

Detection of acoustic repetition for very long stochastic patterns

RICHARD M. WARREN, JAMES A. BASHFORD, JR.,
JEFFREY M. COOLEY, and BRADLEY S. BRUBAKER
University of Wisconsin, Milwaukee, Wisconsin

Guttman and Julesz (1963) employed recycling frozen noise segments (RFNs) as model stimuli in their classic study of the lower limits for periodicity detection and short-term auditory memory. They reported that listeners can hear iteration of these stochastic signals effortlessly as “motorboating” for repetition periods ranging from 50 to 250 msec and as “whooshing” from 250 msec to 1 sec. Both motorboating and whooshing RFNs are global percepts encompassing the entire period, as are RFNs in the pitch range (repetition periods shorter than 50 msec). However, with continued listening to whooshing (but not motorboating) RFNs, individuals hear recurrent brief components such as clanks and thumps that are characteristic of the particular waveform. Experiment 1 of the present study describes a cross-modal cuing procedure that enables listeners to store and then recognize the recurrence of portions of frozen noise waveforms that are repeated after intervals of 10 sec or more. Experiment 2 compares the relative salencies of different spectral regions in enabling listeners to detect repetition of these long-period patterns. Special difficulty was encountered with the 6-kHz band of RFNs, possibly due to the lack of fine-structure phase locking at this frequency range. In addition, a similarity is noted between the organizational principles operating over particular durational ranges of stochastic patterns and the characteristics of traditional hierarchical units of speech having corresponding durations.

In their classic study dealing with the temporal limit for detection of acoustic periodicity, Guttman and Julesz (1963) used recycling frozen noise segments (RFNs) as model stimuli. They found that repetition of these stochastic patterns could be detected with ease for waveforms having durations as long as 1 sec (and with difficulty for durations up to 2 sec). For RFN periods from 1 sec to 250 msec, Guttman and Julesz described the detection of repetition as the effortless detection of a smooth “whooshing,” and for RFNs from 250 msec to the transition to pitch occurring at 50 msec (20 Hz), repetition was described as a staccato “motorboating.” Both whooshing and motorboating are global percepts encompassing the entire waveform, with different RFNs having the same period being distinguished by their individual timbres. However, with continued listening to whooshing, but not motorboating RFNs, features such as clanks and thumps characteristic of the particular waveform emerge.

The findings of Guttman and Julesz (1963) concerning the effortless detection of RFN repetition in the whooshing and motorboating ranges have been replicated and extended in other laboratories (e.g., Kaernbach, 1992; Limbert & Patterson, 1982; Pollack, 1975; Warren & Bash-

ford, 1981). Although it has been suggested that RFNs can provide a measure of the durational limits of auditory memory relatively uncontaminated by verbal encoding or identification of specific signals (Guttman & Julesz, 1963; Pollack, 1972), little work has been reported on the ability to detect repetition for RFNs with periods greater than 1 or 2 sec. There have been conflicting informal reports concerning the ability to detect long-period RFNs. The 2-sec limit for detecting RFN repetition described by Guttman and Julesz was increased by Julesz and Hirsh (1972), who noted in passing that repetition could be perceived “with great effort” up to 4 sec. This upper limit was also cited by Limbert (1984, p. 3) in his doctoral dissertation (also without giving direct experimental evidence). He stated that “with samples about 4000 msec long, the percept is continuous noise only; features never emerge to indicate the repetition rate.” However, Warren and Bashford (1981, note 2) had observed informally that 1 highly trained individual was able to tap to the repetition of a prominent feature in RFNs with periods of 10 sec.

The present formal experiments were undertaken to help resolve this conflict in informal reports and to determine the ability of listeners to tap to the recurrence of portions of frozen noise waveforms repeated after intervals of 10 sec or more under a number of different conditions. Experiment 1 employed laboratory personnel as subjects. These individuals had considerable experience listening to motorboating and whooshing range RFNs (repetition periods from 50 msec to 1 sec), and they were given

This study was supported by a grant from the National Institute of Deafness and Other Communication Disorders (DC 00208) to R.M.W. Correspondence should be addressed to R. M. Warren, Department of Psychology, University of Wisconsin, P. O. Box 413, Milwaukee, WI 53201 (e-mail: rmwarren@csd.uwm.edu).

prior cuing to help them identify segments of the repeating waveform. Experiment 2 employed college students who had never heard repeated frozen noise prior to their single experimental session. They were presented with bandpassed long-period RFNs with different center frequencies, one group attempting to detect repetition with the aid of prior cuing, and the other group attempting the same task without cuing.

EXPERIMENT 1

A preliminary report by Bashford, Brubaker, and Warren (1993) claimed that recognition of repetition for a long-duration RFN could be facilitated (at least for experienced laboratory personnel) by a prior pairing with a repeating light flash that recurred in synchrony with one portion of the acoustic waveform. Experiment 1A of the present study tested the ability of listeners to tap to the recurrence of a portion of a long-period RFN that had previously been paired with synchronous light flashes. The flash cuing of Experiment 1A not only highlighted a portion of the periodic waveform, but also identified the fixed interval separating the occurrences of that RFN segment. Experiment 1B was undertaken to eliminate the possibility of anticipating the time of recurrence of an identifiable segment: Flash cuing was also used, but successive presentations of frozen noise were interleaved with non-recurrent, variable duration segments of on-line noise.

Experiment 1A

Method

Subjects. Two of the authors of this paper served as subjects. Both were experienced laboratory personnel with normal hearing for their ages (30s for B.B. and 40s for J.B.).

Stimuli and Procedure. The frozen noise signals were derived from white noise that was bandpass filtered from 100 Hz to 10 kHz (48 dB/octave slopes) and recorded digitally with 22-kHz sampling and 16-bit quantization. Fifteen frozen noise samples (five samples each at segment durations of 10, 15, and 20 sec) were presented in recycling form, with separate sessions lasting about 15 min each employed for each sample.

Testing was conducted in a sound-attenuating chamber with stimuli delivered diotically through headphones at a level of 75 dBA. During the training phase of a session, the listeners monitored both the RFN and a light panel that provided a 200-msec flash that they

knew was synchronized with a randomly selected fixed portion of the frozen noise waveform. This cross-modal cuing regime is illustrated in Figure 1. The listeners took as much time as needed to clearly identify a repetitive auditory feature and then tapped to three recurrences of that feature. The RFN and the light flash were then turned off for 5 sec. Following this pause, the same RFN was presented for testing, but without the light flash. Listeners were told that the second presentation of the RFN would begin at a random point in the periodic waveform, so that the rhythm established during the learning phase would be disrupted. Listeners had the option of aborting a series of taps once started, but this option was never used. The test session ended when they had tapped nine times to consecutive occurrences of the selected feature.

Results and Discussion

For each of the three RFN periods, light flashes were effective in cuing the detection of auditory features occurring near the time of the light flash. Table 1 shows the average number of repetitions required to initiate tapping during the learning phase and during the subsequent test phase for 10-, 15-, and 20-sec RFNs. Averaged across RFN periods, about 13 repetitions occurred before initiation of tapping during the training phase, and only about two repetitions were required for the resumption of tapping during the test phase of trials.

During the test phase listeners were not only quick to respond but also were highly accurate in their tapping to these long-duration RFNs. Table 2 presents the mean period and mean position of taps (relative to the location of the flash presented during the prior training phase) averaged across listeners and noise samples. The largest deviation of the listeners' average tap period from the true RFN period was only 0.4% (80 msec) at the longest duration (20 sec). The average positioning of taps indicates that the features recognized and selected as targets for tapping by our listeners most frequently occurred soon after the point in the RFN waveform that had been concurrent with the light flash. On average, listeners' taps were within 1 sec of that point.

The ability of listeners to tap accurately to RFNs having repetition periods up to 20 sec does not conflict with the 1-sec limit described by Guttman and Julesz (1963) for the recognition of RFN repetition as a whooshing sound encompassing the entire period. Our listeners reported tapping to a particular brief event occurring in the

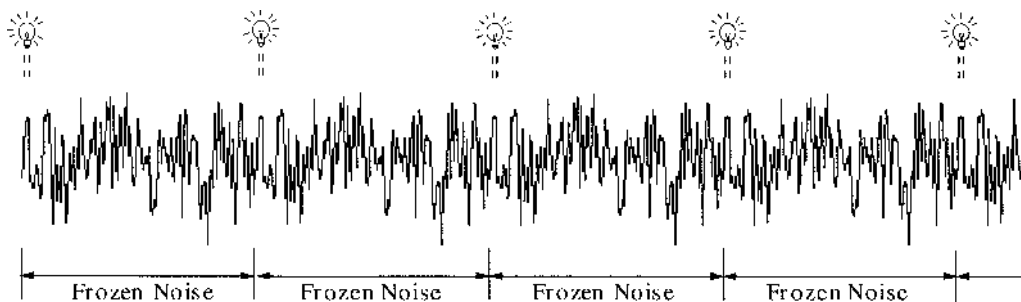


Figure 1. Diagrammatic representation of stimuli used in Experiment 1A. Repeating frozen noise segments were 10, 15, or 20 sec in duration and their repetitions were cued by a 200-msec light flash occurring at a fixed randomly selected waveform location.

Table 1
Mean Number (and Standard Errors) of Frozen Noise Repetitions Required for the Initial Tapping Response in the Learning Phase (Light Cue Present) and for the Subsequent Test Phase (Light Cue Absent) in Experiment 1A

RFN Period	Learning Phase		Test Phase	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
10 sec				
B.B.	12.33	0.95	0.71	0.26
J.B.	9.61	1.34	1.29	0.54
15 sec				
B.B.	18.61	4.18	0.72	0.42
J.B.	10.95	1.62	4.58	2.00
20 sec				
B.B.	16.95	4.07	1.09	0.40
J.B.	9.45	2.23	2.63	0.66

Note—Results for each of the 2 listeners are averaged over the five independent samples of recycling frozen noise (RFN) presented at each of three repetition periods (i.e., noise segment durations).

long-duration RFN, with the remainder of the pattern sounding like on-line noise. This is consistent with the hypothesis that the ability to hear repetition of these longer RFNs does not depend upon the signal's periodicity. It was reasoned that with the aid of cuing, listeners might be able to detect repetition when segments of frozen noise recur aperiodically, alternating with variable duration segments of nonrecurrent noise. This hypothesis was tested directly in Experiment 1B using a new set of frozen noise segments.

Experiment 1B

Method

Subjects. The same 2 listeners employed in Experiment 1A served as subjects.

Stimuli and Procedure. During each of the five 20-min sessions in Experiment 1B, the listener was presented with a different 3-sec segment of a frozen noise that recurred 60 times in an aperiodic fashion, with successive presentations of the fixed-waveform segment separated by segments of nonrecurrent "on-line" noise that varied randomly in 1-sec steps between 7 and 13 sec. These unpredictable repetition intervals ensured that individual intertap intervals were based upon the occurrence of the acoustic pattern rather than estimation of a fixed repetition period. As in Experiment 1A, both listeners received the same segments of frozen noise. Figure 2 illustrates the nature of the aperiodic target as well as the cuing regime in this experiment, which paired a light flash with the onset of the 3-sec frozen noise.

Listeners were instructed to monitor both the noise and the light flash until they recognized some recurrent auditory feature, and to tap whenever that aperiodic feature recurred. After the listener's third tap to the auditory feature, the light cue was turned off, and the listener continued to tap until the stimulus ended. The instructions emphasized the importance of precision in tapping to the same portion of the recurring auditory feature, and listeners were allowed to cancel a preceding tap by rapidly pressing their response key three times and then resuming the tapping to successive appearances of the target feature. A computer recorded the position of taps relative to the aperiodic onset of the frozen noise segment.

Results and Discussion

The results are presented in Table 3 and Figure 3. After the light cue was turned off, an average of eight presen-

tations of the target segments occurred before listeners commenced tapping. The option of canceling a tap was exercised only once by B.B. and twice by J.B. In addition, some occurrences of the target were not tapped to. Out of the total of 300 presentations of the target to each listener in the five sessions (this total includes those heard prior to tapping), 195 tap judgments were made by B.B. and 177 were made by J.B. Turning to the individual sessions, in which there were 60 presentations, the number of listener-accepted tap responses varied between 20 and 48. Hence, the mean tap position and standard errors given in Table 3, as well as the tap-placement histograms of Figure 3, are based upon the first 20 tap responses in each session. Out of the total of the 200 taps represented in Table 3, none were outside the 3-sec frozen noise segment.

Taken together, Experiments 1A and 1B demonstrate that, at least for highly trained laboratory personnel aided by cross-modal cuing, the durational limit for recognition of repeated frozen noise fragments extends well beyond the 1-sec limit for effortless global recognition of a featureless whooshing. It is also well beyond the 4-sec limit stated by Limbert (1984) for detection of repetitive features in RFNs. However, the present results are consistent with estimates in the literature for the upper limit for auditory storage obtained from paradigms using other types of nonverbal as well as verbal stimuli (for a review, see Cowan, 1984, 1987).

EXPERIMENT 2

After completion of Experiments 1A and 1B, a preliminary experiment in our laboratory (Brubaker, Warren, & Bashford, 1994) indicated that the ability to detect repetition of long-period RFNs was not limited to experienced laboratory personnel. Following a simple training procedure, listeners who had never heard RFNs before were able to tap in synchrony with a particular portion of a waveform for RFNs having durations well beyond the

Table 2
Mean Tap Period and Mean Tap Location and Standard Errors for the Nine Taps Made to Recycling Frozen Noise (RFN) During the Test Phase of Trials (Light Cue Absent) in Experiment 1A

RFN Period	Tap Period		Tap Location	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
10 sec				
B.B.	9.96	0.04	1.06	0.19
J.B.	10.01	0.01	0.82	0.37
15 sec				
B.B.	14.99	0.05	1.12	0.20
J.B.	14.99	0.08	0.40	0.15
20 sec				
B.B.	20.10	0.13	1.39	0.27
J.B.	20.01	0.06	0.93	0.30

Note—Tap locations in the table are relative to that portion of the frozen noise concurrent with the onset of the light flash during the learning phase. Results for the 2 listeners are averaged across the five RFN samples presented at each of the three repetition periods.

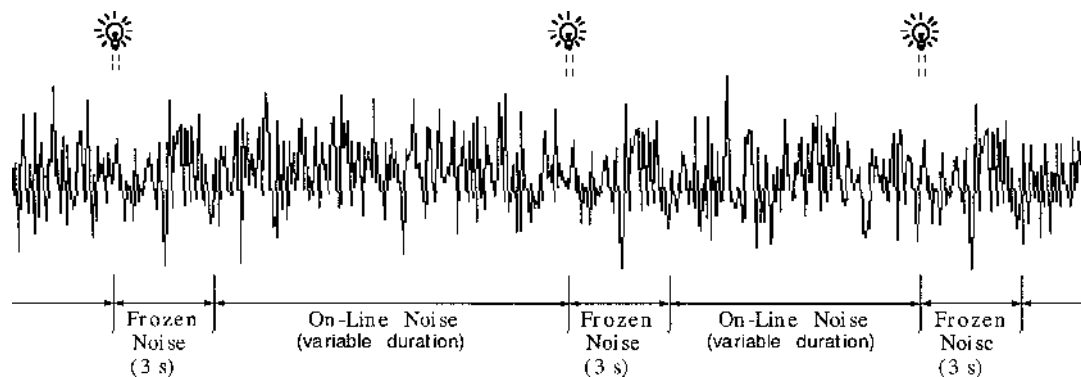


Figure 2. Diagrammatic representation of stimuli used in Experiment 1B. Frozen noise segments repeated aperiodically, with intervening segments of nonrepeating noise varying randomly from 7 to 13 sec in duration. Repeating noise segments were cued at onset by a 200-msec light flash.

whooshing range. This successful training was accomplished during a half-hour session in which subjects were presented first with a 1-sec whooshing range RFN accompanied by a light flash occurring in synchrony with a randomly selected portion of the acoustic waveform. When the same RFN was presented subsequently without the flash, listeners could tap in synchrony with the previously highlighted segment of the waveform. The process was then repeated with a series of increasing RFN periods. It was found that listeners could continue to tap accurately to the same segment when the light cue was no longer present for RFNs of as much as 10 sec in duration. As a control, a separate group of subjects from the same pool of students heard the same series of RFNs having increasing periods, but without the prior cross-modal cuing. Surprisingly, these listeners also could tap in synchrony with the RFN period for repetition frequencies up to 10 sec. However, their performance was less accurate than that of the group provided with cuing, and the taps by the uncued group were not clustered about any single portion of the RFN waveform. On the basis of these preliminary observations, Experiments 2A and 2B were undertaken with inexperienced listeners to determine the relative salience of various spectral regions in the recognition of acoustic patterns within long-period stochastic waveforms. Both flash-cued and uncued conditions were employed.

Method

Subjects

The 90 subjects (45 serving in Experiment 2A and 45 in Experiment 2B) were recruited from introductory courses in psychology. Subjects were in their late teens or 20s and reported having no known hearing problems. They were given either money or course credit for participation.

Stimuli

The RFNs employed as stimuli were derived from white noise and consisted of $\frac{1}{2}$ -octave bands having slopes of 96 dB/octave and center frequencies of 370, 1,500, and 6,000 Hz, respectively. Each of these three frequency bands was used to construct four RFN samples (two of which were used in Experiment 2A and two in Ex-

periment 2B) at each of the following six durations: 1, 2, 3, 5, 7, and 10 sec for a total of $3 \times 4 \times 6$, or 72 RFN samples.

Experiment 2A Procedure (Prior Light-Flash Cuing)

Stimuli were delivered diotically through headphones at a level of 70 dBA to the listener seated in a sound-attenuating chamber. The 45 subjects were divided into three groups of 15 subjects. Each group received bandpassed RFNs having one of the three center frequencies (370, 1,500, or 6000 Hz). A 60-msec flash cue used for cuing prior to recognition testing was produced by a light-emitting diode and had exactly the same repetition period as the accompanying RFN. The flash occurred at a different randomly determined portion of the RFN for each subject, so that any enhancement of waveform recognition in the vicinity of the flash cue could not be attributed to the chance presence of a salient acoustic feature. Subjects received two independent RFNs accompanied by the light cue for each of the six repetition periods. The 1-sec RFNs were presented first as practice stimuli. (Since these were in the whooshing range, all listeners perceived repetition with ease, and results were not analyzed as part of the formal experiment.) As soon as the listeners informed the experimenter that they could recognize a com-

Table 3
Responses of 2 Listeners Instructed to Tap to Each Repetition of Some Recognizable Feature Within a 3-sec Frozen Noise Segment Recurring Aperiodically in Experiment 1B

	Frozen Noise Sample				
	1	2	3	4	5
Listener B.B.					
Repetitions heard prior to tapping	8	12	9	8	9
Mean tap location	2.28 sec	1.18 sec	2.62 sec	1.80 sec	2.22 sec
Standard error	.023	.039	.020	.018	.019
Listener J.B.					
Repetitions heard prior to tapping	6	5	8	6	9
Mean tap location	1.75 sec	1.74 sec	2.46 sec	0.67 sec	2.12 sec
Standard error	.015	.015	.028	.011	.012

Note—Variable durations (7–13 sec) of on-line noise separated frozen noise repetitions. Included for each listener and noise sample are the number of pattern repetitions heard prior to the initial tapping response (light cue present at the onset of the frozen noise) and the mean tap location relative to the onset of the 3-sec frozen noise, along with the standard error of the mean, for 20 taps made during the subsequent test phase of trials with the light cue absent.

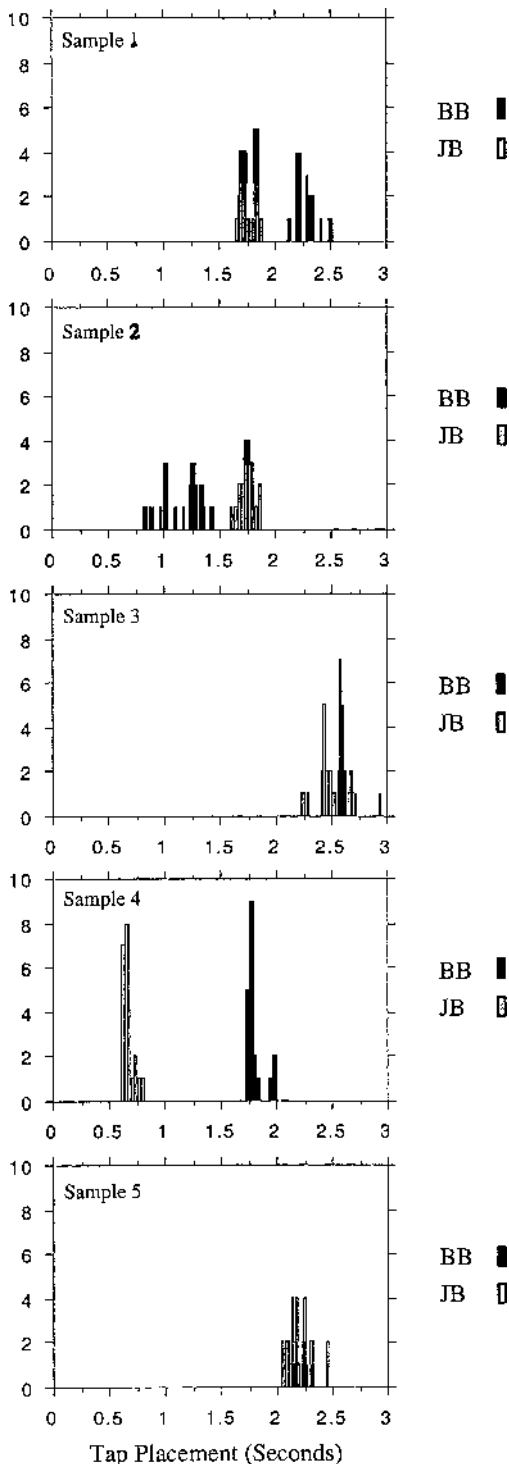


Figure 3. Tap locations relative to the onset of 3-sec frozen noise segments presented aperiodically in Experiment 1B.

ponent of the repeating stimulus coincident with the flash, the visual and auditory stimuli were terminated. They were told that they would be hearing the same repeating sound again, but this time without the flash. Subjects were instructed to tap on a contact key in synchrony with the recurrence of the portion of the sound that

had occurred along with the previous light flashes. The RFN was then presented by itself. Taps during the first 20 sec were not scored, and the next five intertap intervals were determined. If the subject clearly missed a tap, but was accurate with the following one, then the tapping was allowed to continue for five successive intertap intervals. Thus, the goal of the trial was to obtain a measure of five consecutive intertap intervals to be used for analysis. The procedure was repeated with the other 1-sec cued RFN and then continued with each pair of RFNs in order of increasing durations up to the two 10-sec stimuli. A computer recorded the tap positions relative to the prior light flash as well as the intertap intervals.

Experiment 2B Procedure (No-Cuing)

Three separate groups of 15 subjects were tested using the same procedure employed in Experiment 2A except for the absence of cross-modal cuing prior to the formal test and the use of a different set of two independent samples for each of the three RFN frequency ranges. Listeners were instructed to listen to the repeating sound until they could hear a recurrent feature, and then to tap in synchrony with that feature. Instead of scoring the tap locations relative to the onset of the light flash, as with the cued groups, a randomly determined position within the RFN waveform was used as an arbitrary landmark for scoring responses. As in Experiment 2A, responses by the uncued groups for the 1-sec whooshing range RFNs were considered as familiarization and practice, and were not counted as part of the formal experiment.

Results and Discussion

In order for a trial to be scored as a successful recognition of repetition, the mean of the subject's five successive intertap intervals was required to be within 10% of the RFN's actual period. All listeners completed some of their 30 trials successfully, but no one was successful on all trials. Table 4 lists the percentage of the 30 trials in Experiments 2A (cued) and 2B (uncued) that were scored as successful for each of the three center frequencies at each of the five RFN periods, as well as the percentages averaged across periods. Figure 4 shows the distribution of taps relative to a fixed reference location on the RFN waveform for successful trials under cued and uncued conditions for each of the three center frequencies summed across RFN repetition periods. The vertical lines in the histograms for cued conditions represent the location of the flash in the prior training phase, and the vertical lines in the histograms for uncued conditions are at a randomly selected fixed reference location. Each bin in the histograms represents 10% of the RFN period: Negative bin numbers designate taps made in the half-period preceding the reference location, and positive bin numbers designate taps made in the following half-period.

The percentage of successful tapping trials for listeners in each condition was subjected to a three-way mixed analysis of variance (ANOVA), with center frequency and cuing as between-subjects variables and RFN duration as a within-subjects variable. The ANOVA revealed that performance was improved both by cuing [$F(1,84) = 22.1, p < .0001$] and by shorter repetition durations [$F(4,84) = 15.2, p < .0001$]. There was also a significant effect of RFN center frequency [$F(2,84) = 28.8, p < .0001$]. Both the 370- and the 1500-Hz conditions had performance levels that were significantly higher than the 6000-Hz condition ($p < .05$ by Tukey HSD test), but there was no

Table 4
Mean Percentage of Cued and Uncued Trials Completed Successfully
in Experiment 2 (Trials Were Scored as Successful if the Mean Intertap
Interval of Completed Trials Was Within 10% of the RFN Period)

RFN Periods	Center Frequencies (Hz)											
	370				1500				6000			
	Cued		Uncued		Cued		Uncued		Cued		Uncued	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
2 sec	93.3	4.5	93.3	4.5	80.0	6.5	63.3	9.1	73.3	8.3	40.0	12.1
3 sec	80.0	8.2	63.3	9.1	76.7	9.6	60.0	8.7	63.3	10.3	26.7	9.6
5 sec	73.3	8.3	63.3	11.4	76.7	8.3	60.0	8.7	40.0	11.1	20.0	9.5
7 sec	73.3	8.3	63.3	10.3	80.0	8.2	40.0	12.1	36.7	11.4	6.7	4.5
10 sec	70.0	9.5	53.3	11.4	56.7	11.8	30.0	8.2	20.0	8.2	0.0	0.0
Averaged	78.0	3.6	67.3	4.5	74.0	4.1	50.7	4.4	46.7	4.9	18.7	4.0

Notes—Percentages are based on a maximum score of 30 successful trials (two trials for each of 15 subjects). Percentage values summed across RFN periods appear at the bottom of the table.

significant difference in performance for the 370- and the 1500-Hz conditions.

Experiment 2 demonstrates that without any previous training, most of the listeners in most of the conditions could detect repetition of long-period narrowband RFNs. The effectiveness of cross-modal cuing in enhancing repetition detection can be seen in both Table 4 and Figure 4. Figure 4 also shows that, for the group receiving prior cuing, the two bins flanking the location of a prior light-flash cue contained the highest number of successful trials. For both the cued and uncued groups, repetition detection for the 6000-Hz center frequency was the most difficult. There is a lack of phase locking of auditory nerve fibers to the waveform fine structure at this high frequency (Palmer & Russell, 1986; Rose, Brugge, Anderson, & Hind, 1968), and this could result in the loss of detailed information concerning the recurring acoustic patterns. A greater salience of pattern information in low and midrange frequency regions was observed previously by Hanna (1984), who used brief (0.1–409.6 msec) bandpass-filtered segments of frozen noise presented as nonrecycling, “one-shot” stimuli in a two-interval forced-choice discrimination task involving *same/different* judgments.

GENERAL DISCUSSION

Recycled Frozen Noises (RFNs) evoke a continuum of detectable periodicity with an upper durational limit of the repeated waveform (as reported in Experiment 1) of at least 20 sec (frequency of 0.05 Hz) and a lower durational limit of roughly 6×10^{-5} sec (corresponding to a fundamental frequency between 15,000 and 20,000 Hz). Because there are no a priori restrictions concerning waveform or spectrum for any given repetition frequency, these stochastic waveforms have been used as generic or model periodic stimuli for both the tonal and infratonal frequency ranges (see Warren, 1999, pp. 73–80; Warren & Bashford, 1981). As with RFNs in the pitch range (frequencies above 20 Hz), RFNs in the motorboating range (approximately 4–20 Hz) are heard as global percepts,

with listeners unable to detect discrete temporal components or events. For RFNs from roughly 1 to 4 Hz, there are two modes of detecting repetition: (1) the global perception of a repetitive whooshing sound that is heard almost immediately, even when RFNs had never been heard before; and (2) the detection of repetitive components such as clanks and thumps occupying limited portions of the waveform (these can emerge slowly with continued listening). For frequencies below roughly 1 Hz (periods greater than 1 sec), a global percept encompassing the entire waveform is no longer possible. However, the present study has shown that not only can experienced laboratory personnel recognize the recurrence of portions of RFNs having 10-sec periods, but also listeners with no prior exposure to RFNs can accomplish the task. While light flashes synchronized with a randomly selected portion of a long-period RFN waveform highlight the segment and enable listeners to recognize it when the flash is no longer present, this cross-modal cuing is not required for repetition detection (although detection scores are generally lower without prior cuing). This study has also shown that when the spectral components of long-period RFNs are limited by bandpass filtering, repetition of low- and mid-range bands is recognized more readily than the repetition of a high-frequency band (center frequency 6000 Hz), possibly due to the phase locking of nerve fiber response to fine structure at the lower frequencies.

The ability to recognize portions of novel stochastic patterns after delays of 10 or even 20 sec indicates that salient segments can be held in auditory memory for intervals encompassing the durations of melodic themes, and long enough to allow the use of remote semantic and syntactic context for lexical retrieval in sentences. It is of interest that Sams, Hari, Rif, and Knuutila, (1993) reported neuromagnetic evidence that the memory trace for a particular tone persists for at least 10 sec, and that Cowan, Lichty, and Grove (1990) found that the memory trace for unattended spoken syllables in running speech also persisted for at least 10 sec.

In keeping with the concept that RFNs represent generic stimuli for examining basic rules governing percep-

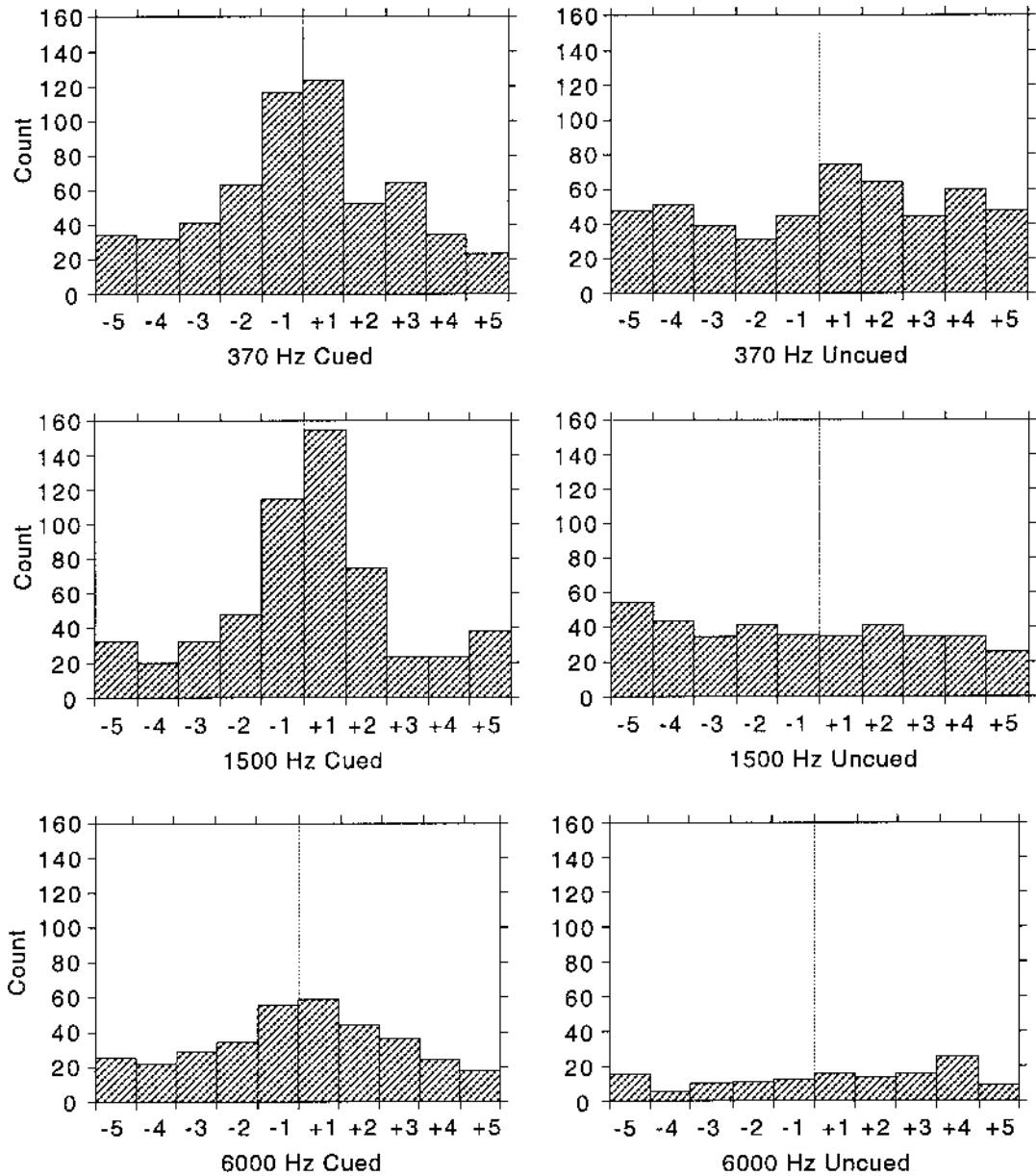


Figure 4. Distribution of listeners' taps to recurrent features heard for $\frac{1}{2}$ -octave bands of RFNs having the center frequencies shown beneath each panel. The panels on the left represent judgments by listeners receiving prior light-flash cuing, and panels on the right are for responses by listeners receiving no prior cuing. Each bin represents 10% of the RFN period, with negative bin numbers designating intervals in the half-period preceding the reference position and with positive numbers representing intervals in the subsequent half-period. The reference position was that of the prior light flash for the cued group and the fixed randomly determined point for the uncued group. Data represent responses summed across the five RFN periods ranging from 2 through 10 sec.

tion of acoustic patterns (including speech), there appear to be similarities in the characteristics exhibited by the hierarchical organizations of speech and those of RFNs having durations corresponding to phonemes, words,

and sentences. Thus, Efron (1963) reported that in normal speech, phonemes have average durations of roughly 80 msec (placing them in the durational range of motor-boating as RFNs). He also reported that the average du-

ration of words in connected speech is approximately 450 msec (placing both monosyllabic and polysyllabic words in the range of whooshing RFNs). Both motor-boating RFNs and phonemes are too brief to be segmented by listeners into component elements or discrete events, and individual exemplars within each type are distinguished by their characteristic spectra and resulting timbres. Whooshing RFNs and words can both be heard as global patterns, and in addition, each can be segmented into briefer components (clanks and thumps within RFNs, and phonemes within syllables and words).

The study of RFNs started with the observations published by Guttman and Julesz (1963). In a retrospective on his career to date, Julesz (1995) pointed out the relation of RFNs to the patterns he employed in his innovative research involving random-dot correlograms. These stochastic optical patterns have provided a quite useful tool for revealing basic principles of visual perception freed from the special features and rules associated with particular organizational patterns. It appears that the stochastic acoustic patterns of RFNs can serve a similar purpose in auditory perception.

REFERENCES

- BASHFORD, J. A., JR., BRUBAKER, B. S., & WARREN, R. M. (1993). Cross-modal enhancement of repetition detection for very long period recycling frozen noise [Abstract]. *Journal of the Acoustical Society of America*, **93**, 2315.
- BRUBAKER, B. S., WARREN, R. M., & BASHFORD, J. A., JR. (1994). Repetition detection by untrained listeners presented with very long period recycled noises [Abstract]. *Journal of the Acoustical Society of America*, **95**, 2966.
- COWAN, N. (1984). On short and long auditory stores. *Psychological Bulletin*, **96**, 341-370.
- COWAN, N. (1987). Auditory memory: Procedures to examine two phases. In W. A. Yost & C. S. Watson (Eds.), *Auditory processing of complex sounds* (pp. 289-298). Hillsdale, NJ: Erlbaum.
- COWAN, N., LICHTY, W., & GROVE, T. R. (1990). Properties of memory for unattended spoken syllables. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 258-269.
- EFRON, R. (1963). Temporal perception, aphasia, and déjà vu. *Brain*, **86**, 403-424.
- GUTTMAN, N., & JULESZ, B. (1963). Lower limits of auditory periodicity analysis. *Journal of the Acoustical Society of America*, **35**, 610.
- HANNA, T. E. (1984). Discrimination of reproducible noise as a function of bandwidth and duration. *Perception & Psychophysics*, **36**, 409-416.
- JULESZ, B. (1995). *Dialogues on perception*. Cambridge, MA: MIT Press.
- JULESZ, B., & HIRSH, I. J. (1972). Visual and auditory perception—An essay of comparison. In E. David & P. Denes (Eds.), *Human communication: A unified approach* (pp. 283-340). New York: McGraw-Hill.
- KAERNBACH, C. (1992). On the consistency of tapping to repeated noise. *Journal of the Acoustical Society of America*, **92**, 788-793.
- LIMBERT, C. (1984). *The perception of repeated noise*. Unpublished doctoral dissertation, Cambridge University.
- LIMBERT, C., & PATTERSON, R. D. (1982). Tapping to repeated noise [Abstract]. *Journal of the Acoustical Society of America*, **71** (Suppl. 1), S38.
- PALMER, A. R., & RUSSELL, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of the inner hair-cells. *Hearing Research*, **24**, 1-15.
- POLLACK, I. (1972). Memory for auditory waveform. *Journal of the Acoustical Society of America*, **52**, 1209-1215.
- POLLACK, I. (1975). Identification of random auditory waveforms. *Journal of the Acoustical Society of America*, **58**, 1262-1271.
- ROSE, J. E., BRUGGE, J. F., ANDERSON, D. J., & HIND, J. E. (1968). Patterns of activity in single auditory nerve fibers of the squirrel monkey. In A. V. S. de Reuk and J. Knight (Eds.), *Hearing mechanisms in vertebrates* (pp. 144-157). London: Churchill.
- SAMS, M., HARI, R., RIF, J., & KNUUTILA, J. (1993). The human memory trace persists about 10 sec: Neuromagnetic evidence. *Journal of Cognitive Neuroscience*, **5**, 363-370.
- WARREN, R. M. (1999). *Auditory perception: A new analysis and synthesis*. New York: Cambridge University Press.
- WARREN, R. M., & BASHFORD, J. A., JR. (1981). Perception of acoustic iterance: Pitch and inrapitch. *Perception & Psychophysics*, **29**, 395-402.

(Manuscript received March 1, 1999;
revision accepted for publication April 12, 2000.)