Increment and decrement detection in sinusoids as a measure of temporal resolution

Andrew J. Oxenham[a)]
Institute for Perception Research (IPO), P.O. Box 513, 5600 MB Eindhoven, The Netherlands

(Received 17 December 1996; revised 30 May 1997; accepted 3 June 1997)

Measuring thresholds for the detection of brief decrements in the level of a sinusoid is an established method of estimating auditory temporal resolution. Generally, a background noise is added to the stimulus to avoid the detection of the ''spectral splatter'' introduced by the decrement. Results are often described in terms of a temporal-window model, comprising a band-pass filter, a compressive nonlinearity, a sliding temporal integrator, and a decision device. In this study, thresholds for increments, as well as decrements, in the level of a 55 dB SPL, 4-kHz sinusoidal pedestal were measured as function of increment and decrement duration in the presence of a broadband background noise ranging in spectrum level from -20 to $+20$ dB SPL. Thresholds were also measured using a 55-dB, 8-kHz pedestal in the absence of background noise. Thresholds for decrements, in terms of the dB change in level (ΔL), were found to be more dependent on duration than those for increments. Also, performance was found to be dependent on background-noise level over most levels tested. Neither finding is consistent with the predictions of the temporal-window model or other similar models of temporal resolution. The difference between increment and decrement detection was more successfully simulated by using a decision criterion based on the maximum *slope* of the temporal-window output. *© 1997 Acoustical Society of America.* [S0001-4966(97)05909-2]

PACS numbers: 43.66.Mk, 43.66.Dc, 43.66.Ba [JWH]

INTRODUCTION

Temporal resolution in the auditory system is often described in terms of the ability to detect brief fluctuations in the level of a stimulus. Many models of temporal resolution assume that the ability to detect such fluctuations is limited by a smoothing process in the auditory system. This process has often been incorporated within a more complete model of temporal resolution comprising a band-pass filter, a rectifier and power-law nonlinearity, a smoothing device (implemented as a low-pass filter or a sliding temporal integrator), and a decision device (Buunen and van Valkenburg, 1979; Viemeister, 1979; Buus and Florentine, 1985; Forrest and Green, 1987; Moore *et al.*, 1988; Plack and Moore, 1991).

Recently, measuring thresholds for brief decrements in the level of a sinusoidal pedestal has become a popular method of estimating temporal resolution, as a function of pedestal frequency (Moore *et al.*, 1993b) and level (Peters *et al.*, 1995; Moore *et al.*, 1996). In these studies, the results from the experiments have been used to derive the parameters of the smoothing device (implemented as a sliding temporal integrator or temporal window) and the subsequent decision device. It is assumed that detection of a decrement occurs when the decrease in the output level of the temporal window equals or exceeds a certain criterion level (in dB), referred to as ΔO. A stated assumption of this model is that increments are detected in essentially the same manner as decrements, namely when the increase in output of the temporal window exceeds the same criterion level (ΔO). However, this assumption has not been directly tested. In all previous temporal-window studies of increment and decrement detection (Moore *et al.*, 1993b, 1996; Peters *et al.*, 1995) thresholds for only one increment duration were measured, so it is not clear whether the same model parameters that describe decrement detection can also account for increment detection as a function of duration. Conversely, it is not clear whether an experiment measuring increment detection as a function of duration would produce the same model parameters as a decrement-detection experiment.

Studies of increment and decrement detection in broadband noise have found that very brief decrements, when measured in terms of the dB change in level (ΔL), are generally less detectable than brief increments of the same duration (Irwin and Purdy, 1982; Forrest and Green, 1987). Irwin and Purdy (1982) suggested that the data could be explained if neural adaptation was taken into account (Smith, 1979), although this approach seems not to have been pursued since. One previous study has investigated the detection of increments and decrements in the level of a 1-kHz sinusoid (de Boer, 1986). There it was also found that decrements were less detectable than increments at the shortest durations. No quantitative modeling was attempted, however. Also, the increments and decrements used by de Boer (1986) were gated on and off instantaneously. While the pedestal was presented in a background notched noise, the large bandwidth of the notch (2 octaves) makes it difficult to rule out cues associated with ''spectral splatter'' due to the onset and offset of the signal (e.g., Leshowitz and Wightman, 1971).

Macmillan (1971) provided data which indicate that the detection of brief increments and decrements may be mediated by a ''change detector,'' rather than an energy detector.

[a)] Electronic mail: oxenham@ipo.tue.nl

He reached this conclusion based on the fact that the discriminability of increments and decrements is less than would be predicted based on their respective detectability. While some of his results may have been due to spectral artifacts (Leshowitz and Wightman, 1972), a further study, controlling for spectral effects, confirmed the trend of the initial findings (Macmillan, 1973). Such a change detector is inconsistent with the assumptions of the temporal-window model, and other models of temporal resolution (e.g., Buus and Florentine, 1985), which assume that detection is due to a maximum in the peak or dip of a smoothed representation of stimuli. However, another decision device, proposed by Viemeister (1979) within a model accounting for amplitude modulation detection, may be compatible with the notion of a change detector: in this model, the standard deviation of the output, regardless of the direction of the change, is used as the decision device.

One purpose of this study is to compare the detection of increments and decrements directly by measuring thresholds for both as a function of duration in the same listeners. The results are used to evaluate the ability of current models of temporal resolution to account for the data. A high pedestal frequency (4 kHz) was chosen in order to allow the measurement of relatively short increments and decrements without the bandwidth of the signal exceeding the bandwidth of the auditory filter at 4 kHz.

A second question studied here is the effect of adding a broadband background noise to the pedestal. This has been done in most previous studies in order to reduce the possibility that listeners detect the spectral splatter associated with the onset and offset of the signal. It is generally assumed that the level of the background noise can be set sufficiently high to mask any off-frequency signal energy, but not so high as to influence on-frequency performance. Purely in terms of the overall noise energy within and beyond the auditory filter centered at the signal frequency, this assumption is reasonable. However, other aspects of the background noise, such as its inherent fluctuations, may play a significant role, either within or across different frequency channels. The effect of adding noise is investigated here by measuring performance over a wide range of background-noise levels. If the assumptions of the temporal-window model are correct, performance should remain constant over a considerable portion of the range of noise levels tested.

I. EXPERIMENT 1. INCREMENT AND DECREMENT DETECTION IN BROADBAND BACKGROUND NOISE

A. Stimuli

Thresholds for detecting increments and decrements in the level of a 4-kHz sinusoidal pedestal were measured as a function of signal (increment or decrement) duration. All stimuli were generated and controlled digitally on a Silicon Graphics Indigo workstation and were played out via the built-in D/A converter at a sampling rate of 32 kHz. The pedestal had a total duration of 500 ms, gated with 50-ms raised-cosine ramps, and was presented three times in each trial, separated by 200-ms interstimulus intervals. The level of the pedestal was always 55 dB SPL. The signal was added to one of the three pedestals at random and was temporally centered in the pedestal. The signal was either added in phase with the pedestal, to produce an increment in the overall level, or in antiphase, to produce a decrement. The signal had a steady-state duration of 2, 6, 14, and 198 ms, and was gated with 2-ms raised-cosine ramps, producing increments and decrements with half-amplitude durations of 4, 8, 16, and 200 ms.

The broadband (0–15 kHz) Gaussian background noise was generated at the beginning of each run and was stored in a 4-s circular buffer. Three 700-ms segments of the noise were chosen at random (with replacement) and added to each of the three pedestals, with each noise onset 200 ms before each pedestal onset. The three segments were then concatenated to produce a single trial of 2.1-s duration. As the noise was broadband and was gated abruptly, there was no discontinuity between the 700-ms segments and so the noise provided a continuous (2.1-s) background during each trial. Thresholds for the increments and decrements were measured for noise spectrum levels of -20, -10, 0, 10, and 20 dB SPL and were also measured in the absence of noise.

The stimuli from the Silicon Graphics D/A converter were passed through a Tucker-Davies programmable attenuator (TDT PA4) and headphone buffer (TDT HB6) before being presented to one ear of the listener via a Beyer DT 990 headset.

B. Procedure

Thresholds were measured using a three-alternative forced-choice method with a three-down one-up adaptive procedure that tracks the 79.4% correct point on the psychometric function. Each listener was tested in 2-h sessions, divided into four blocks of runs of approximately 20 min each. Within each block only increments or only decrements were tested.

At the beginning of each run the level of the signal was set to the same level as the pedestal. This resulted in a complete gap in the case of decrements and a 6-dB increase in level in the case of increments. The task of the listener was to select which of the three intervals contained the increment or decrement. Responses were made via a computer keyboard, and feedback was provided via a computer monitor. In each run the signal level was decreased after three consecutive correct responses and was increased after each incorrect response. A change from increasing to decreasing signal level, or vice versa, defines a reversal. Initially the signal step size was 8 dB. This value was halved after every second reversal until it reached its minimum value of 2 dB, after which it remained constant. A run was terminated after 14 reversals and the threshold was defined as the median signal level at the last 10 reversals. For the decrements, the signal level was not permitted to exceed 0 dB, relative to the level of the pedestal. If the adaptive procedure required a higher level, the level was maintained at 0 dB and testing continued. If 30 trials within a run required a nominal level greater than 0 dB, the run was terminated. Any runs which included nominal levels of more than 0 dB within the test phase (last 10 reversals) were discarded. This procedure was designed to allow listeners some practice with conditions in which they initially could not detect the decrement.

Three threshold estimates were obtained for each listener and condition, and transformations of the mean of the three estimates are reported here. In the case of decrements, if more than one of the three estimates was discarded (see above), that condition was deemed not to be detectable. If only one of the three estimates had been discarded, a further estimate was made. If this was also discarded, the condition was deemed not to be detectable. If not, the mean of the three remaining conditions was recorded. In practice, conditions where the decrement was initially inaudible generally remained so. Thus a further estimate was necessary for only one data point of one listener.

Listeners were tested individually in a single-walled sound-attenuating booth, situated in a sound-attenuating room.

C. Subjects

Four normally hearing listeners participated as subjects. One was the author. The other three were male students who were paid an hourly wage for their participation. The ages of the listeners ranged from 22 to 27 years. None of the three students had previous experience in psychoacoustic tasks, and all four listeners were given 2-h practice, divided equally between increments and decrements, before data were collected. All listeners had audiometric thresholds of 15 dB HL or less at octave frequencies between 250 and 8000 Hz in the test ear. No practice effects were noted during the course of the experiment.

D. Results

Results were reasonably consistent both within and across listeners. Typically, standard deviations across the three threshold estimates for each condition and listener were less than 2 dB. Figure 1 shows the individual results from the four listeners. For each data point, the mean signal level at threshold, together with ±1 standard error of the mean, was calculated and then transformed into the level of the increment or decrement, ΔL, defined as $20 \log[(p+\Delta p)/p]$, where p is the sound pressure of the pedestal, and Δp and $-\Delta p$ are the sound pressures of the signal for increments and decrements, respectively. As performance was fairly similar across listeners, mean thresholds across listeners were also calculated and then transformed, as described above. These mean data are shown in Fig. 2.

Consider first the effect of background noise in Fig. 2. For the decrements, thresholds seem to be dependent on background-noise level across the entire range of levels tested, especially at the shorter durations. For the increments, the effect seems less pronounced at the lower noise levels. However, increment thresholds for noise levels above −10-dB spectrum level are also strongly dependent on background-noise level. A comparison of increment with decrement thresholds in units of ΔL shows that at the longest signal duration (200 ms), thresholds seem fairly symmetric. This is consistent with the assumption of the temporal-window model that the criterion level, ΔO, is the same whether the task is to detect an increment or a decrement. At shorter durations, performance in these units is asymmetric, both at low and high noise levels. For instance, even with no noise, the mean increment threshold at 4 ms is 2.2 dB, while the mean decrement threshold is −7.0 dB. At the highest noise level tested (20 dB), only the 200-ms decrement was detectable at all.

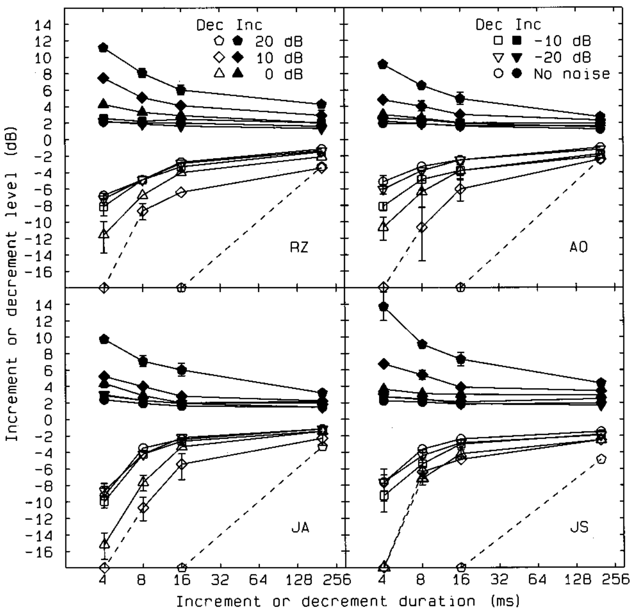

FIG. 1. Individual thresholds for detecting increments (filled symbols) and decrements (open symbols) in the level of a 55-dB, 4-kHz pedestal, in terms of ΔL. Error bars represent ±1 standard error of the mean and are omitted if smaller than the symbol. Different symbols represent different background-noise spectrum levels, as shown in the upper panels. Symbols lying on the abscissa and joined by dashed lines represent conditions where no threshold could be measured.

The units of ΔL were chosen to be consistent with previous temporal-window studies (e.g., Moore *et al.*, 1993b;

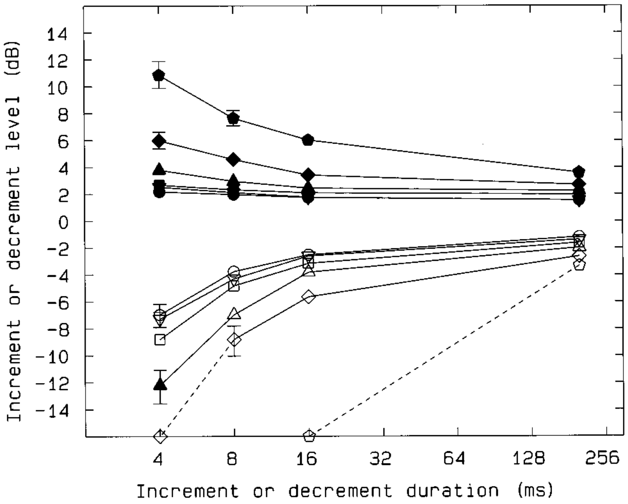

FIG. 2. Mean thresholds for detecting increments and decrements. Different symbols represent different background-noise spectrum levels, as shown in Fig. 1. Error bars represent ±1 standard error of the mean across listeners. The single filled symbol in the decrement condition represents the mean of three, rather than four, listeners (a threshold for JS could not be measured).

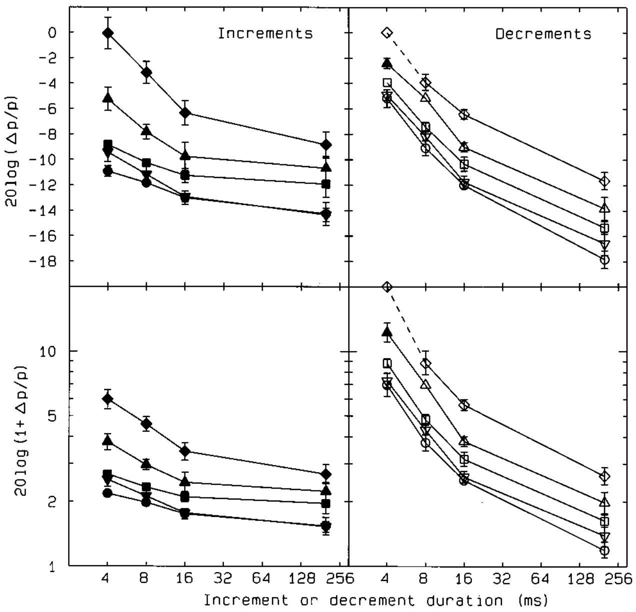

FIG. 3. Mean data from Fig. 2, replotted using two alternative level measures. The upper two panels represent increment and decrement thresholds in terms of signal level, relative to pedestal level [$20\log(\Delta p/p)$], and the lower two panels plot ΔL, or $20\log(1+\Delta p/p)$, on a logarithmic scale. Error bars represent ± 1 standard error of the mean.

Peters *et al.*, 1995). As pointed out by, for instance, Forrest and Green (1987), these units are highly compressive at small values. Thus small but significant differences may not be readily visible. Also, a different choice of units can affect any apparent asymmetries in the data (Forrest and Green, 1987). For this reason, the mean data from Fig. 2 are replotted in Fig. 3, using two alternative signal measures. Data from the highest noise level are omitted in this figure, as only one of the decrement conditions was measurable. The upper panels of Fig. 3 show the results in terms of signal level relative to pedestal level; the lower panels show the results in units of ΔL, but plotted on a logarithmic scale, as suggested by Buus and Florentine (1991) for level discrimination.

The asymmetry between increments and decrements at short signal durations remains apparent in Fig. 3, especially at the lower noise levels. Interestingly, in these plots the effect of the background noise is apparent for both increments and decrements, even at the longest signal duration. Another aspect, not readily apparent in Fig. 2, is that there is a tendency for thresholds for the longest decrements to be *lower* than for the longest increments, thus reversing the asymmetry at shorter durations. A similar trend was noted by Forrest and Green (1987) using broadband noise, but not by Irwin and Purdy (1982). De Boer (1986), in his study using sinusoidal pedestals, plotted his results using the same linear scale of ΔL used in Fig. 2, making comparisons between increments and decrements at long durations difficult. Nevertheless, no asymmetry at long durations is readily apparent in his data. As the plots of Fig. 3 seem to yield more information, all further figures are shown in units of ΔL on a logarithmic scale, as in the lower panels of Fig. 3.

In summary, using all three types of units, brief decrements are less detectable than brief increments of the same duration. This conclusion is also consistent with the studies of Irwin and Purdy (1982) and Forrest and Green (1987), both of which used broadband noise as a pedestal and signal. The fact that increment thresholds at low noise levels are much less dependent on duration is also consistent with the results of Oxenham (1997). Presumably, at the highest noise levels, the noise becomes the dominant masker and temporal integration is observed which is similar to that for a signal in broadband noise alone (e.g., Gerken *et al.*, 1990). This was confirmed for listener AO: detection threshold for the 4-ms signal in the 20-dB noise alone was 59.5 dB SPL, which matches closely with the threshold signal level of 60.3 dB in the increment-detection condition with the 20-dB background noise. The pedestal itself, however, was clearly audible in all conditions.

E. Discussion

The large effect of background noise cannot be accounted for by a single-channel model, such as the temporal-window model, in which only the overall noise power within the channel is taken into account. For instance, the 10-dB noise, which renders the shortest decrement undetectable for all listeners, has an effective level of only 36 dB SPL within the equivalent rectangular bandwidth (ERB) of an auditory filter centered at 4 kHz (Glasberg and Moore, 1990). This is nearly 20 dB lower than the pedestal level (55 dB SPL), and so is not predicted to affect thresholds significantly.

One possible explanation for the effect of the background noise is that performance was based on the detection of spectral splatter at all but the highest noise levels. In this case, performance would be expected to deteriorate with increasing noise level, as the splatter would be increasingly masked.

The detectability of spectral splatter in the absence of noise can be assessed as follows. If the pedestal can be represented as a line spectrum, the phase of the signal, relative to the pedestal only affects the amplitude spectrum at the pedestal frequency (Leshowitz and Wightman, 1971). Thus all off-frequency energy due to the signal onset and offset is independent of whether the signal phase is 0 or 180 degrees. This means that the off-frequency energy for a given signal level is the same for both increments and decrements. This in turn leads to the conclusion that, even with no noise present, increment detection was *not* achieved through detection of spectral splatter. If this were the case, then increment and decrement thresholds for the shortest duration should be equal in terms of signal level (Fig. 3, upper panels). As they are not equal, it is possible to conclude that *increment* detection is based on cues other than spectral splatter.

The shortest signal employed in this experiment had a total duration of 6 ms and a 3-dB bandwidth of about 210 Hz. This is much less than the estimated ERB of the auditory filter centered at 4 kHz (456 Hz). The first side lobes of the signal's power spectrum are 17.5 dB lower in level than the main lobe and are also spaced less than 1 ERB from the main lobe, at 3660 and 4340 Hz. No side lobes beyond the region of 3–5 kHz are at a level greater than -50 dB relative to the main lobe. It might therefore be expected that spectral splatter for *decrement* detection was also not detectable in the absence of noise. Informal simulations using a gammatone filterbank (Patterson *et al.*, 1992) supported this conjecture: it was found that for filters where the output would have been above absolute threshold, the transient increase in response due to the onset and offset of the decrement was generally smaller than the minimum detectable increment in quiet. The question of the detectability of splatter is examined more empirically in experiment 3.

Whether or not thresholds for decrement detection in the absence of noise are mediated by spectral splatter, the effect of the background noise at higher levels remains to be explained. One possibility relates to the upward spread of excitation of the pedestal and signal (Zwicker, 1956). As the noise level increases, more of the pedestal and signal's excitation pattern would fall beneath masked threshold, resulting in a deterioration in performance. Moore *et al.* (1996) tested this directly in the context of increment and decrement detection. They found that performance at 4 kHz in the presence of a low-pass background noise with a cutoff frequency of 5 kHz was superior than performance in the presence of a broadband noise with the same spectrum level. However, the lowest pedestal level tested by Moore *et al.* (1996) was 70 dB SPL. This is somewhat higher than the 55 dB used here, and the benefit of the upward spread of excitation would be expected to diminish with decreasing pedestal level. Another possibility is that the inherent fluctuations of the noise directly mask, or distract from, the fluctuations introduced by the signal. In other words, the modulation frequencies introduced by the signal are masked by the modulation frequency spectrum of the noise. Much evidence has been presented recently for the role of masking in the modulation domain, both within channel (Bacon and Grantham, 1989; Houtgast, 1989; Dau *et al.*, 1997a,b; Oxenham, 1997) and across channels (Yost and Sheft, 1989; Mendoza *et al.*, 1995). The effect of the background noise was studied further in experiment 2 by measuring thresholds in noise with various spectral characteristics.

II. EXPERIMENT 2. EFFECTS OF BACKGROUND-NOISE SPECTRUM

This experiment investigated the effect of changing the spectral composition of the background noise. If the effect of the noise can be attributed to the masking of the pedestal's spread of excitation, then using a low-pass noise with a cutoff frequency near the pedestal frequency should improve performance relative to the broadband condition. Using a high-pass noise should then produce a smaller (if any) improvement. On the other hand, if on-frequency masker energy (or fluctuations) are responsible for the effect of the background noise, then using a noise with a spectral notch centered around the pedestal frequency should improve performance. Performance in these three conditions (low-pass, high-pass, and notched noise) were measured for both increment and decrement detection.

A. Method

Thresholds for increments and decrements in the 55-dB, 4-kHz pedestal were measured for half-amplitude signal durations of 4, 8, and 16 ms. The spectrum level of the background noise was held constant at 10 dB SPL and three different conditions were tested: low-pass noise with a cutoff frequency of 4400 Hz; high-pass noise with a cutoff frequency of 3600 Hz; and notched noise with cutoff frequencies of 3600 and 4400 Hz. The noises were generated by taking the Fourier transform of a 4-s broadband Gaussian noise buffer and setting the amplitudes of the spectral components outside the desired passband to zero. As the noise was bandlimited, concatenating two independent noise samples from the buffer, as was done in experiment 1, would have led to audible clicks at the interval boundaries. To avoid this, the noise was gated with 0.5-ms raised-cosine ramps. These were long enough to make the splatter inaudible, but not so long as to introduce audible gaps between the intervals. Thus as in experiment 1, the percept was of a 2.1-s noise present throughout a single trial. All other aspects of the stimuli and procedure were the same as in experiment 1.

Three of the four listeners from experiment 1 participated in experiment 2; listener JA was not available.

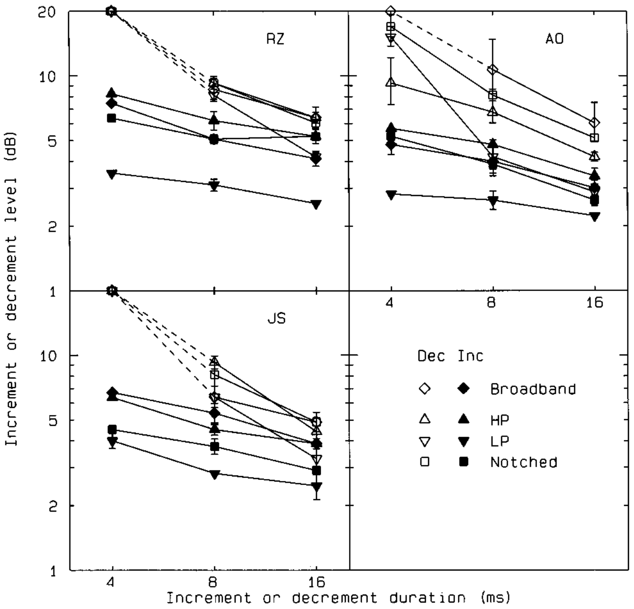

FIG. 4. Individual thresholds, in terms of ΔL, for increment and decrement detection in the presence of high-pass (upward-pointing triangles), low-pass (downward-pointing triangles), and notched noise (squares) at a spectrum level of 10 dB. Thresholds in broadband noise (diamonds) are replotted from Fig. 1. Error bars represent ±1 standard error of the mean.

B. Results

Results from the three listeners are shown in Fig. 4. Data from the broadband condition (diamonds) are taken from Fig. 1. If the effect of the noise were due to the masking of the upward spread of excitation, performance should be better in the low-pass noise condition than in the other noise conditions. This can be seen clearly in the increment-detection data from all three listeners: the down-pointing triangles tend to be lowest. The data from the decrement-detection conditions are less clear cut. Relative to thresholds in the broadband noise, listeners RZ and JS show an im-

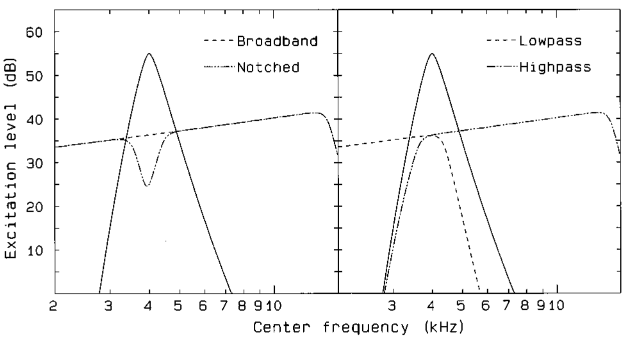

FIG. 5. Excitation patterns for the pedestal (solid curves) and the different background noises used in experiments 1 and 2.

provement in the presence of the low-pass noise only at the longest decrement duration (16 ms), while AO seems to show some improvement at all three durations. All thresholds in the presence of the low-pass noise are, however, still higher than those found with no noise present. It is in principle possible that the low-pass noise still produces some upward spread of masking. This possibility was evaluated using the excitation pattern model, proposed by Glasberg and Moore (1990). Figure 5 shows the excitation patterns of the different maskers and the pedestal. Assuming that this is a reasonable representation of the individual maskers, it seems that the upward spread of excitation due to the low-pass noise should not have a large effect on thresholds. Thus the residual effect of the low-pass masker seems not to be due to its masking of the upward spread of excitation.

The difference between thresholds in broadband noise and in the other two conditions (high-pass noise and notched noise) is dependent on the listener. For AO, decrement detection seems somewhat improved by the two new conditions, while increment detection is relatively unaffected. Conversely, listener JS shows a possible improvement due to the notched noise in the increment conditions, but shows no clear improvement in the decrement conditions. For listener RZ, performance in neither increment nor decrement conditions is improved by the notched or the high-pass noise, relative to the broadband condition.

According to the excitation pattern model, the excitation produced by the notched noise at the pedestal frequency is about 10 dB lower than that produced by the broadband noise (see Fig. 5). Therefore, if performance were limited by on-frequency noise energy, performance in the presence of notched noise should resemble that found for the 0-dB spectrum level broadband masker (Fig. 1). Instead, thresholds are generally higher, indicating that the effects of the noise are probably not solely due to on-frequency energy either.

The fact that, for listeners RZ and JS, none of the conditions substantially improved decrement detection suggests that the noise is playing a masking role which cannot be accounted for purely in terms of its spectral characteristics. It may be that the noise fluctuations in off-frequency channels mask the fluctuations introduced by the decrement, in a manner similar to that observed in modulation detection interference (MDI) (e.g., Yost and Sheft, 1989). Whatever the precise mechanisms, it is clear that a single-channel model of temporal resolution, such as the temporal-window model, cannot account for the results.

In summary, the effects of the noise are not clearly understood and cannot be predicted by current models of auditory processing. Therefore, including a background noise when estimating temporal resolution poses the problem of which the ''correct'' level to use should be. It would clearly be preferable to dispense with the noise altogether. As mentioned earlier, it is likely that at 4 kHz spectral splatter was not audible, even in the absence of noise. The following experiment tests this conjecture.

III. EXPERIMENT 3. EFFECTS OF SIGNAL PHASE

In this experiment, the detectability of spectral splatter in the absence of noise was tested. As stated earlier, the detection of spectral splatter should be independent of the phase relationship between the pedestal and the signal. This is because the pedestal can be represented by a line spectrum, meaning that the pedestal and signal only interact at their common center frequency. The assumption that the detectability of splatter is independent of signal phase is tested empirically in the Appendix and is found to hold for the three listeners tested there.

The present experiment compared performance with the signal added in antiphase (180 degrees; as for the decrements in experiment 1) with performance when the signal was added with a phase shift of 120 degrees. When the signal and pedestal are at the same level, adding an antiphasic signal results in a complete gap, while adding a signal at 120 degrees results in no change of level, except during the onset and offset portions of the signal. At lower signal levels, the difference between the two conditions diminishes, but the antiphasic signal continues to produce the greater decrement. Nevertheless, both signals produce the same amount of off-frequency energy (see above and the Appendix).[1] Thus if detection is mediated by the largest temporal ''dip'' in an on-frequency filter, performance should be worse for the 120-degree condition than for the antiphasic (180-degree) condition. If detection is mediated by spectral splatter, both conditions should produce the same threshold signal level. Only the shortest signal duration was tested, as spectral splatter is most likely to play a role at that duration.

A. Method

Thresholds for a 4-kHz, 4-ms (half-amplitude duration) signal, temporally centered in a 4-kHz, 500-ms pedestal, were measured in the absence of background noise. The phase of the signal was either 120 or 180 degrees, relative to the pedestal. The 180-degree condition was a replication of the no-noise decrement-detection task of experiment 1. For both phases, the signal level was not permitted to exceed 0 dB, relative to the pedestal level of 55 dB SPL. All four listeners from experiment 1 participated. After the initial results were collected, the experiment was repeated using three new normally hearing listeners (for reasons discussed below) with ages of 26, 30, and 31 years.

TABLE I. Signal thresholds (dB), relative to the pedestal level, for the detection of a signal with a phase of either 120 or 180 degrees, relative to the pedestal. Estimated standard deviations are given in parentheses.

Listener	180 degrees	120 degrees
AO	−7.3 (1.2)	−3.0 (1.0)
JS	−4.7 (0.6)	−2.3 (0.6)
JA	−3.7 (0.6)	−2.0 (0.0)
RZ	−5.3 (0.6)	−7.3 (0.6)
JB	−7.0 (1.0)	−4.7 (1.2)
TD	−13.3 (1.5)	−3.7 (0.6)
SP	−9.0 (1.0)	−3.0 (1.0)

B. Results

Signal levels at threshold, relative to the pedestal level (as in the upper panels of Fig. 3), are given in Table I for the 120- and 180-degree conditions. For three of the four original listeners (AO, JS, and JA), thresholds for the 120-degree condition are higher than for the 180-degree (decrement) condition. This difference is significant for all three listeners $[t(4)>4.9;\ p<0.01]$, and suggests that spectral splatter was not detected in the decrement conditions of experiment 1. However, for listener RZ, thresholds for the 120-degree condition are *lower* by 2 dB. This difference in the opposite direction is also significant $[t(4)=4.25;\ p<0.05]$. This exception is puzzling. If listener RZ could detect the spectral splatter, both conditions should be equally detectable. However, there seems to be no ready explanation for why the 120 degree should be more detectable. Also, RZ seems much more sensitive to detection in the 120-degree condition than the other three listeners. This unexpected result led to the experiment being repeated using three new listeners, all of whom had extensive experience in psychoacoustic tasks. Their results are presented in the lower three rows of Table I. All show the same (significant) trend as three of the four original listeners: the 180-degree condition produces lower thresholds than the 120-degree condition. Listener TD was considerably more sensitive than any other listener in the 180-degree condition. Note also that while there were no measurable practice effects for the listeners of experiment 1, the more experienced listeners (AO, JB, TD, and SP) generally achieved lower thresholds in the 180-degree condition than the less experienced listeners (JA, JS, and RZ).

The results indicate that, for three of the four original listeners and for all three new listeners, spectral splatter was probably not detectable for the stimuli of experiment 1, even in the absence of noise. However, the difference in thresholds between the two conditions was generally small, and one listener (RZ) showed the opposite effect. Auditory filters are thought to widen with increasing center frequency. Therefore, spectral splatter should become less detectable with increasing pedestal frequency. In the following experiment increment and decrement thresholds were measured using a pedestal and signal frequency of 8 kHz. At this frequency, it was possible to omit the background noise without any risk of introducing audible spectral splatter. Temporal resolution is thought to be approximately independent of center frequency, at least above about 1 kHz (Moore *et al.*, 1993b). Thus it was expected that the results at 8 kHz should be very similar to those at 4 kHz.

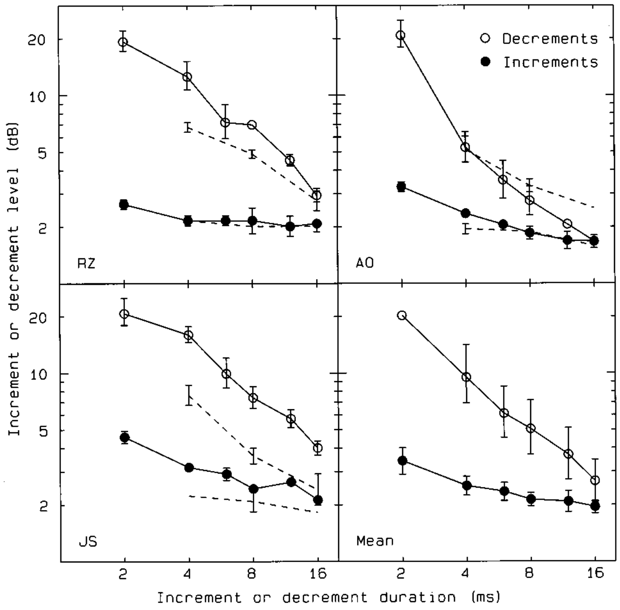

FIG. 6. Individual and mean thresholds for increment and decrement detection with an 8-kHz, 55 dB SPL sinusoidal pedestal. Dashed lines show thresholds with the 4-kHz sinusoidal pedestal, replotted from Fig. 1. Error bars represent ±1 standard error of the mean.

IV. EXPERIMENT 4. INCREMENT AND DECREMENT DETECTION AT 8 kHz

A. Method

Thresholds for increments and decrements were measured in the absence of noise at a frequency of 8 kHz. Again, the stimuli and procedure were the same as in experiment 1, with the following exceptions: thresholds were measured for half-amplitude signal durations of 2, 4, 6, 8, 12, and 16 ms; the initial step size of the adaptive procedure was 4 dB and the minimum step size was 1 dB. The minimum step size was reduced in order to allow more accurate estimates of decrement thresholds at very short signal durations. Three listeners (AO, JS, and RZ) participated in the experiment.

B. Results

Individual and mean thresholds, in terms of ΔL, are shown in Fig. 6. Dashed lines show the corresponding thresholds at 4 kHz. As with the results at 4 kHz, there is a large asymmetry between increment and decrement detection. Again, the results of AO are somewhat different from those of JS and RZ. While AO shows little or no effect of pedestal frequency, JS and RZ generally require larger increments and decrements for detection at 8 kHz than at 4 kHz. The reason for this is not clear; few other studies of temporal resolution have included frequencies as high as 8 kHz. The deterioration may be related to the deterioration of intensity discrimination observed at very high frequencies if slow amplitude modulation is used as a measure (Riesz, 1928): an increase in the intensity difference limen would result in a roughly parallel upward shift in thresholds, as is observed for

JS and RZ. The following section examines the results both from 4 and 8 kHz in light of current models of temporal resolution.

V. PREDICTIONS USING THE TEMPORAL-WINDOW MODEL

A. Model structure

The temporal-window model used here is very similar to that used in many previous studies (Moore *et al.*, 1993b, 1996; Oxenham and Moore, 1994; Peters *et al.*, 1995). It is assumed that stimuli are band-pass filtered (simulating the auditory filters), rectified, compressed, and passed through a sliding temporal integrator (temporal window). A decrement or increment in level is ''detected'' if the decrease or increase in the output of the temporal window reaches a certain criterion value, ΔO (in dB).

All stimuli were represented by their temporal envelopes. When included, the background noise was represented by a flat temporal envelope with a level equal to the effective level of the noise within the auditory filter centered at the pedestal frequency. As in previous studies, the noise and pedestal were assumed to add in a way equivalent to quadrature phase. Initially, the stimuli were compressed by raising the envelopes to the power n, where $n=0.7$, as in Peters *et al.* (1995) and Moore *et al.* (1996).

The shape of the temporal window was defined by a pair of exponential functions, described by

$$W(t)=\exp(t/T_b), \quad t\leqslant 0 \tag{1}$$

and

$$W(t)=\exp(-t/T_a), \quad t>0, \tag{2}$$

where t is time, and T_b and T_a are the time constants determining the sharpness of the function for times before and after the peak, respectively (all in ms). For convenience, the equivalent rectangular duration (ERD) of the function is defined as T_b+T_a, although this is strictly only true for the case where $n=2$. The value of T_b was assumed to be 1.5 times the values of T_a, reflecting the fact that the decay of backward masking is steeper than that of forward masking (Oxenham and Moore, 1994). Thus the time constants provide the model with one free parameter, the ERD.

The second free parameter is the value of the decision criterion, ΔO which is defined as 10 log(S/D) for decrements and as 10 log(P/S) for increments, where D is the minimum value at the output of the temporal integrator in response to a decrement, P is the maximum response to an increment, and S is the response during the steady-state portion of the stimulus (pedestal alone). The units (10 log) were chosen to conform with previous studies (e.g., Moore *et al.*, 1993b). The best-fitting values for ΔO and the ERD were estimated from a given data set as described in Moore *et al.* (1993b) and Peters *et al.* (1995).

B. Model predictions in the absence of background noise

The results from experiment 3 indicate that at 4 kHz, spectral splatter was probably not audible in the absence of background noise, for at least three of the four listeners. For this reason, only the data from the no-noise conditions were examined initially. Data from increment- and decrement-detection conditions were fitted separately. As in the study of Moore *et al.* (1993b), the longest-duration signals (in this case 200 ms) were not included in the fit, as detection in these conditions may be enhanced by a ''multiple-looks'' strategy (Viemeister and Wakefield, 1991) not accounted for within this model. Both individual data and mean data were fitted.

TABLE II. Best-fitting model parameters for the individual and mean data in the no-noise condition at 4 kHz. ERDs are given in ms.

	Decrements		Increments	
Listener	ERD	ΔO	ERD	ΔO
AO	11.1	0.62	2.0	0.60
JS	5.4	0.84	2.2	0.67
JA	16.2	0.49	3.7	0.58
RZ	7.9	0.87	1.4	0.70
Mean	8.9	0.72	2.3	0.64

Table II shows the resulting parameters from the 4-kHz pedestal conditions. As noted previously (Buus and Florentine, 1985; Peters *et al.*, 1995; Moore *et al.*, 1996), there is a trade-off between the values of the ERD and ΔO. Thus a large value for the ERD with a small ΔO can sometimes produce similar results to a small ERD with a large ΔO. Something of this relationship can be seen in the best-fitting values for the decrement data (left two columns of Table II). For instance, the largest value of the ERD (16.2 ms) is found for listener JA, together with the smallest value of ΔO (0.49). This trade-off can account for much of the individual differences in parameters for decrement detection. The ERD of 8.9 ms for the mean data is in good agreement with previous estimates of the temporal-window shape both from nonsimultaneous masking and decrement-detection experiments (Oxenham and Moore, 1994; Peters *et al.*, 1995). In contrast, the best-fitting ERDs for the increment conditions (right two columns) are much smaller than those for the decrement conditions. This is true for all four listeners and the mean data. The difference cannot be accounted for in terms of the trade-off between the ERD and ΔO. In fact, for three of the four listeners, ΔO is also smaller in the increment conditions. For the mean data, the ERD is smaller for increments than for decrements by a factor of nearly 4.

A similar pattern of results can be seen for the 8-kHz data in Table III. Again, estimated ERDs are much greater

TABLE III. Best-fitting model parameters for the individual and mean data at 8 kHz.

	Decrements		Increments	
Listener	ERD	ΔO	ERD	ΔO
AO	18.0	0.33	3.0	0.59
JS	6.9	1.42	3.1	0.82
RZ	8.5	1.04	1.2	0.73
Mean	8.56	0.88	2.34	0.72

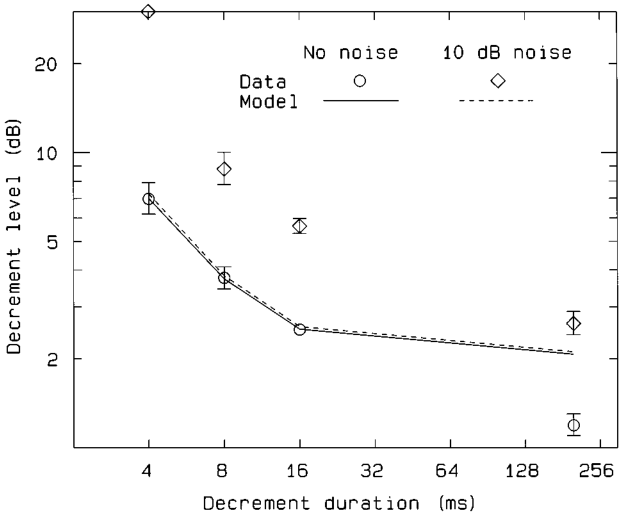

FIG. 7. Effects of a 10-dB spectrum level background noise on the predictions of the temporal-window model (curves), compared with the mean results from experiment 1 (symbols).

for decrement than for increment conditions. Thus in contrast to the assumptions of the temporal-window model, increment- and decrement-detection data produce very different estimates of temporal resolution.

C. Predicted effects of background noise

Within the temporal-window model, the background noise has generally been simulated by a sinusoid, added in quadrature phase to the pedestal, at a level equal to the effective level of the noise falling within the passband of the auditory filter centered at the pedestal frequency. The effect of the noise simulation on model predictions was compared with the effect of noise on listeners' performance in decrement detection. Using the best-fitting parameters for the mean 4-kHz data (see Table II), thresholds in the decrement-detection task were predicted both without noise and with the effect of a 10-dB spectrum level noise taken into account. A 10-dB spectrum level noise has an overall level within the ERB around 4 kHz of 36.6 dB SPL. The resulting predictions, together with the relevant mean data (taken from Fig. 3) are shown in Fig. 7. Thresholds in the absence of noise are predicted well (solid curve). This is to be expected, as the parameters were derived from this data set (excluding the 200-ms point). However, as shown by the dashed curve, predictions are hardly affected by the addition of the simulated 10-dB spectrum level noise. This contrasts strongly with the data which show a marked deterioration due to the noise. Thus the temporal-window model fails to account for the effects of background noise.

In a second approach, mean decrement thresholds in the presence of different noise levels were fitted independently with the temporal-window model. Again, signal durations of 4, 8, and 16 ms were used. For the purposes of modeling it was assumed that listeners could just detect a complete 4-ms gap in the 10-dB noise decrement condition. The best-fitting parameters at each noise level are given in Table IV. As can be seen, the main change is an increase in the detection criterion, ΔO, with increasing noise level; the ERD stays reasonably constant. This provides some support for the temporal-window model: the noise may have a distracting effect which increases the minimum detectable change in the output of the temporal integrator, but the properties of the integrator itself are not affected. Nevertheless, no quantitative way of accounting for the effects of noise is currently available.

TABLE IV. Best-fitting model parameters for the mean decrement data at 4 kHz using background-noise spectrum levels between −20 and 10 dB.

Noise spectrum level	ERD	ΔO
−20 dB	9.3	0.75
−10 dB	8.1	0.92
0 dB	9.3	1.08
10 dB	7.2	1.71

D. Influence of compression on predicted thresholds

While the temporal-window model cannot accurately predict both increment and decrement detection using the same parameters, *some* of the asymmetry apparent in the data is predicted. The size of the predicted asymmetry depends on the amount of compression used. In the simulations described so far the stimulus envelopes were raised to the power $n=0.7$, which is equivalent to raising stimulus intensity to the power 0.35. This value was also used in previous studies (Moore *et al.*, 1993b, 1996; Peters *et al.*, 1995), but it is somewhat less compressive than most psychophysical estimates of peripheral compression, which have centered around $n=0.4$ (e.g., Oxenham and Moore, 1994, 1995; Oxenham and Plack, 1997; Oxenham *et al.*, 1997). On the other hand, earlier work on temporal resolution often operated on (uncompressed) amplitude ($n=1$; Viemeister, 1979; Forrest and Green, 1987) or intensity-like ($n=2$; Moore *et al.*, 1988; Plack and Moore, 1990) quantities.

Simulations showed that reducing the value of n (i.e., increasing the amount of compression) decreased the predicted asymmetry, making predictions worse. Eliminating compression and integrating an intensity-like quantity ($n=2$) increased the predicted asymmetry, but not sufficiently to match the data. This is illustrated in Table V. There, the best-fitting parameters for the mean data at 4 kHz are given using values of n of 0.3 and 2. While the ERDs for the increments and decrements are less disparate using $n=2$, they still differ by more than a factor of 2.

In summary, using a different compression exponent cannot resolve the discrepancy between the data and the model predictions. In fact, a more ''realistic'' compression exponent than was initially used produces a greater discrepancy.

TABLE V. Best-fitting model parameters for the mean data at 4 kHz, using compression exponents of $n=0.3$ and $n=2$.

	Decrements		Increments	
n	ERD	ΔO	ERD	ΔO
0.3	10.7	0.29	2.2	0.27
2.0	5.8	3.45	2.5	1.82

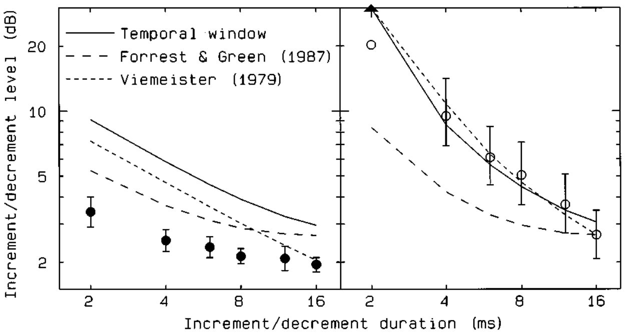

FIG. 8. Predictions of three models of temporal resolution, compared with the mean data from experiment 4, using an 8-kHz pedestal. Increment and decrement thresholds are shown in the left and right panels, respectively. The upward-pointing arrow in the right panel represents a decrement duration at which the model simulations joined to the arrow predicted that a complete gap was undetectable.

VI. ALTERNATIVE MODELS OF TEMPORAL RESOLUTION

In all the following simulations the mean 8-kHz data are used as the reference. It seems certain that spectral splatter did not play a role at 8 kHz and, furthermore, thresholds were measured for a larger number of short durations at 8 kHz than at 4 kHz.

A. Two other models from the literature

Viemeister (1979) proposed a model of temporal resolution to account for the detection of sinusoidal amplitude modulation in a broadband noise carrier. This model comprises the following stages: a bandpass ‘‘pre-detection’’ filter; a half-wave rectifier; a first-order low-pass filter; and a decision device based on the standard deviation of the output of the low-pass filter. A similar model was used by Forrest and Green (1987) to account for both gap detection and modulation detection in a broadband noise. The primary difference between the two models lies in the detection device, which for Forrest and Green is based on the ratio between the maximum and minimum output of the low-pass filter within the observation interval. This is referred to as the max/min decision device and is very similar to the ΔO of the temporal-window model. The pre-detection filter was omitted from the present simulations, as the bandwidth of the stimuli was much less than the 2000–4000 Hz bandwidth usually assumed for that filter. As in Forrest and Green (1987), the low-pass filter had a time constant of 3 ms ($\text{cutoff frequency} = 53\ \text{Hz}$).

First, the max/min decision device was used. The criterion max/min ratio of 2.65 dB was chosen to correctly predict the mean decrement threshold for the 16-ms decrement. Predictions are shown as long-dashed curves in Fig. 8. Next, the model of Viemeister (1979) was tested. Using the standard deviation as the decision statistic presents a potential problem, in that the predictions are dependent on the total duration of the analysis window. However, as long as the window is more than twice the duration of the longest signal, predictions are at least monotonic with increasing signal duration. The analysis window duration was set to 400 ms, which was the length of the steady-state portion of the pedestal in the experiments. Again, the ‘‘threshold’’ value of the standard deviation was set to correctly predict the mean experimental threshold for the 16-ms decrement. The predictions of this model are shown as short-dashed curves in Fig. 8. For comparison, the predictions of the temporal-window using the best-fitting parameters derived from the 8-kHz decrement data (see Table III), are shown as solid curves.

Predictions from all three models fail to mirror the asymmetry observed in the data. Also, compressing the signal prior to the low-pass filtering has the same effect as for the temporal-window model.

B. Smoothed onset detection

The failure of the above models to account satisfactorily for the data may be due to inappropriate preprocessing, an inappropriate decision device, or both. One early approach was to include an approximation of neural adaptation in the preprocessing stages (Irwin and Kemp, 1976; Irwin and Purdy, 1982). While this approach was fairly successful, some important assumptions of the model are open to question. For instance, it is assumed that the onset response to an increment is 2.5 times higher than the steady-state response, regardless of prior stimulation by the pedestal. This is clearly not the case physiologically (Smith and Zwislocki, 1975). Also contrary to available data (e.g., Smith *et al.*, 1985), adaptation was assumed not to affect the response to decrements.

Another approach is to alter the decision device. A number of workers have suggested that in some circumstances *changes* (i.e., rapid onsets or offsets) may provide a more salient cue than overall level (Macmillan, 1971, 1973; Laming, 1986; Oxenham, 1997). As mentioned in the Introduction, Macmillan measured the detectability and discriminability of increments and decrements. The shortest signal duration tested by Macmillan was 15 ms and consequently little or no asymmetry was observed between increment and decrement detection. Also, a simple change detector based, say, on the absolute value of the first derivative of the envelope would not be able to predict the asymmetry observed in the present data. In fact, the present data could be interpreted as indicating that the auditory system is more sensitive to onsets than offsets. Brief increments may be detected by virtue of their onsets, which are independent of increment duration. Brief decrements may be detected by the positive-going slope at the end of the decrement. At very short decrement durations, this slope may be ‘‘masked’’ by the prior pedestal, leading to a stronger dependence of thresholds on decrement duration.

This hypothesis was evaluated using a variant of the temporal-window model. The only change to the model was that the decision device based on ΔO was replaced by a decision device based on the maximum (positive) *slope* of the temporal-window output. It was therefore assumed that detection occurred when the slope of the output exceeded a certain criterion value, relative to the steady-state output of the temporal window. Simulations of the 8-kHz data were run using exponents of $n = 0.7$ (compression) and $n = 2.0$ (intensity-like processing). The model parameters (ERD and

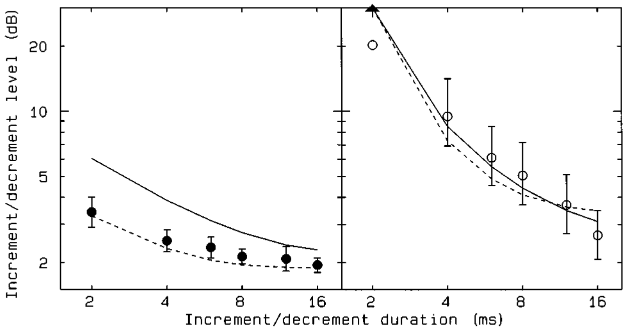

FIG. 9. Predictions of the temporal-window model, using a ‘‘maximum slope’’ decision device, compared with the mean data from experiment 4. Increment and decrement thresholds are shown in the left and right panels, respectively. The solid curves represent predictions using an exponent $n=0.7$ (compression), and the dashed curves represent predictions using an exponent $n=2.0$ (intensity-like processing). The upward-pointing arrow in the right panel denotes a decrement duration at which both models predicted that a complete gap would not be detectable.

criterion slope value) were derived using data from the 8-kHz decrement-detection condition. The resulting ERDs were 12.1 and 5.9 ms for the $n=0.7$ and $n=2$ conditions, respectively. The results from these simulations are shown in Fig. 9. The predictions using the compressive nonlinearity (solid curves) still overestimate thresholds in the increment condition, although the difference is less than for the original decision device (see solid curve in Fig. 8). The predictions using the square-law device are better, and generally lie within one standard error of the mean data.

A growing body of evidence suggests that it is important to include a compressive nonlinearity in models of auditory processing (Penner, 1980; Penner and Shiffrin, 1980; Oxenham and Moore, 1994, 1995; Oxenham and Plack, 1997; Moore and Oxenham, 1997). It is therefore somewhat surprising that the best fit was achieved using a square law, rather than a compressive nonlinearity. Nevertheless, the simulations show that a straightforward change in the decision criterion can improve the fit considerably for both types of nonlinearity.

The temporal-window model has been successfully used to account for a fairly wide range of psychoacoustic data, including the temporal decay of forward and backward masking (Moore *et al.*, 1988), the additivity of nonsimultaneous masking (Oxenham and Moore, 1994, 1995), decrement detection (Plack and Moore, 1991) and gap detection (Moore *et al.*, 1993a). Figure 9 suggests that modifying the decision device may enable the model to also predict at least some aspects of the data presented in this study.

VII. SUMMARY

Thresholds for detecting increments and decrements in the level of a 55-dB sinusoid were measured at 4 kHz as a function of signal duration over a range of background-noise levels. Thresholds were also measured at 8 kHz in the absence of noise. The results were compared with simulations using a temporal-window model and other models of temporal resolution. The following conclusions were reached:

(1) At the shortest durations there was a marked asymmetry between increment and decrement detection, when measured in terms of ΔL or signal level $20\log(\Delta p/p)$: for a given duration, decrements were less easily detected than increments.

(2) Thresholds were dependent on background-noise level over all noise levels tested. The results are not due to the masking of spectral splatter.

(3) Neither aspect of the data mentioned above could be accounted for by the temporal-window model in its standard form. While the decrement-detection data produced estimates of the ERD comparable to those of previous studies (around 9 ms), the increment-detection data were best described by an ERD of less than 2.5 ms, which is smaller by nearly a factor of 4. Two other popular models of temporal resolution (Viemeister, 1979; Forrest and Green, 1987) also failed to account for the data.

(4) Manipulations of the spectral content of the background noise revealed no simple relationship between noise spectrum and signal thresholds. It is possible that off-frequency noise fluctuations interfere in the detection of fluctuations caused by the signal.

(5) A different decision criterion, based on the maximum positive slope of the temporal-window output, produced a better fit to the data in the absence of noise. This may indicate that positive-going slopes (or onsets) are of more perceptual importance than is attributed to them in current models of temporal processing.

ACKNOWLEDGMENTS

This work was supported by a Wellcome Trust Research Fellowship (0044215/Z/95/Z). Torsten Dau, Armin Kohlrausch, Brian Moore, Steven van de Par, Chris Plack, and two reviewers provided helpful comments on an earlier version of this manuscript.

APPENDIX: SIGNAL PHASE AND SPECTRAL SPLATTER

It has been claimed that spectral splatter should be independent of the phase relationship between a long-duration sinusoidal masker and a brief signal of the same frequency (Leshowitz and Wightman, 1971). This is because it is assumed that the masker can be represented by a line spectrum. Thus all off-frequency spectral components of the signal do not interact with the masker. This assertion, important to the interpretation of all the experiments presented here, was tested directly using three listeners. Two listeners (AO and JB) also participated in other experiments in this study; the third (JV) was a student who had some previous experience in psychoacoustic tasks. The measurement procedure was the same as that used in experiment 1. The pedestal was also the same as in experiment 1. The signal was temporally centered in the pedestal, had a total duration of 4 ms, and was gated abruptly. Its starting phase was always zero. The phase relationship between the masker and the signal was the independent variable and was set to 0, 120, and 180 degrees. This was achieved by setting the starting phase of the pedestal to 0, 240, or 180 degrees. As the signal was very brief and was gated abruptly, it was expected that performance would be limited by the detection of spectral splatter. If, as expected, spectral splatter is independent of the masker-signal phase relationship, then thresholds from all three conditions should not be significantly different from each other.

This prediction was confirmed in the results. Thresholds for each condition and listener were repeated 3 times. The mean thresholds, relative to pedestal level, were -17.9, -17.0, and -17.3 dB for the 0-, 120-, and 180-degree conditions, respectively. A within-subjects analysis of variance confirmed that there was no significant effect of phase $[F(2,4)=1.96;\ p>0.25]$. Thus it seems reasonable to assume that when splatter is detectable, performance is independent of the phase relationship between signal and pedestal.

---

[1] While of off-frequency energy is independent of the phase relationship between the pedestal and the signal, this is not true for the phase relationship between the signal window and the signal itself. However, in this case, the ramps of the window (2 ms) are large compared with the signal period (0.25 ms), meaning that the change in off-frequency energy with changing window-signal phase is negligible. The signals used in this study did not exceed 55 dB SPL. Thus any part of the signal spectrum which was more than 55 dB below the peak was certainly not detectable. (In fact, absolute thresholds for a 4-ms signal rarely fall below 20 dB SPL.) Using a 55-dB dynamic range for analysis, the maximum difference in the spectra of the 4-ms signals with 0 and 120 degree starting phases was less than 0.06 dB.

Bacon, S. P., and Grantham, D. W. (**1989**). ‘‘Modulation masking: Effects of modulation frequency, depth and phase,’’ J. Acoust. Soc. Am. **85**, 2575–2580.

Buunen, T. J. F., and van Valkenburg, D. A. (**1979**). ‘‘Auditory detection of a single gap in noise,’’ J. Acoust. Soc. Am. **65**, 534–537(L).

Buus, S., and Florentine, M. (**1985**). ‘‘Gap detection in normal and impaired listeners: The effect of level and frequency,’’ in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, New York), pp. 159–179.

Buus, S., and Florentine, M. (**1991**). ‘‘Psychometric functions for level discrimination,’’ J. Acoust. Soc. Am. **90**, 1371–1380.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). ‘‘Modeling auditory processing of amplitude modulation. I. Detection and masking with narrow-band carriers,’’ J. Acoust. Soc. Am. (submitted).

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). ‘‘Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,’’ J. Acoust. Soc. Am. (submitted).

de Boer, E. (**1986**). ‘‘On thresholds of short-duration intensity increments and decrements,’’ in *Auditory Frequency Selectivity*, edited by B. C. J. Moore and R. D. Patterson, (Plenum, New York), pp. 429–436.

Forrest, T. G., and Green, D. M. (**1987**). ‘‘Detection of partially filled gaps in noise and the temporal modulation transfer function,’’ J. Acoust. Soc. Am. **82**, 1933–1943.

Gerken, G. M., Bhat, V. K. H., and Hutchinson-Clutter, M. (**1990**). ‘‘Auditory temporal integration and the power function model,’’ J. Acoust. Soc. Am. **88**, 767–778.

Glasberg, B. R., and Moore, B. C. J. (**1990**). ‘‘Derivation of auditory filter shapes from notched-noise data,’’ Hearing Res. **47**, 103–138.

Houtgast, T. (**1989**). ‘‘Frequency selectivity in amplitude-modulation detection,’’ J. Acoust. Soc. Am. **85**, 1676–1680.

Irwin, R. J., and Kemp, S. (**1976**). ‘‘Temporal summation and decay in hearing,’’ J. Acoust. Soc. Am. **59**, 920–925.

Irwin, R. J., and Purdy, S. C. (**1982**). ‘‘The minimum detectable duration of auditory signals for normal and hearing-impaired listeners,’’ J. Acoust. Soc. Am. **71**, 967–974.

Laming, D. (**1986**). *Sensory Analysis* (Academic, London).

Leshowitz, B., and Wightman, F. L. (**1971**). ‘‘On-frequency masking with continuous sinusoids,’’ J. Acoust. Soc. Am. **41**, 1180–1190.

Leshowitz, B., and Wightman, F. L. (**1972**). ‘‘On the importance of considering the signal’s frequency spectrum: Some comments on Macmillan’s ‘Detection and recognition of increments and decrements in auditory intensity’ experiment,’’ Percept. Psychophys. **12**, 209–210.

Macmillan, N. A. (**1971**). ‘‘Detection and recognition of increments and decrements in auditory intensity,’’ Percept. Psychophys. **10**, 233–238.

Macmillan, N. A. (**1973**). ‘‘Detection and recognition of intensity changes in tone and noise: The detection-recognition disparity,’’ Percept. Psychophys. **13**, 65–75.

Mendoza, L., Hall, J. W., and Grose, J. H. (**1995**). ‘‘Modulation detection interference using random and sinusoidal amplitude modulation,’’ J. Acoust. Soc. Am. **97**, 2487–2497.

Moore, B. C. J., and Oxenham, A. J. (**1997**). ‘‘Psychoacoustic consequences of compression in the peripheral auditory system,’’ Psychol. Rev. (in press).

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1993a**). ‘‘Detection of temporal gaps in sinusoids: Effects of frequency and level,’’ J. Acoust. Soc. Am. **93**, 1563–1570.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1993b**). ‘‘Effects of frequency on the detection of decrements and increments in sinusoids,’’ J. Acoust. Soc. Am. **94**, 3190–3198.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1996**). ‘‘Detection of decrements and increments in sinusoids at high overall levels,’’ J. Acoust. Soc. Am. **99**, 3669–3677.

Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (**1988**). ‘‘The shape of the ear’s temporal window,’’ J. Acoust. Soc. Am. **83**, 1102–1116.

Oxenham, A. J. (**1997**). ‘‘Temporal integration at 6 kHz as a function of masker bandwidth,’’ J. Acoust. Soc. Am. (submitted).

Oxenham, A. J., and Moore, B. C. J. (**1994**). ‘‘Modeling the additivity of nonsimultaneous masking,’’ Hearing Res. **80**, 105–118.

Oxenham, A. J., and Moore, B. C. J. (**1995**). ‘‘Additivity of masking in normally hearing and hearing-impaired subjects,’’ J. Acoust. Soc. Am. **98**, 1921–1934.

Oxenham, A. J., and Plack, C. J. (**1997**). ‘‘A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing,’’ J. Acoust. Soc. Am. **101**, 3666–3675.

Oxenham, A. J., Moore, B. C. J., and Vickers, D. A. (**1997**). ‘‘Short-term temporal integration: Evidence for the influence of peripheral compression,’’ J. Acoust. Soc. Am. **101**, 3676–3687.

Patterson, R. D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C., and Allerhand, M. (**1992**). ‘‘Complex sounds and auditory images,’’ in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, Oxford).

Penner, M. J. (**1980**). ‘‘The coding of intensity and the interaction of forward and backward masking,’’ J. Acoust. Soc. Am. **67**, 608–616.

Penner, M. J., and Shiffrin, R. M. (**1980**). ‘‘Nonlinearities in the coding of intensity within the context of a temporal summation model,’’ J. Acoust. Soc. Am. **67**, 617–627.

Peters, R. W., Moore, B. C. J., and Glasberg, B. R. (**1995**). ‘‘Effects of level and frequency on the detection of decrements and increments in sinusoids,’’ J. Acoust. Soc. Am. **97**, 3791–3799.

Plack, C. J., and Moore, B. C. J. (**1990**). ‘‘Temporal window shape as a function of frequency and level,’’ J. Acoust. Soc. Am. **87**, 2178–2187.

Plack, C. J., and Moore, B. C. J. (**1991**). ‘‘Decrement detection in normal and impaired ears,’’ J. Acoust. Soc. Am. **90**, 3069–3076.

Riesz, R. R. (**1928**). ‘‘Differential intensity sensitivity of the ear for pure tones,’’ Phys. Rev. **31**, 867–875 .

Smith, R. L. (**1979**). ‘‘Adaptation, saturation, and physiological masking in single auditory-nerve fibers,’’ J. Acoust. Soc. Am. **65**, 166–178.

Smith, R. L., Brachman, M. L., and Frisina, R. D. (**1985**). ‘‘Sensitivity of auditory-nerve fibers to changes in intensity: A dichotomy between decrements and increments,’’ J. Acoust. Soc. Am. **78**, 1310–1316.

Smith, R. L., and Zwislocki, J. J. (**1975**). ‘‘Short-term adaptation and incremental responses in single auditory-nerve fibers,’’ Biol. Cybern. **17**, 169–182.

Viemeister, N. F. (**1979**). ‘‘Temporal modulation transfer functions based on modulation thresholds,’’ J. Acoust. Soc. Am. **66**, 1364–1380.

Viemeister, N. F., and Wakefield, G. H. (**1991**). ‘‘Temporal integration and multiple looks,’’ J. Acoust. Soc. Am. **90**, 858–865.

Yost, W. A., and Sheft, S. (**1989**). ‘‘Across-critical-band processing of amplitude-modulated tones,’’ J. Acoust. Soc. Am. **85**, 848–857.

Zwicker, E. (**1956**). ‘‘Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs,’’ Acustica **6**, 356–381.