

A second way that the room A experience might exert an effect after apparently being replaced by room B is to produce faster buildup during a re-buildup period. The question

is whether there would be any savings if room A were re-introduced following 11 repetitions of room B. This gets to the heart of the question of whether we hold on to and store the model of a room so that it can be readily re-activated. In other words, would buildup occur more quickly in room A if the listener had previously experienced buildup in that same space?

III. EXPERIMENT 2

Experiment 2 investigated how the experience in room A continued to exert an effect when followed by presentations of room B. Schematic diagrams of exemplars of all stimulus conditions are illustrated in Fig. 4. In one test we simulated a new room, room C, in which lead clicks were presented from the right and lag clicks from the left. This simulated a reflecting surface on the left side of the room, the opposite of room A. Unlike the presentation of five room A clicks prior to the test click (A5|A1), which built up echo threshold in experiment 1, condition C5|A1 should not raise echo threshold for the test click, A1, above that for NC, the test click presented in isolation. Substituting room C for room A before room B input (CB) provides a control condition for AB from experiment 1 to determine whether the elevated thresholds for the test click were due specifically to the room A buildup.

The second test for a lingering effect of room A was re-exposure to this stimulus after room B experience had apparently obliterated its influence. The model for room A appears to be completely gone after 11 presentations of room B (A5|B11 in experiment 1). We tested whether buildup would take place more quickly the second time by comparing increasing levels of re-exposure to room A after prior buildup/breakdown (ABA) versus exposure to two different rooms (CBA) and prior exposure to room B alone (BA). The latter condition provides a baseline for the effect of no prior exposure to either room A or C. It is possible that prior buildup to room A will have a priming effect such that re-buildup to room A occurs more quickly compared to prior exposure to either room C or room B. This result would suggest that even though echo threshold was very low after B11, some effect of room A in memory served to boost a second buildup quickly. On the other hand if re-exposure to room A follows a similar trajectory regardless of what came before B, this would suggest that the “slate is wiped clean” when there has been sufficient exposure to a new acoustic environment.

A. Methods

The experiment was run with the same four listeners who participated in experiment 1. All equipment and procedures were identical to those used in experiment 1, and echo thresholds were calculated in the same manner. Table II lists all conditions tested in this experiment, and Fig. 4 displays example stimuli from the different conditions. The first four rows of the table show control conditions involving room C. Each condition began with five presentations of room C (the reverse of room A). Results for condition C5|A1 were compared to those for condition A5|A1 (from experiment 1), and

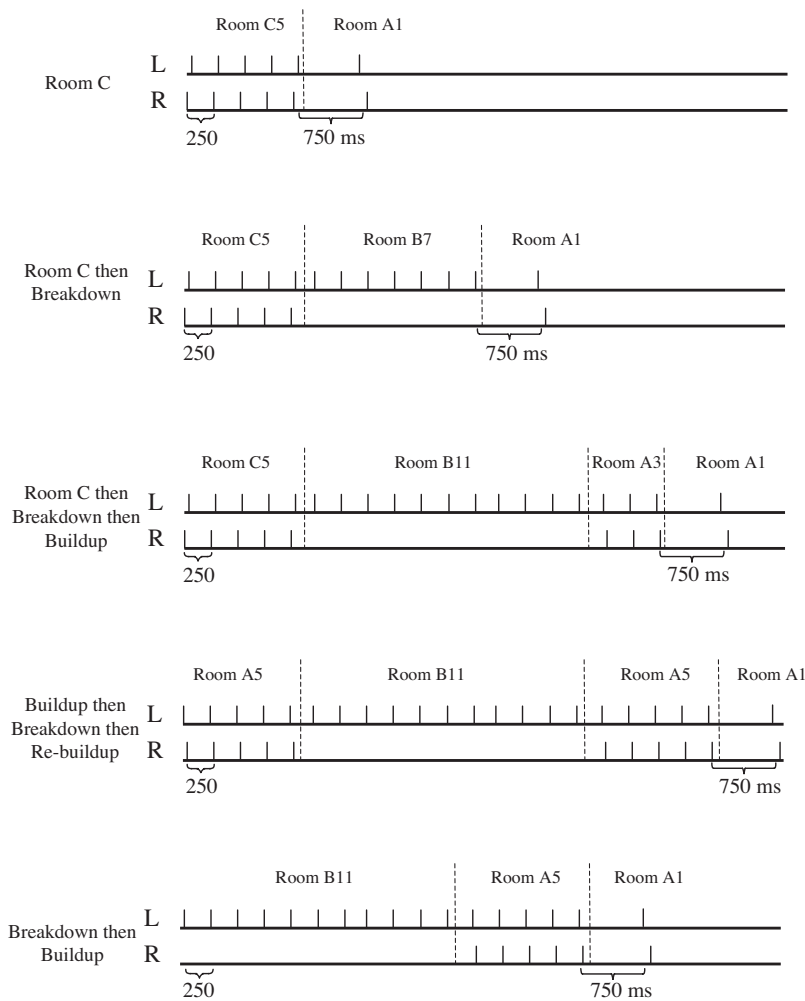


FIG. 4. Schematic illustrations of examples of stimuli used in experiment 2.

three levels of room B exposure (3, 7, and 11) were interspersed between room C input and A1, the test click, for comparison to A5|Bn|A1 from experiment 1. The conditions that tested re-buildup of echo threshold after 11 presentations of room B are shown in the remainder of Table II. The basic condition was five presentations of room A then B11, followed by varying amounts of room A before the test click (A5|B11|An|A1). Values of An were 3, 5, 7, 9, and 11 click pairs. Comparison conditions were C5|B11|An|A1 and B11|An|A1.

Each condition was tested with the same seven lag-click delays (2, 4, 6, 8, 10, 12, and 14 ms) as in experiment 1. All conditions were run in 35-trial blocks as before, with five repetitions of each delay presented in a randomized order within a block. For all conditions there were four blocks per condition across sessions, for a total of 20 trials for each delay and condition combination. The conditions were presented in six experimental sessions. The first four sessions consisted of the conditioning trains A5|B11|An and B11|An, a total of ten conditions. All ten blocks, one for each condition, were presented in a randomized order in each session. The last two sessions consisted of all the conditions involving room C, seven in total. Each of the seven conditions was presented twice in a randomized order during each 14-block session.

B. Results

As expected, hearing five presentations of room C before the test click did not increase echo threshold for room A. As shown in Fig. 5, left panel, the solid circle (mean = 8.0 ms for C5|A1) is virtually the same value as the open circle for the isolated test click (mean = 8.1 ms for NC). The left panel also displays the effect of the breakdown produced by single-source clicks of room B, and the right panel displays the buildup that occurred with repeated presentations of room A following the breakdown. With the exception of the circles, the data in the left panel are replotted from Fig. 3. We first tested whether the response to the test click differed for conditions A and C. A 2 (condition) \times 4 (exposure level to room B) ANOVA tested the decline in echo threshold as room B input increased from zero (A5|A1 and C5|A1) to 11 (A5|B11|A1 and C5|B11|A1). The effect of the number of room B clicks before the test click was significant [$F(3,9) = 20.16$, $p < 0.0001$], as was the linear trend [$F(1,3) = 22.69$, $p < 0.018$]. The curves for room C and room A began at different levels, but once input from room B began, they both progressed downward until the means were similar at B11 (mean = 5.7 ms for A5|B11|A1 and mean = 5.8 ms for C5|B11|A1). There appears to be a general (but rapidly disappearing) effect of having heard five presentations of a delayed sound that elevates echo threshold above those for

TABLE II. List of conditions for experiment 2.

Segment 1	Segment 2	Segment 3	Segment 4	Name/description
Train	Train	Train Room C5	Test Room A1	Room C C5 A ₁
Train	Train Room C5 Room C5 Room C5	Train Room B3 Room B7 Room B11	Test Room A1 Room A1 Room A1	Room C, breakdown C5 B3 A1 C5 B7 A1 C5 B11 A1
Train Room A5 Room A5 Room A5 Room A5 Room A5	Train Room B11 Room B11 Room B11 Room B11 Room B11	Train Room A3 Room A5 Room A7 Room A9 Room A11	Test Room A1 Room A1 Room A1 Room A1 Room A1	Buildup, breakdown, re-buildup A5 B11 A3 A1 A5 B11 A5 A1 A5 B11 A7 A1 A5 B11 A9 A1 A5 B11 A11 A1
Train Room C5 Room C5 Room C5	Train Room B11 Room B11 Room B11	Train Room A3 Room A7 Room A11	Test Room A1 Room A1 Room A1	Room C, breakdown, buildup C5 B11 A3 A1 C5 B11 A7 A1 C5 B11 A11 A1
Train	Train Room B11 Room B11 Room B11 Room B11 Room B11	Train Room A3 Room A5 Room A7 Room A9 Room A11	Test Room A1 Room A1 Room A1 Room A1 Room A1	Breakdown, buildup B11 A3 A1 B11 A5 A1 B11 A7 A1 B11 A9 A1 B11 A11 A1

the anechoic room B condition. It does not appear to matter which direction the delayed sound came from, as room A versus C had no effect after 11 presentations of room B clicks.

The effect of re-exposure to room A following 11 presentations of B can be stated simply: it does not appear to matter whether A or C came before, at least with the statistical power available with a small N . Although the B11|An curve appears to be lower than the other curves, an analysis testing whether the three curves in the right panel of Fig. 5 were different showed that they were not. Thresholds for condition A5|B11|An did not differ from those for

C5|B11|An [$F(1,3) < 1.0$, N.S.] or from B11|An [$F(1,3) = 4.06$, $p > 0.10$]. Level of exposure (or re-exposure) to room A was, of course, highly significant for both follow-up ANOVAs [$F(3,9) = 9.14$, $p < 0.004$ for the ABA and CBA comparison and $F(4,12) = 11.55$, $p < 0.0001$ for the ABA and BA comparison].

IV. DISCUSSION

In the course of a normal day, people will experience not two or three acoustic spaces, but dozens, as they move about their environment. In two experiments we have shown that

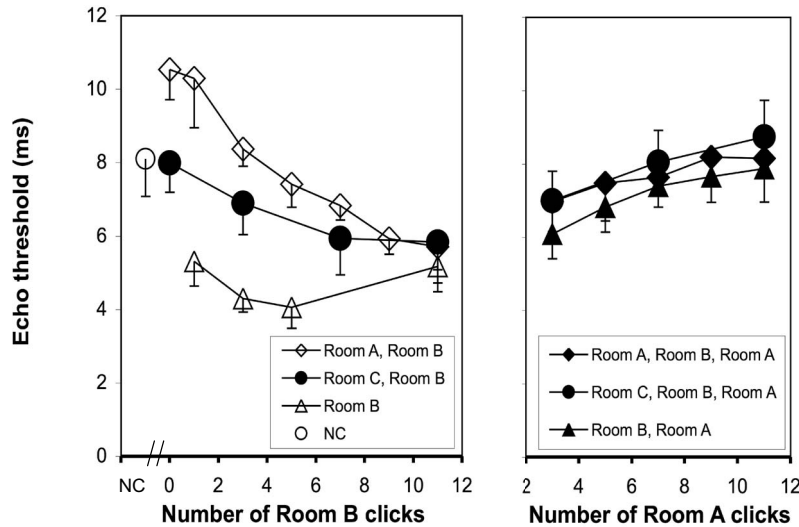


FIG. 5. Mean echo thresholds as a function of the number of clicks presented with (diamonds, circles) or without (triangles) an initial train of five clicks. Error bars indicate one standard error of the mean. Data in the left panel show the effect of breakdown with increasing number of room B clicks and are, with the exception of the filled symbols, replotted from experiment 1. The right panel shows data from experiment 2 indicating buildup/re-buildup from the additional room A clicks that followed the room B clicks. See Table II for a description of all conditions.

people form models of an acoustic space when exposed to sound and its attendant delayed reflections in that space. Furthermore, the model for a particular space is disrupted when the listener is exposed to a new space having different acoustic properties. The processes of building up and breaking down models are dependent on the amount of input from each spatial configuration of sound. In experiment 1 we charted the course of disruption of a room model once built up. Following five presentations of sounds in room A (which was sufficient to build up echo threshold to about 10.5 ms), we inserted varying amounts of exposure to room B (1, 3, 5, 7, 9, and 11 inputs). Echo threshold for room A steadily decreased so that after 11 presentations of room B, threshold was depressed to about 5 ms, which was comparable to one exposure to room A preceded only by presentations of room B. It was as if room A had never been experienced. This process can be thought of as a competition between two room models, with the more recently experienced room B model ultimately winning out.

In experiment 2 we tested whether there were savings in the buildup process by re-exposing listeners to room A for a second time. This was done after 11 presentations of room B had driven echo threshold for room A down to a low level. We found that buildup to room A progressed similarly under three conditions: prior exposure to room A (ABA), prior exposure to a space with different echoes (CBA), or only anechoic exposure (BA). In other words, buildup was independent of whether the listener had previously experienced that room.

Although the rooms we simulated are unusual and would not be encountered very often outside the laboratory, their very simplicity allows us to draw certain conclusions about how the auditory system handles exposure to the varied spaces we do inhabit. First, just as there is buildup in echo threshold when one hears sound produced within a space that has reflections, there is breakdown of that threshold when the listener enters a new space. The breakdown does not occur immediately upon entry into the new space (Djelani and Blauert, 2001; Freyman and Keen, 2006) but depends on the amount of input. Buildup and breakdown are incremental processes, with each instance of new input having a cumulative effect. The processes can take place rapidly and reach asymptote quickly because only 9–11 inputs are required to reach maximum echo threshold for buildup (Freyman *et al.*, 1991) or to reach lowest echo threshold for breakdown, as in the current data set. Our experience in real rooms can seem like an immediate adjustment to echoes because most sounds (speech, music, and noise) have numerous onsets occurring rapidly. It is only when the listener is exposed to punctate, countable sounds like brief clicks that buildup and breakdown in echo threshold become noticeable and can be measured. We can slow the process down by presenting click pairs slowly; at rates of 1/s maximum echo threshold is reached after about 9 or 10 s, whereas fast rates of 8/s or 16/s produce seemingly instantaneous buildup (Freyman *et al.*, 1991). This rapid buildup/breakdown is at the core of how the nervous system forms and discards models of acoustic space. Our data indicate that models are formed rapidly from brief exposure and are discarded with

equal rapidity. There is little cost to rapid abandonment of a model because it is quickly reformed on subsequent exposure. Storing old models seems useless and inefficient and would only be valuable if acquiring new models were slow or difficult. It is the ultimate “throwaway” economy, perhaps a necessary one in light of the numerous acoustic spaces we experience every day.

And yet, the possibility of memory for past acoustic spaces is real and plausible. The usefulness of long-term memory for a familiar acoustic space is obvious for informing a listener about whether that space had changed, perhaps in some dangerous way. Although we know of no research that has tested listeners’ sensitivity to changes in a familiar room, discrimination of different rooms’ reverberant properties has been investigated. Robart and Rosenblum (2005) tested whether listeners could discriminate among four different rooms whose reflective properties varied widely. The same sounds were recorded in a gymnasium, a classroom, a rest room, and a small laboratory. Untrained listeners looked at photographs of the four rooms while listening to sound recorded in one of the rooms. Subjects were able to select the correct room on 78 of 100 trials with no feedback. Several studies have demonstrated humans’ ability when blindfolded to detect objects by means of reflected sound (for reviews see Rice, 1967; Stoffregen and Pittenger, 1995). This prior work, as well as research on binaural room simulation as reviewed in Blauert (1997), supports our contention that listeners are highly sensitive to the structure of sound produced by numerous reflective surfaces in a room, and experience in different environments may generate memory for common acoustic spaces.

The idea that localization of sound in a new space is influenced by comparison of the current acoustic cues with the listener’s acquired knowledge of spatial cues was proposed by Plenge (1974). In discussing how dummy-head recordings presented over earphones were first heard intracranially but after further experience were heard extracranially, Plenge (1974) hypothesized that subjects used cues from prior experience to localize sound. He suggested that knowledge of how sound behaves in spaces is acquired over a lifetime of experience and that we carry “stored stimulus patterns” for comparison to new stimuli. It appears that Plenge (1974) proposed a more generic form of stored knowledge than retention of specific room parameters; as he stated: “Only a short-term storage of such knowledge of sound sources and room conditions is useful” (p. 951). He more clearly described the process as follows: “The moment the hearer leaves the room, the stored information concerning its peculiarities is cleared, and information concerning the new situation is stored” (p. 951). A better description of the essence of our current data would be hard to find, although Plenge (1974) wrote this more than 3 decades ago.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute on Deafness and Other Communicative Disorders (Grant No. DC01625). The authors would like to thank Jamie Chevalier, Ackland Jones, Leah Novotny,

Amanda LePine, and April Teehan for their assistance with the project. The authors gratefully acknowledge the helpful comments of D. Wesley Grantham, an anonymous reviewer, and Associate Editor Brian Moore.

ANSI. (1996). ANSI S3.6-1996, "Specifications for audiometers," American National Standards Institute, New York.

Bech, S. (1995). "Timbral aspects of reproduced sound in small rooms. I," *J. Acoust. Soc. Am.* **97**, 1717–1726.

Bech, S. (1996). "Timbral aspects of reproduced sound in small rooms. II," *J. Acoust. Soc. Am.* **99**, 3539–3549.

Bech, S. (1998). "Spatial aspects of reproduced sound in small rooms," *J. Acoust. Soc. Am.* **103**, 434–445.

Benade, A. (1976). *Fundamentals of Musical Acoustics* (Oxford University Press, London).

Blauert, J. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge, MA).

Blauert, J., and Col, J. P. (1992). "A study of temporal effects in spatial hearing," in *Auditory Physiology and Perception*, edited by Y. Cazal, L. Demany, and K. Korner (Pergamon, Oxford), pp. 531–538.

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., and Freyman, R. L. (1997). "The precedence effect: Beyond echo suppression," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. B. Anderson (Lawrence Erlbaum, Hillsdale, NJ), pp. 233–255.

Clifton, R. K., Freyman, R. L., Litovsky, R. Y., and McCall, D. (1994). "Listener expectations about echoes can raise or lower echo threshold," *J. Acoust. Soc. Am.* **95**, 1525–1533.

Clifton, R. K., Freyman, R. L., and Meo, J. (2002). "What echoes tell us about the auditory environment," *Percept. Psychophys.* **64**, 180–188.

Djelani, T., and Blauert, J. (2000). "Some new aspects of the buildup and breakdown of the precedence effect," in *Physiological and Psychophysical Bases of Auditory Function*, edited by D. J. Breebart, A. J. Houtsma, A. Kohlrausch, V. F. Prijs, and R. Schoonhoven (Shaker, Maastricht, The

Netherlands), pp. 200–207.

Djelani, T., and Blauert, J. (2001). "Investigations into the build-up and breakdown of the precedence effect," *Acta Acust. Acust.* **87**, 253–261.

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

Freyman, R. L., and Keen, R. (2006). "Constructing and disrupting listeners' models of auditory space," *J. Acoust. Soc. Am.* **120**, 3957–3965.

Grantham, D. W. (1996). "Left-right asymmetry in the buildup of echo suppression in normal-hearing adults," *J. Acoust. Soc. Am.* **99**, 1118–1122.

Hartmann, W. M., and Rakerd, B. (1989). "Localization of sound in rooms IV: The Franssen effect," *J. Acoust. Soc. Am.* **86**, 1366–1373.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (1999). "The precedence effect," *J. Acoust. Soc. Am.* **106**, 1633–1654.

McCall, D. D., Freyman, R. L., and Clifton, R. K. (1998). "Sudden changes in room acoustics influence the precedence effect," *Percept. Psychophys.* **60**, 593–601.

Plenge, G. (1974). "On the differences between localization and lateralization," *J. Acoust. Soc. Am.* **56**, 944–951.

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.

Rice, C. E. (1967). "Human echo perception," *Science* **155**, 656–664.

Robart, R. L., and Rosenblum, L. D. (2005). "Hearing space: Identifying rooms by reflected sound," in *Studies in Perception and Action XIII*, edited by H. Heft and K. L. Marsh (Lawrence Erlbaum Associates, Inc., Hillsdale, NJ).

Sanders, L. D., Joh, A. S., Keen, R. E., and Freyman, R. L. (2008). "One sound or two? Object-related negativity indexes echo perception," *Percept. Psychophys.* **70**, 1558–1570.

Stoffregen, T. A., and Pittenger, J. B. (1995). "Human echolocation as a basic form of perception and action," *Ecological Psychol.* **7**, 181–216.

Yang, X., and Grantham, D. W. (1997). "Echo suppression and discrimination suppression aspects of the precedence effect," *Percept. Psychophys.* **59**, 1108–1117.

Yost, W. A., and Guzman, S. (1996). "Auditory processing of sound sources: Is there an echo in here?" *Curr. Dir. Psychol. Sci.* **5**, 125–131.