Excitation-based and informational masking of a tonal signal in a four-tone masker[a]

Lori J. Leibold[b] and Jack J. Hitchens
Department of Allied Health Sciences, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Emily Buss
Department of Otolaryngology/Head and Neck Surgery, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

Donna L. Neff
Boys Town National Research Hospital, Omaha, Nebraska 68131

(Received 22 July 2009; revised 31 December 2009; accepted 5 January 2010)

This study examined contributions of peripheral excitation and informational masking to the variability in masking effectiveness observed across samples of multi-tonal maskers. Detection thresholds were measured for a 1000-Hz signal presented simultaneously with each of 25, four-tone masker samples. Using a two-interval, forced-choice adaptive task, thresholds were measured with each sample fixed throughout trial blocks for ten listeners. Average thresholds differed by as much as 26 dB across samples. An excitation-based model of partial loudness [Moore, B. C. J. *et al.* (1997). J. Audio Eng. Soc. **45**, 224–237] was used to predict thresholds. These predictions accounted for a significant portion of variance in the data of several listeners, but no relation between the model and data was observed for many listeners. Moreover, substantial individual differences, on the order of 41 dB, were observed for some maskers. The largest individual differences were found for maskers predicted to produce minimal excitation-based masking. In subsequent conditions, one of five maskers was randomly presented in each interval. The difference in performance for samples with low versus high predicted thresholds was reduced in random compared to fixed conditions. These findings are consistent with a trading relation whereby informational masking is largest for conditions in which excitation-based masking is smallest.
[DOI: 10.1121/1.3298588]

PACS number(s): 43.66.Dc, 43.66.Lj [CJP] Pages: 2441–2450

I. INTRODUCTION

Studies of simultaneous masking have shown large detrimental effects of masker-frequency uncertainty, often created by randomizing the frequency content of a multi-tonal masker each time it is presented (e.g., Watson *et al.*, 1975; Neff and Green, 1987; Kidd *et al.*, 1994; Oh and Lutfi, 1998; Alexander and Lutfi, 2004; Richards and Neff, 2004; Durlach *et al.*, 2005). For example, simultaneous-masking studies with fixed-frequency sinusoidal signals and random-frequency multi-tonal maskers have reported as much as 50 dB of masking for trained listeners (e.g., Neff and Green, 1987). The "informational" masking produced by varying the spectral content of the masker can be observed even when the frequency components that comprise the masker are restricted from falling within a presumed auditory filter centered on the signal. The use of a "protected region" centered on the signal frequency is intended to reduce the potential contributions of peripheral (i.e., energetic) masking. Thus, whereas energetic masking is believed to reflect interactions between the signal and masker at the auditory periphery, informational masking is believed to reflect interactions occurring within the central auditory system.

The processes underlying informational masking are incompletely understood and may well vary across stimuli, tasks, and listeners. One potential contributor to informational masking is a failure of sound source segregation (e.g., Bregman, 1990)—the process by which acoustic components are identified as coming from one or more sources. Evidence for this comes from reports of significant reductions in informational masking when cues that promote sound source segregation are introduced (e.g., Kidd *et al.*, 1994; Neff, 1995; Richards and Neff, 2004). Manipulations that decrease the similarity of the target relative to the masker can also yield substantial release from informational masking (e.g., Kidd *et al.*, 2002; Durlach *et al.*, 2003). Note that many cues believed to affect sound source segregation also influence target-masker similarity. Another related mechanism which likely contributes to informational masking is selective auditory attention—the ability to attend to a relevant target sound and ignore irrelevant, interfering sounds. Results from early studies of informational masking (e.g., Watson *et al.*, 1976; Spiegel *et al.*, 1981; Neff and Green, 1987) were interpreted as indicating that listeners were not able to attend only to information in the signal frequency region. Consistent with this explanation, Lutfi and colleagues (e.g., Lutfi, 1993; Lutfi

---

[a] Portions of these results were presented to the 32nd MidWinter Meeting of the Association for Research in Otolaryngology in Baltimore, MD, February, 2009.

[b] Author to whom correspondence should be addressed. Electronic mail: leibold@med.unc.edu

et al., 2003, Alexander and Lutfi, 2004) modeled informational masking in terms of the number and frequency range of auditory filters monitored when a listener is asked to detect a tonal signal at a fixed frequency in the presence of a random-frequency, multi-tonal masker. In that approach, data with little evidence of informational masking were modeled as resulting from a highly frequency-selective process, characterized in terms of a very narrow attentional filter. On the other hand, data with evidence of extensive informational masking were modeled as resulting from a combination of auditory filter outputs, characterized in terms of a wide attentional filter.

Interpreting effects of masker-frequency uncertainty is complicated by the observation of large within- and between-subjects variability in performance for multi-tonal masking conditions, even when the degree of masker-spectral uncertainty is modest (e.g., Neff and Callaghan, 1988; Neff and Dethlefs, 1995; Wright and Saberi, 1999; Alexander and Lutfi, 2004; Richards and Neff, 2004; Durlach *et al.*, 2005). Several studies have shown that thresholds in conditions with little or no masker-frequency uncertainty (e.g., masker samples fixed across intervals of each trial or across the entire block of trials) are often considerably higher than the absolute threshold for the signal in quiet (e.g., Neff and Callaghan, 1988; Neff and Dethlefs, 1995; Wright and Saberi, 1999; Alexander and Lutfi, 2004; Richards and Neff, 2004; Durlach *et al.*, 2005). Moreover, multi-tonal masker samples can differ widely in masking effectiveness when particular samples are selected at random from the pool of samples used for testing with minimal uncertainty (e.g., Neff and Callaghan, 1987; Wright and Saberi, 1999). For example, Wright and Saberi (1999) measured detection threshold for a 1000-Hz pure tone in the presence of a ten-tone masker; they found a range of more than 20 dB in average threshold across the ten masker samples tested. Differences in peripheral excitation patterns across the masker samples may be responsible for the substantial variability in masked threshold observed within a given subject.

In contrast to differences in performance within a listener, individual differences in susceptibility to informational masking appear to play a role in the substantial between-subjects variability in masked threshold observed for a fixed masker sample (e.g., Neff and Callaghan, 1987; Alexander and Lutfi, 2004; Durlach *et al.*, 2005). For example, Alexander and Lutfi (2004) reported thresholds ranging from 15- to 52 dB sound pressure level (SPL) across 16 normal-hearing listeners asked to detect a 2000-Hz pure tone in the presence of a ten-tone, fixed-frequency, simultaneous masker. Alexander and Lutfi (2004) noted that this variability in performance was inconsistent with expectations based on excitation-based masking and suggested that some listeners may have had difficulty perceptually segregating the signal from the masker even in the absence of stimulus uncertainty. Similarly, Durlach *et al.* (2005) reported average thresholds for a 1000-Hz signal ranging from 33 to 48 dB SPL across ten listeners in the presence of an eight-tone, fixed-frequency, simultaneous masker. Although individual differences do not in themselves indicate informational masking, the magnitude of these differences contrasts sharply with the well-established stability across listeners of thresholds in broadband noise and suggests non-peripheral contributions elevating thresholds for at least the poorer performers.

The current study examined the degree to which differences in both excitation-based and informational masking contribute to the differences in masking observed across multi-tonal masker samples and across listeners. In the first set of *Fixed* conditions, thresholds were measured for a 1000-Hz signal presented simultaneously with each of 25 different four-tone masker samples. The model of partial loudness from Moore *et al.* (1997) was used to estimate the contribution of peripheral, excitation-based masking to thresholds observed across masker samples. The relative contributions of excitation-based and informational masking to performance for conditions with high masker-spectral uncertainty were examined in a subsequent set of *Random* conditions, in which one of five masker samples was selected at random for each interval throughout a block of trials. Performance was compared across two random conditions: one in which the five masker samples were those with the smallest estimates of excitation-based masking, and one in which the five samples were those with the largest estimates of excitation-based masking.

II. METHOD

A. Listeners

Ten adults (19–37 years) with normal-hearing sensitivity participated in all conditions, including authors LL (L5) and JH (L121). All listeners had air-conduction thresholds less than 20 dB hearing level (HL) (re: ANSI, 2004) at octave frequencies from 250 to 8000 Hz and reported no known history of chronic ear disease. None of the listeners had more than 2 years of musical training, with the exception of author JH (L121) who is a trained musician.

All listeners had previously participated in similar psychoacoustic experiments using multi-tonal maskers. Although few studies have systematically examined training effects for these conditions, the available data suggest that some individual listeners may improve over time (Neff and Callaghan, 1988; Neff and Dethlefs, 1995). The rationale for including listeners with previous experience using similar multi-tonal maskers is that evidence of training effects appears to be limited to approximately the first 600 trials (Neff and Callaghan, 1988).

B. Stimuli and conditions

The signal was a 300-ms, 1000-Hz sinusoid, including 5-ms onset/offset ramps (raised cosine). Twenty-five masker samples were randomly generated prior to the experiment and stored to disk. The same 25 masker samples were used for all listeners. Maskers were multi-tonal complexes with four sinusoidal components, presented simultaneously with the signal (when present) for 300 ms. Masker frequencies were drawn randomly from a uniform distribution on a linear frequency scale with a range of 300–3000 Hz, excluding 920–1080 Hz. The frequency range from 920 to 1080 Hz extends beyond the equivalent rectangular bandwidth centered on 1000 Hz (Glasberg and Moore, 1990). Masker start-

TABLE I. Component frequencies (Hz) and predicted masked thresholds (dB SPL) for the 25 masker samples based on Moore *et al.* (1997), ordered from lowest to highest predicted threshold. The subscripts indicate samples used in the *RanLow* and *RanHigh* conditions.

	Masker components (Hz)				
Masker sample	Frequency 1	Frequency 2	Frequency 3	Frequency 4	Predicted threshold (dB SPL)
16_{low}	311	1758	2246	2805	6.3
11_{low}	447	1870	2276	2716	6.3
12_{low}	308	411	1638	2350	6.3
24_{low}	512	521	1914	2430	8.2
7_{low}	464	468	605	1443	12.8
9	585	659	1372	1973	19.9
17	491	627	1320	2737	19.4
6	314	719	1440	2067	23.1
20	309	372	726	1446	23.7
4	413	474	479	1205	25.3
3	772	2348	2478	2586	27.4
21	312	343	465	1159	29.4
14	505	1140	1311	1728	32.1
25	448	826	2094	2961	32.6
18	489	826	1447	2573	33.6
8	717	1177	2142	2173	34.1
22	364	677	830	854	37.6
2	835	1236	1858	1989	37.7
19	704	1128	1214	1279	38.1
10	704	1110	1513	2069	39.4
5_{high}	787	836	877	2564	39.4
23_{high}	396	847	903	1560	41.3
13_{high}	348	884	1148	2920	44.1
1_{high}	511	919	1213	2675	44.8
15_{high}	344	867	1109	1868	45.0

ing phases were drawn from a uniform distribution with a range of $0-2\pi$. Each masker tone was equal amplitude, and the four-tone complex was presented at an overall level of 60 dB SPL (54 dB SPL per component). Table I shows the four frequencies that comprised each masker sample. The column on the right-hand side lists the corresponding Moore *et al.* (1997) model threshold predictions.

Stimuli were digitally summed and played through a 24-bit digital-to-analog converter (Digital Audio Labs, Chanhassen, MN) at a sampling rate of 20 kHz. Stimuli were presented monaurally to the listener's left ear via Sennheiser HD-25 earphones. The presentation of stimuli was controlled by a computer using custom software.

In *Fixed* conditions, a single masker sample was used on every presentation throughout a block of trials. All 25 samples were tested in separate conditions without any masker randomization. In two *Random* conditions, the masker presented in each interval was randomly selected with replacement from a subset of five masker samples. In the *RanLow* condition, the five samples with the lowest predicted thresholds comprised the pool of randomly selected maskers. In the *RanHigh* condition, the five samples with the highest predicted thresholds were used. Masker samples used in the *RanLow* and *RanHigh* conditions are indicated by the subscripts "low" and "high" in Table I, respectively.

C. Procedure

Thresholds were measured using a two-interval, forced-choice adaptive procedure that estimated 70.7% correct on the psychometric function (Levitt, 1971). The signal occurred in either interval with equal *a priori* probability. Each trial consisted of two, 300-ms observation intervals separated by a 400-ms interstimulus interval. A 300-ms feedback interval followed the listener's response, visually indicating the interval that contained the signal. The starting level of the adaptive track was 10–15 dB above the expected threshold. The initial step size was 4 dB, followed by a step size of 2 dB after the second reversal. Testing continued until ten reversals were obtained, and threshold was computed as the average signal level at the last eight reversals.

Listeners were tested individually in a double-walled, sound-treated room (Industrial Acoustics, Bronx, NY) in 1-h sessions that included regular breaks. An average of seven sessions per listener was required to complete the conditions. Four threshold estimates were obtained for each listener and condition. Listeners completed testing for the *Fixed* conditions prior to data collection for the *Random* conditions. For both *Fixed* and *Random* conditions, a complete randomized set of conditions was tested before moving to the next repetition of the conditions. Conditions were independently randomized for each listener. Data reported include threshold averages and estimates of the standard error (SE) across the four replications per condition.

III. RESULTS

A. Predicted thresholds using the Moore *et al.* (1997) excitation-based model of partial loudness

Thresholds in quiet for the 1000-Hz signal ranged from −7 to 17 dB SPL across listeners (average=2.9). Threshold for the 1000-Hz tone in the presence of each of the 25 masker samples was predicted using an excitation-based model of partial loudness (Moore *et al.*, 1997). The Moore *et al.* (1997) model is based on loudness as a function of frequency, taking into account changes in spread of excitation with level. This model was selected to generate threshold predictions based on the power spectrum of each masker. In applying the model, predicted thresholds were obtained by finding the level at which the partial loudness of the signal component was equal to 2 phons. Several investigators have shown that predictions generated using this general approach provide a reasonable account of observed masked thresholds across a range of masking paradigms (Van Der Heijden and Kohlrausch, 1994; Jesteadt *et al.*, 2007). Note, however, that this approach assumes that the long term power spectrum is the sole determinant of threshold, and that the temporal properties of the stimulus play no role in detection. Although there are counterexamples to these assumptions in the literature (e.g., Zwicker, 1976; Green, 1988; Moore and Glasberg, 1987; Richards, 1992; Richards and Nekrich, 1993), the degree of agreement between predicted and observed thresholds for the different masker conditions provides a framework within which to examine the contribution of peripheral excitation to differences in observed threshold across masker samples.

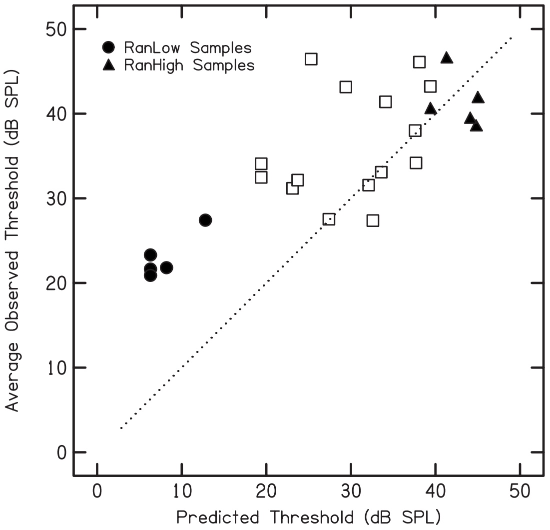

FIG. 1. Scatterplot of average observed threshold across listeners as a function of predicted threshold for the 25 masker samples. Filled symbols indicate data for samples used in the *RanLow* (circles) and *RanHigh* (triangles) conditions. Data on or near the dotted diagonal line indicate average observed thresholds that were well predicted by the model.

Thresholds predicted by the model for the 1000-Hz signal for the 25 masker samples are shown in Table I, rank ordered from lowest to highest predicted threshold. Despite the absence of masker components within 80 Hz of the signal frequency, predicted thresholds ranged from 6.3 to 45.0 dB SPL across masker samples. Predicted thresholds for *RanLow* maskers ranged from 6.3 to 12.8 dB SPL (mean =8.0), whereas predictions for *RanHigh* maskers ranged from 39.4 to 45.0 dB SPL (mean=42.9). Thus, the average difference in predicted threshold for *RanHigh* and *RanLow* conditions was approximately 35 dB. A relation between the proximity in frequency of the signal and neighboring masker tones is evident in Table I, with more masking for maskers with components closer to 1000 Hz.

B. Fixed conditions

Average masked thresholds across listeners for the 1000-Hz signal in the presence of each masker sample fixed across blocks are shown in Fig. 1, plotted as a function of predicted thresholds. Filled symbols indicate data for samples later used in the *RanLow* (circles) and *RanHigh* (triangles) conditions. Data on or near the dotted diagonal line indicate average observed thresholds that were well predicted by the model. Consistent with previous studies (e.g., Neff and Callaghan, 1987; Wright and Saberi, 1999), masker samples varied widely in masker effectiveness. Average masked thresholds ranged from 20.9 dB SPL (sample 11) to 46.6 dB SPL (sample 23), a range of approximately 26 dB. The preponderance of data above the dotted diagonal line in Fig. 1 indicates that threshold predictions based on excitation patterns often underestimated average observed thresholds. Nonetheless, there appears to be a relation between predicted and average observed thresholds.

Individual differences in the relationship between the predictions and the data were evident. Individual data for three listeners are shown in Fig. 2 to illustrate the range of results. No relation between the data and model predictions was found for L21 ($R^2=0.19$; $p>0.05$). In contrast, predicted thresholds for L121 accounted for half of the variance in observed thresholds ($R^2=0.51$; $p<0.01$). Finally, a strong relation between predicted and observed thresholds was observed for L99 ($R^2=0.73$; $p<0.01$).

Individual differences in masked threshold were more pronounced for some masker samples than for others. Figure 3 shows the difference between the observed and predicted thresholds for each masker sample and each listener, rank ordered by predicted threshold as in Table I. Data above the solid line indicate observed thresholds that were under-predicted by the model and data below the solid line indicate observed thresholds that were over-predicted by the model. Even without masker randomization, large individual differences in masked threshold were observed for many of the maskers, with thresholds spanning a range of up to 41 dB across listeners. Individual differences were most pronounced for samples predicted to produce little excitation-based masking. A Spearman's rank order correlation was computed to assess the relationship between predicted thresholds and the magnitude of individual differences, quantified as the standard deviation across threshold estimates. This correlation was −0.76 ($p<0.0001$), confirming that individual differences are greatest for masker samples with the lowest predicted thresholds. For example, the lowest predicted threshold (6.3 dB SPL) was found for masker samples 11, 12, and 16. Data-model differences as large as 35 dB were observed for these samples (L123). In contrast, the highest predicted threshold (45 dB SPL) was associated with

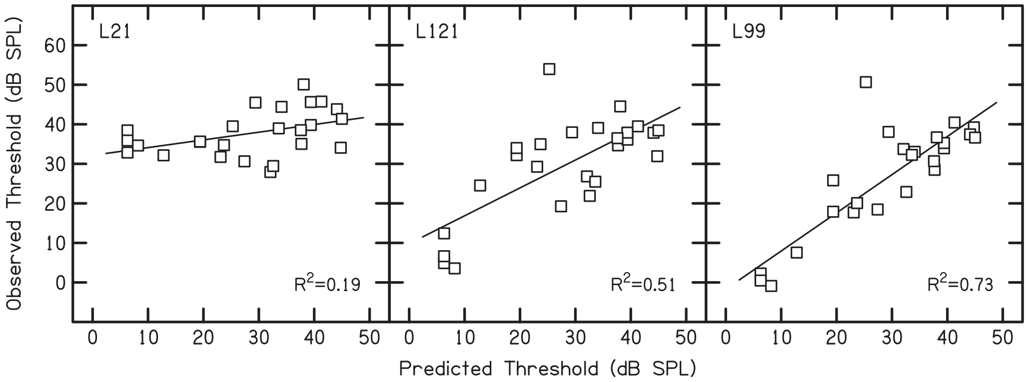

FIG. 2. Scatterplots of observed threshold as a function of predicted threshold for the 25 masker samples for three listeners (L21, L121, and L99). The solid line represents the best least-squares fit to all data points for each of the three listeners.

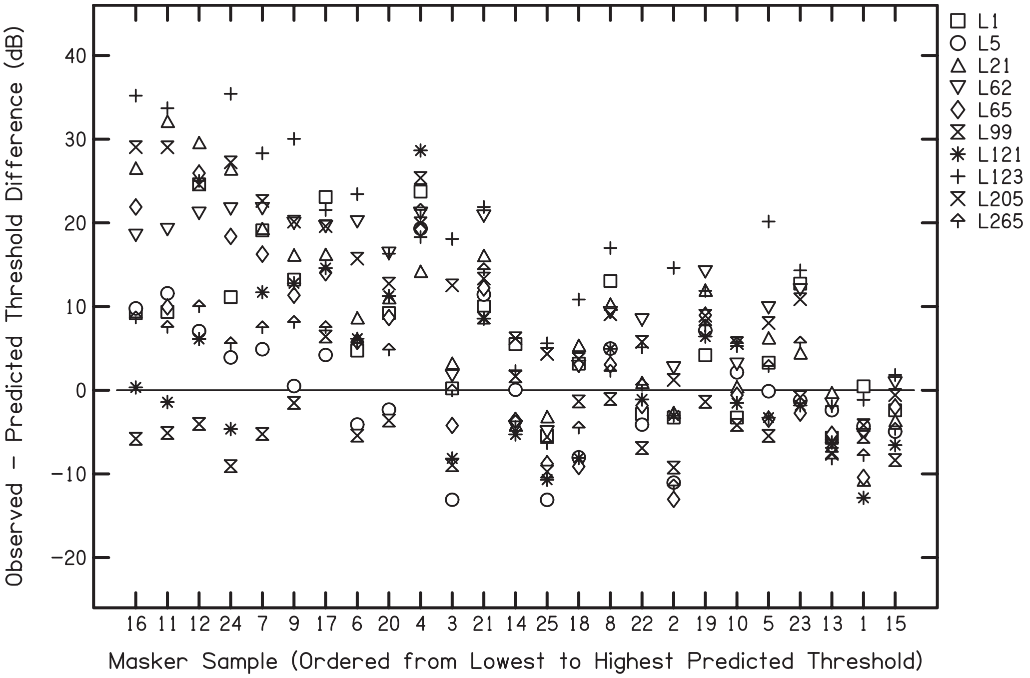

FIG. 3. Difference between observed and predicted thresholds as a function of masker sample for all listeners, ordered from lowest to highest predicted threshold. Data at or above the solid line indicate observed thresholds that were higher than the model predictions.

masker sample 15. The largest discrepancy between the data and model for any listener for this masker sample was −8 dB (L99).

Note that no evidence of significant practice effects was observed in the data. The average improvement in masked threshold across the first and fourth blocks of trials ranged from 0 to 6 dB across listeners (mean=2.5 dB). A within-subjects linear regression of threshold as a function of block number indicated no significant improvement in threshold with increasing block number $[F(1,99)=1.2;\ p=0.3]$.

C. Random conditions

The effect of masker-frequency uncertainty was estimated by comparing thresholds obtained when each masker sample was presented alone for a block of trials to thresholds obtained when the masker samples were drawn randomly on each presentation from the pool of five samples within each block. Figure 4 presents individual masked thresholds and the average across listeners, with listeners ordered on the abscissa by thresholds in the *FixLow* condition. Open circles

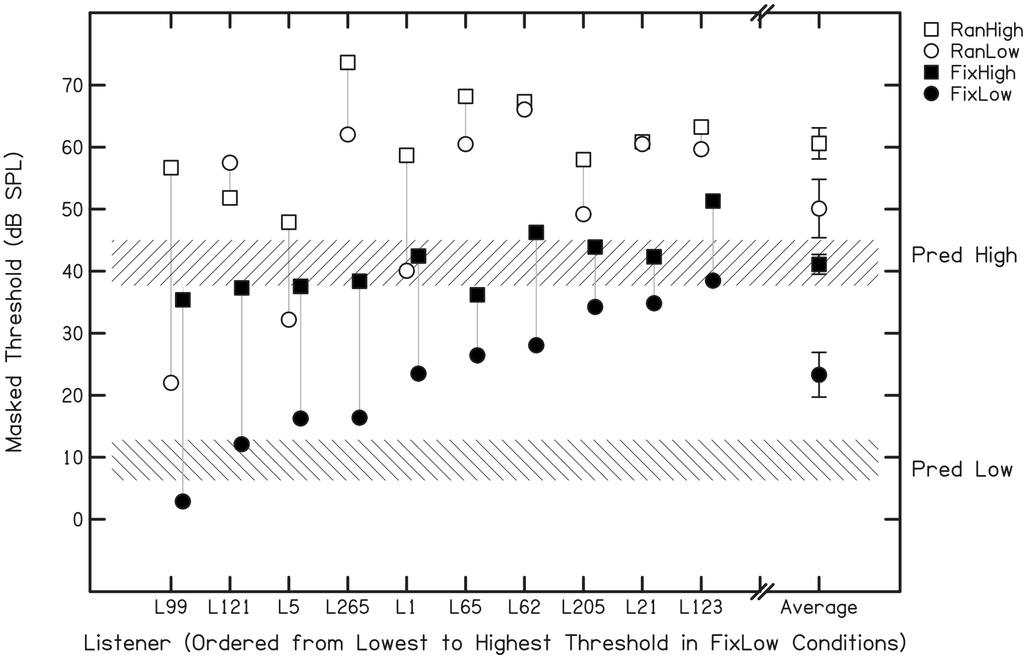

FIG. 4. Masked thresholds are plotted for individual listeners and for the average across listeners (with SEs) for *RanLow* (open circles) and *RanHigh* (open squares) conditions. Filled circles and squares show the average thresholds obtained in the fixed-presentation conditions across the five samples used for *RanLow* and *RanHigh*, respectively. The ranges of predicted thresholds for the five samples used for the *RanLow* and those for the *RanHigh* conditions are indicated with hatched shading.

and squares indicate *RanLow* and *RanHigh* conditions, respectively. Filled circles and squares indicate thresholds obtained in the associated fixed-presentation conditions. Error bars represent ±1 SE for the average data. The gray vertical lines emphasize the difference in threshold between conditions with low and high predicted thresholds for both random and fixed presentations. The ranges of masked thresholds predicted by the Moore *et al.* (1997) model for the five samples used for the *RanLow* samples and five *RanHigh* samples are shown by the hatched areas. Thus, the degree to which the data fall within these ranges provides an indication of whether or not the model provides a good account of observed thresholds.

Figure 4 clearly shows that the excitation-based model does not successfully predict the data for either random condition for any listener. Moreover, the estimated threshold for L121 in the *RanLow* condition is higher than in the *RanHigh* condition, offering the most serious violation of the model. Consistent with earlier reports (e.g., Neff and Callaghan, 1987; Wright and Saberi, 1999; Durlach *et al.*, 2005), however, masker-frequency uncertainty increased masked thresholds for all ten listeners. Thresholds were always higher for the *Random* conditions, even with only five samples randomly selected across presentations, than for the corresponding *Fixed* conditions. For data averaged across listeners, the threshold in the *RanLow* condition was 27.6 dB higher than the corresponding *FixLow* conditions (compare circles in Fig. 4). Similarly, the average threshold in the *RanHigh* condition was 19.5 dB higher than that for the *FixHigh* conditions (compare squares in Fig. 4). Note that for every listener the *RanLow* threshold was higher than the worst single *FixLow* threshold. Similarly, the *RanHigh* threshold was higher than the worst single *FixHigh* threshold. A two-way, repeated-measures analysis-of-variance, with two main factors of Uncertainty (random versus fixed) and Predicted Masking (low versus high), indicated a significant effect of Uncertainty $[F(1,9)=66.3;\ p<0.0001]$. The main effect of Predicted Masking was significant $[F(1,9)=25.8;\ p<0.01]$, with higher observed thresholds for masker samples with high predicted thresholds compared to low predicted thresholds. The Uncertainty × Predicted Masking interaction was also significant $[F(1,9)=6.9;\ p<0.05]$, indicating that the difference in performance for samples with low versus high predicted thresholds was smaller when masker samples were randomly interleaved on an interval-by-interval basis compared to fixed throughout a block of trials. That is, the extent to which listeners failed to achieve excitation-based threshold predictions was greatest when predicted thresholds were low.[1]

IV. DISCUSSION

A. Masking in the absence of masker-spectral uncertainty (*Fixed* conditions)

The present results demonstrated a wide range of thresholds across the 25 masker samples when the same masker sample was used on every presentation throughout trial blocks. Frequency components comprising the masker samples were excluded from falling within a 160-Hz region centered on the 1000-Hz signal. Nonetheless, average thresholds across listeners differed by as much as 26 dB across masker samples. As outlined in the Introduction, substantial differences in masked threshold across fixed-frequency multi-tonal masker samples have previously been reported in the literature (e.g., Neff and Callaghan, 1987; Wright and Saberi, 1999). The variability in masked thresholds observed in the present and earlier studies may reflect, in part, differences in peripheral excitation across masker samples. For example, sample 11 used in the current study was comprised of one tone more than an octave below the signal frequency (447 Hz) and three tones 800-Hz or more above the signal frequency (1870, 2276, and 2716 Hz). Both the predicted (6.3 dB SPL) and average observed (20.9 dB SPL) thresholds were lowest for this sample. In contrast, higher predicted (45 dB SPL) and average observed (41.9 dB SPL) thresholds were found for sample 15. Two of the four tones comprising this sample were close in frequency and flanked the 1000-Hz signal (867 and 1109 Hz).

Differences in predicted peripheral excitation across masker samples were systematically examined by computing thresholds for the 1000-Hz signal in the presence of each of the 25 samples using the excitation-pattern model of partial loudness proposed by Moore *et al.* (1997). Similar to observed thresholds, predicted thresholds differed widely across masker samples (range=6.3–45.0 dB SPL). Moreover, the model predicted a significant portion of the variance in the average observed thresholds for several listeners. These observations are inconsistent with the assumption that sparsely sampled multi-tonal maskers with a spectral gap centered on the signal frequency produce minimal energy-based masking. Instead, the variability in threshold across samples may reflect significant contributions from peripheral auditory processes.

It is difficult, however, to provide a complete account of listeners' thresholds in terms of differences in peripheral excitation patterns across the 25 masker samples. Average thresholds for some samples were not well predicted by the model. Whereas the model and data were in good agreement for samples with high predicted thresholds, the model tended to underestimate performance for samples with low predicted thresholds. This pattern of results suggests substantial contributions of informational masking for at least some fixed-frequency masker samples. In particular, informational masking appears to be largest for conditions in which effects of excitation-based masking are smallest.

The current results extend those of developmental studies showing remote-frequency masking under conditions of minimal stimulus uncertainty during infancy (Werner and Bargones, 1991; Leibold and Werner, 2006) and childhood (Leibold and Neff, 2007). For example, the presence of two, fixed-frequency tones remote from the signal frequency can produce significant amounts of informational masking of a 1000-Hz signal for infants and most children (Leibold and Werner, 2006; Leibold and Neff, 2007). Note also that several adults tested as control subjects in each of the developmental studies appeared to be susceptible to some degree of informational masking without masker-frequency uncertainty.

The presence of large individual differences, on the order of 40 dB for some masker samples, is also atypical of masking produced by peripheral mechanisms. Moreover, these individual differences do not appear to reflect listening strategies that can be improved with practice. Despite thousands of intervening trials, limited improvements in performance across listeners or across masker samples were observed between the first and fourth blocks of trials. The large individual differences and resistance to training suggest contributions of informational masking for at least the poorer performers (e.g., Neff and Dethlefs, 1995). Inconsistency across listeners in masking effectiveness for particular multitonal samples fixed across presentations has been noted in previous studies (Neff and Callaghan, 1987; Wright and Saberi, 1999; Alexander and Lutfi, 2004; Durlach *et al.*, 2005; Leibold and Neff, 2007). For example, Neff and Callaghan (1987) measured thresholds for a 1000-Hz tone in each of 50, ten-tone masker samples fixed across a block of trials. Maskers were then ranked in terms of effectiveness, and the top and bottom 10 selected in terms of amount of masking for each listener. There were four samples even at these extremes of the distributions that fell in the top category for one listener but the bottom for another.

The wide range of thresholds across listeners, most pronounced for samples with low predicted masking, suggests that listeners may have adopted different strategies to perform the detection task. Alternatively, all listeners may be pursuing the same strategy but with different degrees of success. A subset of listeners appeared limited in their ability to resolve the signal from the masker at the level of the auditory periphery. Observed thresholds for these listeners closely followed the model predictions. For example, L99 (see Fig. 2) used an effective strategy resulting in little evidence of informational masking for *Fixed* conditions. In contrast, other listeners appeared to adopt a non-optimal strategy for the detection task. These listeners showed relatively large discrepancies between the model and the data for some masker samples. For example, no relation between the model and data was found for L21. Observed thresholds exceeded predictions by as much as 30 dB for this listener (Fig. 2). As with the average data, differences between the model and data for L21 were largest for samples with low predicted thresholds.

Neff *et al.* (1993) suggested that the large range in performance across listeners with random-frequency multi-tonal maskers may be indicative of individual differences in the ability to listen "analytically." In the context of the current study, analytic listening refers to the extent to which listeners attended to the 1000-Hz target signal and ignored information at masker frequencies. Whereas listeners with little or no informational masking are described as analytic, listeners with substantial informational masking are described as holistic or synthetic listeners. Presumably, holistic listeners integrate information across frequency even though it is disadvantageous to do so. Several researchers have applied quantitative methods to evaluate these apparent individual differences in listening strategies with random-frequency maskers (e.g., Neff *et al.*, 1993; Lutfi, 1993; Richards *et al.*, 2002; Alexander and Lutfi, 2004; Durlach *et al.*, 2005). The most comprehensive model to date is the component-relative entropy (CoRE) model proposed by Lutfi (1993). Similar to models of energy detection, the CoRE model assumes that detection is based on the output of energy at the auditory filter centered on the signal frequency. The CoRE model differs from traditional energy-detection models, however, because it considers potential contributions from auditory filters that do not contain information about the presence of the signal. Using this model, Lutfi and colleagues observed a relation between amount of informational masking and both the number and frequency range of monitored filters (e.g., Lutfi, 1993; Alexander and Lutfi, 2004).

In a related approach, listeners' decision weights for the detection task are compared to weights of an ideal observer (e.g., Berg, 1989; Lutfi, 1995; Richards and Zhu, 1994). For detection of a fixed-frequency pure tone in the presence of a random-frequency masker, an ideal observer assigns weight exclusively to the output of auditory filters representing the signal. No weight is assigned to auditory filter outputs dominated by masker stimuli. Results from studies using this approach have shown that low-threshold listeners have weighting functions that more closely approximate those of an ideal listener, whereas high-threshold listeners assign significant weight to masker components (e.g., Alexander and Lutfi, 2004; Richards *et al.*, 2002). In addition, a significant negative relation between weighting efficiency and amount of informational masking has been observed (e.g., Alexander and Lutfi, 2004). These results are consistent with the idea that the individual differences in informational masking observed for random-frequency maskers reflect perceptual differences in analytic listening strategies.

Similar differences in the ability to listen analytically might be responsible for the wide range of thresholds observed across listeners for the current *Fixed* conditions. To examine individual listening strategies, an "attentional-filter" analysis was performed following the general approach described by Neff *et al.* (1993). Neff *et al.* (1993) compared thresholds for a 1000-Hz tone embedded in a random-frequency, multi-tonal masker while parametrically varying the width of the protected region around the 1000-Hz signal. Four of eight listeners showed substantial informational masking even for conditions with large spectral gaps in the maskers. In contrast, much smaller effects of masking were observed for the remaining four listeners across all conditions. Although notched-noise measures of auditory filter width (Patterson *et al.*, 1982) were similar across the two groups, high-threshold listeners had wider attentional filters and poorer estimates of processing efficiency compared to low-threshold listeners. That is, the listeners most susceptible to informational masking appeared to be unable to ignore the irrelevant masker energy.

Attentional filters were fitted to individual listeners' data in *Fixed* conditions following methods similar to those of Neff *et al.* (1993). The filter was defined as a two-parameter rounded-exponential (roex) model,

$$W(g) = (1 - r)(1 + pg)e^{-pg} + r,$$

where p defines the width of the filter, r is the dynamic range, and g is the frequency offset relative to the filter cen-

TABLE II. Results of fitting "attentional filters" with the roex model (see text) for individual listeners' data in *Fixed* conditions. Listeners are ordered based on lowest to highest mean threshold in the five *FixLow* conditions, as in Fig. 4.

Listener	p	r (dB)	K (dB)	VAC
L99	26.34	−49.19	−10.62	0.83
L121	16.73	−43.79	−14.19	0.63
L5	30.24	−36.91	−7.37	0.71
L265	19.98	−35.10	−10.80	0.73
L1	12.34	−51.72	−12.58	0.59
L65	6.81	−50.60	−20.50	0.34
L62	10.79	−31.66	−10.50	0.53
L205	11.30	−15.80	−10.82	0.55
L21	16.46	−14.27	−11.95	0.29
L123	9.80	−14.98	−7.21	0.27

ter frequency, defined as a ratio. An additional variable specifies the signal-to-noise ratio at the output of the filter that is associated with threshold. This value, described as efficiency (K), is assumed to be constant across frequency. In this analysis the spectra of the 25 masker samples were weighted in power by function W, and a separate least-squares fit was made to each individuals' *Fixed* masker thresholds. Resulting parameter estimates are reported in Table II, along with the percentage of variance accounted for by each fit. Data for each listener appear in a separate row, and listeners are ordered according to the averaged threshold in the five *FixLow* conditions, as in Fig. 4. There was a significant correlation between this rank ordering and estimates of both p and r ($p < 0.05$), with better sensitivity being associated with narrower frequency resolution (larger values of p) and wider dynamic range (smaller values of r). There was a nonsignificant association between rank order based on sensitivity and K ($r=0.11$, $p=0.77$).

Contrary to the results for the subjects in Neff *et al.* (1993), these results on attentional filters indicate that individual differences for listeners in the present study can primarily be attributed to differential frequency selectivity rather than efficiency. What remains unclear, however, is to what extent these individual differences in filter width estimates reflect differences in frequency selectivity at peripheral or central levels within the auditory system. One interpretation of these results is that true individual differences in peripheral filtering are considerably larger than previously estimated using traditional notched-noise measures of auditory filter width (Patterson *et al.*, 1982). Alternatively, the magnitude of estimates for some listeners may suggest that these differences reflect "attentional" filtering at more central levels within the auditory system. Future studies are required to systematically examine the mechanisms responsible for these individual differences.

One approach that might be used to disentangle contributions of peripheral and central processes is to introduce an acoustic cue that has been shown to provide a substantial release from informational masking produced by random-frequency multi-tonal maskers and examine whether thresholds improve for listeners with relatively high thresholds in the *Fixed* conditions. We used this general approach to examine performance for one listener (L205) who exhibited relatively high thresholds for masker samples predicted to produce low levels of excitation-based masking. Following Neff (1995), a temporal cue believed to promote sound source segregation and reduce informational masking was provided. Masker duration was increased to 400 ms, resulting in a 100-ms masker fringe preceding the 300-ms signal; stimulus component starting phase was adjusted in the fringe conditions so that the 300-ms listening interval was identical in conditions with and without the preceding 100-ms fringe. If central, rather than peripheral, filtering underlies this listener's poor performance with masker samples predicted to produce minimal energetic masking, introducing this temporal fringe cue should produce large reductions in threshold. Thresholds were collected for L205 with the masker fringe for the five masker samples with the lowest predicted energetic masking (samples 16, 11, 12, 24, and 7), with conditions completed in random order. Mean threshold without fringe was 34.5 dB SPL (range=30.9–35.4 dB SPL), similar to *Fixed* thresholds for these maskers and this listener in the main experiment. In contrast, mean threshold with the temporal fringe was 15.2 dB SPL (range=10.9–23.3 dB SPL). Thus, temporal fringe reduced thresholds by 18.4 dB, consistent with a substantial effect of informational masking in the *Fixed* conditions for masker samples with low predicted thresholds in the main data set.

B. Effects of masker-spectral uncertainty (*Random* conditions)

Previous studies have established that randomizing masker spectra for multi-component maskers can produce substantial informational masking for many listeners (e.g., Neff and Dethlefs, 1995). The current observation that masked thresholds were elevated for *Random* relative to the corresponding *Fixed* conditions is in agreement with this earlier work. External stimulus uncertainty was not required, however, to produce informational masking for many listeners. Thresholds were higher for all listeners when one of five masker samples was randomly selected on each presentation (*Random* conditions) compared to thresholds for the same samples fixed on every presentation throughout trial blocks (*Fixed* conditions). Of particular interest, even listeners with little evidence of informational masking for *Fixed* conditions appeared to be susceptible to informational masking when masker-frequency uncertainty was introduced. For example, threshold for L99 was approximately 14-dB higher in the *RanLow* condition compared to this listener's highest threshold (sample 7) for the same samples presented in the *Fixed* conditions.

Differences in the stimuli used in the current and previous works limit comparisons of the effect of masker-frequency uncertainty across studies. Masker-frequency uncertainty is often produced by drawing component frequencies completely at random from a specified frequency range (e.g., Neff and Green, 1987). In the current *Random* conditions, the masker presented in each interval was randomly selected from a subset of five masker samples. Thus, listeners may have relied on their memory of the five samples used for each *Random* condition and/or used a different decision strategy than if maskers had been drawn from a larger or completely random set. Data from Richards *et al.* (2002), however, suggest that listeners adopt consistent strategies for small and large sets of masker samples. Richards *et al.* (2002) examined measured detection thresholds for a 1000-Hz tone in the presence of a random-frequency, six-tone masker. The pool of masker samples was varied across conditions, resulting in masker set sizes ranging from 3 to 24. Performance was also assessed for a completely random masker set. The results were interpreted as showing that many listeners remembered individual masker samples, even when the set size was 24. However, there was no evidence that listeners changed their decision strategy as the masker set size was increased.

It is generally accepted that some portion of the masking produced by random-frequency multi-tonal maskers is informational in that it appears to arise from mechanisms other than those modeled by energy detection in peripheral auditory filters centered at the signal frequency. Effects of masker-frequency uncertainty appear to be greater, however, when samples with low compared to high predicted masking were randomly selected on each presentation. The mechanisms responsible for this apparent difference in the effect of masker-frequency uncertainty are not understood. One possible explanation is that strategies based on frequency-specific cues support the best sensitivity, though the success of these strategies differs across listeners and across masker samples. In cases where frequency-specific detection cues are not used effectively, listeners might rely on differences in overall loudness across the two intervals at higher signal levels, effectively placing an upper limit on thresholds. Note that a strategy based on overall loudness does not require perceptual segregation of the signal and masker.

C. Implications of the present results for understanding informational masking

The majority of results reported here are not unique to the current study. For example, a number of previous studies have compared performance across fixed- and random-frequency multi-tonal maskers (e.g., Neff and Callaghan, 1988; Neff and Dethlefs, 1995; Wright and Saberi, 1999; Alexander and Lutfi, 2004; Richards and Neff, 2004; Durlach *et al.*, 2005; Leibold and Neff, 2007). Similarly, previous investigations have provided estimates of the amount of energetic masking that might be expected for given multi-tonal masker samples (e.g., Lutfi, 1993; Durlach *et al.*, 2005) and the effects of energetic masking on magnitude of informational masking associated with stimulus frequency uncertainty (Neff *et al.*, 1993). The novel approach used in the current study was to examine contributions of peripheral excitation and informational masking to the variability in masking effectiveness in *Fixed* conditions observed across samples of multi-tonal maskers using a combination of these previously reported approaches.

A common metric for determining informational masking related to masker-frequency uncertainty is the difference in performance between conditions using random-frequency multi-tonal maskers and either quiet thresholds or thresholds in conditions using equal power, broadband-noise maskers. We have recently argued, however, that a multi-tonal masker sample should be used as a reference condition to estimate the contribution of masker-frequency variability to informational masking (Leibold and Werner, 2006; Leibold and Neff, 2007). In this approach, the fixed-frequency multi-tonal masker sample is matched to the random-frequency multi-tonal masker in as many aspects as possible except spectral variability, as in the minimal-uncertainty conditions first described by Watson *et al.* (1976). The present data indicate that estimates of informational masking based on comparison across fixed and random masker conditions depend critically on selecting conditions with comparable energetic masking. Data based on stimuli with low predicted thresholds would result in large estimates of variability-based masking, whereas stimuli with high predicted thresholds would result in small estimates of variability-based masking.

The observation of informational masking in the absence of stimulus uncertainty complicates efforts to define informational masking. Informational masking is often defined as masking that occurs in excess of energetic masking. As Kidd *et al.* (2008) stated, "…it would be helpful in attempting to quantify informational masking if there were a precise model of energetic masking that could accurately predict performance for a wide range of stimuli or a measurement procedure in which one could be certain that only energetic masking was present" (p. 145). The Moore *et al.* (1997) model provides a useful framework for conditions with simultaneous multi-tonal maskers, providing estimates of masked threshold based on excitation patterns (e.g., Jesteadt *et al.*, 2007). Applying this framework to the current data for *Fixed* conditions indicates contributions of both energetic and informational masking to the masking produced by four-tone masker samples for many listeners despite the absence of external stimulus uncertainty. In addition, the relatively close correspondence between threshold predictions and behavioral thresholds from the best-performing listener (L99) lends some credibility to this model of energetic masking for these stimuli. The current data are consistent with Durlach *et al.* (2003) who argued that stimulus variability is not required to produce informational masking, but that informational masking may also be determined by the degree of similarity between the signal and the masker.

V. SUMMARY AND CONCLUSIONS

Consistent with previous studies (e.g., Neff and Callaghan, 1987; Wright and Saberi, 1999), individual masker samples varied widely in masking effectiveness. Average thresholds in conditions in which particular samples were fixed across blocks differed by as much as 26 dB across masker samples.

There were marked individual differences in masked threshold, on the order of 40 dB for some fixed samples. Observed thresholds also differed substantially from predicted thresholds in some cases. This model-data discrepancy was most pronounced for samples with low predicted thresholds.

When masker samples were randomly selected on a trial-by-trial basis, the difference in performance for samples with low versus high predicted thresholds was reduced.

The mechanisms responsible for the individual differences in performance across listeners for specific masker samples require further investigation. Masking remains an operational definition with multiple contributing mechanisms even for stimuli and conditions in which sensitivity is often assumed to be limited by peripheral mechanisms. The current results are consistent with the view that informational masking can be affected by multiple factors, including stimulus uncertainty and availability of cues aiding sound segregation for signals and maskers.

ACKNOWLEDGMENTS

This work was supported by NIH NIDCD Grant Nos. R03 DC008389 and R01 DC007391. We thank Walt Jesteadt and Hongyang Tan for use of custom programs. Brian Glasberg and Brian Moore have generously made their software openly available.

---

[1] There was no difference in the basic form of the data or conclusions drawn when amount of masking was analyzed instead of masked thresholds.

Alexander, J. M., and Lutfi, R. A. (**2004**). "Informational masking in hearing-impaired and normal-hearing listeners: Sensation level and decision weights," J. Acoust. Soc. Am. **116**, 2234–2247.

ANSI (**2004**). ANSI S 3.6-2004, "American national standard specification for audiometers" (American National Standards Institute, New York).

Berg, B. G. (**1989**). "Analysis of weights in multiple observation tasks," J. Acoust. Soc. Am. **86**, 1743–1746.

Bregman, A. S. (**1990**). *Auditory Scene Analysis* (MIT Press, Cambridge, MA).

Durlach, N. I., Mason, C. R., Gallun, F. J., Shinn-Cunningham, B., Colburn, H. S., and Kidd, G., Jr. (**2005**). "Informational masking for simultaneous nonspeech stimuli: Psychometric functions for fixed and randomly mixed maskers," J. Acoust. Soc. Am. **118**, 2482–2497.

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, G., Jr. (**2003**). "Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity," J. Acoust. Soc. Am. **114**, 368–379.

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

Green, D. M. (**1988**). *Profile Analysis* (Oxford University Press, Oxford).

Jesteadt, W., Khaddam, S., Tan, H., and Leibold, L. J. (**2007**). *Prediction of Behavioral Thresholds Using a Model of Partial Loudness*, in Abstracts of the 30th Annual MidWinter Research Meeting of the Association for Research in Otolaryngology. Abstract No. 936.

Kidd, G., Jr., Mason, C. R., and Arbogast, T. L. (**2002**). "Similarity, uncertainty, and masking in the identification of nonspeech auditory patterns," J. Acoust. Soc. Am. **111**, 1367–1376.

Kidd, G., Jr., Mason, C. R., Deliwala, P. S., Woods, W. S., and Colburn, H. S. (**1994**). "Reducing informational masking by sound segregation," J. Acoust. Soc. Am. **95**, 3475–3480.

Kidd, G., Jr., Mason, C. R., Richards, V. M., Gallun, F. J., and Durlach, N. I. (**2008**). "Informational masking," in *Auditory Perception of Sound Sources*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer, New York), pp. 143–190.

Leibold, L. J., and Neff, D. L. (**2007**). "Effects of masker-spectral variability and masker fringes in children and adults," J. Acoust. Soc. Am. **121**, 3666–3676.

Leibold, L. J., and Werner, L. A. (**2006**). "Effect of masker-frequency variability on the detection performance of infants and adults," J. Acoust. Soc. Am. **119**, 3960–3970.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lutfi, R. A. (**1993**). "A model of auditory pattern analysis based on component-relative-entropy," J. Acoust. Soc. Am. **94**, 748–758.

Lutfi, R. A. (**1995**). "Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks," J. Acoust. Soc. Am. **97**, 1333–1334.

Lutfi, R. A., Kistler, D. J., Oh, E. L., Wightman, F. L., and Callahan, M. R. (**2003**). "One factor underlies individual differences in auditory informational masking within and across age groups," Percept. Psychophys. **65**, 396–406.

Moore, B. C. J., and Glasberg, B. R. (**1987**). "Factors affecting thresholds for sinusoidal signals in narrow-band maskers with fluctuating envelopes," J. Acoust. Soc. Am. **82**, 69–79.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). "A model for the prediction of thresholds, loudness, and partial loudness," J. Audio Eng. Soc. **45**, 224–237.

Neff, D. L. (**1995**). "Signal properties that reduce masking by simultaneous, random-frequency maskers," J. Acoust. Soc. Am. **98**, 1909–1920.

Neff, D. L., and Callaghan, B. P. (**1987**). "Psychometric functions for multicomponent maskers with spectral uncertainty," J. Acoust. Soc. Am. **81**, S53.

Neff, D. L., and Callaghan, B. P. (**1988**). "Effective properties of multicomponent simultaneous maskers under conditions of uncertainty," J. Acoust. Soc. Am. **83**, 1833–1838.

Neff, D. L., and Dethlefs, T. M. (**1995**). "Individual differences in simultaneous masking with random-frequency, multicomponent maskers," J. Acoust. Soc. Am. **98**, 125–134.

Neff, D. L., Dethlefs, T. M., and Jesteadt, W. (**1993**). "Informational masking for multicomponent maskers with spectral gaps," J. Acoust. Soc. Am. **94**, 3112–3126.

Neff, D. L., and Green, D. M. (**1987**). "Masking produced by spectral uncertainty with multicomponent maskers," Percept. Psychophys. **41**, 409–415.

Oh, E. L., and Lutfi, R. A. (**1998**). "Nonmonotonicity of informational masking," J. Acoust. Soc. Am. **104**, 3489–3499.

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (**1982**). "The deterioration of hearing with age: Frequency selectivity, the critical ration, the audiogram, and speech threshold," J. Acoust. Soc. Am. **72**, 1788–1803.

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

Richards, V. M., and Neff, D. L. (**2004**). "Cuing effects for informational masking," J. Acoust. Soc. Am. **115**, 289–300.

Richards, V. M., and Nekrich, R. D. (**1993**). "The incorporation of level and level-invariant cues for the detection of a tone added to noise," J. Acoust. Soc. Am. **94**, 2560–2574.

Richards, V. M., Tang, Z., and Kidd, G. D., Jr. (**2002**). "Informational masking with small set sizes," J. Acoust. Soc. Am. **111**, 1359–1366.

Richards, V. M., and Zhu, S. (**1994**). "Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients," J. Acoust. Soc. Am. **95**, 423–434.

Spiegel, M. F., Picardi, M. C., and Green, D. M. (**1981**). "Signal and masker uncertainty in intensity discrimination," J. Acoust. Soc. Am. **70**, 1015–1019.

Van Der Heijden, M., and Kohlrausch, A. (**1994**). "Using an excitation-pattern model to predict auditory masking," Hear. Res. **80**, 38–52.

Watson, C. S., Kelly, W. J., and Wroton, H. W. (**1976**). "Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty," J. Acoust. Soc. Am. **60**, 1176–1186.

Watson, C. S., Wroton, H. W., Kelly, W. J., and Benbassat, C. A. (**1975**). "Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals," J. Acoust. Soc. Am. **57**, 1175–1185.

Werner, L. A., and Bargones, J. Y. (**1991**). "Sources of auditory masking in infants: Distraction effects," Percept. Psychophys. **50**, 405–412.

Wright, B. A., and Saberi, K. (**1999**). "Strategies used to detect auditory signals in small sets of random maskers," J. Acoust. Soc. Am. **105**, 1765–1775.

Zwicker, E. (**1976**). "Masking-period patterns of harmonic complex tones," J. Acoust. Soc. Am. **59**, 166–175.