The role of masker fringes for the detection of coherent tone pips

Virginia M. Richards

Department of Cognitive Sciences, University of California, Irvine, California 92697-5200

Daniel E. Shub

NIHR National Biomedical Research Unit in Hearing, School of Psychology, Nottingham University, Nottingham NG7 2RD, United Kingdom

Eva Maria Carreira^{a)}

Department of Cognitive Sciences, University of California, Irvine, California 92697-5200

(Received 3 September 2010; revised 20 June 2011; accepted 23 June 2011)

Three experiments investigated the role of pre/post exposure to a masker in a detection task with complex, random, spectro-temporal maskers. In the first experiment, the masker was either continuously presented or pulsed on and off with the signal. For most listeners, thresholds were lower when the masker was continuously presented, despite the fact that there was more uncertainty about the timing of the signal. In the second experiment, the signal-bearing portion of the masker was preceded and followed by masker “fringes” of different durations. Consistent with the findings of Experiment 1, for some listeners shorter-duration fringes led to higher thresholds than long-duration fringes. In the third experiment, the masker fringe (a) preceded, (b) followed, or (c) both preceded and followed, the signal. Relative to the middle signal conditions, a late signal yielded lower thresholds and the early signal yielded higher thresholds. These results indicate that listeners can use features of an ongoing sound to extract an added signal and that listeners differ in the importance of pre-exposure for efficient signal extraction. However, listeners do not appear to perform this comparison retrospectively after the signal, potentially indicating a form of backward masking.

© 2011 Acoustical Society of America. [DOI: 10.1121/1.3613701]

PACS number(s): 43.66.Dc, 43.66.Mk [LD]

Pages: 883–892

I. INTRODUCTION

For the detection of a tone added to noise, performance can depend on whether the noise is continuous or pulsed on and off with the signal. In monaural conditions (e.g., N_0S_0), whether the noise is continuous or pulsed has a small effect on thresholds, but in binaural conditions (e.g., N_0S_π) the presentation mode has a substantial effect such that thresholds are approximately 10 dB lower with the continuous masker (e.g., McFadden, 1966; Yost, 1985; Zwicker and Zwicker, 1984). Presumably, for the binaural conditions, the non-signal bearing portions of the masker provide a reference from which the binaural system can detect a “change” when the signal is added (e.g., McFadden, 1966; Yost, 1985). By this argument, the difference in the pulsed-continuous contrast for monaural and binaural presentations reflects differences in the degree to which the auditory system can detect a change in state associated with the added signal.

The aim of the current experiments is to determine whether the non-signal bearing portion of the masker can provide a reference for the detection of a signal when the added signal is defined in terms of a high-level perceptual organization that emerges as the sound stream unfolds. Just as the background of an image can provide a reference for detecting a visual signal, there may be locations in time and

frequency that listeners rely on when detecting a target in the presence of competing acoustical elements. Here, we employ a multi-burst random (MBR) masker (similar to that used by Kidd *et al.* 1995 and Huang and Richards, 2006, 2008) to evaluate linear models of the detection of sound patterns in time and frequency. The left panel of Fig. 1 is a schematic spectrogram for a signal added to a continuous MBR masker. The signal to be detected is a series of 1000-Hz tone pips and the masker is a sequence of “bursts” composed of multiple tone pips with both the number of tone pips and the frequencies of the tone pips drawn at random. The “burst” nature of the masker forms a natural metric for the extraction of relative weights in time to evaluate linear models. Further, the burst nature decreases the likelihood that listeners can rely on spectral splatter associated with signal onset and offsets for detection.

There are three features of the signal and MBR masker that limit signal detection on the basis of simple changes at the auditory periphery. First, the signal and masker are both composed of tone pips, and these tone pips are turned on and off synchronously. Consequently, there is no unique “onset” or “offset” event associated with the added signal. Second, the tone pips that comprise the masker are drawn at random on each trial. This means that masker components sometimes fall in the region of the signal frequency leading to large variations in energy at the signal channel. As a result, a comparison of energy before and during the signal is a non-optimal decision strategy. Third, for any one burst, the number of

^{a)}Author to whom correspondence should be addressed. Electronic mail: ecarreira@uci.edu

FIG. 1. Schematic of signal stimuli for Experiments 1–3 with an MBR masker. The left panel shows an example of a signal composed of three 60-ms tone pips (dark bars) added to a continuous MBR masker (Experiment 1). The right panel shows both a pressure waveform (top) and spectrogram (panel) for a signal (dark bars) added to an MBR masker with two fringe bursts (Experiment 2). This figure defines the component tone pip, bursts (vertical line in right-hand panel), and indicates the fringe bursts using gray filled schematic pressure waveform.

tones that comprise the burst is random, meaning the addition of the signal tone pip does not provide a reliable increase in the number of tone pips in a burst.

If listeners can make comparisons across time to detect a signal stream embedded in an MBR masker, one would expect thresholds to be lower when the masker is continuous than when the masker is pulsed. Several studies suggest this result is unlikely. First, [Huang and Richards \(2008\)](#) and [Shub and Richards \(2009\)](#), who used MBR-like maskers, found that the duration over which listeners accumulate information is longer than the signal duration. Second, [Bonino and Leibold \(2008\)](#) found large effects of temporal uncertainty with continuous MBR-like maskers. Third, unlike a continuous masker, for a pulsed stimulus there is an opportunity to “review” or reconsider the stimulus in memory before making a decision.

However, contrary to these expectations, the results of Experiment 1 indicate that for the MBR masker, and for some listeners, sensitivity is superior for a continuous than a pulsed masker. For the remaining listeners, sensitivity was nearly equal across the two masker presentation modes. In that experiment the signal is presented at random times throughout a 5.1-min presentation and listeners are instructed to press a button when they hear the signal. Presumably, listeners relied on the statistics/features of the random masker preceding and/or following the signal to detect a relatively high-level change associated with the addition of the signal. For example, a reduction in the variance of energy over time would indicate the presence of the signal, as would a change in the modulation spectrum. In an effort to clarify the results of this first experiment, two additional experiments were run.

In Experiments 2 and 3, the non-signal bearing portion of the MBR masker is systematically altered. In Experiment 2 the time between the masker onset and offset, relative to the signal, is varied to determine the fringe length required to mimic a continuous masker. Experiment 3 examines the impact on thresholds of the portions of the masker preceding and those following the signal-bearing portion of the MBR masker.

These experiments allow the exploration of two basic hypotheses and allied predictions. First, to the degree the masker provides a reference for the detection of a change in features associated with an added signal, in parallel with the results from binaural conditions ([Trahiotis and Dolan, 1972](#); [Zwicker and Zwicker, 1984](#)), one would expect a change to be detected whether the masker precedes the signal and is turned off with the signal or the masker is turned on with the signal and continues beyond the signal offset. Second, the mechanisms tapped by this stimulus set, potentially reflecting components of sound source segregation, may depend on the unfolding soundscape to isolate new sound sources, and if so, be sensitive to contrasts that are strongly directional from past to future.

In addition to examining thresholds as a function of signal location in the masker, Experiment 3 measures relative weights associated with a linear model. Recall that the MBR masker forms a time-by-frequency random-tone display, allowing the estimation of linear relative weights in *both* time and frequency. As will be described below, the resulting relative weights emphasized the region of the signal frequency, and so the analysis concentrates on the pattern of relative weights as a function of time.

II. EXPERIMENT 1: THE DETECTION OF EQUAL-FREQUENCY TONE PIPS WITH CONTINUOUS AND PULSED MASKERS

A. Methods

1. Participants

Nine listeners, five male and four female, ranging in age from 19 to 31, participated in this experiment. All listeners had thresholds in quiet of 20 dB HL or less at octave frequencies ranging from 250 to 8000 Hz in both ears. Listeners L1–L5 will be referred to as Group A and L6–L9 will be referred to as Group B. The members of Group A ran the experiment individually in a double-walled soundproof booth, while Group B ran the experiment individually in a sound-isolated room situated in a quiet hallway. All listeners were reimbursed for their participation.

2. Stimuli and conditions

The stimuli were generated and presented at a sample rate of 48,828 Hz using Matlab with TDT RP2 and HB7 modules connected to a desktop PC. The stimuli were attenuated using both digital and analog attenuation to achieve desired stimulus levels and presented diotically using Sennheiser HD410SL headphones.

For Group A, 8 conditions were tested, resulting from the crossing of three independent variables: type of masker (noise or MBR masker), burst duration (30- or 60-ms), and presentation mode (continuous or pulsed). Group B ran the experiment after Group A had completed the experiment. Based on the results for Group A, only conditions with burst durations of 60 ms were tested for Group B.

The MBR masker was composed of multiple sequential bursts, with each burst being the sum of tone pips of different frequencies. The bursts were turned on and off using 5-ms raised cosine ramps, and the burst durations were either 30-

or 60- ms. There was no temporal gap between the offset of one burst and the onset of the next. The number of tone pips comprising each burst was drawn at random from a Poisson distribution with a parameter/mean of 20 tone pips. The frequency of each tone pip in each burst was independently drawn at random from a range of 200 to 5000 Hz using a uniform distribution on a logarithmic frequency scale. The level of each tone was 50 dB SPL yielding an overall level of 63 dB SPL. The noise, which was digitally generated by shaping the power spectral density of Gaussian noise using 1/f decay, ranged in frequency from 200 to 5000 Hz. This “pink” noise also had an average overall level of 63 dB SPL and a long-term spectrum equivalent to the MBR masker. The noise masker, however, did not form bursts, i.e., it was not turned on and off except at the onset and offset of the masker.

The signal was composed of a sequence of six 1000-Hz tone pips for the 30-ms condition and three 1000-Hz tone pips for the 60-ms condition. The tone pips that comprised the signal were gated synchronously with the MBR masker bursts. For the noise masker, the signal was turned on and off at the same point in time relative to the onset and offset of the masker.

The presentation mode was either continuous or pulsed. The left panel of Fig. 1 is a schematic of a portion of a continuous MBR masker. For the continuous mode, the masker was present throughout the 5.1-min run. For the pulsed mode the masker preceded and followed the 180-ms signal-bearing portion of the masker by one burst, either 30- or 60-ms depending on the duration of the bursts being tested. Otherwise, for the pulsed mode, the time between “trials” was silent.

The probability of the signal being presented on a trial was 0.8. This relatively high rate of signal trials was adopted so as to increase the rate at which signals occurred across the 5-min run (see below). For the continuous mode, the trials were *not* indicated to the listener. For the pulsed mode, the masker onset indicated the presence of a trial, thereby reducing temporal uncertainty with regard to when the signal could be present. For each signal trial, the signal level tested was chosen at random from a range extending from 5 dB below to 5 dB above the mean signal level.

3. Procedures

Detection trials were presented at times assigned by random draws from a geometric distribution across the 5.1-min runs. For the continuous maskers, the time at which a trial was initiated was known to the experimenter but not to the listener. For the pulsed trials, the trial onset and offset were indicated to the listener because the masker was turned on/off in an otherwise quiet environment. For the 30-ms conditions, a geometric distribution with an expected rate of 1 s and a no-trial dead time of 2.5 s following each trial was enforced. For the 60-ms conditions, the expected rate was 2 s and the no-trial dead time was 5 s. These differences between the 30- and 60-ms conditions were not intended, and reflect a programming error. Participants were instructed to press a response button when they heard a signal. Ongoing, but delayed, feedback was provided using a GUI display as follows. After the first 10 s of an experimental run, the bottom line of the screen presented times at which a

signal occurred (filled circles), times at which responses occurred (asterisks), and times at which no-signal trials occurred (unfilled circles). Time was indicated from left to right. After the second 10 s of the run, the initial feedback graph scrolled up one line, and the bottom line was replaced with feedback for the second 10 s of the experiment. This updating procedure was continued throughout the 5.1-min run, yielding a full raster plot of feedback. Next, summary statistics of hit, false alarm, and miss rates were provided as well as an estimate of d' . Listeners were encouraged not to miss any signals at the cost of increasing false alarms.

The members of Group A (L1 – L5) initially practiced two or more hours to become familiar with the task. During this time, the signal level was chosen to provide an overall value of d' of approximately 1.5. Additionally, in most instances, listeners practiced half an hour prior to data collection in a newly encountered condition. For members of Group B (L6-L9) listeners practiced for 2 – 3 h on several conditions prior to data collection. When a new condition was introduced, listeners completed at least three 5.1-min runs prior to data collection.

During data collection listeners completed one condition before starting another. Listeners L1 and L2 completed the 30-ms burst conditions prior to starting the 60-ms burst conditions and for each duration the four conditions (two maskers and two presentation modes) were tested in a random order. For the remaining listeners, L3–L9, the order in which the conditions were completed (duration by masker by presentation mode) was random.

For each condition data were collected over two consecutive sessions, yielding approximately 20–30 5.1-min runs per condition. The smallest data set analyzed was composed of 15, 5.1-min runs.

For the members of Group B conditions from this experiment were intermixed with conditions from a pilot study. Listener L6 did not complete the continuous presentation mode for the noise masker.

4. Data analysis

Because the trials were not clearly delineated within the sound presentation procedure in the continuous masker conditions, a classification scheme was required to define hits, misses, correct rejections and false alarms (Egan *et al.*, 1961). Keep in mind that signal and no-signal trials were included even if their temporal onsets and offsets were known only to the experimenter (Watson and Nichols, 1976). The trial was coded as a hit regardless of whether the listener was aware that a detection trial was underway, if the listener responded “signal” no earlier than 0.05 s after the signal onset and no later than the 2 s following the signal onset. A failure to respond within 2 s of the signal onset indicated a miss. No-signal trials were treated in the same manner (i.e., a window was positioned according to where the signal would have begun, had there been a signal). For no-signal trials, false alarms occurred only when the listener responded “signal” within the 1.95 s window that followed the onset of the “virtual” signal.

For each condition, data from the 5.1-min runs were combined. Recall that the signal level was chosen at random from a 10-dB range on each trial. The signal levels were accumulated into seven bins with approximately equal number of signal trials within each bin. Then, for each bin, the hit rate was estimated. The estimate of false alarms was based on all no-signal trials. The resulting psychometric function relating d' ($d' = Z_{\text{hit}} - Z_{\text{FA}}$) and signal level was fitted using a weighted linear least-squares fit. The threshold was estimated at the signal level associated with an estimated value of $d' = 1.5$. The 68% confidence intervals were estimated using a bootstrap simulation with 500 replicate threshold estimates based on re-draws of the original data set of hits and false alarms. The 68% confidence intervals provide a proxy for the standard error of the threshold estimate.

B. Results and discussion

Figure 2 shows thresholds for pulsed and continuous presentation modes for each listener. The left and central panels show the results for Group A (L1–L5) for the 30-ms and 60-ms burst conditions, respectively. The right panel shows thresholds for Group B (L6–L9) with 60-ms bursts. For each panel the pulsed and continuous conditions are plotted separately, and the dashed and solid lines are for the noise and MBR masker conditions, respectively. Error bars indicate an estimate of the standard error of the threshold and reasonably reflect the quality of the fits to the psychometric functions. Of the 55 fitted psychometric functions, all but six accounted for over 75% of the variance in the corresponding data. For the 30-ms burst conditions (left panel), there is less intersubject variability for the continuous presentation mode than for the pulsed presentation mode, particularly for the MBR masker (solid lines). For the 60-ms burst data (Groups A and B), this trend is amplified, along with an overall increase in threshold with the MBR masker.

The results for the noise masker indicate a relatively small effect of whether a pulsed or continuous presentation

mode is used or whether the duration of the signal pips is 30 or 60 ms. Averaged across all listeners, and both the 60- and 30-ms conditions, there is an approximately 2.5 dB advantage in thresholds for the continuous presentation mode relative to the pulsed presentation mode. This result is consistent with results reported by Yost (1985), even though there are substantial differences in methods. In our method, based on the method of free response (Egan *et al.*, 1961), the detection interval is not demarcated in the continuous presentation mode, which required sustained vigilance on the part of the listener for the continuous relative to the pulsed presentation mode. The general agreement in the magnitude of the advantage suggests that subjects were able to maintain their focus over the entire 5.1-min block, which is consistent with the findings of Shub and Richards (2009). The difference in thresholds for the 30 ms minus the 60 ms signal pips, averaged across listeners and the pulsed and continuous condition, is approximately 1 dB. Thus, for the detection of the signal pips added to noise, and given that the total signal duration is the same, thresholds are roughly the same regardless of whether the signal is composed of six 30-ms tone pips or three 60-ms tone pips.

Next consider the MBR conditions. On average, performance was worse with the MBR than with the noise masker, and worse for the pulsed than the continuous presentation mode. For the 60-ms bursts, listeners L2, L5, and L9 have very high thresholds (as well as poorly estimated thresholds, note the larger error bars) for the pulsed presentation mode whereas for listeners L4, L6, L7, and L8 the thresholds in the pulsed and continuous presentation modes lie within a 3 dB range. These trends reflect larger individual differences in thresholds for the pulsed mode than the continuous mode.

Despite the fact that in the continuous presentation mode the 60-ms duration tone pips had more temporal uncertainty than the 30-ms duration tone pips,¹ changing from the pulsed to the continuous presentation mode resulted in a greater benefit for the 60-ms than the 30-ms bursts. For the former a change from pulsed to continuous maskers yielded

FIG. 2. Thresholds for individual listeners are plotted as a function of the presentation mode (pulsed vs continuous). Error bars indicate the 68% confidence interval as an indication of the standard error of the threshold estimate. Results for the noise masker are plotted using dashed lines and results for the MBR masker are plotted using solid lines. The left panel is the 30-ms bursts and the center and right panels show results for the 60-ms bursts. The left and center panels are results for Group A and the right panel shows results for Group B.

a 11.7 dB reduction in thresholds, and for the 30-ms bursts thresholds improved by 5.1 dB.

Two *t*-tests (paired two-tailed) were run to examine (a) the effect of burst duration for the MBR masker and (b) the effect of presentation mode for the MBR masker. The former, which was restricted to the members of Group A, indicated that thresholds were significantly higher for the 60-ms than the 30-ms bursts ($p < 0.005$). The latter, which was applied to the members of both Groups A and B, indicated thresholds were significantly lower for the continuous condition than the pulsed condition ($p < 0.025$).

To summarize, for noise and MBR maskers composed of 30-ms bursts, thresholds are relatively consistent across listeners. For the MBR maskers composed of 60-ms bursts, thresholds are more variable across listeners and on average higher for the pulsed than continuous presentation mode even though there is more uncertainty in the latter than the former stimulus condition (compare with Bonino and Leibold, 2008). There are at least two factors that might be contributing to the relatively poor performance for the pulsed relative to the continuous conditions when the burst duration is 60 ms. First, listeners may integrate across a wide frequency region at the stimulus onset, thereby increasing the effective amount of masking in the pulsed condition (e.g., Scharf *et al.*, 2008, although see Reeves and Scharf, 2010, for a counter example). Second, the continuous masker may provide a reference against which the change associated with the added signal can be detected (e.g., McFadden, 1966). Whatever the case, these results suggest that for some listeners the statistics/features of an ongoing MBR masker contributes to efficient detection of a series of equal-frequency signal pips.

III. EXPERIMENT 2: FRINGE EFFECTS FOR YES/NO TRIALS

The results from Experiment 1 indicate that thresholds were typically lower for the continuous than the pulsed MBR masker. In Experiment 2 the number of bursts preceding and following the signal-bearing portion of the MBR masker is systematically varied to determine the number of preceding and following bursts required to reach asymptotic threshold levels. Note that for Experiments 2 and 3, the term “masker” will refer to the masker bursts during which the signal might be added, and the term “fringe” will refer to bursts that precede and follow the masker.

A. Methods

1. Participants

A subset of listeners in Group B (L7–L9) and two newly recruited listeners (L10 and L11) participated in this experiment. The two newly recruited listeners, one male and one female, were aged 22 and had thresholds in quiet of 20 dB HL or less at octave frequencies ranging from 250 to 8000 Hz in both ears. Those from Group B ran the experiment in a sound-isolated room situated in a quiet hallway and the newly recruited participants ran the experiment in a double-walled sound proof booth. All listeners were reimbursed for their participation.

2. Stimuli and conditions

The stimuli were generated and presented using methods essentially identical to the pulsed MBR, 60-ms burst condition of Experiment 1. The signal was a sequence of three 60-ms tone pips turned on and off synchronously with the MBR masker bursts. The average masker level was 63 dB SPL and the stimuli were presented diotically using Sennheiser HD410SL headphones. The detection trials were presented using a yes/no procedure, reducing the amount of temporal uncertainty regarding the timing of the signal compared to Experiment 1. The number of temporal “fringe” bursts that preceded and followed the masker was 0 (no fringe), 1 (similar to the pulsed MBR condition of Experiment 1), 2, or 4. As an example, if the number of fringes is two, two bursts preceded and followed the three masker bursts for a total of seven bursts. The right panel of Fig. 1 shows an example of a signal trial with a fringe of two.

3. Procedures and data analysis

The participants from Group B completed Experiment 1 and a pilot experiment prior to beginning Experiment 2. The newly recruited participants completed the pilot experiment before beginning Experiment 2. Prior to data collection, at least ten blocks of 50 trials were run for each of the four conditions. These practice data were used to determine the average signal level to be tested. Data were collected in one condition before moving onto the next one, and listeners were informed of the conditions. The order in which the different conditions were tested was randomly chosen for each listener, with the constraint that the sequence was unique across listeners.

Data were collected in 50-trial blocks. On each trial, the signal was as likely to be presented as not, and if present, the signal level was chosen at random from a 10-dB range. Listeners indicated their response, no signal or signal, by pressing a keypad. Correct-answer feedback was presented after every trial. Summary statistics were provided after each block. For each condition, listeners completed fifteen consecutive blocks of 50 trials each, with breaks following every five blocks. After completing all of the conditions, listeners ran an additional fifteen 50 trial blocks for each condition in reverse order. The final 1300 trials were analyzed for each listener in each condition.

Psychometric functions relating d' and signal level were estimated using the methods described for Experiment 1 except that there was no “response window.” Briefly, false alarms were estimated using all no-signal trials and hit rates were estimated by partitioning the signal levels into seven bins, and estimating a hit rate for each bin. The resulting psychometric function relating d' and signal level was fitted using a weighted linear least-squares fit. The threshold was estimated as the signal level associated with an estimated value of $d' = 1.5$. The 68% confidence intervals were estimated using a bootstrap simulation and used as an estimate of the standard error of the threshold estimate.

B. Results and discussion

Figure 3 shows estimated thresholds as a function of the number of fringe bursts for each listener. Across the 20 fitted

FIG. 3. Thresholds as a function of the number of fringe bursts preceding and following the central three-burst masker. Results are shown individually for the different listeners, and error bars indicate the 68% confidence interval for each threshold estimate.

psychometric functions (five listeners, four threshold estimates each), all but two accounted for over 88% of the variance in the corresponding data. For two of the listeners, L7 and L11, there is little impact of changes in the number of fringes, but for L8, L9, and L10 thresholds fall as the number of fringes increase.

Comparing thresholds estimated in the continuous presentation mode of Experiment 1 and the thresholds estimated with four fringe bursts in the current experiment for L7–L9, thresholds are on average 2 dB lower in this experiment. This near-equivalence of thresholds suggests that as long as there is a fringe of approximately four bursts, thresholds are approximately the same whether there is temporal uncertainty regarding the placement of the signal (Experiment 1) or not (Experiment 2). Put another way, it is unlikely that the results of Experiment 1 with the continuous presentation mode include a measurable effect of uncertainty regarding the signal's temporal location on thresholds. Taken together, the results of Experiments 1 and 2 indicate that for the detection of a sequence of tone pips, there is little, if any, effect of signal temporal uncertainty on thresholds, and for some subjects, there is a marked advantage to detection of including fringe bursts. In Experiment 3, below, this result is extended by determining whether fringes before, after, or both, mediate the reduction in threshold sometimes observed with the addition of the fringe.

The results from Experiments 1 and 2 stand in stark contrast to the results from [Bonino and Leibold \(2008\)](#) which demonstrated large effects of temporal uncertainty on detectability with MBR-like maskers. There are four notable differences in experimental procedures which may account for this discrepancy. First, their bursts contained only two masker tones. Second, their “bursts” were 120 ms in duration. Third, the signal consisted of a single “tone pip” of 120 ms. These three differences in stimulus presentation method may have several consequences on the outcomes of the experiments. For example, when the signal is presented in a single burst, there is no opportunity for the signal to form a

stream. Moreover, listeners can discriminate between two (masker) and three (signal-plus-masker) tones ([Neff and Green, 1987](#)), which might have encouraged a very different strategy for their listeners. A fourth important difference between these two studies is that in the [Bonino and Leibold \(2008\)](#) study temporal uncertainty was reduced by allowing the listener to initiate a trial, even when the masker was continuously present. Potentially, the listener chose moments in the masker stream during which the statistics of the masker were particularly well established; thereby maximizing the likelihood the target signal could be detected.

IV. EXPERIMENT 3: DETECTION OF SIGNALS OCCURRING EARLY OR LATE IN THE STIMULUS STREAM

The results of Experiments 1 and 2 indicate that for some listeners there are advantages to preceding and following a three-burst masker by a fringe composed of two or more bursts. In Experiment 3A the contribution of the leading and lagging fringes on thresholds is considered separately by estimating thresholds for signals that occur either early or late in the stimulus. There are a variety of potential outcomes for this experiment. As one example, if the fringe provides an opportunity for listeners to learn about the statistics of the masker, providing statistical contrast against which the subsequent unique signal stream is emphasized, one might expect a late signal/early fringe to support lower thresholds than an early signal/late fringe. As a second example, if as is true for the binaural masking level difference, the fringe provides the basis for detecting change, one would expect the early and late signals to yield nearly equal thresholds (e.g., [Trahiotis et al., 1972](#); [Zwicker and Zwicker, 1984](#); [Yost, 1985](#)).

The role of early and late fringes (late and early signals) is further studied in Experiment 3B by estimating relative weights as a function of time. Because the components of the MBR masker are similar to the signal, on occasion the masker might form a series of tone pips with frequencies similar to those of the signal. Additionally, the “burst” nature of the masker provides a natural segmentation of the stimulus in time. These features, along with the random nature of the masker, allow an estimate of relative weights using a logistic regression to model listeners' responses. The resulting relative weights indicate listeners' dependence on different spectral-temporal regions for detection.

A. Methods

1. Participants

Listeners L10–L15 participated in Experiment 3A and listeners L10–L12 and L16 participated in Experiment 3B. The newly recruited listeners, L12–L16, included three females and two males ranging in age from 19 to 23 with thresholds in quiet of 20 dB HL or less at octave frequencies ranging from 250 to 8000 Hz in both ears except for L12 whose threshold in quiet at 8000 Hz was 25 dB HL in both ears. Listeners were individually seated in a double-walled soundproof booth and were reimbursed for their participation.

Listeners L10–L12 completed Experiment 3A before beginning Experiment 3B.

2. Stimuli and conditions

The stimuli in Experiments 3A and 3B were generated and presented using methods essentially identical to those used in Experiment 2. The average masker level was 63 dB SPL and the stimuli were presented diotically using Sennheiser HD410SL headphones. The signal was a sequence of three 60-ms, 1000-Hz tone pips turned on and off synchronously with the MBR masker bursts. In Experiment 3A a total of 10 conditions were tested. Four of the conditions were similar to those run in Experiment 2, testing fringes of 0, 2, 4, and 8. These conditions will be referred to as the “middle signal” conditions because the signal, when present, was located at the temporal “middle” of the stimulus. In addition, there were three “late signal” conditions in which 2, 4, or 8 fringe bursts preceded, but did not follow, the masker (the final three bursts). Finally, there were three “early signal” conditions in which the masker comprised the first three bursts of the stimulus, followed by a fringe of 2, 4, or 8 bursts.

In Experiment 3B only two conditions were tested, early and late signal conditions with a fringe of four.

3. Procedures and data analysis

In Experiment 3A, listeners initially practiced for 250 trials in two of the ten conditions. Then, for each condition, 100 trials were run during which the signal level was adjusted to an appropriate level for testing. Data were collected separately for the early, middle, and late signal conditions in different orders for different listeners. Within each of the conditions, the order in which the different number of fringes was tested varied across listeners. Listeners completed five 50-trial blocks for each of the different number of fringes in turn, and then ran five 50-trial blocks in the opposite order. One observer demonstrated improved sensitivity for the second block, and for her the data were re-collected.

On each trial, the signal was as likely as not to be present, and if present, the signal level was chosen from a 10-dB range. Listeners indicated their response, no signal or signal, by pressing a keypad, correct-answer feedback was presented after every trial, and summary statistics were provided after each 50-trial block. Thresholds were estimated using the same procedures used for Experiment 2.

For Experiment 3B, L10–L12 had completed Experiment 3A, and so additional practice was not required. For the newly recruited listener L16, approximately one hour of practice was required to achieve stable levels of performance. Listeners L10 and L12 ran the early signal condition first, and L11 and L16 ran the late signal condition first. A total of 40 50-trial blocks were run for each condition. In all other regards the procedures were as in Experiment 3A.

In the initial data analysis for Experiment 3B, thresholds ($d' = 1.5$) were estimated using a weighted linear least-squares fit to psychometric functions relating d' to signal level. Next, using only the no-signal trials, the stimuli were

represented using a vector with seven elements, one for each burst (four fringe bursts, three masker bursts). The entries for each element indicated the masker energy (on a log scale) that fell within $\pm 1/8$ th of an octave surrounding the 1000-Hz signal. The scale set the level of a single tone at 50 dB, and the absence of energy was set to 0 dB. Then, for each condition and each listener, the listeners' responses were regressed against the stimulus vector using a logistic regression² (*glmfit* in MATLAB). In addition to providing relative weights, the regression analysis provided estimates of the standard error of those weights. The relative weights and the standard errors of the estimates were then scaled so that for each listener in each condition, the square root of the sum of the squared relative weights was unity.

B. Results and discussion of experiment 3A

Figure 4 plots thresholds as a function of number of fringe bursts for the early (circle), middle (triangle), and late (square) signal conditions. The data for the different listeners are plotted in different panels, and the bottom right panel shows the averaged data. Error bars indicate 68% confidence intervals as the standard error of threshold estimate, and the standard error of the mean, for the individual and averaged data, respectively. Across the 30 fitted psychometric functions, half accounted for between 75% and 89% of the variance in the corresponding data; the other half of the fitted functions accounted for over 90% of the variance in the corresponding data.

On average, thresholds are lowest for the late signal, highest for the early signal, and intermediate for the middle signal (Fig. 4, bottom right panel). The pattern is observed for four of the six observers, with the results of L13 and L15 indicating deviations from this pattern. A within-subject, repeated measures analysis of variance indicates that the changes in threshold with signal position (averaged across number of fringe bursts) are statistically significant [$F(2,10) = 28.4$, $p < 0.001$]. With regard to individual data sets, it should be noted that L10's threshold in this experiment is lower than expected based on his thresholds in Experiment 2 (unfilled squares in Fig. 3, upper left panel in Fig. 4). This shift might suggest that substantial practice can reduce fringe effects, but the listeners in Group A of Experiment 1 had substantial experience, yet some of those listeners maintained a large effect in that experiment (e.g., L2 and L5), and for most listeners, there was little difference in thresholds, Experiment 2 (4 fringe bursts) versus Experiment 1 (continuous masker).

When considering the pattern of results displayed in Fig. 4, it should be remembered that in the middle-signal condition, a fringe both precedes and follows the masker. Thus, the observation that performance is better for the middle signal conditions than for the early signal conditions (with one exception: L15, fringe = 8), and that performance is better for the late signal conditions than for the no fringe condition, indicates a *benefit* of a preceding fringe, regardless of the presence of a following fringe. In contrast, there is an increase in thresholds associated with a later fringe; for five of the six listeners, performance in the middle signal

FIG. 4. Thresholds as a function of the number of fringe bursts are plotted for conditions in which the fringe precedes (late signal, squares), follows (early signal, circles), or surrounds (middle signal, triangles) the signal epoch. The symbols for the middle signal condition set are shifted slightly to the left to allow clearer viewing of the result. Results are shown individually for the different listeners and for the averaged data in the bottom right hand panel. Error bars indicate the 68% confidence intervals as an indication of the standard error of the threshold estimate for each listener, and the standard error of the mean for the averaged data.

conditions is worse than in the late signal conditions, and for four of the six listeners performance in the early signal condition is worse than in the no fringe condition. This pattern of results indicates a *cost* associated with the presence of energy following the signal. Potentially, this loss of sensitivity reflects an “overwriting” of information in working memory or a form of long-lived backward masking.

In Experiment 3B relative weights are estimated as a function of time in an effort to clarify potential differences in processing/decision rules associated with early and late signal presentations.

C. Results and discussion of experiment 3B

Table I shows the thresholds estimated for the early and late signal conditions from Experiments 3A and 3B. For L10–L12, who participated in both experiments, thresholds estimated here (Experiment 3B) are similar to thresholds in Experiment 3A. In Experiment 3B, on average thresholds in the late signal conditions are 7 dB lower than in the early signal conditions.

Figure 5 plots the relative weights associated with the 1/4-octave-wide bin logarithmically centered at the signal

TABLE I. Thresholds (dB SPL) are shown for the early- and late-signal conditions, fringe = 4, for Experiments 3A and 3B.

	Experiment 3A		Experiment 3B	
	Early signal	Late signal	Early signal	Late signal
L10	62.9	54.4	59.8	54.5
L11	55.9	51.2	57.5	50.6
L12	57.3	52.3	56.1	52.7
L16			63.4	49.7
AVG	58.7	52.6	59.2	51.9
SEM	2.1	0.9	1.6	1.1

frequency as a function of the burst number. The results for different listeners are plotted using different symbols. The upper panel is for the late signal condition (the signal occurs in the final three bursts) and the lower panel is for the early signal condition (the signal occurs in the initial three bursts). Error bars indicate the standard errors of the estimates. As a point of comparison, an ideal energy observer would have positive and equal weights at the time of the signal, and zero-valued weights at the times the signal is not present.

First consider the late signal condition (upper panel of Fig. 5). Across listeners, the pattern is similar to that of the ideal energy observer: relative weights are uniformly large for the final three temporal bursts. In addition, negative weights are observed for the first four bursts, the fringe bursts. This suggests a contrast or cancelation between the early and late portions of the stimulus.

Next consider the early signal condition (lower panel). Here the results diverge from expectations associated with the ideal energy observer; only for L12 are *any* of the first three relative weights more than two standard errors of the

FIG. 5. Relative weights for the three masker bursts and four fringe bursts are plotted for late (upper) and early (lower) signals for each listener. The relative weights are estimated using only no-signal trials, where the energy in the 1/4 octave surrounding the signal frequency forms the dependent variable. Error bars indicate the standard error of the estimate. The data points are jittered slightly to allow clearer viewing.

estimates removed from zero. There is a lack of consistency in the relative weights across listeners. For L11 and L16 the average of the first three relative weights is *smaller* than the average of the last three relative weights; the pattern is the opposite for L10 and L12.

Clearly the pattern of relative weights is vastly different for early and late signal conditions. For the late signal condition, the relative weights suggest a strong dependence on the energy in the frequency region of the signal at temporal epochs during which the signal occurs, and a small, but consistent, dependence in the frequency region of the signal at temporal epochs prior to the signal. For the early signal condition, this result is not observed, and indeed, no consistent pattern emerges across listeners. Nonetheless, the results for the early signal condition are meaningful. These data reject the hypothesis that listeners are simply integrating energy across too many bursts, yielding an increase in the variance of a decision variable. Such a model would be indicated by positive weights beyond the signal offset. The absence of a clear pattern of relative weights in the early signal condition might mean that the model used to evaluate the relative weights, which is based on energy in the region of the signal for each burst, is inappropriate for this condition.

An alternative explanation for the benefit of bursts preceding the masker is that on a substantial number of trials no masker element occurred at the signal frequency just prior to signal onset, thereby providing psychophysical enhancement (e.g., Viemeister, 1980; Wright *et al.*, 1993; Richards and Neff, 2004; Richards *et al.*, 2004; Nelson and Young, 2010). This possibility was evaluated by re-examining the stimuli and responses made by listeners in Experiment3B with the late signal. First, the stimuli were “binned” to include masker elements that fell within 1/4 of an octave of the signal frequency. Then, the stimuli were separated into two groups: those with and without energy at the signal frequency in the burst preceding the signal burst. Here, the “enhancement” trials are those in which there was no energy preceding the signal burst. Then, d' was estimated for the two groups of stimuli and associated responses. Across the four listeners, for two the value of d' was higher for the enhancement subset of trials and for two the opposite result occurred. On average, the value of d' was larger for the “enhanced” trials, but the advantage was only 0.03 units. When the criterion for enhancement required that two bursts prior to the signal had no energy, for three of the four listeners the values of d' for the enhancement trials were larger. Again, however, the difference in values of d' conditional on the presence/absence of energy prior to the signal was very small (0.01 units). Overall, the current data do not support a strong role of psychophysical enhancement for the observed benefit associated with a preview of the random masker in the late signal conditions.

V. SUMMARY AND CONCLUSIONS

The current experiments examine the formation of a signal stream embedded in a random masker. The stimuli have the properties that (a) the added signal could not be detected on the basis of simple changes such as spectral splatter at the

signal onset and offset, (b) the masker fringe could be systematically varied, and (c) the masker provided a natural metric on which relative weights in time could be estimated. The resulting stimuli allow an examination of the role of early and late “fringes” for mechanisms associated with perceptual organization as it emerges over time, which is the dimension across which information is provided to the auditory system.

Consistent with past research (e.g., Yost, 1985), Experiment 1 shows that thresholds for the detection of a sequence of three equal-frequency 60-ms tone pips (or six equal-frequency 30-ms tone pips) added to a noise masker are approximately 3 dB lower when the masker is continuous relative to when it is turned on just before the signal onset and just after the signal offset (pulsed masker). For MBR maskers, however, for some listeners thresholds are much lower for the continuous than the pulsed masker; across listeners the differences in thresholds ranged from 1 to 24 dB. The advantage of the continuous masker held even though the detection of the added signal required sustained vigilance to overcome the uncertainty regarding the time at which the signal would be presented. The observed individual differences and relatively high detection thresholds with the MBR masker suggest a contribution of informational/central masking, probably due to the onset and offset ramps shared by the signal and masker.

The results of Experiment 1 suggest at least two potential explanations. First, listeners detect a “change” associated with the signal onset and/or offset (e.g., as for the BMLD, McFadden, 1966), and second, listeners take advantage of the continuous masker to learn the statistics of the masker and detect a change in the statistics when the signal is added. The former explanation predicts that masker bursts either before the signal onset or after the signal offset would suffice to detect a change (e.g., as for the BMLD, Trahiotis *et al.*, 1972). The latter explanation predicts a temporal direction such that past sounds are used to detect changes at present.

The intent of the second and third experiments was to clarify the fringe effects observed in Experiment 1. The second experiment indicates that for those listeners whose thresholds are high for short fringes, thresholds drop as the number of fringes increased. More relevant to the processes supporting perceptual organization, the third experiment shows that for most comparisons across listeners and conditions, there are a *benefit* of a forward fringe before the masker and a *cost* of a backward fringe following the masker. Unlike the reasonable/reliable patterns of relative weights obtained for signals presented late in the stimulus, when the signal occurred at the stimulus onset the patterns of relative weights are inconsistent across listeners.

To summarize, the experiments indicate that (a) for some listeners a substantial amount of pre-exposure to the ongoing sound is needed for efficient signal extraction, (b) preexposure to an MBR masker aides signal detection, (c) continuing a postexposure to the MBR after the signal presentation increases detection thresholds, and (e) for MBR maskers there is little cost of uncertainty regarding the timing of the signal onset.

ACKNOWLEDGMENTS

This work was supported by NIH NIDCD grants R01 DC02012 and R21 DC010058. We thank Dr. Yi Shen and two anonymous reviewers for helpful comments on earlier drafts.

¹As indicated previously, a programming error with the 60-ms burst-duration stimuli led to a doubling, relative to the 30-ms condition, of both the duration of the masker/signal burst and the time between signal onsets. Therefore the expected time between signals and the standard deviation of the time between signals for 60-ms burst-duration stimuli is twice that for 30-ms burst duration stimuli.

²Data for one of the listeners was used to evaluate whether estimation of relative weights in both time and frequency would yield superior/more accurate fits to the data. The results of those analyses suggested little advantage in modeling the data using a larger number of degrees of freedom (e.g., estimating relative weights for more frequencies, thereby providing a time-by-frequency map of relative weights).

- Bonino, A. Y., and Leibold, L. J. (2008). “The effect of signal-temporal uncertainty on detection in bursts of noise or a random-frequency complex,” *J. Acoust. Soc. Am.* **124**, EL321–EL327.
- Egan, J. P., Greenberg, G. G., and Schulman, A. I. (1961). “Operating characteristics, signal detectability, and the method of free response,” *J. Acoust. Soc. Am.* **33**, 993–1007.
- Huang, R., and Richards, V. M. (2006). “Coherence detection: Effects of frequency, frequency uncertainty, and onset/offset delays,” *J. Acoust. Soc. Am.* **119**, 2298–2304.
- Huang, R., and Richards, V. M. (2008). “Estimates of internal templates for the detection of sequential tonal patterns,” *J. Acoust. Soc. Am.* **124**, 3831–3840.
- Kidd, G. Jr., Mason, C., and Dai, H. (1995). “Discriminating coherence in spectro-temporal patterns,” *J. Acoust. Soc. Am.* **97**, 3782–3790.
- McFadden, D. (1966). “Masking-level differences with continuous and burst masking noise,” *J. Acoust. Soc. Am.* **40**, 1414–1419.
- Neff, D. L., and Green, D. M. (1987). “Masking produced by spectral uncertainty with multicomponent maskers,” *Percept. Psychophys.* **41**, 409–415.
- Nelson, P. A., and Young, E. D. (2010). “Neural correlates of context-dependent perceptual enhancement effects in Inferior Colliculus,” *J. Neurosci.* **30**, 6577–6587.
- Reeves, A., and Scharf, B. (2010). “Auditory frequency focusing is very rapid,” *J. Acoust. Soc. Am.* **128**, 795–803.
- Richards, V. M., and Neff, D. L. (2004). “Cuing effects for informational masking,” *J. Acoust. Soc. Am.* **115**, 289–230.
- Richards, V. M., Huang, R., and Kidd, G., Jr. (2004). “Masker-first advantage for cues in informational masking,” *J. Acoust. Soc. Am.* **116**, 2278–2288.
- Scharf, B., Reeves, A., and Giovanetti, H. (2008). “Role of attention in overshoot: Frequency certainty vs uncertainty,” *J. Acoust. Soc. Am.* **123**, 1555–1561.
- Shub, D. E., and Richards, V. M. (2009). “Psychophysical spectro-temporal receptive fields in an auditory task,” *Hear. Res.* **251**, 1–9.
- Trahiotis, C., Dolan, T. R., and Miller, T. H. (1972). “Effect of ‘backward’ masker fringe on the detectability of pulsed diotic and dichotic tonal signals,” *Percept. Psychophys.* **12**, 335–338.
- Viemeister, N. F. (1980). “Adaptation of masking,” in *Psychophysical, Physiological, and Behavioral Studies in Hearing*, edited by G. van den Brink and F. S. Bilsen (Delft University Press, Delft, The Netherlands), pp. 190–198.
- Watson, C. S., and Nichols, T. L. (1976). “Detectability of auditory signals presented without defined observation intervals,” *J. Acoust. Soc. Am.* **59**, 655–668.
- Wright, B. A., McFadden, D., and Chaplin, C. A. (1993). “Adaptation of suppression as an explanation of enhancement effects,” *J. Acoust. Soc. Am.* **94**, 72–82.
- Yost, W. A. (1985). “Prior stimulation and the masking-level difference,” *J. Acoust. Soc. Am.* **78**, 901–907.
- Zwicker, U. T., and Zwicker, E. E. (1984). “Binaural masking level difference as a function of mask and test-signal duration,” *Hear. Res.* **13**, 215–219.