Published in final edited form as:

J Acoust Soc Am. 2003 December ; 114(6 Pt 1): 3273–3282.

Psychometric functions for informational masking

Robert A. Lutfi^{a)}, Doris J. Kistler^{b)}, Michael R. Callahan^{c)}, and Frederic L. Wightman^{b)}
Waisman Center, University of Wisconsin, Madison, Wisconsin 53705

Abstract

The term *informational masking* has traditionally been used to refer to elevations in signal threshold resulting from masker uncertainty. In the present study, the method of constant stimuli was used to obtain complete psychometric functions (PFs) from 44 normal-hearing listeners in conditions known to produce varying amounts of informational masking. The listener's task was to detect a pure-tone signal in the presence of a broadband noise masker (low masker uncertainty) and in the presence of multitone maskers with frequencies and amplitudes that varied at random from one presentation to the next (high masker uncertainty). Relative to the broadband noise condition, significant reductions were observed in both the slope and the upper asymptote of the PF for multitone maskers producing large amounts of informational masking. Slope was affected more for some listeners and conditions while asymptote was affected more for others; consequently, neither parameter alone was highly predictive of individual thresholds or the amount of informational masking. Mean slopes and asymptotes varied nonmonotonically with the number of masker components in a manner similar to mean thresholds, particularly when the estimated effect of energetic masking on thresholds was subtracted out. As in past studies, the threshold data were well described by a model in which trial-by-trial judgments are based on a weighted sum of levels in dB at the output of independent auditory filters. The psychometric data, however, complicated the model's interpretation in two ways: First, they suggested that, depending on the listener and condition, the weights can either reflect a fixed influence of masker components on each trial or the effect of occasionally mistaking a masker component for the signal from trial to trial. Second, they indicated that in either case the variance of the underlying decision variable as estimated from PF slope is not by itself great enough to account for the observed changes in informational masking.

I. INTRODUCTION

In recent years there has been a growing interest in the detection of signals in uncertain noise backgrounds. The interest is largely motivated by a desire to understand detection in real-world settings where background noise varies unexpectedly from one moment to the next. Typically in studies, the effect of noise uncertainty has been expressed as an elevation in threshold for the signal relative to a "minimal uncertainty" condition in which the noise is either fixed or varies little from one presentation to the next. The term *informational masking* was first used in this context to describe the elevations in threshold resulting from masker uncertainty (Pollack, 1975; Watson and Kelly, 1981; Lutfi, 1990). Because informational masking has been commonly associated with the effect of attentional processes on threshold, it is often distinguished from *energetic masking*, which is attributed

© 2003 Acoustical Society of America.

^{a)}Also at Department of Communicative Disorders, University of Wisconsin—Madison.

^{b)}Current address: Dept. of Psychological and Brain Sciences, University of Louisville, Life Sciences Bldg., Louisville, KY 40292.

^{c)}Current address: Epic Systems Corp. 5301 Tokay Blvd., Madison, WI 53711.

PACS numbers: 43.66.Ba, 43.66.Fe [MRL]

to the competition of signal and noise at the auditory periphery (i.e., cochlea or auditory nerve). Some authors, in fact, have argued that the term informational masking might be better used to refer to any masking that is not energetic (Durlach *et al.*, 2003). In this paper we use the term according to its original definition inasmuch as the focus is on conditions of masker uncertainty. However, the two definitions are not mutually exclusive, and the actual amount of informational masking we report in different conditions of this study is an estimate of the amount of masking that is not energetic (cf. Oh and Lutfi, 1998).

What is currently known about informational masking comes almost exclusively from results of adaptive psychophysical procedures that estimate signal strength corresponding to a single level of performance. Much has been learned from these data. However, for understanding the nature of the decision process it is more useful to know how performance varies over a range of signal levels, as given by the psychometric function (PF). Two features of the PF are of particular interest: its slope and its upper asymptote. The slope or spread of the PF reflects variability of the underlying decision process. Values of slope can, therefore, be used to provide strong tests of models that assume a particular form for the decision variable. Asymptotic performance is also of interest in cases wherein the listener's attention might stray from the psychophysical task, or wherein the noise may simply be "confused" for the signal on some number of trials. In such cases, asymptotic performance provides a measure of lapse rate or frequency of confusions. Lapses and confusions also affect PF slope to the extent that they add variance to the underlying decision variable (cf. Green, 1995).

Both parameters of the PF are relevant to the understanding of informational masking. To date, however, no clear picture has emerged regarding the effects of masker uncertainty on these parameters. Perhaps the most extensive data come from a study by Allen and Wightman (1995). These investigators obtained PFs from 17 children (3–4 years of age) and 13 adults in a task involving the detection of a fixed-frequency, pure-tone signal in the presence of noise and a random-frequency, pure-tone masker. They used a logistic function to fit PFs to the data of individual listeners and reported only the results of the fits. On average, masker uncertainty appeared to have the effect of decreasing the slope of the PF, though the results were variable; many listeners actually showed shallower PFs in the absence of the masker. One problem with this study was that the fitted functions were forced to have an upper asymptote of 100 percent. Hence, the estimates of PF slope were to some extent confounded with asymptotic performance. Other tone-detection studies reporting PFs have used a random-frequency multitone masker instead of a pure tone; the results of these studies are mixed. Neff and Callaghan (1987), using simultaneous masker–tone complexes, and Watson and Kelly (1981), using sequential masker–tone complexes (see their Fig. 3.6), report a systematic decrease in the slope of the PF with increasing masker uncertainty. Tang and Richards (2002) and Wright and Saberi (1999), however, report significant individual differences in the effect of masker uncertainty on slope for simultaneous-masker complexes. These studies used many fewer listeners and also did not estimate upper asymptotic performance in conjunction with slope. Moreover, the latter two studies used adaptive procedures that can potentially lead to bias in estimates of slope (Leek, 2001). Finally, results from discrimination and identification studies using nonadaptive procedures have generally reported a reduction in PF slope with increasing masker uncertainty (Arbogast, Mason, and Kidd, 2002; Kidd *et al.*, 1994; Kidd, Mason, and Arbogast, 2002), although here too there are some discrepancies (Watson and Kelly, 1981; see their Fig. 3.3). To our knowledge, no study to date has reported systematic estimates of upper asymptotic performance in conditions of informational masking.

The goal of the present study was twofold: to provide more extensive documentation than past studies of the effects of masker uncertainty on the PF, and to use measures of slope and

asymptote to test two interpretations of informational masking. To this end, PFs from 44 normal-hearing listeners were measured in a minimal-uncertainty condition and in comparable conditions known to produce high levels of informational masking. The task of the listener was to detect a pure-tone signal in the presence of a broadband noise masker (minimal uncertainty condition) and in the presence of a multitone masker with frequencies that varied at random from one presentation to the next. Broadband noise is commonly identified as a “minimal uncertainty” masker in such cases, as the spectra of individual samples appear qualitatively more similar to one another than do the spectra of the individual multitone maskers (Neff and Green, 1987; and for a quantitative account see Lutfi, 1990; and Oh and Lutfi, 1998).

With regard to the test of interpretations, the data are analyzed within the context of a model that successfully describes thresholds in these conditions (Lutfi, 1993; Oh and Lutfi, 1998). According to the model, listeners are assumed to make trial-by-trial judgments based on the weighted sum of levels (in dB) at the outputs of independent auditory filters. Here, the weights are intended to reflect the role of attentional factors that are believed to play a major role in informational masking. They have, however, at least two different interpretations. The first is that they represent a constant proportional contribution of each auditory filter to the decision variable on each trial. We refer to this as the fixed-weight hypothesis. Fixed-weight models have been used successfully to predict threshold in a variety of conditions of informational masking (Lutfi, 1992, 1993; Oh and Lutfi, 1998; Wright and Saberi, 1999), as well as to draw inferences regarding internal noise in these conditions (Berg, 1990; Richards, Tang, and Kidd, 2002; Tang and Richards, 2002). These models are the most common in the literature on informational masking. An equally viable interpretation, however, is that the weights represent the exclusive contribution of a single but different auditory filter on each trial. We refer to this as the variable-weight hypothesis. In effect, the variable-weight hypothesis proposes that a nonsignal filter is mistaken for the signal filter on some proportion of trials. Lutfi (1992, 1993) has analyzed the predictions of this hypothesis, treating the weights as a vector of random binomial variables rather than as a vector of constants.¹ Evidence in support of the hypothesis comes from the observation that high levels of informational masking often occur in conditions where the signal and masker are perceptually similar to one another and so are more likely to be confused (Arbogast *et al.*, 2002; Brungart, 2001; Kidd *et al.*, 2001; Kidd *et al.*, 1994; Kidd *et al.*, 2002; Neff, 1995). Particularly relevant in this regard is the report of plateaus in the PF (asymptotic-like behavior) where signal and masker components are of near-equal level (Brungart, 2001).

The fixed- and variable-weight hypotheses make similar predictions for signal thresholds; however, they make different predictions for the psychometric function. For fixed weights the random-frequency multitone masker must produce greater variance in the decision variable than does the broadband noise. For the broadband noise the output of any auditory filter will change little from one presentation to the next, whereas for the multitone masker it will vary greatly depending on whether or not a tone falls in the bandpass of the filter. The fixed-weight hypothesis, therefore, predicts that the slopes of the PFs for multitone maskers will be much shallower than those for the broadband noise. The fixed-weight hypothesis also makes a strong prediction for how the slopes of the PFs will change with the number of tones in the multitone masker. If the range of potential frequencies is fixed, there should be some critical number of masker components for which the probability is 0.5 that one or more components will fall within a given filter. At this point the variance of filter outputs with fixed weights will approach a maximum and the slope of the PF will have its lowest value. For greater or fewer masker components the variance of filter outputs will be less and

¹Lutfi (1992) provides an analytic development of two possible versions of a variable-weight model, one treating the weights as a binomial random variable, the other as a normal random variable. Only the former is tested here.

the slopes of the PF will increase. The fixed-weight hypothesis, therefore, predicts a nonmonotonic relation between the slope of the PF and the number of masker components similar to what is seen in the threshold data (Neff and Green, 1987; Oh and Lutfi, 1998).² The variable-weight hypothesis also predicts some effect on slope, but differs fundamentally from the fixed-weight hypothesis in that the predominant effect is predicted to be a change in upper asymptote. Because the outputs of nonsignal auditory filters are assumed to be independent of the output of the signal filter, mistaking a nonsignal filter for a signal filter has much the same effect as a momentary lapse in attention. Hence, while the variable-weight hypothesis predicts some differences in PF slopes for multitone and broadband noise maskers, it predicts clearly lower asymptotic performance for multitone maskers. These specific predictions are tested in the following experiment.

II. METHODS

A. Stimuli

The signal was a 1000-Hz pure tone presented simultaneously with the maskers. Both the multitone maskers and the broadband noise maskers consisted of frequency components for which the distribution of component amplitudes was Rayleigh and the distribution of component phases was rectangular. Hence, as the number of components of the multitone masker increased the multitone maskers began to approximate samples of the broadband noise masker. The number of components comprising the multitone maskers ($m=2,10,20,40,200,400,906$) was fixed within blocks of trials but varied across blocks. For each presentation, the multitone maskers were created by choosing m components (represented as complex numbers) at random from the fast-Fourier transforms (FFTs) of a sample of Gaussian noise. The frequencies were sampled uniformly over the range of 0.1 to 10 kHz, excluding the frequencies between 920–1080 Hz. Sampling was restricted such that no two neighboring frequencies were separated by less than 11 Hz. The multitone maskers were then synthesized by computing the inverse FFT of the m randomly selected components. The duration of all maskers was 370 ms. Hence, for the broadband noise the effective number of masker components was $m=3608$, corresponding to a resolution of 2.7 Hz (1/370 ms). To be consistent with past studies using similar conditions, the average total power of all maskers was kept constant at 60 dB SPL regardless of the number of masker components. The signal and maskers were gated on and off together with 10-ms, cosine-squared onset and offset ramps. All stimuli were played at a sampling rate of 44.1 kHz using 16-bit, digital-to-analog conversion (Tucker Davis Technologies DD1). Stimuli were presented monaurally through Sennheiser model HD-414 headphones to individual listeners seated in a double-walled, IAC sound-attenuation chamber.

B. Procedure

Psychometric functions for detection of the signal in the multitone maskers, in the noise masker, and in quiet were obtained using a cued, two-interval, forced-choice (2IFC), staircase procedure. The listener heard three sounds on each trial separated by 1 half-second. The first sound was the signal presented at 60 dB SPL in isolation as a cue to aid detection. The subsequent two sounds both contained maskers but only one contained a signal. The listener's task was to identify the sound containing the signal, which was equally likely to be either sound. Feedback indicating the correct response was given after each trial. In this regard the procedure was typical of adaptive procedures commonly used in past studies of informational masking. The primary difference was that the level of the signal on each trial did not depend on the listener's response, but rather was predetermined for each block. The

²The relation is predicted to be similar to but not identical to that of the threshold data because of the effect of energetic masking (cf. Oh and Lutfi, 1998), a point that we revisit in the Results and Discussion sections.

sequence began with the signal at a clearly audible level; for the next three trials it was reduced in 8-dB steps, then for the next three trials it was increased in 8-dB steps so as to return to the starting level. This up-down pattern was repeated five times for each block, yielding 40 trials per block. Practice trials were given until the listener appeared familiar with the task (almost always one block of trials). Listeners then completed three or more trial blocks for each condition, yielding at least 15 trials each for the highest and lowest signal levels, and 30 trials for each intermediate signal level. The starting level of the signal was selected such that performance for the intermediate levels fell between chance and best performance. If this criterion was not met an additional block of trials was obtained with the starting level adjusted to achieve the desired range of performance. Also, in rare cases where the slope of the PF appeared to be unusually shallow, an additional block of trials was collected to provide a more reliable estimate of upper asymptotic performance. An additional block of trials was necessary for approximately 5 to 10 percent of all runs, and then typically only for the first run of a condition. The order of conditions was quiet, followed by broadband noise, followed by the seven multitone masker conditions. The multitone masker conditions were presented in random order. Listeners participated in 1.5-h sessions on each day with breaks and required approximately 5 min to complete each block of 40 trials. Completion of all conditions required 3–4 sessions, depending on the listener.

C. Listeners

The data analyzed in this study are a subset of data previously collected as part of a larger study of individual differences in informational masking (see Lutfi *et al.*, 2003). The data are from 44 listeners (14 males and 30 females) of a total of 84 listeners who had participated in that study.³ Of these 84 listeners, 30 were not included because their data were collected using an adaptive procedure rather than the method of constant stimuli. Another ten were not included because they were of an age (less than 6 years) for which thresholds in these conditions tend to be significantly elevated relative to other age groups (cf. Lutfi *et al.*, 2003). All of the listeners had pure-tone thresholds less than 15 dB HL at the octave frequencies from 250 to 4000 Hz, as determined by standard audiometric tests (ANSI, 1989). Sixteen of the listeners were adults (19–30 years), 16 were late school-aged (11–16 years), and 12 were early school-aged (6–10 years). Though age is not a factor under consideration in this study, it was anticipated, based on previous work (e.g., Oh *et al.*, 2001), that within- and between-listener variability in slope and asymptote estimates would be large, and that a large pool of listeners would therefore be necessary to observe significant differences across conditions. The pooling of data across age seemed justified inasmuch as Lutfi *et al.* (2003) showed only small, statistically significant differences in mean thresholds across the age groups selected here. Moreover, that study concluded that the same factor responsible for differences across age groups is responsible for individual differences within age groups.

D. The psychometric function

The data relating proportion of correct responses p to signal level L was used to estimate psychometric functions for each listener and each condition using a maximum-likelihood criterion (cf. Wichmann and Hill, 2001). The assumed form of the PF between chance and asymptotic performance was a logistic. This form is commonly used in studies of auditory detection and typically yields excellent fits. The complete expression for the PF is

$$\hat{p} = 0.5 + (0.5 - \lambda)(1 + e^{-(L - \alpha)/\beta})^{-1}, \quad (1)$$

³Neff, Kessler, and Dethlefs (1996) report larger amounts of informational masking in female listeners.

where $\hat{p}$ denotes the estimate of p , $0 \leq \lambda \leq 0.5$ is the lapse parameter determining upper asymptotic performance, α is a threshold parameter corresponding to the signal level yielding performance halfway between chance and asymptotic performance, and $\beta > 0$ is the slope parameter (larger values corresponding to smaller slopes). With the exception of the specified bounds for λ and β , no constraints were placed on any of the free parameters.

Figure 1 shows representative fits from two listeners for the $m=20$ multitone masker condition (unfilled symbols) and the broadband noise minimal uncertainty condition (filled symbols). These fits demonstrate how λ and β capture the effect of masker uncertainty in terms of the change in upper asymptote and slope of the psychometric function, respectively. The PFs of the first listener (top panel) are consistent with the fixed-weight hypothesis; the PFs of the second listener (bottom panel) are consistent with the variable-weight hypothesis. For the purposes of this study the analysis focuses on how λ and β change with the number of components in the multitone masker, and whether one or the other of these parameters is significantly greater for the multitone maskers than for the broadband noise. The variable-weight hypothesis predicts changes in both β and λ across conditions, but predicts that the greatest changes will occur in λ . The fixed-weight hypothesis predicts significant changes in β alone. Moreover, the fixed-weight hypothesis predicts a nonmonotonic relation between β and the number of masker components mirroring that obtained for the masked thresholds.

III. RESULTS

A. Goodness of fit

Goodness of fit was evaluated using deviance, a measure that is a monotonic transformation of the likelihood of obtaining a particular set of data given the presumed form of the PF is correct (cf. Wichman and Hill, 2001). For the present application the deviance measure is

$$D = 2 \sum_{i=1}^K \left\{ n_i p_i \log \left(\frac{p_i}{\hat{p}_i} \right) + n_i (1 - p_i) \log \left(\frac{1 - p_i}{1 - \hat{p}_i} \right) \right\}, \quad (2)$$

where K is the number of data points for each PF, and n_i is the number of trials for the i th data point (15 or 30). Slightly greater than 5 percent of the fitted functions were estimated to have a likelihood of less than 5 percent assuming the present model is correct. Only one of these functions had estimated parameters that identified it as a clear outlier. Eliminating these data did not essentially change the pattern of results, so they were included in the analysis to offset any bias toward underdispersion in the data.

In addition to estimating the deviance for each fitted function, maximum-likelihood estimates of confidence intervals were obtained for each estimate of α ('threshold') and β . Confidence intervals were not obtained for λ as these would be rather meaningless given the strict upper and lower bounds on this parameter. The procedure used to estimate confidence limits is described by Wichmann and Hill (2001). The approach is to use a "bootstrapping" technique to estimate the sampling distributions of α and β , and then to take confidence limits of the estimated distributions. Using the values of percent correct at each signal level from each fitted function, and assuming binomially distributed percent correct values, a simulated percent correct at each signal level was randomly drawn and a logistic was fit to these values. This was repeated 10 000 times to provide 10 000 estimates of each of the three logistic parameters. The 2.5% and 97.5% points from the sampling distributions of α and β were then taken as the 95% confidence limits. Figure 2 and Figure 3 give scatterplots of these values for each listener (different points) and each condition (different panels). The intent of these figures is to show the distribution of estimates for the upper and lower

bounds, rather than to identify individual confidence intervals. Nonetheless, individual intervals can often be identified by a pair of points located horizontally from one another. Figure 2 shows the confidence intervals to be roughly symmetric about α , changing in size with number of masker components. For 2–40 masker components the size of the intervals is in the vicinity of 20 dB; for the quiet threshold condition (0 components) and for maskers comprised of 200 or more components they are in the neighborhood of 10 dB. Figure 3 shows the confidence intervals for β to be highly asymmetric, with a dramatic increase in upper bound as β increases. This is not surprising given the 0 lower bound for this parameter and the inverse exponential relation between β and PF slope; because of this relation, large increases in β have little impact on slope when β is already large. Again, it is because of the difficulty associated with getting reliable estimates of β when PF slope is shallow that we have chosen in this study to evaluate effects by obtaining estimates of β averaged over a large number of listeners.

B. Analysis of mean parameter values

Figure 4 shows the mean thresholds across listeners (filled symbols) plotted against the number of masker components. Thresholds correspond to the 71-percent-correct point of the individually fitted PFs; 95-percent confidence limits are also shown. We report threshold values corresponding to 71 percent correct to permit comparison to past studies using adaptive procedures; however, the values obtained were essentially identical to those obtained for α .⁴ A two-way ANOVA of the individual thresholds revealed a significant main effect of age group [$F(2,41) = 8.48, p < 0.001$], with younger listeners having on average higher thresholds than adults, and a significant interaction between age group and number of masker components [$F(16,328) = 293, p < 0.001$], with younger listeners having higher mean thresholds when the number of masker components was small ($m \leq 20$). These effects were quite small, however, so only the mean thresholds across age groups are shown (for further information on age effects on threshold the reader is referred to the paper by Lutfi *et al.*, 2003). The main effect of number of masker components was significant [$F(8,328) = 151, p < 0.001$] and replicates that obtained for very similar conditions using adaptive psychophysical procedures (cf. Neff and Green, 1987; Oh and Lutfi, 1998). In particular, the data show the characteristic nonmonotonic relation between threshold and number of masker components with highest thresholds for maskers comprised of 20–40 components. The highest thresholds are 12–15 dB greater than thresholds for the broadband noise condition, which also compares favorably with the mean magnitude of effects observed in past studies. The dashed line in Fig. 4 gives the estimate of energetic masking for these conditions using a roex auditory filter model (see Oh and Lutfi, 1998 for details regarding this derivation). The unfilled symbols give the difference between the obtained thresholds and the estimated amount of energetic masking. Hence, the unfilled symbols are estimates of informational masking alternatively defined as the amount of masking that is not energetic.

Figure 5 plots the mean β values in similar fashion to Fig. 4. A two-way ANOVA of the individual β 's revealed no significant main effect for age group [$F(2,41) = 0.381, p = 0.686$] and no significant interaction of age group and number of masker components [$F(16,328) = 0.732, p = 0.761$]; hence, only the mean β 's are shown. The absence of a significant age effect might appear inconsistent with the results of Allen and Wightman (1995) except that the children in Allen and Wightman's study were much younger (3–4 years versus greater than 6 years in the present study). Again, we intentionally chose our listeners so as to minimize the age effect on thresholds and on β based on the results of Lutfi *et al.* (2003). The main effect of number of masker components was significant [$F(8,328) = 20.27$,

⁴Note, α marks the midpoint of the PF, the point about which PF slope is symmetric. It is for this reason that confidence intervals were estimated for individual values of α rather than for individual values of threshold defined by the 71%-correct point.

$p < 0.001$], with the broadband noise yielding significantly smaller values of β than the multitone maskers comprised of less than 400 components. The largest values of β occur for the 10–20-component maskers, unlike the highest thresholds, which occur for 20–40 component maskers. Also, the curve for β is nearly symmetric about the maximum, whereas the curve for threshold is elevated when maskers are comprised of 40 or more components. These differences are to be expected as the added effect of energetic masking (dashed curve in Fig. 4) is predicted to elevate thresholds with little effect on β when maskers are comprised of 40 or more components (cf. Oh and Lutfi, 1998). The mean values of β agree much more favorably with the estimates of informational masking given by the unfilled symbols in Fig. 4—this in agreement with the fixed-weight hypothesis.

Figure 6 shows the mean values of λ (continuous curve) again plotted in the fashion of Fig. 4 and Fig. 5. The conversion to upper asymptotic performance as percent correct is $100(1-\lambda)$ and the conversion to lapse rate or confusion rate as a percentage of trials is 200λ . Again, confidence intervals are not shown here because of the upper and lower bounds on λ . Instead, we give the mean (continuous curve) and individual values for each condition (note values have been jittered along the abscissa for visibility). The range of mean values is quite small relative to the range of individual values, 0–0.05 versus 0–0.25 for the individual values. The range for the individual values corresponds to upper asymptotic performance between 75 and 100 percent correct or confusion rates between 0 and 50 percent. Once again there was no significant main effect for age group [$F(2,41) = 1.654, p = 0.204$] and no significant interaction of age group and number of masker components [$F(16,328) = 0.677, p = 0.817$]. The mean λ 's also changed significantly with the number of masker components in a manner similar to that of the mean thresholds [$F(8,328) = 7.454, p < 0.0001$] and the estimated amount of informational masking. In this respect, the data are also consistent with the variable-weight hypothesis.

C. Analysis of individual parameter values

Since the mean values of β and λ , and the estimated amount of informational masking all appeared to vary in a similar fashion with the number of masker components, we wished to determine whether there might be significant correlations between the individual values of these quantities. One question relevant to the fixed- versus variable-weight distinction is which parameter of the PF, β or λ , correlates more highly with the amount of informational masking. Figure 7 and Figure 8 give the individual values of β and λ plotted against the estimated amount of informational masking for all listeners and all conditions. The amount of energetic masking in this calculation was assumed to be the same for all listeners as given by the dashed curve in Fig. 4. Solid circles and unfilled squares represent the multitone masker, and broadband noise masker conditions, respectively. The figures show the values β and λ clustering at or near zero, as might be expected since zero is the lower bound of these parameters. The largest estimates of β and λ , and the greatest variability in these estimates, occur for the multitone masker conditions. Interestingly, the figures reveal little or no discernible relation between the estimated amount of informational masking and either of the two parameters of the PF. This result would be difficult to explain except for the possibility that changes in informational masking are mediated differently for different listeners and/or conditions. In particular, no strong relation with either parameter would be expected if for some listeners and conditions the weights are variable, predominantly affecting asymptote, while for other listeners and conditions the weights are fixed, affecting only slope.

Figure 9 shows the individual estimates of β plotted against λ for all listeners and conditions. The multitone masker and broadband noise conditions are represented by the filled circles and unfilled squares as before; the quiet condition, which was not shown in Fig. 7 and Fig. 8, is represented by the unfilled triangles. Taken across all conditions, the two parameters of

the PF appear nearly orthogonal to one another (Pearson correlation of -0.11) consistent with the interpretation of mutually exclusive effects. Moreover, the orthogonal relation is most evident for the multitone maskers for which thresholds vary most widely. Within-listener variability in the estimates of β and λ was too great and the effect of number of masker components on these parameters was too small to discern any systematic dependence of this result on the number of masker components for individual listeners. Moreover, individual listener data were not always characterized exclusively by changes in one or the other parameter.

IV. DISCUSSION

The results of the present study reveal significant reductions in both the slope and upper asymptote of the PF for conditions of informational masking. Variability in estimates was too large to discern any systematic dependence of slope and asymptote on the number of masker components for individual listeners. However, the mean values of these parameters taken across listeners varied nonmonotonically with the number of masker components in a manner similar to that of the mean thresholds, particularly when the estimated effect of energetic masking was subtracted from thresholds. Despite the common relation of mean slope and asymptote to mean threshold, there was little correlation between the individual values of slope and asymptote across listeners and conditions. Indeed, when the individual values of these parameters were plotted against one another, the two parameters appeared nearly orthogonal to one another. The results indicate that, depending on the particular listener and condition, masker uncertainty has primarily one of two effects on the PF: a reduction in slope with little change in asymptote, or a reduction in asymptote with comparatively little change in slope. An important consequence of this result is that neither the values of slope alone nor the values of asymptote alone are generally predictive of the amount of informational masking.

The results are consistent with a model that assumes trial-by-trial judgments are based on a weighted sum of levels at the output of independent auditory filters. They suggest, however, that the weights are determined differently for different listeners in different multitone masker conditions. For some listeners and conditions they appear to reflect a fixed contribution of signal and nonsignal auditory filters on each trial (fixed-weight hypothesis). In such cases, the effect of masker uncertainty is a reduction in slope. For other listeners and conditions the weights appear to reflect some likelihood that a nonsignal auditory filter will be confused for the signal filter on any given trial (variable-weight hypothesis). In these cases, the effect of masker uncertainty is predominantly a reduction of upper asymptote. Both effects occurred with near-equal frequency for the different listeners and multitone masker conditions of this experiment. The results suggest that confusions are a significant component of informational masking. Consistent with this view are the results of numerous studies identifying perceptual similarity between the signal and masker as an important factor in informational masking (Arbogast *et al.*, 2002; Kidd *et al.*, 1994, 2001, 2002; Neff, 1995). Perhaps the most compelling evidence of this type comes from a study by Lutfi and Alexander (2002), in which some listeners were shown to make significant numbers of errors on randomly intermixed trials in which the signal was presented alone.

The effect of confusions and, more generally, the distinction between fixed- and variable weights has important implications for predicting informational masking. In fixed-weight models the listener's response is largely stimulus driven such that, if the weights are known, one should be able to predict the listener's response on any given trial from the parameters of the signal and masker on that trial (cf. Alexander and Lutfi, 2002). In variable-weight models the trial-by-trial responses are not so easily predicted since it is not possible to know on any given trial whether the listener is likely to make a confusion. This distinction allows

us to get a sense of the relative importance of these two processes by determining the degree to which the data can be strictly predicted from the values of stimulus parameters. Following Oh and Lutfi (1998), we summarize the effect of fixed weights with a single free parameter n representing the number of nonsignal auditory filters contributing equally to the decision variable on each trial. This greatly simplifies the approach while still providing accurate predictions for threshold. Accordingly, the variance in the decision variable for the 2IFC task is

$$\frac{n\sigma^2 + \sigma_e^2}{2} = \beta^2, \quad (3)$$

where $\frac{1}{2}\sigma_e^2$ is an internal noise variance given by β^2 for the quiet-threshold condition, and σ^2 is the variance in level at the output of a typical nonsignal filter. The value of σ^2 is the stimulus-driven component of the fixed-weight model; it is computed directly from the spectra of the maskers on each trial and varies with the number of masker components (cf. Oh and Lutfi, 1998). The dashed curve of Fig. 5 gives σ , while the continuous curve is a prediction for β obtained according to Eq. (3) with n as the only free parameter; n in this case has a value of 0.3. (Note: a value less than 1 implies that the combined weight for nonsignal filters is less than that for the signal filter.) The continuous curve of Fig. 4 gives the fixed-weight model prediction for threshold. The prediction is $E + d' \sqrt{n}\sigma$, where E is an estimate of energetic masking (dashed curve) and $d' = 0.78$ is the index of sensitivity corresponding to 71-percent correct in the 2IFC task (for complete details regarding this derivation see Oh and Lutfi, 1998). In this case the best-fitting value of n is 3.5, in agreement with past data (cf. Oh and Lutfi, 1998). The two estimates of n should agree if the fixed-weight model were correct. That the two estimates are largely discrepant from one another suggests that there is another factor affecting threshold that is not tied to variance in the stimulus. To evaluate the extent to which this second factor is the effect of confusions we repeated the analysis, this time using only data for which λ was less than 0.001 (for which there were few if any confusions). If confusions alone are responsible for the discrepancy the two estimates should agree this time. The resultant estimate of n for thresholds remained essentially the same, 3.6, while the estimate for β increased to 0.6. This result suggests that confusions are, at best, only partly responsible for the discrepancy. More importantly, it suggests that while changes in PF slope appear on average to covary with the amount of informational masking, the changes in slope are not large enough to account for informational masking in terms of the fixed-weight model.

The clear failure of the fixed-weight model in this case is noteworthy given that many past studies have nonetheless made accurate predictions for individual and mean thresholds assuming fixed weights. The outcome can be understood, however, in light of the fact that past predictions have been made for thresholds without constraints imposed by the underlying psychometric functions. When making predictions for threshold alone, the variance of the decision variable, $n\sigma^2$, need only be scaled by the free parameter n to yield the best fit. However, when additionally evaluating the PFs, the variance of the decision variable must agree with PF slope. Remarkably, the data from the present study do show good agreement between the two, except for the scalar n . This result complicates the interpretation of fixed-weight models despite their demonstrated ability to accurately describe individual thresholds in many studies. This outcome, together with the finding of less than perfect asymptotic performance for many PFs, suggests that fixed-weight assumptions are unlikely to yield meaningful estimates of other parameters of the decision process in conditions of informational masking (cf. Berg, 1990; Richards, Tang, and Kidd, 2002; Tang and Richards, 2002).

V. CONCLUSIONS

1. Conditions of informational masking produce significant reductions in both the slope and the upper asymptote of the PF. Slope is affected more for some listeners and conditions, while asymptote is affected more for other listeners and conditions.
2. Individual values of slope and upper asymptote are nearly orthogonal; hence, neither parameter alone is highly predictive of individual thresholds or the amount of informational masking.
3. Mean slope and asymptote parameters of the PF vary nonmonotonically with the number of masker components in a manner similar to the amount of informational masking, where the amount of informational masking is given as the difference between masked threshold and the estimated amount of energetic masking.
4. Masked thresholds are well described by a fixed-weight model. However, the effect of confusions and the failure to obtain similar values of model parameters from threshold and psychometric data complicate the interpretation of this model.

Acknowledgments

The research was supported by grants from the NIDCD (R01 DC1262-10 and R01 HD23333-08). The authors wish to acknowledge the helpful comments of Joshua Alexander, Dr. Marjorie Leek, and two anonymous reviewers.

References

- Alexander JM, Lutfi RA. Sensation level and decision weights as factors underlying informational masking in hearing-impaired listeners. *Abstracts of the Assoc. for Research in Otolaryngology* 2002;Vol. 25:180.
- Allen P, Wightman F. Effects of signal and masker uncertainty on children's detection. *J. Speech Hear. Res* 1995;38:503–511. [PubMed: 7596115]
- Arbogast TL, Mason CR, Kidd GK Jr. The effect of spatial separation on informational and energetic masking of speech. *J. Acoust. Soc. Am* 2002;112:2086–2098. [PubMed: 12430820]
- Berg BG. Observer efficiency and weights in a multiple observation task. *J. Acoust. Soc. Am* 1990;88:149–158. [PubMed: 2380443]
- Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am* 2001;109:1101–1109. [PubMed: 11303924]
- Durlach NI, Mason CR, Kidd G Jr, Arbogast TL, Colburn HS, Shin-Cunningham B. Note on informational masking. *J. Acoust. Soc. Am.* 2003 (in press).
- Green DM. Maximum-likelihood procedures and the inattentive observer. *J. Acoust. Soc. Am* 1995;97:3749–3760. [PubMed: 7790653]
- Kidd G Jr, Arbogast TL, Mason CR, Walsh M. Informational masking in listeners with sensorineural hearing loss. *J. Assoc. Res. Otolaryngol* 2001;3:107–119. [PubMed: 12162362]
- Kidd G Jr, Mason CR, Arbogast TL. Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns. *J. Acoust. Soc. Am* 2002;111:1367–1376. [PubMed: 11931314]
- Kidd G Jr, Mason CR, Deliwalla PS, Woods WS, Colburn SH. Reducing informational masking by sound segregation. *J. Acoust. Soc. Am* 1994;95(6):3475–3480. [PubMed: 8046139]
- Kidd G Jr, Mason CR, Rohtla TL, Deliwalla PS. Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns. *J. Acoust. Soc. Am* 1994;104(1):422–431. [PubMed: 9670534]
- Leek MR. Adaptive procedures in psychophysical research. *Percept. Psychophys* 2001;63(8):1279–1292. [PubMed: 11800457]
- Lutfi RA. How much masking is informational masking? *J. Acoust. Soc. Am* 1990;88:2607–2610. [PubMed: 2283433]

- Lutfi RA. Informational processing of complex sound. III. Interference. *J. Acoust. Soc. Am* 1992;91:3391–3401. [PubMed: 1619116]
- Lutfi RA. A model of auditory pattern analysis based on component-relative entropy. *J. Acoust. Soc. Am* 1993;94:748–758. [PubMed: 8370881]
- Lutfi RA, Alexander JM. Informational masking without maskers. *J. Acoust. Soc. Am* 2002;111:2470.
- Lutfi RA, Kistler DJ, Oh EL, Wightman FL, Callahan MR. One factor underlies individual differences in auditory informational masking within and across age groups. *Percept. Psychophys* 2003;65(3): 396–406. [PubMed: 12785070]
- Neff DL, Callahan BP. Psychometric functions for multicomponent maskers with spectral uncertainty. *J. Acoust. Soc. Am. Suppl. 1* 1987;81:S53.
- Neff DL, Green DM. Masking produced by spectral uncertainty with multicomponent maskers. *Percept. Psychophys* 1987;41:409–415. [PubMed: 3601622]
- Neff DL, Kessler CJ, Dethlefs TM. Sex differences in simultaneous masking with random-frequency maskers. *J. Acoust. Soc. Am* 1996;100:2547–2550. [PubMed: 8865658]
- Oh EL, Lutfi RA. Nonmonotonicity of informational masking. *J. Acoust. Soc. Am* 1998;104:3489–3499. [PubMed: 9857508]
- Oh EL, Wightman F, Lutfi RA. Children's detection of pure-tone signals with random multitone maskers. *J. Acoust. Soc. Am* 2001;109:2888–2895. [PubMed: 11425131]
- Pollack I. Auditory informational masking. *J. Acoust. Soc. Am* 1975;57:S5.
- Richards VM, Tang Z, Kidd G Jr. Informational masking with small set sizes. *J. Acoust. Soc. Am* 2002;111:1359–1366. [PubMed: 11931313]
- Tang Z, Richards VA. Internal noise invariance across two informational masking tasks. *J. Acoust. Soc. Am* 2002;111:2470.
- Watson, CS.; Kelly, WJ. The role of stimulus uncertainty in the discrimination of auditory patterns. In: Getty, DJ.; Howard, JH., Jr, editors. *Auditory and Visual Pattern Recognition*. Hillsdale, NJ: Erlbaum; 1981.
- Watson CS, Kelly WJ, Wroton HW. Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty. *J. Acoust. Soc. Am* 1976;60:1176–1188. [PubMed: 977844]
- Wichmann FA, Hill NJ. The psychometric function. I. Fitting, sampling, and goodness of fit. *Percept. Psychophys* 2001;63(8):1293–1313. [PubMed: 11800458]
- Wright BA, Saberi K. Strategies to detect auditory signals in small sets of random maskers. *J. Acoust. Soc. Am* 1999;105:1765–1775. [PubMed: 10089600]

Fig. 1.

Representative PFs from two listeners fitted according to Eq. (1). Data are from the $m = 20$ -multitone masker condition (unfilled symbols) and the broadband noise minimal uncertainty condition (filled symbols). The top panel (listener SUM) shows data consistent with the fixed-weight hypothesis. The effect of masker uncertainty in this case is a reduction in the slope of the PF for the multitone masker. The second panel (listener SVB) shows data consistent with the variable-weight hypothesis. The effect of masker uncertainty in this case is a reduction in upper asymptote of the PF for the multitone masker.

Fig. 2.

Scatterplot of the estimated upper and lower confidence limits of α obtained by “bootstrapping” technique (see the text for details). Each pair of points above and below the diagonal represents the upper and lower limit of the estimated confidence interval for a single listener and a single condition, respectively. The number of masker components for each condition is given in the upper-left-hand corner of each plot.

Fig. 3.
Scatterplot of the estimated upper and lower confidence limits of β plotted as in Fig. 2.
Arrows indicate values falling above the upper limit of the plot at 60 dB.

Fig. 4.

Mean signal thresholds corresponding to the 71%-correct point on the PF are plotted as a function of the number of masker components m , where $m = 0$ corresponds to the quiet-threshold condition and $m = 3608$ corresponds to the broadband-noise condition. The data are means of 44 listeners with vertical bars representing the 95%-confidence limits. The dashed curve is an estimate of energetic masking for these conditions, and the continuous curve is a prediction for the total amount of masking based on the standard deviation σ of level at the output of a typical nonsignal auditory filter, as computed from the trial-by-trial stimuli (see the text for details).

Fig. 5.

Same as Fig. 4 except data are the mean slope parameter (β) estimates of the 44 listeners. The dashed curve is the standard deviation σ of the level at the output of a typical nonsignal auditory filter as computed from the trial-by-trial stimuli. The continuous curve is a

prediction for $\beta = \sqrt{(n\sigma^2 + \sigma_e^2)/2}$, where σ_e is an internal noise estimate corresponding to β for the quiet-threshold condition.

Fig. 6.

Same as Fig. 4 except individual mean values of the lapse parameter (λ) represented by the continuous curve are accompanied by the individual values of this parameter. Values have been jittered along the abscissa for visibility.

Fig. 7. Scatterplot of individual slope parameter (β) estimates against individual amounts of informational masking. Filled circles designate multitone-masker conditions; unfilled squares, the broadband noise (minimal uncertainty) condition.

Fig. 8. Scatterplot of individual lapse parameter (λ) estimates against individual amounts of informational masking. As in Fig. 7, filled circles designate multitone-masker conditions; unfilled squares, the broadband noise (minimal uncertainty) condition.

Fig. 9.

Scatterplot of individual slope parameter (β) and lapse parameter (λ) estimates for all conditions and all listeners. As in Fig. 7, filled circles designate multitone-masker conditions; unfilled squares, the broadband noise (minimal uncertainty) condition; and unfilled triangles, the threshold in quiet condition.