

Masking of a brief probe by sinusoidal frequency modulation

Brent W. Edwards^{a)} and Neal F. Viemeister

Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455

(Received 7 August 1995; revised 12 September 1996; accepted 7 October 1996)

Contrary to level detection models, the thresholds for a brief-duration probe masked by a sinusoidal frequency modulation (FM) masker increases as the modulation index (β) of FM increases [Zwicker, *Acustica* **31**, 243–256 (1974)]. In this paper the reason for this phenomenon is investigated. In experiment 1, a 10-ms, 1-kHz probe was detected in the presence of an FM masker centered at 1 kHz and sinusoidally modulated at 16 Hz. Thresholds increased by over 15 dB with increasing β , consistent with Zwicker's findings. In experiment 2, the instantaneous frequency changes of the masker used in experiment 1 were clipped and the resulting thresholds indicated that detection was determined primarily by the masker's total frequency excursion rather than by its instantaneous sweep rate. In experiment 3, the FM maskers from the first two experiments were passed through a roex filter centered at 1 kHz and the resulting envelope was used to amplitude modulate a 1-kHz tone, producing approximately the same effective envelope at 1 kHz as the FM maskers. Threshold functions for the amplitude modulated (AM) maskers were similar to those for their corresponding FM maskers. Thresholds increased by over 15 dB while the total energy of the AM masker decreased by over 10 dB, again contrary to standard level-detection models. The results from these experiments can be explained, at least qualitatively, by a model based on envelope shape discrimination: similarities between the envelopes of the masker alone and masker-plus-probe at the output of an auditory filter centered on the frequency of the probe are primarily responsible for the observed masking, particularly at large β 's. © 1997 Acoustical Society of America. [S0001-4966(97)03902-7]

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

INTRODUCTION

An old, and common, assumption is that the detectability of a pure-tone signal is determined by the amount of masker passed by an auditory filter, or critical band, centered at the frequency of the signal. This notion dates back at least to Fletcher (1940) and is described in more recent formulations such as the energy detector model (Green and Swets, 1966) and the power spectrum model (Moore, 1989). There is evidence that in certain circumstances this notion is incorrect, that information other than a simple change in the overall filter output ("level detection," henceforth) is the basis for monaural detection. For example, adding "comodulated" noise to frequencies far removed from the signal frequency can lead to an improvement in pure-tone detection in narrow-band noise (Hall *et al.*, 1984). Until recently such discrepancies generally have been considered minor and level-detection schemes have been taken as useful approximate descriptions of the detection process. Indeed, psycho-physical measurements of auditory frequency selectivity, most of which rely heavily on such assumptions (e.g., Patterson and Moore, 1986), are considered to be reasonably valid.

Recent evidence indicates, however, that level detection may be the exceptional case. Several studies have shown, for example, that randomizing the noise level over a wide range has little effect on detection performance when the signal-to-noise ratio is fixed (Gilkey, 1987; Kidd *et al.*, 1989; Richards *et al.*, 1991). Such presentation-by-presentation "roving"

would wreak havoc on the performance of a simple single-channel level detector. Current research has moved in the direction of examining other possible cues used in detection and has returned, in a more sophisticated way, to earlier concerns about this fundamental problem (Jeffress, 1964; Ahumada and Lovell, 1971).

A striking discrepancy with the level-detection scheme antedates many of these recent studies but, surprisingly, has been addressed only in passing. The discrepancy is shown in studies on the masking of a brief probe by sinusoidal frequency modulation (FM) and by frequency sweeps. These studies (Zwicker, 1974; Smoorenburg and Coninx, 1980; Kemp, 1982) show that when the instantaneous frequency of the masker is equal to the probe frequency, an increase in the sweep rate may cause an *increase* in threshold.

Figure 1 plots a spectrographic representation of the masker and probe for two conditions similar to those investigated by Zwicker. A 10-ms, 1-kHz probe is presented when the instantaneous frequency of the masker increases through 1 kHz. The probe detection threshold when the masker's total frequency excursion (TFE) equals 1024 Hz [Fig. 1(a)] is more than 15 dB *higher* than when the TFE is 64 Hz [Fig. 1(b)]. Clearly, the masker energy in a narrow-band centered at 1 kHz is less for the larger TFE. For example, the masker energy, or average power, passed by a typical auditory filter centered at 1 kHz is 10 dB less for the masker with a TFE of 1024 Hz than the masker with 64-Hz TFE. The simple level detector therefore would show a 10 dB lower threshold for the larger TFE; the discrepancy with the data is 25 dB.

Although Zwicker investigated the effects of modulation

^{a)}Current address: ReSound Corp., 220 Saginaw Dr., Redwood City, CA 94063.

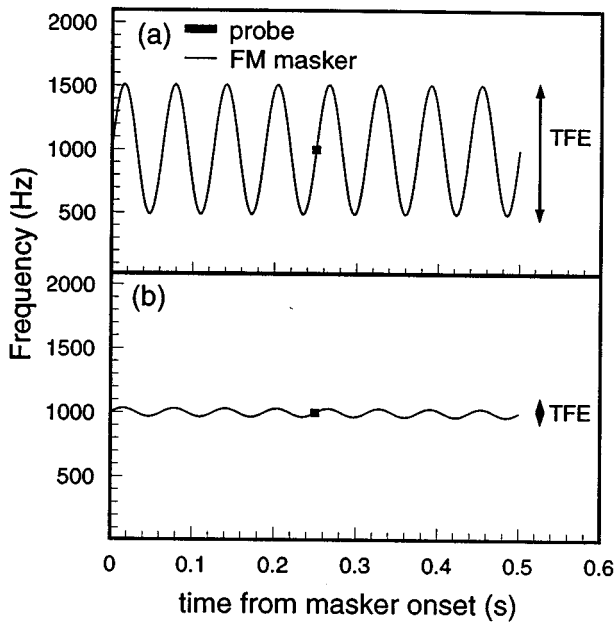


FIG. 1. Spectrographic representations of stimuli demonstrating the masking phenomenon. The total frequency excursions (TFE) of the FM masker is (a) 1024 Hz, and (b) 64 Hz. While the masker in (b) has 10 dB more energy through a critical band centered at 1 kHz than the masker in (a), 15 dB more masking is produced by the masker in (a).

frequency and later Kemp (1982) investigated the effects of modulator phase, neither discussed the reason for the unexpected effect of increased thresholds with increased TFE. This paper investigates this phenomenon and presents a possible explanation. In experiment 1, conditions similar to those used by Zwicker (1974) and by Kemp (1982) are run with a 16-Hz modulation frequency. Results from this experiment are used as a comparison for the results obtained in the next two experiments. In experiment 2, the FM masker is processed to separate correlated characteristics of the masker, specifically, the frequency sweep rate and TFE, and to determine the relevancy of each to the masking of the probe. In experiment 3, a filter model is used to account for the data and thresholds are obtained with a single-channel representation of the stimuli used in the first and second experiments.

I. EXPERIMENT 1: SINUSOIDAL FM MASKER

A. Stimuli

The FM masker was 500 ms in duration with 20-ms raised-cosine on and off ramps. The steady-state portion of the masker was generated using

$$A \sin[2\pi f_c t - \beta \cos(2\pi f_m t)], \quad (1)$$

where f_c is the carrier frequency, f_m is the modulation frequency, and A is a constant set such that the level of the masker was 60-dB sound-pressure level (SPL). The TFE of the masker is $2\beta f_m$. This dependence of the TFE on modulation frequency is a result of the way β is typically defined and is one reason why FM detection thresholds decrease as a function of β but remain relatively constant as a function of TFE (for small f_m). Throughout the following experiments, f_m was fixed at 16 Hz and TFE was changed by varying β .

The carrier frequency was 1 kHz for all experiments.

The probe was a 1-kHz tone with 5-ms raised-cosine ramps and a total duration of 10 ms. Probe levels were defined as the SPL at the maximum portion of the probe. The temporal center of the probe coincided with the temporal center of the masker, i.e., the onset of the probe was 245 ms after the onset of the masker. The instantaneous frequency of the masker was increasing through 1 kHz at the temporal center of the probe (as shown in Fig. 1). This condition provides maximum masking of the probe (Zwicker, 1974) and was used for all of the present experiments.

Since the phase of the FM masker changes during presentation of the probe, it is not clear what probe phase should be used. The problem is compounded because the phase changes in the masker depend on TFE, a variable of primary interest in the present studies. Therefore, to minimize systematic phase-related effects on probe threshold, the phase of the probe was randomly chosen (rectangular distribution, $0-2\pi$) for each probe presentation and for all conditions (including TFE=0). Pilot data indicated that there was little difference between threshold functions obtained with random phases and threshold functions obtained with phase fixed at 90° (the orthogonal condition for TFE=0).

The stimuli were generated digitally on a NeXT computer using its 16-bit D/As and a sampling rate of 44.1 kHz. The NeXT's programmable attenuators were used to increase the effective dynamic range of the D/A. Stimuli were presented monaurally using Sony MDR-V6 headphones.

B. Procedure

A 3IFC task was used, with each 500-ms observation interval separated by 500 ms of silence. An identical masker was presented in each interval, and the probe was added to the masker in one of the intervals chosen at random with equal probability.

A two-down, one-up tracking procedure was used to estimate the 70.7% correct point on the psychometric function. The level of the probe was adjusted in 3-dB steps until four reversals were achieved, then the probe was adjusted in 1-dB steps. The procedure terminated after a total of 12 reversals; the first four reversals were ignored and the levels of the last eight reversals were averaged to estimate the threshold for that block of trials. Threshold estimates were discarded if the standard deviation of the reversals within a block exceeded 4 dB.

The masker was fixed within a block and the masker TFE was randomized across blocks. Thresholds from four blocks with the same TFE were averaged to estimate the final threshold for a condition.

Visual cues denoted the observation intervals (for the entire 500 ms) and provided correct answer feedback.

C. Subjects

The three subjects had normal hearing (within 10 dB of laboratory norms). S1 was an experienced subject; S2 and S3 received five hours of training using a variety of conditions from experiment 1, during which their probe thresholds stabilized.

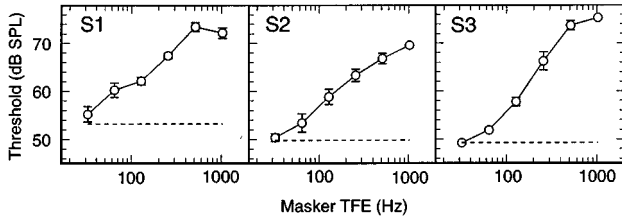


FIG. 2. Probe threshold as a function of masker TFE for experiment 1.

D. Results

Figure 2 shows the probe detection thresholds as a function of the masker TFE. Each panel represents a different subject, and the dashed line represents detection threshold when the masker is a 1-kHz tone, i.e., TFE=0. Error bars denote the standard deviation of the threshold estimates.

For the condition which results in the lowest threshold, TFE=32 Hz, the masker sweeps through less than 16 Hz in the 10-ms window during possible probe presentation; for the condition which results in the highest threshold, TFE =1024 Hz, the masker sweeps through more than 490 Hz during this window. These results demonstrate the general phenomena which is contradictory to the simple level-detection scheme: probe-detection thresholds increase while the masker energy near the frequency of the probe decreases. The data are consistent with those of Zwicker (1974) and Kemp (1982) which demonstrate the same phenomenon at different modulation frequencies.

II. EXPERIMENT 2: CLIPPED FM MASKER

While the results from the previous experiment indicate that the masking of a tonal probe by an FM signal increases as the TFE of the masker increases, it is unclear what underlying property of the masker causes the threshold increase. Specifically, as the masker TFE increases, so does the rate of change of frequency during the probe presentation. When the instantaneous frequency of the masker is 1 kHz, occurring at the temporal center of the probe, the rate of change of the masker's frequency, or sweep rate, is proportional to the TFE of the masker:

$$\begin{aligned} \frac{d}{dt} f(t)|_{1000 \text{ Hz}} &= 2\pi f_m^2 \beta \text{ Hz/s} \\ &= \pi f_m \times \text{TFE} \text{ Hz/s.} \end{aligned} \quad (2)$$

The results from experiment 1 demonstrate that thresholds increase as masker TFE increases. Because of the relationship between masker TFE and sweep rate shown in Eq. (2), it is unclear if the increase in threshold is due only to the TFE *per se* or, as hypothesized by Smoorenburg and Coninx (1980) for linear-sweep FM, results from the increase in the masker's sweep rate during the probe presentation. Experiment 2 was intended to isolate these two effects.

A. Stimuli

The probe and masker were similar to those in the previous experiment except that the instantaneous frequency of the masker was "clipped" such that the masker's instantane-

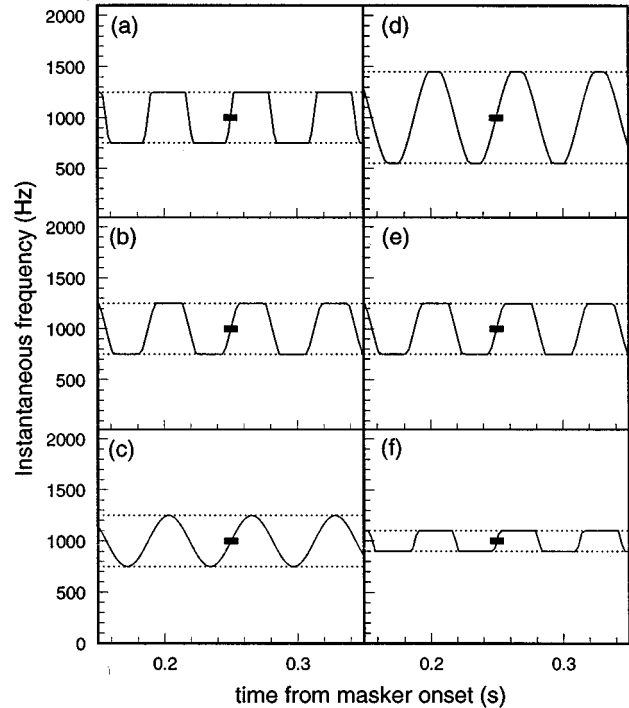


FIG. 3. Panels (a)–(c) demonstrate stimuli conditions for experiment 2 where the masker TFE is held constant while the sweep rate during probe presentation varies; panels (d)–(f) demonstrate conditions where the masker sweep rate is held constant while the TFE varies.

ous frequency deviation from 1 kHz did not exceed a predefined limit, f_{cl} . Before clipping, the masker was defined in the same manner as experiment 1. For a fixed TFE (or fixed β), the instantaneous frequency of the masker was then restricted such that

$$\begin{aligned} f(t) &= f_c + f_{cl}, & \text{for } f_c + f_m \beta \sin(2\pi f_m t) > f_c + f_{cl}, \\ &= f_c - f_{cl}, & \text{for } f_c + f_m \beta \sin(2\pi f_m t) < f_c - f_{cl}, \\ &= f_c + f_m \beta \sin(2\pi f_m t), & \text{otherwise.} \end{aligned} \quad (3)$$

By fixing f_{cl} , the masker's actual TFE was held constant while the sweep rate was changed by varying β in Eq. (3) [or by varying the TFE before clipping; see Eq. (2)]. Figure 3(a)–(c) demonstrates this situation: the sweep rate during the probe presentation in each figure is different while the masker TFE is the same. Figure 3(d)–(f) displays the opposite case, where the TFE in each figure is different while the sweep rate during probe presentation remains unchanged. It should be noted that clipping the instantaneous frequency did not result in the spectral splatter that would result from a similar clipping of the envelope. In addition, care was taken not to introduce any discontinuities in the phase of the stimuli which would result in splatter and noticeable clicks.

B. Procedure

The procedure was the same as in experiment 1, with the following exceptions. Within a set of blocks, the masker always had the same TFE. Different blocks within a set had different masker sweep rates. TFE and sweep rate were varied in a pseudorandom manner and at least four threshold

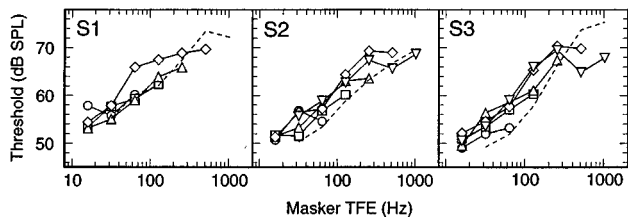


FIG. 4. Data for the conditions shown in Fig. 3. Each symbol represents a different sweep rate during the probe presentation: 1024 (circles), 2048 (squares), 4096 (up-triangles), 8192 (diamonds), and 16 384 (down-triangles) Hz/s. The abscissa represents the TFE of the masker after its instantaneous frequency has been clipped. Thus, different data points but with the same symbol correspond to panels (d)–(f) in Fig. 3. The dashed line plots the data from Fig. 2 as a function of sweep rate.

estimates were obtained for each condition. Final thresholds were obtained for each condition by averaging the four threshold estimates.

C. Subjects

The subjects from experiment 1 participated in this experiment.

D. Results

The results are shown in Fig. 4, which plots probe threshold versus actual (post-clipped) masker TFE. The parameter is the masker sweep rate when the frequency is increasing through 1 kHz. Also shown in Fig. 4 with the dashed lines are the thresholds from Fig. 2, i.e., thresholds when the masker was unclipped. To simplify the figures, standard deviations are not shown, but were all less than 3 dB.

One result obvious from Fig. 4 is that when the masker sweep rate is fixed, thresholds still increase with increasing masker TFE. The thresholds are similar to the thresholds obtained using unclipped maskers with the same TFE (dashed line). If thresholds from experiment 1 were dependent upon the sweep rate and not the TFE, then (1) the lines would be flat as a function of TFE, and (2) the lines would be ordered according to the sweep rate they represent, with lines corresponding to higher sweep rates falling above lines corresponding to lower sweep rates. This last property is only seen with a single-condition for subject 1, where thresholds from the fastest sweep rate are higher than all the thresholds from all the other sweep rates.

Figure 5 further demonstrates the roles of TFE and sweep rate. Here, the thresholds from Fig. 4 are replotted as a function of sweep rate, with masker TFE as the parameter. If the data were dependent upon sweep-rate and not post-clipped TFE, then all of the lines would coincide. Instead, thresholds for a given post-clipped TFE are nearly constant; the lines are separated and ordered, with a few exceptions, such that lines corresponding to higher clipped TFEs have higher thresholds. This figure further illustrates that the sweep rate has little effect on threshold.

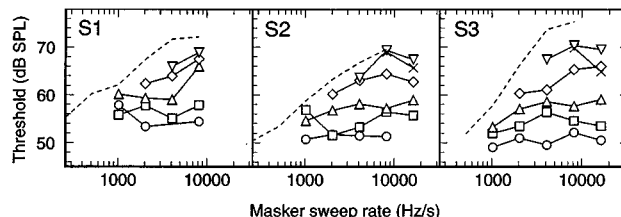


FIG. 5. The data from Fig. 4, replotted in a different manner. The abscissa represents the frequency sweep rate of the masker during probe presentation. The parameter is the actual masker TFE: 16 (circles), 32 (squares), 64 (up-triangles), 128 (diamonds), 256 (down triangles), and 512 (crosses) Hz. Different data points but with the same symbol correspond to (a)–(c) in Fig. 3.

III. EXPERIMENT 3: EQUIVALENT AM MASKERS

The results of experiment 2 indicate that the masker’s TFE, and not the masker’s sweep rate, is the primary factor in determining the amount of masking of the probe. Why thresholds increase with increasing TFE, the fundamental problem motivating this paper, is still not clear.

One possible explanation is based on the following. As the masker sweeps through an auditory filter, a brief increase occurs in the envelope of the signal at the output of the filter. This envelope “pulse” is similar to that caused by a brief probe whose frequency is at the center frequency of the filter. It is possible that the listener has difficulty detecting the envelope fluctuation produced by adding the probe because the masker is causing similar fluctuations. That is, we hypothesize that the subject is basing decisions on differences between the temporal structure of the envelopes; the task is effectively envelope-shape discrimination. As the envelope of the filtered masker becomes more similar in shape to that of the probe, the envelope of the masker-plus-probe will be, at least roughly, more similar to that of the masker and detectability will deteriorate.

To examine this hypothesis, the FM maskers from the previous two experiments are transformed into AM maskers to simulate the form of the envelope fluctuations that the masker produces at the output of an auditory filter. A carrier of 1 kHz is used so the envelope seen at the output of an auditory filter due to the AM masker will be approximately the same as that due to an FM masker.

A. Stimuli

The probe was the same as in the previous two experiments. The masker was obtained by passing the FM maskers from the previous two experiments through a roex filter (Patterson *et al.*, 1982) with a transfer function defined by

$$H(f) = (1 + p|f_c - f|/f_c) \exp(-p|f_c - f|/f_c), \quad (4)$$

where f_c is the center frequency of the filter (1 kHz), and p is a constant which determines the attenuation rate of the filter (fixed here at 30; Rosen *et al.*, 1992).

Figure 6 demonstrates the signal processing of the masker. Once an FM signal was filtered, the envelope was extracted using a Hilbert transform. The envelope was then used to amplitude modulate a 1-kHz tone, creating an AM signal with no frequency modulation. The envelope of the

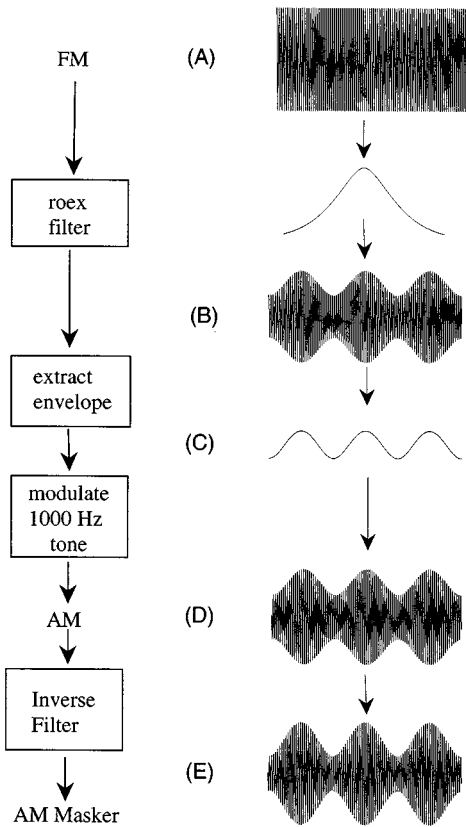


FIG. 6. Processing done on FM masks from experiments 1 and 2 to obtain their equivalent AM masks for experiment 3.

AM signal (point D in Fig. 6) is identical to the envelope of the FM masker after it is filtered (point B). Finally, the AM signal was filtered by the inverse of Eq. (4) to negate the effect of further filtering by the listener's own auditory filter. The intent of this processing scheme is that the envelope at the output of the listener's auditory filter due to the AM masker will be similar to the envelope of the auditory filter's output due to the corresponding FM masker. After the AM masker was generated, the probe was then added such that the temporal center of the probe corresponded to the temporal center of the masker, i.e., the point equivalent to the probe's location in the previous two experiments.¹

Again, the general goal is to evaluate whether the observed increases in probe threshold with increasing TFE are at least qualitatively consistent with a single-channel, envelope-extraction scheme. Although it might be desirable to measure filter functions for each subject and to use those in the processing, such an effort seems unnecessary at this stage of inquiry. Thus, the simulated auditory filter and subsequent processing were the same for all subjects.

B. Procedure

The procedure was similar to that of experiments 1 and 2, and the same subjects participated.

C. Results

Figure 7 shows the results obtained when the sinusoidal FM masks from experiment 1 were used to generate the

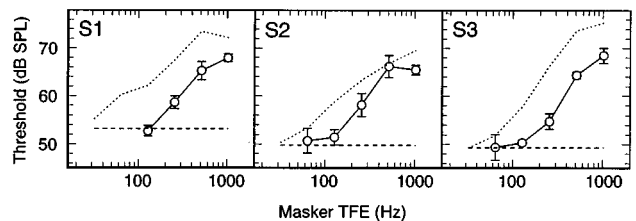


FIG. 7. Probe detection thresholds for AM masks that were created by processing the FM masks that were used in experiment 1. The abscissa represents that TFE of the masker before it was processed. The AM masks have the same envelope fluctuations at 1 kHz as their FM equivalents from experiment 1. The dashed line represents threshold when the masker was a 1-kHz pure tone. The data from Fig. 2 are replotted with dotted lines.

AM masks. The abscissa represents the TFE of the masker before it was transformed to an AM signal (i.e., the TFE at point A in Fig. 6); the actual masker presented to the subjects had no instantaneous frequency fluctuations.² The dashed line represents threshold when the masker is an unmodulated tone, i.e., TFE=0. The dotted line represents the average of the data from Fig. 2, i.e., data from the FM masks used to create the AM equivalents. The relationship between this line and the data from experiment 3 will be discussed in the next section.

As can be seen, thresholds increased as the pre-processed³ TFE of the masker increased. It should be noted that the total energy of the masker decreased as the pre-processed TFE increased. This is because more of the FM signal is filtered by the roex filter as more of the FM signal's energy occurs outside of the filter's bandwidth. Thus, thresholds in Fig. 7 are increasing while the total energy of the AM masker is decreasing, once again contradictory to the level-detection scheme.

Figure 8 shows the thresholds obtained when the clipped-FM masks of experiment 2 are used to generate the AM masker. The TFE of the post-clipped, pre-processed masker is plotted along the abscissa; the pre-processed sweep rate is the parameter. The dashed line is the data from Fig. 7. Most of the lines tend to coincide, except for S1's where the data parallel the dashed line but are shifted to the left. Consistent with experiment 2, the results indicate that the TFE of the FM masker has more effect on probe threshold than the sweep rate.

This is confirmed in Fig. 9, which replots the data from

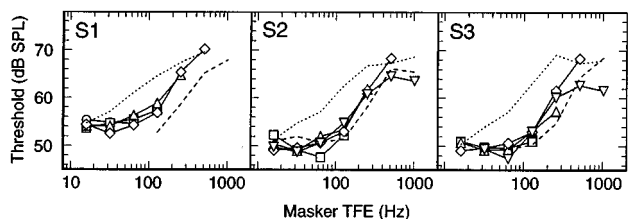


FIG. 8. Probe-detection thresholds for AM masks that were created by processing the FM masks that were used in experiment 2. The abscissa represents the TFE of the FM masker after its frequency was clipped but before it was processed. The symbols represent the same frequency sweep rates as in Fig. 4. The data from Fig. 7 are replotted with the dashed line. The averages of the data in Fig. 4 are plotted with dotted lines.

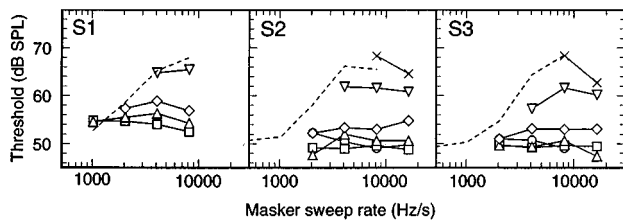


FIG. 9. Replotting of data in Fig. 8, but with the pre-processed frequency sweep rate as the abscissa and the pre-processed TFE as the parameter; the symbols represent the same TFEs as in Fig. 5. The dashed line replots the data from Fig. 7.

Fig. 8 using pre-processed sweep rate as the abscissa and the post-clipped, pre-processed TFE as the parameter; this plot is similar to the plot in Fig. 5. As in Fig. 5, the data indicate the dominance of FM TFE over FM sweep rate. The dotted lines in Figs. 8 and 9 represent the average of the data in Figs. 4 and 5, respectively.

IV. DISCUSSION

Experiment 1 replicated the phenomenon that is the focus of this paper: the increased masking of a probe with increased masker TFE. The results are consistent with those of Zwicker.

Experiment 2 was designed to determine which properties of the masker were related to the increased thresholds. Two previously correlated properties of the masker, sweep rate during probe presentation and TFE, were independently manipulated and thresholds were found to be much more dependent upon TFE than sweep rate. This is somewhat surprising since it had been previously thought that increased sweep rate was responsible for the increased thresholds. Smoorenberg and Coninx (1980) noted that the increased sweep rate increased the frequency region which is excited by the masker during the probe presentation; since they demonstrated that the loudness of a frequency sweep increases with the width of the region swept (Coninx and Smoorenberg, 1979), they suggested that the increased thresholds were due to the increased loudness of the masker.

Since thresholds are relatively unaffected by the sweep rate of the masker during probe presentation, the properties of the masker either before or after probe presentation are responsible for the increased thresholds, possibly suggesting forward or backward masking. The results are in contrast to standard frequency effects for these types of masking, however, in that the most masking in our experiment is produced for maskers that reach frequencies most distant from the probe's frequency.

A. Detection of energy splatter

Although a simple account in terms of masking seems untenable, a consideration of the possible role of energy splatter may make such an account more plausible. Specifically, it may be that the thresholds obtained at small-masker TFEs are relatively low because the subject is listening "off frequency" and is detecting energy splatter produced by adding the probe. As the TFE increases this strategy becomes less advantageous, partly because the signal-to-masker ratio

at the nominal frequency of the probe becomes increasingly large, and partly because the splatter is being masked. We at first discounted this possibility because, even though the probe is brief (10 ms), its equivalent rectangular bandwidth is only 150 Hz and all but 1 dB of its energy is in a 146-Hz band centered at 1 kHz. However, subsequent observations seem consistent with splatter detection at low TFEs. First, varying the phase of addition of the probe to a zero-TFE masker (increment detection) had little effect on threshold. If the subject was basing decisions on the output of a filter centered at the probe frequency, this phase manipulation might be expected to produce a significant change in threshold; the task would be increment detection, a situation where the relative phase of addition would have a large effect. If, however, the subject were detecting splatter then, because of the similarity in the spectra, little phase effect might be seen. Second, adding a broadband noise whose level was sufficient to raise the probe-alone threshold by 20 dB, and thus presumably to mask the splatter, raised the probe threshold more at low masker TFEs than at high. This is also consistent with the detection of splatter at low TFEs. Finally, reducing the masker level from 60 to 35 dB reduced the probe threshold less at low TFEs than at high. Again this is consistent with the splatter hypothesis: the thresholds are "artificially" low at the higher masker level and low TFE because of splatter detection.

The question raised by these observations is whether the splatter detection at low TFEs, together with simple level detection at large TFEs, can account for all of the TFE effect. This is very unlikely. In both the noise and 35-dB masker conditions, the threshold at a TFE of 512 Hz was still at least 10-dB higher than at a TFE of zero. Thus, there is still a substantial effect of TFE even when splatter detection is almost certainly not involved.⁴ Also, it seems that level detection cannot account for the basic phenomenon: the large amount of masking seen for FM maskers with large TFEs. Even with *no* auditory filtering, the average level of the masker is approximately 5–7 dB lower than that expected based on level detection.⁵ In summary, splatter detection may be involved in the present experiments, but we think that it affects only the magnitude of the TFE effect, not the presence of the effect itself.

B. Envelope fluctuations

The results from experiment 2 discounted some possible causes of the general masking phenomenon shown in experiment 1, but they did not suggest an account for the data. Experiment 3 examined the hypothesis that the TFE effect reflects changes in the envelope fluctuations at the output of an auditory filter centered at the frequency of the probe. The AM maskers used in experiment 3 attempted to simulate *only* the envelope fluctuations due to the masker that might occur at the output of such a filter.

Threshold changes obtained with AM maskers in experiment 3 were similar to those obtained with their corresponding FM maskers. Thresholds increased by as much as 20 dB and the slopes of the threshold functions were similar to those found in experiments 1 and 2. We have suggested that detection of energy splatter may partially account for the

relatively low thresholds at small TFEs. Such would also be the case in experiment 3: In the limiting case of $TFE=0$, where splatter detection is most likely, the maskers used in all the experiments are identical.

It should be noted that the point at which the thresholds begin to increase above the pure-tone masker threshold is at a larger TFE in experiment 3 than experiment 2 (compare data with dotted lines in Figs. 7–9), i.e., the thresholds obtained with the AM masker are generally lower than those obtained with the FM masker. This could be for a number of reasons. Although the masker energy at 1 kHz is the same for both the FM masker and its AM equivalent, the AM masker has less energy at the frequency locations above and below 1 kHz because of the processing done to transform the masker from a wideband FM signal to a narrow-band AM signal.⁶ Thus, excitation or splatter from the probe that has spread to these other frequency regions is not masked as well by the AM masker and thresholds will be lower than for the FM masker. Another possibility, of course, is that the roex filter used to obtain the AM equivalent is not correct. At this point, we consider the discrepancy in thresholds to be of secondary importance: the important aspect is the similarity in the form of the functions for FM and the AM equivalent, specifically those relating probe threshold to TFE and to pre-processed TFE.

C. Envelope-shape discrimination

Consistent with the hypothesis that motivated experiment 3, we propose that in the present experiments detection of the probe, particularly at large TFEs, involves detecting differences between the envelopes produced by the masker and the masker-plus-probe. Such envelope-shape discrimination qualitatively accounts for the differences in detectability shown for maskers with different TFEs. Figure 10 illustrates the notion of envelope-shape discrimination, here when clipped FM was used to produce the AM masker. In Fig. 10 the envelopes have been normalized to have equal total areas and a dc shift has been added in the upper panel (one of the curves was lowered) to facilitate comparison of the envelope shapes. The solid lines show one period of the envelopes of two different AM maskers used in experiment 3; the dashed lines show the envelopes of the masker plus probe⁷ over the same interval. Clearly, the addition of the same probe produces a larger change in envelope shape for the masker with the smaller TFE (upper panel). This should lead to higher detectability of the probe or, consistent with the data, a lower probe threshold.

D. Envelope spectra

Another possibility is suggested by Green *et al.* (1992) who argue that in certain circumstances detection may be based on differences in the amplitude spectrum of the envelope. Envelope spectral-shape discrimination also provides a qualitative account of the present data. Figure 11 shows the normalized spectra of the envelopes from Fig. 10. Clearly,

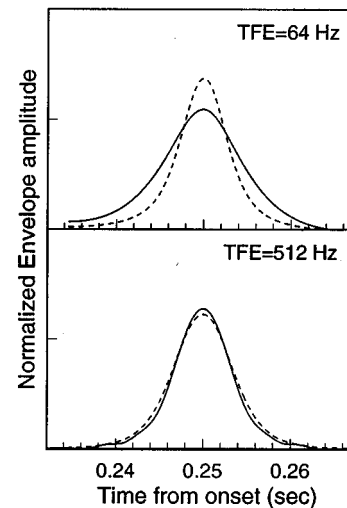


FIG. 10. The response of a roex filter centered at 1 kHz to stimuli used in experiments 1 and 2. The lines represent one period of the response's envelope. The solid lines are responses to only maskers. The dashed line represents the average response to masker+probe, where the envelopes were averaged across probe phase. The probe in both panels was at 65 dB. The pre-clipped TFE for both panels is 1024 Hz. The masker in the upper panel has an actual TFE of 64 Hz while the masker in the lower panel has an actual TFE of 512 Hz. The envelopes have been normalized to unit area for comparison purposes.

the spectra shown in the upper figure differ more than those shown in the lower figure and, consistent with the data, should yield lower probe thresholds.

Although both the envelope-shape and the envelope amplitude spectrum provide a basis of accounting for the present data, we believe that waveshape is more realistic. The argument against a role for envelope spectrum is similar to that regarding the long-term amplitude spectrum of sounds. For example, sounds with identical amplitude spectra, such as a waveform and its time-reversed version, generally sound considerably different. Similarly, sounds with identical envelope-amplitude spectra but different phase

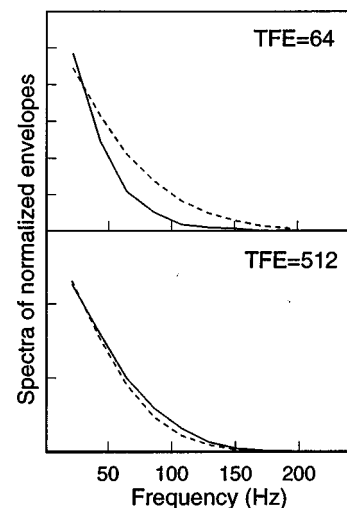


FIG. 11. The spectra of the envelopes in Fig. 10. The solid line corresponds to the spectra of the masker envelope; the dashed line corresponds to the spectra of the masker+probe envelope.

spectra may sound considerably different. A concrete example is the recent work by Patterson (1994a,b) on ramped versus damped sinusoids in which the envelopes are time-reversed versions of each other. The envelope amplitude spectra are thus identical but the sounds can be readily discriminable. We argue that this indicates sensitivity to the temporal structure of the envelope.

E. Final observations

The envelope-shape model is preliminary and needs considerable refinement and quantitative implementation. In a sense the present model is rather trivial: It argues that the more dissimilar the envelopes of the waveforms, the more discriminable they will be. A visually based argument, such as we have done in discussing Fig. 10, is not satisfactory. Clearly, "envelope shape" needs more precise specification. One difficulty in such refinement stems from the nonadditivity of envelopes: the envelope of the masker-plus-probe is generally not the sum of the envelopes of the masker and of the probe. This appears to preclude an analytic examination of various decision strategies, such as envelope cross-correlation, and thus makes such an examination a rather tedious simulation exercise. More importantly, the data from the present experiments are not ideal for developing the model. First, there is the complication of splatter detection. Second, the envelopes used in experiment 3 are not simple and, in the context of a model, derivation of the envelopes, especially for the FM maskers, requires assumptions about the characteristics of the auditory filter. This is somewhat secondary to the basic notion of the model. Other experiments, perhaps using amplitude modulated broadband noise, may provide a basis for examining envelope-shape discrimination models with minimal dependence on assumptions about the shape of the auditory filter.

As a qualitative account, however, envelope-shape discrimination is quite appealing. It provides a description of the counter-intuitive results of the present experiments. It also offers the possibility of describing the problematic results, described in the Introduction, obtained with swept-frequency maskers. Finally, the basic notion of the importance of envelope shape generalizes to other detection and discrimination tasks and, using the strategy of experiment 3, provides a testable alternative hypothesis.

lead;1V. SUMMARY AND CONCLUSIONS

(1) Probe detection thresholds increased substantially as the amount of FM in the masker increased, consistent with Zwicker's (1974) findings.

(2) Thresholds increased in the same manner when masker frequency sweep rate was fixed and the total frequency excursion (TFE) increased, while thresholds remained approximately constant when TFE was fixed and sweep rate increased. Thus, TFE, not sweep rate, primarily determines the amount of masking.

(3) Detection of energy splatter at small TFEs appears to account for part of the TFE effect. When splatter detection is unlikely, however, the thresholds at large TFEs were still

substantially higher (10 dB) than those at small TFEs and were too large to be explained by simple level detection.

(4) When the FM maskers from the first two experiments were transformed to AM maskers in such a way that the simulated envelope fluctuations at 1 kHz were the same for both, thresholds increased similarly with TFE.

(5) The results are consistent with the concept that detection of a brief probe in an FM masker involves envelope-shape discrimination.

ACKNOWLEDGMENTS

This research was supported by research Grant No. DC00683 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. The authors would like to thank the anonymous reviewers for their helpful comments on improving this paper.

¹The probe was added at the end of the processing and not at the beginning because the inverse filter would improve the ability of the listener to use off-frequency masking: Significant energy in the signal interval would exist distant from 1 kHz, and since the gain of the inverse filter is much larger for these frequencies, the probe would be enhanced at frequencies distant from 1 kHz.

²The inverse filter does not introduce any instantaneous frequency fluctuations in the AM signal because the filter is symmetric around 1 kHz and has zero phase response.

³The term "pre-processed" will be used from here on to indicate the FM signal that is used to generate the AM masker before filtered by the roex filter, i.e., at point (A) in Fig. 6.

⁴There are two strong indications that splatter detection was not occurring when the masker level was 35 dB. First, the expected change in threshold with probe phase (approximately 7 dB) was observed with a TFE of zero, indicating that simple level detection (or envelope-shape discrimination), and not splatter detection, was occurring. Second, the probe threshold observed with zero-TFE, 90° probe phase, was only 6 dB above detection threshold for the probe in quiet. At such a relatively low probe level, splatter almost certainly would have been inaudible.

⁵The prediction based on level detection requires an estimate of the Weber fraction obtained when splatter detection is not involved. The estimates used were based on: (1) increment detection in the 20-dB noise ($10 \log \Delta I/I = 5$ dB); (2) intensity discrimination with a pedestal gated identically to the 10-ms probe ($10 \log \Delta I/I = 4$ dB), which was corrected for the effects of gating with the raised-cosine window; (3) increment detection with a 35-dB masker ($10 \log \Delta I/I = 3$ dB). With TFE=512 Hz the observed threshold for this subject was 68 dB. Thus, assuming level detection, the effective masker level was 65–67 dB. The level of the (unfiltered) masker was 60 dB, 5–7 dB lower than that based on level detection. The discrepancy will increase, of course, if the masker is filtered, as it certainly is by the auditory system.

⁶The inverse filter was bandlimited around 1 kHz because the enormous gain of the inverse filter at frequencies distant from 1 kHz would unnecessarily amplify noise and negligible frequency components.

⁷Since phase of the probe was randomized during the experiments, the probe and masker were added 72 different times, with the phase of the probe incremented by 5° each time, spanning the full range of 360°. The envelopes of each summation were calculated using the Hilbert transform and the 72 envelopes were averaged to produce the envelopes displayed with the dashed lines in Fig. 10.

Ahumada, Jr., A., and Lovell, J. (1971). "Stimulus features in signal detection," *J. Acoust. Soc. Am.* **49**, 1751–1756.

Coninx, F., and Smoorenburg, G. F. (1979). "Masking of short probe sounds by tone bursts with a sweeping frequency," *J. Acoust. Soc. Am. Suppl.* **1** **66**,

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

Gilkey, R. H. (1987). "Spectral and temporal comparisons in auditory masking," in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ).

- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics* (Wiley, New York).
- Green, D. M., Berg, B. G., Huanping, D., Eddins, D. A., Onsan, Z., and Nguyen, Q. (1992). "Spectral shape discrimination of narrow-band sounds," *J. Acoust. Soc. Am.* **92**, 2586–2597.
- Hall, J. W., Haggard, M. P., and Fernandes, M. A. (1984). "Detection in noise by spectro-temporal pattern analysis," *J. Acoust. Soc. Am.* **76**, 50–56.
- Jeffress, L. A. (1964). "Stimulus-oriented approach to detection," *J. Acoust. Soc. Am.* **36**, 766–774.
- Kemp, S. (1982). "Masking period patterns of frequency modulated tones of different frequency deviations," *Acustica* **50**, 63–69.
- Kidd, Jr., G., Mason, R., Brantley, M. A., and Owen, G. A. (1989). "Roving-level tone-in-noise detection," *J. Acoust. Soc. Am.* **86**, 1310–1317.
- Moore, B. C. J. (1989). *An Introduction of the Psychology of Hearing* (Academic, London).
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram, and speech threshold," *J. Acoust. Soc. Am.* **72**, 1788–1803.
- Patterson, R., and Moore, B. C. J. (1986). "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, edited by B. C. J. Moore (Academic, London), pp. 123–177.
- Patterson, R. D. (1994a). "The sound of a sinusoid: Spectral models," *J. Acoust. Soc. Am.* **96**, 1409–1418.
- Patterson, R. D. (1994b). "The sound of a sinusoid: Time-interval models," *J. Acoust. Soc. Am.* **96**, 1419–1428.
- Richards, V. M., Heller, L. M., Green, D. M. (1991). "Detection of a tone in a narrow band of noise: The energy model revisited," *Q. J. Psychol.* **43A**, 481–501.
- Rosen, S., Baker, R. J., and Kramer, S. (1992). "Characterizing changes in auditory filter bandwidth as a function of level," in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demany, and K. Horner (Pergamon, London), pp. 171–177.
- Smoorenburg, G. F., and Coninx, F. (1980). "Masking of short probe sounds by tone bursts with a sweeping frequency," *Hear. Res.* **3**, 301–316.
- Zwicker, E. (1974). "Mithörschwellen und Erregungsmuster stark frequenz-modulierter töne," *Acustica* **31**, 243–256.