Temporal integration at 6 kHz as a function of masker bandwidth

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

(Received 14 October 1996; revised 21 August 1997; accepted 26 August 1997)

Thresholds were measured for a 6-kHz sinusoidal signal presented within a 500-ms masker. The masker was either a bandpass Gaussian noise of varying bandwidth, or a sinusoid of the same frequency as the signal. The spectrum level of the noise masker was kept constant at 20 dB SPL, and the level of the sinusoidal masker was 40 dB SPL. Thresholds for signal durations between 2 and 300 ms were measured for masker bandwidths ranging from 60 to 12 000 Hz. The masker was spectrally centered around 6 kHz. For masker bandwidths less than 600 Hz, the slope of the temporal integration function decreased with decreasing masker bandwidth. The results are not consistent with current models of temporal integration or temporal resolution. It is suggested that the results at narrow bandwidths can be understood in terms of changes in the power spectrum of the stimulus *envelope* or modulation spectrum. According to this view, the onset and offset ramps of the signal introduce detectable high-frequency components into the modulation spectrum, which provide a salient cue in narrowband maskers. For broadband maskers, these high-frequency components are masked by the inherent rapid fluctuations in the masker envelope. Additionally, for signal durations between 7 and 80 ms, signal thresholds decreased by up to 5 dB as the masker bandwidth increased from 1200 to 12 000 Hz. The mechanisms underlying this effect are not yet fully understood. [S0001-4966(97)04112-X]

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

INTRODUCTION

The term temporal integration is often used to describe the way signal thresholds decrease with increasing signal duration. Generally, for a sinusoidal signal in quiet, or in the presence of broadband noise, thresholds decrease by between 8 and 10 dB per decade duration for durations between about 10 and 200 ms (Hughes, 1946; Garner and Miller, 1947; Plomp and Bouman, 1959; Florentine *et al.*, 1988; Gerken *et al.*, 1990).

Most models of temporal integration assume either that the signal intensity is integrated within the auditory system over a duration of around 200–300 ms (Garner and Miller, 1947; Green *et al.*, 1957; Plomp and Bouman, 1959) or that some other transformation of the signal (including compression and/or adaptation) is integrated to give an overall response (Zwislocki, 1960, 1969; Penner, 1978). The signal-to-masker ratio at the output of the integrator is assumed to govern performance. An alternative approach, described by Viemeister and Wakefield (1991), assumes that no such physical integration occurs over durations longer than about 5–10 ms. Instead it is suggested that the auditory system combines information from across independent ''multiple looks'' at brief segments of a given signal. Nevertheless, for a given segment, it is still assumed that the signal-to-masker ratio provides the relevant decision criterion (although this criterion could be altered without changing the fundamental principle of a multiple-looks hypothesis).

Both types of model (true integration and multiple looks) rely on long- or short-term level cues. This may seem inconsistent with some recent data on the detection of tones gated simultaneously with maskers of equal duration (Kidd *et al.*, 1989; Richards, 1992; Kidd *et al.*, 1993). In these studies, within-channel energy cues were rendered unreliable by roving the overall level of the stimuli. It was shown that the performance of listeners is better than that predicted on the basis of energy cues alone for both narrow- and broadband maskers. In the case of broadband maskers it seems likely that listeners detect a change in the spectral shape of the stimulus by performing a comparison of levels across frequency channels, rather than across trials (e.g., Kidd *et al.*, 1989). For narrowband maskers, other cues, such as changes in the distribution of the envelope or fine structure, probably play a role (Kidd *et al.*, 1989; Richards, 1992; Green *et al.*, 1992; Kidd *et al.*, 1993). Nevertheless, in the case of narrowband maskers, it seems that roving the overall presentation level does produce somewhat higher thresholds than are measured in a constant-level paradigm (e.g., Kidd *et al.*, 1989; Richards, 1992). This may indicate that listeners use level-related cues in most circumstances, and that other cues become dominant only when level cues are no longer reliable.

In most studies of temporal integration, the signal is presented in a longer, or continuous, masker of a constant level. In such situations it is also possible for listeners to detect the signal based on a within-interval *and* within-channel level comparison, by comparing the level of the stimulus before the signal onset with the level within the time period which may contain the signal. Thus, at present the assumption of

[a)]Present address: Communication Research Laboratory, Department of Speech-Language Pathology and Audiology (133FR), Northeastern University, Boston, MA 02115.

most models of temporal integration, that detection is based on some measure of (integrated) level, seems reasonable for most conditions. Whether levels are compared across trials, across frequency channels, or across time intervals within one trial probably depends on the exact stimulus configuration.

Neither of the integration models mentioned so far predicts an effect of masker bandwidth (BW) on the *slope* of the temporal integration function, as detection is based purely on the signal-to-noise ratio. The prediction of invariant slopes seems to be supported by the two studies that have directly investigated temporal integration as a function of masker BW (Hamilton, 1957; van den Brink, 1964): neither study found that the slope of the integration function varied with masker BW. However, both these early studies used an 800-Hz sinusoidal signal. This frequency may not have been the optimal choice for two reasons. First, this frequency is too low to measure temporal integration at signal durations of less than about 20 ms because the bandwidth of the signal then begins to exceed that of the auditory filter at 800 Hz. For masker BWs narrower than that of the auditory filter, detection may be based on spectral cues associated with the onset and offset of the signal, known as ''spectral splatter.'' Van den Brink (1964) overcame this problem by filtering both the noise and the signal with a narrow bandpass filter before presentation. However, in this case if the signal bandwidth exceeds the bandwidth of the filter, the temporal characteristics of the signal are affected by the filter, making it difficult to measure true temporal integration. Second, temporal processing is thought to be influenced by the effects of peripheral filtering at frequencies below about 1000 Hz (Moore *et al.*, 1993, 1996). Thus, at 800 Hz the effects of temporal integration, which are thought to be of central origin (Zwislocki, 1960), may be difficult to separate from the influence of peripheral filtering.

In the present study temporal integration was measured as a function of masker BW using a signal frequency of 6 kHz. At this frequency, the equivalent rectangular bandwidth (ERB) of the auditory filter is thought to be about 670 Hz (Glasberg and Moore, 1990), meaning that the bandwidth of even brief signals is less than the ERB. For instance, a Hanning-windowed signal with a half-amplitude duration of only 2 ms has a 3-dB bandwidth of less than 400 Hz. This stimulus configuration therefore enables accurate measurement of temporal integration at much shorter signal durations than was possible in the previous studies using narrowband maskers (Hamilton, 1957; van den Brink, 1964).

I. METHOD

A. Stimuli

The signal was a 6-kHz sinusoid, gated with 2-ms raised-cosine ramps. Thresholds were measured for signal durations of 2, 7, 20, 80, and 300 ms, defined in terms of the half-amplitude duration. The signal was temporally centered around the 300-ms point of a 500-ms masker. The masker was either a band-limited Gaussian noise or a 6-kHz sinusoid. In the latter case, the masker and signal were added in quadrature phase. The BW of the noise masker was varied between 60 and 12 000 Hz and, unless otherwise stated, was arithmetically centered around 6 kHz. The spectrum level of the noise masker was 20 dB SPL and the level of the sinusoidal masker was 40 dB SPL. As the spectrum level was kept constant, the overall level of the noise masker increased with increasing BW. The possible confounding effect of increasing overall level is addressed in the Discussion and the Appendix. To reduce the possibility of detection of spectral splatter at narrow masker BWs, a background noise with a spectrum level of -5 dB SPL in its passband was gated on and off with the masker. The background noise was broadband (up to 15 kHz) with a spectral notch 1200 Hz wide arithmetically centered around 6 kHz. The masker and background noise were also gated with 2-ms raised-cosine ramps. Both noise stimuli were obtained by generating a 2-s circular buffer of wideband Gaussian noise, performing a discrete Fourier transform, setting the amplitude of the components outside the desired passbands to zero, and applying an inverse Fourier transform. A random starting point within the resulting noise buffers was selected on each presentation interval. All stimuli were generated digitally at a 32-kHz sampling rate, and were played out using the built-in 16-bit D/A converter and reconstruction (antialiasing) filter of a Silicon Graphics workstation. Stimuli were passed through a programmable attenuator (TDT PA4) and a headphone buffer (TDT HB6) before being presented to the left ear of subjects via a Beyer DT990 headset.

B. Procedure

Thresholds were measured using a three-interval forced-choice method with a two-down one-up adaptive procedure, which tracks the 70.7%-correct point of the psychometric function. Each trial consisted of three intervals containing the masker and the background noise. The interstimulus interval was 400 ms. The signal occurred randomly in one of the three intervals and subjects were required to select the signal interval. The signal level was initially adjusted in steps of 8 dB. After every two reversals, the step size was halved until a minimum step size of 2 dB was reached. The run terminated after a further ten reversals. Threshold was defined as the median level at the last ten reversals. For every subject, four such threshold estimates were made for each condition, and the mean and standard deviation of the four estimates were recorded. Subjects were tested in 2-h sessions, including short breaks. Within a session thresholds for up to six different masker BWs were measured. All signal durations for each masker BW were tested contiguously in either ascending (2–300 ms) or descending (300–2 ms) order. The presentation order of the BWs was neither systematic nor completely randomized, but was selected independently for each subject in such a way as to avoid the repetition of a BW within one session. In this way, the four estimates for each data point were collected on four separate days. Responses were made via a computer keyboard, and feedback was provided via a computer monitor. Subjects were tested in a single-walled sound-attenuating chamber, which was situated in a sound-attenuating room.

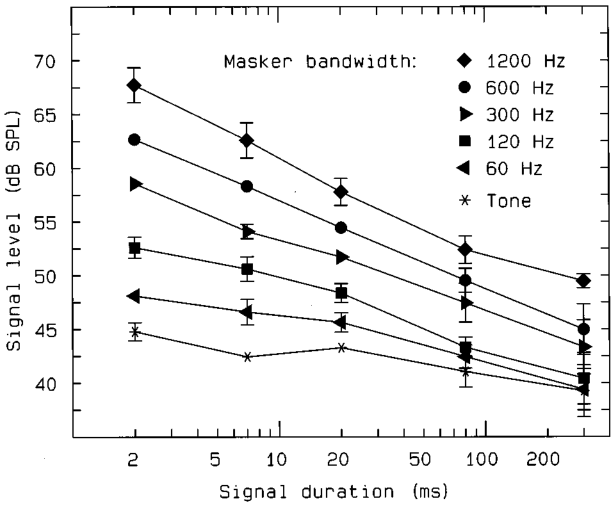

FIG. 1. Signal thresholds as a function of signal duration, with masker bandwidth as the parameter. The signal frequency was 6 kHz and the masker spectrum level was 20 dB SPL (40 dB SPL overall level for the sinusoidal masker). The masker duration was always 500 ms, and the signal was temporally centered around the 300-ms point in the masker. Symbols represent the mean of three subjects. The error bars denote ± 1 standard deviation across the three subjects and are omitted if smaller than the symbol.

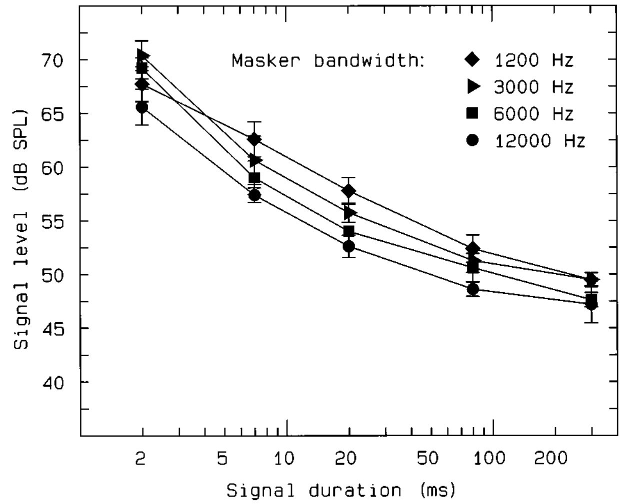

FIG. 2. As in Fig. 1, but for masker bandwidths between 1200 and 12 000 Hz.

C. Subjects

Three subjects took part in the experiment. One subject (the author) had considerable experience in psychoacoustic tasks. The other two were students who had no previous experience in psychoacoustic tasks, and were paid an hourly wage for their participation. All subjects had absolute thresholds of less than 15 dB HL at all octave frequencies between 250 and 8000 Hz, and were given at least 2 h practice before data were collected. No consistent improvements in performance were noted during the course of the experiment; the mean difference between the first and last estimates across all conditions was not significantly different from zero for any of the three subjects.

II. RESULTS

Results across subjects were very similar. Average within-subject standard deviations for each data point were between 1.3 and 1.5 dB for all three subjects. As threshold values were also very similar, only the mean data are shown. Figure 1 shows the data for masker BWs of 1200 Hz and less. Error bars represent ± 1 standard deviation across subjects. The uppermost curve (1200-Hz masker) shows a typical decrease in threshold with increasing signal duration; the difference in threshold between the 2-ms and the 20-ms signal is just under 10 dB, corresponding with previous estimates in quiet and in broadband noise. Reducing the masker BW to 600 Hz produces a decrease in threshold at all durations, resulting in a roughly parallel downwards shift of the integration function. For reductions in masker BW below 300 Hz, however, there is a marked flattening of the integration function, especially for durations of 20 ms and less. For the 60-Hz masker, increasing the signal duration by a factor of 10, from 2 to 20 ms, produces only a 2.5-dB decrease in threshold. For the tonal masker, the difference is even less (1.5 dB). Thresholds for brief signals at narrow masker BWs are therefore lower than would be predicted based on masker energy. It seems that for narrow masker BWs the slope of the integration function is strongly dependent on masker BW.

Results for the masker BW of 1200 Hz are replotted in Fig. 2, together with the results from the other maskers with BWs greater than the ERB. At the longest signal duration of 300 ms, thresholds are not strongly dependent on masker BW. Taken on their own, the data from the 300-ms signal would probably be interpreted as providing additional support for the idea that thresholds are generally independent of masker BW for BWs beyond the ‘‘critical band.’’ At other signal durations the changes with BW are somewhat more marked. For instance, for durations between 7 and 80 ms there is a monotonic *decrease* in threshold as the BW of the masker increases from 1200 to 12 000 Hz. For the shortest duration, the pattern is again somewhat different, and thresholds seem to reach a maximum for a masker BW of 3000 Hz. Thus, at least for signal durations between 7 and 80 ms, increasing the BW of the masker from 1200 to 12 000 Hz, while keeping the spectrum level constant, resulted in a decrease in thresholds, by as much as 5 dB. This pattern of results was observed for all three subjects. A two-way repeated-measures analysis of variance (ANOVA) was carried out on the results for all signal durations and for BWs of 1200 Hz and greater, in order to assess the significance of the observed changes with BW. The factors were BW and signal duration, and the means of the four threshold estimates for each subject were used as the dependent variable. The main effects of duration and BW were both highly significant [$F(4,8) = 232.6$, $p < 0.001$ and $F(3,6) = 33.69$, $p < 0.001$, respectively]. There was also an interaction between BW and duration [$F(12,24) = 4.28$, $p < 0.005$], reflecting the somewhat different pattern of results at the different signal durations.

In Fig. 3, the data from signal durations of 2, 20, and 300 ms are replotted to show signal level at threshold as a function of masker BW. Consider the data from the 300-ms signal (triangles). Consistent with previous studies, thresh-

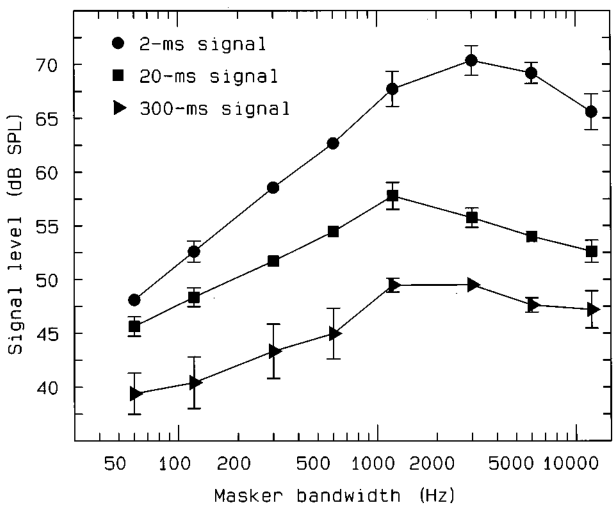

FIG. 3. Data replotted from Figs. 1 and 2, as a function of masker bandwidth, with signal duration as the parameter.

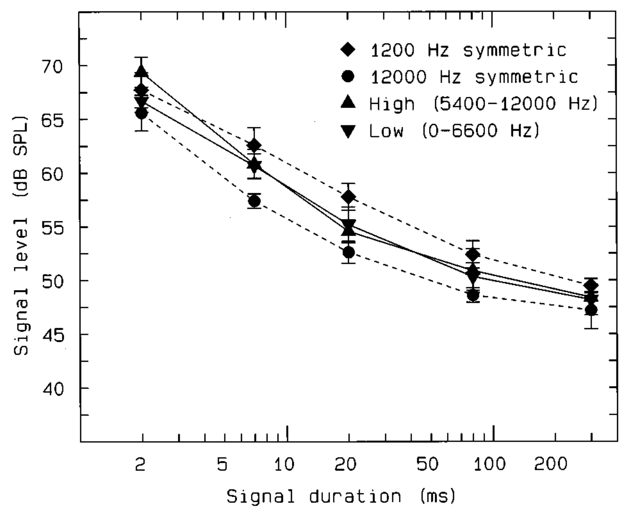

FIG. 4. Additional data using maskers asymmetrically centered around the signal frequency. Symbols joined with solid lines denote signal thresholds in the high-frequency (up-pointing triangles) and low-frequency (down-pointing triangles) asymmetric conditions. Cutoff frequencies for the two maskers are given in the legend. The signal frequency was again 6 kHz, and the masker spectrum level was 20 dB SPL. Data from the 1200-Hz and 12 000-Hz conditions are replotted from Fig. 2.

olds increase with increasing masker BW up to a ‘‘critical bandwidth’’ (in this case around 1200 Hz) beyond which the observed change is fairly small. The slope of the function can be measured in terms of signal level (dB) as a function of 10 log [bandwidth (Hz)]. In these units, the slope for BWs between 60 and 1200 Hz is 0.75. This value lies between the slope of unity predicted if only masker power within the critical band is taken into account (Fletcher, 1940) and the slope of 0.5 predicted if performance is limited only by the statistical fluctuations in the level of the noise masker (e.g., Green *et al.*, 1957; Bos and de Boer, 1966). The slope of the mean 20-ms function is 0.92 and so also lies between the predictions from the two models. In contrast, the slope of the 2-ms function is 1.5 and is therefore not in accordance with either model. In analyzing the slopes, it was assumed that there were no significant differences between the subjects, and the three subject thresholds for each condition were treated as three independent estimates of the same variable. Under this assumption a comparison of regression test was performed (Snedecor and Cochran, 1967, p. 432). This confirmed that signal duration had a significant effect on the slope of the function between 60 and 1200 Hz [$F(2,39) = 27.51$, $p < 0.001$]. This effect was due mainly to the 2-ms data, as the slopes from the 20-ms and 300-ms data were not significantly different from each other [$F(1,26) = 2.04$, $p > 0.1$].

Consider next the data for BWs of 1200 Hz and greater. The slope of the function for the 20-ms signal for BWs from 1200 to 12 000 Hz is -0.52. This is significantly different from zero [$F(1,10) = 62.63$, $p < 0.001$] and confirms that thresholds decrease with increasing BW between 1200 and 12 000 Hz for the 20-ms signal. The slope between 1200 and 12 000 Hz for the 300-ms signal is less steep (-0.26), but is also significantly different from zero [$F(1,10) = 10.95$, $p < 0.01$]. Thus, even in the case of the long-duration signal, the results for a signal frequency of 6 kHz are not consistent with the idea that noise power falling outside the critical band has no effect on threshold (Fletcher, 1940). Thresholds for the 2-ms signal seem to peak at a masker BW of 3000 Hz and decrease on either side of this.

The difference between the 1200-Hz and the 12 000-Hz conditions is surprising. It seems that adding additional (uncorrelated) noise energy outside the critical band can enhance detection, especially for brief signals. Without yet addressing the underlying mechanisms, the effect could be due to the additional high-frequency masker components, the additional low-frequency components, or a combination of both. In order to narrow the possibilities, two further conditions were tested after the main experiment was completed. All three of the original subjects participated, using the same procedure as in the main experiment. The two additional maskers had BWs of 6600 Hz and were asymmetrically placed around 6 kHz. One had cutoff frequencies of 0 and 6600 Hz (low-frequency condition), and the other had cutoff frequencies of 5400 and 12 000 Hz (high-frequency condition). In this way each masker shared one cutoff frequency with the 1200-Hz masker and one with the 12 000-Hz masker. Results are shown in Fig. 4, together with the replotted results from the 1200-Hz and 12 000-Hz conditions. Both high and low spectral regions seem to contribute about equally to the difference in threshold between the 1200-Hz and the 12 000-Hz conditions at all signal durations except 2 ms. Thus, neither the high- nor the low-frequency portion of the 12 000-Hz masker is solely responsible for the decrease in thresholds.

III. DISCUSSION

A. Masker bandwidths less than 1200 Hz: Role of envelope fluctuations

The main finding of this study is that the slope of the integration function at 6 kHz is strongly dependent on masker BW, for BWs less than 600 Hz. These data cannot be accounted for by traditional energy detection (integration) models. For instance, in the case of the sinusoidal masker, a more than tenfold increase in signal energy, as the duration is increased from 2 to 20 ms, leads to only a 1.5-dB decrease in threshold. The simplest explanation for the lack of observed integration would be that, at narrow masker BWs, subjects were able to detect the spread of energy into remote spectral regions due to the onset and offset of the signal. A second possibility, mentioned in the Method section, is related to the fact that the overall masker power decreases with decreasing BW. Oxenham *et al.* (1997) recently measured thresholds for a 6.5-kHz sinusoidal signal in a broadband-noise masker. They found that the slope of the temporal integration function at a masker spectrum level of 20 dB SPL was steeper than at masker levels of -10 and 50 dB by a factor of nearly 2. It could therefore be that the decrease in slope with decreasing masker BW is simply due to the decrease in overall masker power. Both these possibilities are examined in the Appendix, and it is concluded that neither can account for the effect. Thus, another explanation is required.

Martens (1982) and Green *et al.* (1992) have proposed that the auditory system is capable of analyzing the power spectrum of the linear envelope, referred to here as the modulation spectrum. The concept of modulation-spectrum analysis is consistent with studies of modulation masking, which have indicated that something analogous to spectral masking occurs in the modulation-spectrum domain (Bacon and Grantham, 1989; Houtgast, 1989). Dau and colleagues have incorporated this concept within a model of auditory processing using a modulation filterbank. Their model can account well for a range of psychoacoustic data on modulation detection and masking (Dau *et al.*, 1996, 1997a,b). So far, such models have been used mainly to account for the detection of steady-state stimuli. It is proposed here that a similar mechanism may account for why brief signals are more easily detectable in narrowband noise than in wideband noise. In the present experiment, the rapid onset and offset of the signal introduce high-frequency components into the modulation spectrum of the stimulus. This could mean that short-duration signals in narrowband maskers are detected by virtue of the high-frequency modulation energy introduced by the ramps. This modulation energy may be resolved from the lower-frequency modulation energy of the narrowband maskers. For broadband maskers, this cue may be masked by the inherent rapid fluctuations of the noise. Subjects may then be forced to rely on the overall increase in stimulus energy due to the addition of the signal. Some simulations are presented in Sec. IV to illustrate how such a scheme might account for the data.

B. Masker bandwidths of 1200 Hz and greater

Increases of masker BW beyond 1200 Hz tended to produce a decrease in thresholds, especially for signal durations between 7 and 80 ms. It is not clear what mechanism could underlie this result, although it should be noted that a similar finding was observed by Bacon and Smith (1991) using a 10-ms, 4-kHz signal. Their Fig. 1 shows that mean thresholds for the signal temporally centered in the noise also decrease by about 2.5 dB for BWs beyond the critical band. A similar trend can be seen in the data of Wright (1997). Mean thresholds shown in her Fig. 3 for a 20-ms (noise) signal centered around 2500 Hz, temporally centered in a noise masker, also decreased by about 2 dB as the masker BW was increased from 1000 to 8000 Hz.

It is possible that some kind of profile analysis, or spectral shape discrimination, improves performance as the noise covers more peripheral frequency channels (Green, 1988). This interpretation would argue for the importance of across-channel comparisons in detecting a tone in a wideband noise (e.g., Kidd and Dai, 1993). Also, the fact that both low- and high-frequency noise components contribute to the effect is at least qualitatively consistent with this hypothesis. Even so, it is not immediately clear why profile analysis should be most effective for short signal durations. One possibility is that at short signal durations, performance becomes increasingly limited by the trial-to-trial variability of the noise-masker level. In such cases, combining information about the level of the masker across adjacent (independent) frequency channels may improve performance. In the case of long-duration signals, performance may be limited more by ''internal'' variability, or internal noise. If a dominant proportion of such noise is added after information from different channels is combined, then additional masker energy in adjacent frequency channels would not improve performance.

IV. SIMULATIONS USING ENVELOPE CUES

In the previous section it was postulated that at narrow masker BWs the detection of brief signals is achieved by virtue of the high modulation frequencies introduced by the onset and offset ramps of the signal. To illustrate this, some simulations were carried out using masker BWs of 60 and 600 Hz. Predictions using a simple weighted intensity-summation (energy) model were compared with those of a model using only information from high modulation frequencies (modulation-filter model). It is assumed that only envelope information, and not fine structure, is available at 6 kHz. This is in accordance with physiological results showing that fine-structure information is not coded at frequencies above 4–5 kHz (Rose *et al.*, 1968). Also, the effects of peripheral filtering are not incorporated as the BWs of both the maskers and the signal were less than the estimated BW of the auditory filter at 6 kHz.

A. Description of the models

The task of resolving rapid from less rapid envelope fluctuations can be achieved with a simple bandpass filter in the modulation domain. This filter could be implemented in a number of different ways. For the purposes of illustration, spectral analysis is used here, although similar results could be achieved using a time-domain representation, as discussed in Sec. IV E.

In the present simulations, the selected cutoff frequencies for the modulation filter were based on the following considerations. In a recent study of sinusoidal amplitude-modulation detection using sinusoidal carriers, Fassel and Kohlrausch (1995) found that at carrier frequencies of 5 kHz and above, thresholds remained roughly constant up to modulation frequencies of around 150 Hz. Beyond 150 Hz thresholds increased sharply before finally decreasing as the sidebands became spectrally resolved. Based on this, it is assumed that the auditory system is equally sensitive to all modulation frequencies up to around 150 Hz at 6 kHz. This value provided the upper cutoff frequency for the model. The lower cutoff frequency of 80 Hz was chosen so that the overall bandwidth (70 Hz) was approximately in line with estimates of modulation-filter bandwidth derived from simulations of modulation detection and masking data (Dau *et al.*, 1997a).

The modulation-filter model was constructed as follows. The Hilbert envelope of each stimulus was calculated, and the magnitude spectrum of the resulting envelope was extracted using a FFT. All the amplitude components were normalized with respect to the 0-Hz component, making the model level-independent *for a given stimulus*. Then, the rms value of the spectral components between 80 and 150 Hz was calculated and used as the decision criterion. Finally, a normally distributed (Gaussian) random variable was added to the rms value to simulate ''absolute threshold'' for the detection of high modulation frequencies. The mean of the variable was zero and its variance was set such that the model correctly predicted thresholds for the 2-ms signal in the presence of the tonal masker. In this case, no high modulation frequencies are present in the masking stimulus, and so thresholds are assumed to be limited by internal noise alone.

The following considerations were made for the energy model. It has been shown that it is possible to construct an intensity integration weighting function that will predict any desired (straight-line) integration function with a slope between 0 and -1 (Penner, 1978). The same is true of the multiple-looks theory: it is possible to select an arbitrary set of weights for successive looks, such that any slope between 0 and -1 can be predicted. Both schemes are, however, independent of masker BW. A slope of -0.8 was selected to provide a good fit to the data for masker BWs of 600 Hz and greater. The value of -0.8 is also in good agreement with previous estimates of temporal integration at high signal frequencies (e.g., Florentine *et al.*, 1988; Gerken *et al.*, 1990).[1] The predictions for the 600-Hz masker are 5 dB higher than those for the 60-Hz masker at a given signal duration. This relationship is provided by a model which takes into account trial-to-trial variations in overall masker level (Green *et al.*, 1957; Bos and de Boer, 1966). The absolute levels of the energy-model predictions were selected to provide a good (visual) fit to the 600-Hz data.

B. Description of the simulation procedure

The simulations for the modulation-filter model were run for masker BWs of 60 and 600 Hz, using the same three-interval adaptive procedure as was used in the experiment. The stimuli were also generated and controlled in the same way as in the experiment with the following exception: the noise maskers were generated by constructing a new 512-ms circular buffer of random Gaussian noise for each test interval. The duration was 512 ms, instead of 500 ms, so that the number of samples was a power of 2, making the FFT more efficient. The use of a circular buffer meant that no high

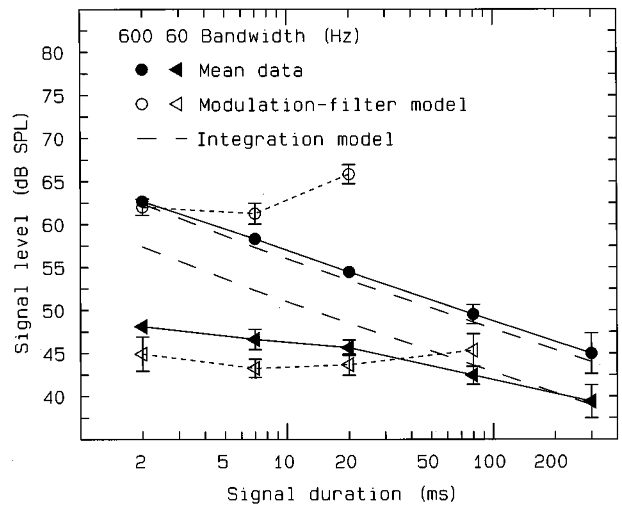

FIG. 5. Predictions using the modulation-filter model (open symbols), together with data replotted from Fig. 1 (filled symbols) for the 60-Hz (triangles) and 600-Hz (circles) conditions. Error bars for the simulations represent ± 1 standard deviation across the 12 estimates for each point. Error bars for the data represent ± 1 standard deviation across the three subjects. The long-dashed lines indicate a decrease in threshold of 8 dB per decade duration.

modulation frequencies were introduced by the onset and offset of the masker, which were instantaneous.

For each trial, three intervals were generated, all containing the masker and one containing the signal. The interval with the largest value for the decision variable was selected by the model. If this interval contained the signal, the answer was deemed ''correct,'' otherwise it was deemed incorrect and the signal level was increased.

Simulations were also carried out using the sinusoidal masker for the purpose of determining the correct variance of the Gaussian random variable, added to the decision variable of the modulation-filter model. After these values had been determined, simulations were performed with masker BWs of 60 and 600 Hz at all the signal durations tested in the experiment. Each condition was run 12 times through the model, and the means and standard deviations of the predicted thresholds are reported.

C. Model predictions

Predictions from both the energy model (dashed lines) and the modulation-filter model (open symbols) are shown in Fig. 5 together with the mean data replotted from Fig. 1 for the 60-Hz and 600-Hz conditions (filled triangles and circles, respectively). Consider first predictions from the energy model (dashed lines). As expected, this model provides a very good fit to the 600-Hz data (upper dashed line), but not to the 60-Hz data (lower dashed line).

Consider next the predictions of the modulation-filter model (open symbols). These are fairly independent of signal duration up to at least 80 ms at 60 Hz, and at least 20 ms at 600 Hz. Beyond these durations the predictions are not shown as they increase to levels which, in most cases, are undefined. This is because, as the signal duration becomes a significant fraction of the duration of the analysis window, the *relative* amount of high-frequency modulation energy actually *decreases* with the addition of the signal, due to the signal's steady-state portion. Thus, for long signal durations, the model consistently selects a nonsignal interval. In such cases, a more suitable criterion would be a reduction in high-frequency modulation energy.

The predictions match the 60-Hz data reasonably well for signal durations up to 20 ms. Note that the predicted thresholds could be increased to match the data better simply by adding an additional internal noise to represent coding inaccuracy in the auditory system. The 600-Hz data are not well described by the predictions of the modulation-filter model; in this condition the energy model provides a better description over the entire range of signal durations.

A number of recent studies have shown that it may be important to take account of peripheral auditory compression when modeling certain psychophysical data (e.g., Oxenham and Moore, 1994, 1995; Moore and Jorasz, 1996; Oxenham and Plack, 1997). Compressing the envelope of a stimulus introduces distortion products into the spectrum that may change the predictions of the model presented here. In order to evaluate the effect of compression, all simulations were repeated with the envelope raised to the power 0.4. This is equivalent to processing intensity raised to the power of 0.2 and is in line with psychophysical estimates of the effective amount of compression (e.g., Oxenham and Moore, 1995). It was found that compression had the effect of raising thresholds by approximately 5 dB for both masker BWs and for all signal durations tested. The pattern of results was, therefore, not affected by the introduction of compression.

In summary, it seems that energy detection provides a reasonable criterion for all signal durations in wideband noise, and for long-duration signals in narrowband noise. For brief signals in narrowband noise, more information may be gained from monitoring increases in high-frequency modulation energy, due to the onset and offset of the signal.

D. Implications for models of temporal resolution

As discussed above, the results from the narrowest BWs cannot be accounted for by current models of temporal integration. Furthermore, current models of temporal *resolution* have time constants which are too long to account for the shallow integration function at the narrowest masker BWs, as shown below.

Two models of temporal resolution were used to predict thresholds at short signal durations in the 60-Hz condition. The first, termed the ''lowpass-filter model,'' is a variant of a model proposed by Viemeister (1979) and adapted by Forrest and Green (1987). In the version used by Forrest and Green (1987), the model comprises an initial bandpass filter with a BW of 4000 Hz, followed by a half-wave rectifier and a first-order lowpass filter with a cutoff frequency of 53 Hz (3-ms time constant). This model has been used to account for amplitude-modulation detection and gap detection in broadband noise. The time constant of the lowpass filter is about two orders of magnitude smaller than that proposed to account for temporal integration (e.g., Plomp and Bouman, 1959). The decision device used by Forrest and Green (1987) was the ratio of the maximum to the minimum instantaneous output of the lowpass filter within one interval (max/min decision device). This decision device was also used in the present simulations: the interval with the largest ratio was selected by the model. No initial bandpass filter was used as all the stimuli were well within the bandpass region of the 4000-Hz-wide filter proposed by Forrest and Green (1987).

The second model is known as the ''temporal-window model.'' It has been used to account for the decay of forward and backward masking (Moore *et al.*, 1988; Plack and Moore, 1990), the additivity of nonsimultaneous masking (Oxenham and Moore, 1994), and decrement detection (e.g., Plack and Moore, 1991; Peters *et al.*, 1995). The temporal-window model comprises a bandpass filter representing peripheral auditory filtering, a rectifying nonlinearity (often including compression), and a sliding temporal integrator or temporal window. The window is assumed to have a double-sided exponential or rounded exponential form (see, e.g., Moore *et al.*, 1996, for further details). In this model, the random fluctuations of noise maskers are generally ignored, and only the expected noise power within one auditory filter is taken into account.

In the present implementation of the temporal-window model, the initial bandpass filter was again omitted because both the signal and the masker fell within the estimated BW of the auditory filter at 6 kHz. The rectified stimuli were raised to the power 0.7 and passed through a sliding temporal integrator consisting of two back-to-back exponential functions, as done previously by Peters *et al.* (1995) and Moore *et al.* (1996). The equivalent rectangular duration (ERD; see Peters *et al.*, 1995) of the window was set to 9.5 ms. This value corresponds closely to values derived for both decrement detection and forward and backward masking at high ($\geqslant 4$ kHz) signal frequencies (Oxenham and Moore, 1994; Peters *et al.*, 1995). The decision device was based on the ratio between the steady-state output of the temporal window in the presence of the masker alone and the maximum output due to the signal and masker together. A constant criterion ratio was selected in these simulations so as to minimize the squared error between the predictions and the mean data for signal durations between 2 and 20 ms.

Thresholds were predicted for signal durations between 2 and 20 ms in the presence of the 60-Hz-wide noise masker. For the lowpass-filter model, simulations were run as they were for the modulation-filter model: 12 estimates were made for each reported data point. Predictions for the temporal-window model, on the other hand, are deterministic, meaning that the adaptive threshold procedure was not necessary. Because of this, a number of additional predicted threshold values were obtained for intermediate durations, in order to define a fairly smooth curve for the temporal-window predictions.

The predictions of both models are shown in Fig. 6, together with the mean experimental data in the 60-Hz condition, for signal durations between 2 and 20 ms. Error bars for the lowpass-filter model (open triangles) represent ± 1 standard deviation across the 12 estimates. Both models show a stronger dependence on signal duration than is observed in the data (filled triangles), although the discrepancy is greater for the temporal-window model (dashed line). The predictions of the lowpass-filter model consistently lie above

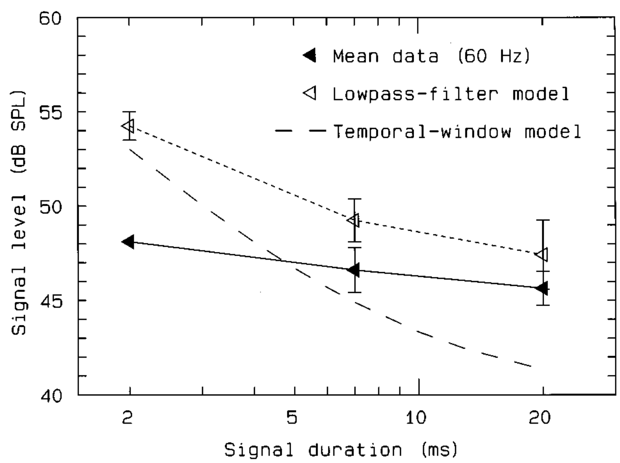

FIG. 6. Predictions for the 60-Hz condition using two models of temporal resolution. Predictions of the lowpass-filter model of Forrest and Green (1987) are shown as open symbols and predictions of the temporal-window model (e.g., Moore *et al.*, 1988) are shown as a long-dashed curve. Mean data (filled symbols) are replotted from Fig. 1.

the data. As no ''internal'' noise was used in the simulations, it would not be possible for the predictions to match the data more closely without altering the processing or the decision device.

Finally, the predictions of both models show no dependence on masker BW, other than a parallel upward shift in thresholds as the masker energy increases (not shown here). As a result, the temporal-window model provides a better fit to the data from the wideband conditions shown in Fig. 2 than does the lowpass-filter model. In summary, neither models of temporal integration nor models of temporal resolution can account for the shallower integration functions at the narrowest BWs.

E. Concluding remarks

The simulations presented here indicate that the data can be understood if it is assumed that at least two distinct cues are used by the auditory system for the detection of tones in masking noise. The first, and more general, is probably based on a transform of overall signal energy, represented by the dashed lines in Fig. 5. The second may be based on high-frequency modulation energy introduced by the onset and offset of a brief signal. Subjects probably base their judgments on whichever cue produces the lowest thresholds in a given condition. Note that the integration mechanism could be implemented as a lowpass filter within a modulation filterbank, meaning that both types of cue would be incorporated within the same overall structure (see Dau *et al.*, 1996).

The modulation-filter model presented here was designed purely for illustrative purposes. The goal was to show that modulation-frequency analysis could provide an explanation for the data at narrow masker BWs, rather than to present any definitive implementation. As mentioned above, the maskers in the simulations were generated in circular buffers to avoid the introduction of high modulation frequencies due to the maskers' onset and offset. This could also have been achieved by using an analysis window with ramps of long duration. If the modulation-filter model as implemented here were taken literally, it would imply that the auditory system has a necessarily long time window over which the modulation spectrum is analyzed. Such an assumption could be easily tested by repeating the narrow BW conditions and introducing, for instance, brief gaps in the masker at times before and after the signal. So long as the gaps occurred within the analysis time-window, the high-frequency modulation energy introduced by the gaps should impair performance. The same impairment would not be expected for wider masker BWs.

Another, perhaps more realistic, method of implementing modulation-frequency analysis would not require such a long analysis window: the modulation filters could be implemented in the time domain, and decisions could be based on maximum short-term peaks in the output of such filters around the expected time of the signal, rather than on the summed output over the whole stimulus interval. The response due to the signal onset would be broadly independent of signal duration. Thus, this type of model would also have the advantage of being able to produce threshold predictions for long-duration signals, which the present modulation-filter model was not able to do.

Finally, analysis in terms of the modulation spectrum may offer a new approach to dealing with detection in situations where ''qualitative'' similarities between masker and signal are thought to play a role. For instance, in forward masking, thresholds for a tonal signal following a narrowband noise can be affected by signal duration and noise BW in ways that are not predicted by any current models using level cues (e.g., Moore and Glasberg, 1982; Neff, 1985): generally, if the signal is in the same spectral region as the masker, and if the temporal envelope of the signal is similar to a single fluctuation in the masker envelope, thresholds are higher than if some extra cue is available to distinguish the signal from the masker. As discussed by Neff (1985), this effect has been ascribed to a lack of ''quality-difference cues'' (Weber and Moore, 1981), and the inability to distinguish masker from signal has been referred to as ''confusion'' (Neff and Jesteadt, 1983). Such confusion may be described by similarities between masker and signal in the modulation spectrum. Thus, using an analysis in the modulation frequency domain may provide a way to quantify such confusion effects, which have so far been described in qualitative terms only.

V. SUMMARY

For a 6-kHz sinusoidal signal, gated with 2-ms raised-cosine ramps, the slope of the integration function is strongly dependent on masker bandwidth below about 600 Hz. The mean difference in threshold between a 2-ms and a 20-ms signal was 8.2 dB for a 600-Hz masker bandwidth, but only 2.5 dB for a 60-Hz masker bandwidth. The results are not consistent with any current model of temporal integration. The lack of dependence of signal threshold on duration for the narrowest masker bandwidths is also not consistent with models of temporal resolution. However, the results can be understood if it is assumed that the auditory system is sensitive to changes in the modulation spectrum, or the power spectrum of the envelope. In this way, the rapid fluctuations introduced by the onset and offset of the signal may be resolved from the slower fluctuations of the narrowband noise. For wideband noise, the inherent rapid fluctuations of the noise may mask those of the signal's onset and offset, forcing subjects to rely on overall energy cues. A further effect was found for wider masker bandwidths: for signal durations between 7 and 80 ms, thresholds decreased by as much as 5 dB as the masker bandwidth was increased from 1200 and 12 000 Hz. The mechanisms underlying this effect are not fully understood.

ACKNOWLEDGMENTS

This work was supported by a Wellcome Trust Travelling Research Fellowship (0044215/Z/95/Z). I thank Torsten Dau for many stimulating discussions. Steven van de Par, Reinier Kortekaas, Armin Kohlrausch, and two reviewers provided helpful comments on an earlier version of this paper.

APPENDIX: POSSIBLE CONFOUNDING FACTORS

The experimental findings reported here suggest that the slope of the temporal integration function at 6 kHz decreases with decreasing masker BW below about 600 Hz. The simplest explanation of this finding would be that subjects are able to detect the spectral spread of energy associated with the onset and offset of the signal at very narrow masker BWs. In this case, as long as the overall energy cue remained less salient than the splatter, no decrease in threshold with duration would be expected. However, as mentioned in the Introduction, the use of a high-frequency signal and the addition of a background noise should have eliminated any possible spectral cues. The mean threshold (across the three subjects) for the 2-ms signal in the presence of the notched noise alone was 29 dB SPL. As this is only 16 and 19.2 dB below the threshold in the presence of the tonal and 60-Hz masker, respectively, it is unlikely that low-level splatter would be detectable. Nevertheless, if performance were dependent on splatter remote from the signal frequency, thresholds would be determined by the masked threshold of the splatter in the background noise. Thus, thresholds should be dependent on the level of the background noise. If splatter were not detectable, and performance were determined by within-channel processes, then the level of the background noise should have little effect on thresholds, so long as the level is not so high as to cause within-channel masking. This was tested by measuring thresholds for the 2-ms signal in the presence of the 60-Hz-wide masker, with the background noise either 5 dB higher or lower in level than in the original experiment. It was found that changing the level of the background noise over a range of 10 dB had no significant effect on masked thresholds: the mean thresholds of the three subjects for background-noise spectrum levels of -10, -5, and 0 dB were 48.6, 48.1, and 49.0 dB SPL, respectively. It therefore appears that detection in narrowband noise was not mediated by spectral splatter.

A second possibility is that the decrease in the slope of the integration function is mediated by the decrease in the overall level of the masker, as the BW is reduced. However, the change in level between the 600-Hz masker and the 60-Hz masker is only 10 dB. Oxenham *et al.* (1997) found that for a broadband noise carrier the slope of the integration function at 6.5 kHz decreased by a factor of nearly 2 as the masker spectrum level was decreased by 30 dB, from 20 to -10 dB SPL. In the present study, the slope of the mean data decreased by a factor of more than 3 as the overall level was changed by only 10 dB. The effect found by Oxenham *et al.* (1997) was therefore much smaller than that observed here. Nevertheless, the possibility that overall masker level was the dominant variable was tested directly using one subject (AO): thresholds in the presence of the 60-Hz masker were measured using a masker level 10 dB higher than in the original experiment, providing the same overall level as the original 600-Hz masker. Except for a 10-dB overall increase in thresholds, the pattern of results remained essentially the same at the higher level. For durations between 2 and 20 ms, the slope of the integration function was -0.28 at 20 dB and -0.3 at 30 dB. This difference was not significant $[F(1,20)=0.01,\ p>0.5]$.

In summary, it appears that neither spectral splatter nor the change in overall level can account for the dependence of the slope of the integration function on masker BW.

[1]For the energy-model predictions, the equivalent rectangular duration (in terms of energy) of each signal duration was used. The equivalent rectangular duration of a raised-cosine ramp of duration t is $3t/8$ (Dallos and Olsen, 1964).

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

Bacon, S. P., and Smith, M. A. (**1991**). ''Spectral, intensive, and temporal factors influencing overshoot,'' Q. J. Exp. Psychol. **43A**, 373–400.

Bos, C. E., and de Boer, E. (**1966**). ''Masking and discrimination,'' J. Acoust. Soc. Am. **39**, 708–715.

Dallos, P. J., and Olsen, W. O. (**1964**). ''Integration of energy at threshold with gradual rise-fall tone pips,'' J. Acoust. Soc. Am. **36**, 743–751.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1996**). ''Modeling modulation perception: Modulation low-pass filter or modulation filterbank?'' in *Psychoacoustics, Speech and Hearing Aids*, edited by B. Kollmeier (World Scientific, Singapore), pp. 45–48.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997a**). ''Modeling auditory processing of amplitude modulation. I. Detection and masking with narrowband carriers,'' J. Acoust. Soc. Am. **102**, 2892–2905.

Dau, T., Kollmeier, B., and Kohlrausch, A. (**1997b**). ''Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration,'' J. Acoust. Soc. Am. **102**, 2906–2919.

Fassel, R., and Kohlrausch, A. (**1995**). ''Modulation detection as a function of carrier frequency and level,'' in *IPO Annual Progress Report 30*, edited by M. D. Brouwer-Janse, D. J. Hermes, W. M. C. J. van Overveld, and H. de Ridder (IPO, Eindhoven).

Fletcher, H. (**1940**). ''Auditory patterns,'' Rev. Mod. Phys. **12**, 47–65.

Florentine, M., Fastl, H., and Buus, S. (**1988**). ''Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking,'' J. Acoust. Soc. Am. **84**, 195–203.

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.

Garner, W. R., and Miller, G. A. (**1947**). ''The masked threshold of pure tones as a function of duration,'' J. Exp. Psychol. **37**, 293–303.

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.

Green, D. M. (**1988**). *Profile Analysis* (Oxford U.P., Oxford).

Green, D. M., Berg, B. G., Dai, H., Eddins, D. A., Onsan, Z., and Nguyen, Q. (**1992**). ''Spectral shape discrimination of narrow-band sounds,'' J. Acoust. Soc. Am. **92**, 2586–2597.

Green, D. M., Birdsall, T. G., and Tanner, W. P. (**1957**). ''Signal detection as a function of signal intensity and duration,'' J. Acoust. Soc. Am. **29**, 523–531.

Hamilton, P. M. (**1957**). ''Noise masked thresholds as a function of tonal duration and masking noise bandwidth,'' J. Acoust. Soc. Am. **29**, 506–511.

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

Hughes, J. W. (**1946**). ''The threshold of audition for short periods of stimulation,'' Proc. R. Soc. London, Ser. B **133**, 486–490.

Kidd, G., and Dai, H. (**1993**). ''A composite randomization procedure for measuring spectral shape discrimination,'' J. Acoust. Soc. Am. **94**, 1275–1280.

Kidd, G., Mason, C. R., Brantley, M. A., and Owen, G. A. (**1989**). ''Roving-level tone-in-noise detection,'' J. Acoust. Soc. Am. **86**, 1310–1317.

Kidd, G., Uchanski, R. M., Mason, C. R., and Deliwala, P. S. (**1993**). ''Discriminability of narrow-band sounds in the absence of level cues,'' J. Acoust. Soc. Am. **93**, 1028–1037.

Martens, J.-P. (**1982**). ''A new theory for multi-tone masking,'' J. Acoust. Soc. Am. **72**, 397–405.

Moore, B. C. J., and Glasberg, B. R. (**1982**). ''Contralateral and ipsilateral cueing in forward masking,'' J. Acoust. Soc. Am. **71**, 942–945.

Moore, B. C. J., and Jorasz, U. (**1996**). ''Modulation discrimination interference and comodulation masking release as a function of the number and spectral placement of narrow-band noise modulators,'' J. Acoust. Soc. Am. **100**, 2373–2381.

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.

Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (**1993**). ''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.

Neff, D. L. (**1985**). ''Stimulus parameters governing confusion effects in forward masking,'' J. Acoust. Soc. Am. **78**, 1966–1976.

Neff, D. L., and Jesteadt, W. (**1983**). ''Additivity of forward masking,'' J. Acoust. Soc. Am. **74**, 1695–1701.

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.

Penner, M. J. (**1978**). ''A power law transformation resulting in a class of short-term integrators that produce time-intensity trades for noise bursts,'' J. Acoust. Soc. Am. **63**, 195–201.

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.

Plomp, R., and Bouman, M. A. (**1959**). ''Relation between hearing threshold and duration for tone pulses,'' J. Acoust. Soc. Am. **31**, 749–758.

Richards, V. M. (**1992**). ''The detectability of a tone added to narrow bands of equal-energy noise,'' J. Acoust. Soc. Am. **91**, 3424–3435.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1968**). ''Patterns of activity in single auditory nerve fibres of the squirrel monkey,'' in *Hearing Mechanisms in Vertebrates*, edited by A. V. S. d. Reuck and J. Knight (Churchill, London).

Snedecor, G. W., and Cochran, W. G. (**1967**). *Statistical Methods* (Iowa State U.P., Ames), 6th ed.

van den Brink, G. (**1964**). ''Detection of tone pulse of various durations in noise of various bandwidths,'' J. Acoust. Soc. Am. **36**, 1206–1211.

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.

Weber, D. L., and Moore, B. C. J. (**1981**). ''Forward masking by sinusoidal and noise maskers,'' J. Acoust. Soc. Am. **69**, 1402–1409.

Wright, B. A. (**1997**). ''Detectability of simultaneously masked signals as a function of masker bandwidth and configuration for different signal delays,'' J. Acoust. Soc. Am. **101**, 420–429.

Zwislocki, J. J. (**1960**). ''Theory of temporal auditory summation,'' J. Acoust. Soc. Am. **32**, 1046–1060.

Zwislocki, J. J. (**1969**). ''Temporal summation of loudness: An analysis,'' J. Acoust. Soc. Am. **46**, 431–441.