Strategies used to detect auditory signals in small sets of random maskers

Beverly A. Wright[a)]
Audiology and Hearing Sciences Program and Institute for Neuroscience, Northwestern University, 2299 North Campus Drive, Evanston, Illinois 60208-3550

Kourosh Saberi[b)]
Division of Biology, Caltech, Pasadena, California 91125

(Received 18 May 1998; accepted for publication 18 November 1998)

Detection performance for a masked auditory signal of fixed frequency can be substantially degraded if there is uncertainty about the frequency content of the masker. A quasimolecular psychophysical approach was used to examine response strategies in masker-uncertainty conditions, and to investigate the influence of uncertainty when the number of different masker samples was limited to ten or fewer. The task of the four listeners was to detect a 1000-Hz signal that was presented simultaneously with one of ten ten-tone masker samples. The masker sample was either fixed throughout a block of two-interval forced-choice trials or was randomized across or within trials. The primary results showed that: (1) When the signal level was low and the masker sample differed between the two intervals of a trial, most listeners based their responses more on the presence of specific masker samples than on the signal. (2) The detrimental effect of masker uncertainty was clearly evident when only four maskers were randomly presented, and grew as the size of the masker set was increased from two to ten. (3) The slopes of psychometric functions measured with the same masker samples differed among the fixed and two random-masker conditions. (4) There were large differences in the influence of masker uncertainty across masker samples and listeners. These data demonstrate the great susceptibility of human listeners to the influence of masker uncertainty and the ability of quasimolecular investigations to reveal important aspects of behavior in uncertainty condition.
[S0001-4966(99)03303-2]

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

INTRODUCTION

This paper concerns how the ability to detect a known auditory signal amidst additional masking sounds is degraded when there is uncertainty as to which masking sounds are to be presented. In most laboratory experiments in hearing, the listener has little or no trial-to-trial uncertainty about the general characteristics of the stimulus. There is, however, considerable uncertainty inherent in normal listening situations, and uncertainty about the characteristics of the signal or masker makes an auditory signal more difficult to detect. The influence of signal uncertainty is generally small, increasing signal thresholds by 3–6 dB when the masker is either fixed or relatively stable and the signal frequency (e.g., Green, 1961; Veniar, 1958a,b), duration (Dai and Wright, 1995), or starting time (Egan *et al.*, 1961) is randomly varied across trials. In contrast, when different masker samples are randomly presented, a signal of fixed frequency can be much harder to detect than would be predicted on the basis of the frequency selectivity of the peripheral auditory system (e.g., Neff and Green, 1987; Watson and Kelly, 1981). For example, maskers consisting of ten tones with frequencies randomly selected on each observation interval can produce as much as 20 dB more masking of a fixed-frequency tone than a broadband noise of equal power (e.g., Neff and Green, 1987). Signal detection is most difficult when there is uncertainty about both the signal and the masker (e.g., Spiegel and Green, 1982). In the most extreme case, a signal that is itself randomly varied in frequency and temporal position within a sequence of randomly selected masker frequencies can be as much as 65 dB more difficult to detect than a signal presented at a fixed frequency and temporal position within a fixed sequence of masking tones (Watson, 1987). The experiments reported here focus on the influence of masker uncertainty, but the approach differs from that used in most earlier studies.

Previous investigations have typically reported the detection threshold for a signal in a masker-uncertainty condition as a single value that was based on the responses across a block of trials in which a large number of different masker samples were presented. Green (1964) used the term ''molar psychophysics'' to describe this class of experiment because the resulting threshold estimate represents an average of the responses over many heterogeneous trials, obscuring the performance on individual trials.

We have instead examined masker uncertainty using what Green (1964) termed a ''quasimolecular'' approach. With this method, signal threshold is determined by the series of responses over multiple presentations of the same masker sample. Unlike the molar technique, this approach yields many threshold values from each block of trials in an

[a)]Electronic mail: b-wright@nwu.edu
[b)]Electronic mail: kourosh@etho.caltech.edu

uncertainty condition, one for each different masker sample presented. The resulting values thus represent an average of the responses over many homogeneous, rather than heterogeneous, trials.

Using a quasimolecular approach, the following basic question can be addressed: Is it harder to detect a known signal that is masked by a particular masker *X* when that masker is randomly intermixed with other maskers than when only masker *X* is presented? The answer is obtained by comparing two measurements: (1) signal threshold in a fixed-masker condition in which only masker *X* is presented, and (2) signal threshold in a random-masker condition estimated from the responses obtained on *only those trials* in which masker *X* is presented. A higher threshold in the random than the fixed condition indicates that uncertainty about which masker is to be presented makes the signal more difficult to detect in masker *X*. To our knowledge, only Pfafflin (1968) and Wright and McFadden (1990) have examined masker uncertainty using this technique. Others who used the quasimolecular approach to study the masking produced by sets of randomly presented reproducible noises did not report performance in the fixed condition (Green, 1964; Gilkey, 1985). Pursuing similar issues with molar psychophysics, Watson and his colleagues (Watson and Kelly, 1981; Spiegel and Watson, 1981; Watson, 1987) reported performance in both fixed and random-masker conditions.

The task of the listeners in the two present experiments was to detect a tonal signal of a fixed and known frequency that was presented simultaneously with one of ten multitonal masker samples. The masker sample was either fixed throughout a block of two-interval forced-choice trials or was randomized across or within trials. The two experiments differed only in the method used to determine the masked threshold of the signal. The results address the following five aspects of the influence of masker uncertainty.

Masker bias: Previous quasimolecular investigations have shown that listeners treat particular masker samples as being more likely to contain the signal when different samples are presented on the two observation intervals of a trial (e.g., Green, 1964; Pfafflin and Mathews, 1966; Wright and McFadden, 1990). Given such a masker bias, measured performance is artificially good when the signal is presented in the favored sample, because then the bias leads to the correct choice of the signal interval even at very low signal levels. Performance is spuriously poor, however, when the signal is presented in the unfavored sample, because then the bias leads to the choice of the nonsignal interval containing the favored masker. Many molar investigations of masker uncertainty have used large numbers of masker samples to try to minimize this problem. Here we further document the existence of masker bias and describe a method by which separate measurements can be made both of this bias and of the sensitivity to the presence of the signal.

Psychometric functions: Molar psychophysical data indicate that psychometric functions measured with random multi-tonal maskers have shallower slopes than those measured with broadband noise (Kidd *et al.*, 1995a, 1998; Neff and Callaghan, 1987) or than those predicted for an ideal observer (Lutfi, 1994). Shallow psychometric slopes have also been reported for individual masker samples presented in random-masker conditions (Neff and Callaghan, 1987). Little, however, is known of the relationship between the slope of psychometric functions in fixed- and random-masker conditions. In apparently the only investigation of this issue, Watson and Kelly (1981) tentatively concluded from a molar analysis that the slope of the psychometric function did not change significantly as the amount of uncertainty was varied from minimal to very high. Here we show that the slopes of psychometric functions measured with the same masker sample differ systematically among fixed-masker and two types of random-masker conditions.

Size of the masker set: In most previous studies of masker uncertainty, a different masker sample was presented on every trial or observation interval in random-masker conditions (e.g., Lutfi, 1994; Neff and Callaghan, 1988; Watson *et al.*, 1976). In the remaining experiments a small, but constant, number of samples were used (e.g., Pfafflin, 1968). There has been no systematic investigation of how the influence of masker uncertainty is affected by the number of possible masker samples to be presented. Here we report that the mean detrimental effects of two types of masker uncertainty grow approximately in parallel as the size of the masker set increases from two to ten.

Sample-specific influence of masker uncertainty: The aim of many experiments using molar psychophysics has been to determine the conditions under which masker uncertainty is most disruptive. Those studies have manipulated such variables as the number and frequency distribution of the tones in each masker sample (Lutfi, 1994; Neff and Callaghan, 1988; Neff *et al.*, 1993; Watson *et al.*, 1976), the repetition pattern of the masker (Kidd *et al.*, 1994, 1995b), and the presentation mode (monotic or dichotic; Kidd *et al.*, 1994; Neff, 1995). There is, however, a scarcity of information on the influence of uncertainty for specific masker samples belonging to the same general category. Only Pfafflin (1968) has reported the influence of uncertainty on the detection of a tonal signal in a set of 12 frozen noises. Her results show relatively small overall effects of uncertainty probably because her noise maskers all sounded quite similar. Nevertheless, some of her masker samples were more affected by uncertainty than others. Here we report marked and consistent differences in the influence of uncertainty across different masker samples.

Individual differences: Individual listeners tested using molar psychophysics show marked threshold differences in random-masker conditions (Leek, 1987; Neff and Dethlefs, 1995). These differences are not paralleled in the performance of the same listeners on the detection of tones in quiet or in measures of peripheral filter width made with notched noise, suggesting that the threshold variations in random-masker conditions are due to the introduction of uncertainty (Neff and Dethlefs, 1995). Here we report large individual differences in performance in both random- and fixed-masker conditions, and show that listeners' reactions to masker uncertainty are not necessarily revealed solely by their performances in random-masker conditions.

I. EXPERIMENT 1: ADAPTIVE TRACKING

A. Method

1. Listeners

Four listeners (two female, JF and WD) ranging in age from 19 to 23 years were paid for their participation. All had hearing within 15 dB of normal between 125 and 8000 Hz as determined by a Bekesy tracking procedure and had previous experience in other psychoacoustic tasks.

2. Stimuli

The signal was a 1000-Hz tone and the masker was one of up to ten different ten-tone complexes. The signal and masker were gated together for a total of 200 ms using a cosine-squared rise/fall time of 16.8 ms. The frequencies of the ten tones in each of the ten masker samples were chosen at random from a uniform distribution ranging from 200–5000 Hz, excluding the region from 800–1200 Hz around the signal. Also, to guarantee that the tones in each 200-ms sample were orthogonal, no two masking components in a given sample were allowed to be closer than 5 Hz (1/duration). The phases of the masking tones were randomly selected from a uniform distribution ranging from 0 to 359 degrees. Once chosen, the frequencies and phases of the ten tones in each of the ten masker samples were fixed throughout the entire experiment. The signal was always presented in zero phase. Table I lists the frequencies and phases comprising each of the masker samples (A–J). The individual masking tones were each presented at 50 dB SPL, producing an overall level of 60 dB SPL. Similar stimulus parameters produced substantial masking in previous experiments on masker uncertainty (Neff and Green, 1987; Neff and Callaghan, 1988; Neff *et al.*, 1993).

The signal and masker were digitally generated in the time domain at a sampling rate of 20 000 Hz using a digital-signal-processing board (TDT AP2). They were delivered separately through two 16-bit digital-to-analog converters (TDT DA1) followed by separate 10-kHz low-pass filters (TDT FLT3, 60 dB attenuation at 11.5 kHz), separate programmable attenuators (TDT PA4), and a single sound mixer (TDT SM3). The listeners were seated in a sound-treated room and listened monaurally through the left earpiece of Sennheiser HD450 headphones.

3. Procedure

The procedure was two-interval forced-choice with feedback. The two observation intervals of a single trial were marked by lights, and were separated by 300 ms. The signal level was adjusted adaptively using the three-down/one-up rule of Levitt (1971) which tracks the 79% correct point on the psychometric function. The step size was 8 dB through the first reversal, 4 dB through the next three reversals, and 2 dB thereafter. The adaptive track was terminated after 60 trials. The first four reversals were discarded, and threshold was calculated as the mean of the remaining reversals. Two conditions employed multiple independent adaptive tracks, as described below. Six to ten blocks of trials were collected from each listener in each condition.

TABLE I. The frequencies (Hz) and phases (degrees) of the ten components in each of the ten masker samples (A–J). The signal was always a 1000-Hz tone presented at zero phase.

	Frequency (Hz)	Phase (deg)		Frequency (Hz)	Phase (deg)
A	1692	157.2	*B*	443	296.4
	1717	108.9		1573	26.0
	2530	25.2		1626	290.7
	2687	66.4		2343	134.0
	2737	325.9		2394	218.5
	2873	165.0		2494	73.9
	3023	308.5		2696	151.3
	3748	336.6		3539	117.1
	4104	197.1		3674	107.8
	4329	219.1		4202	79.7
C	458	299.2	*D*	352	179.3
	514	2.6		364	165.7
	2320	66.4		1317	324.2
	2445	327.2		2293	213.8
	2584	120.8		2334	283.1
	2666	253.1		2454	217.4
	2692	258.1		2714	104.7
	3665	269.5		3815	230.3
	3905	294.9		4334	158.1
	4891	31.0		4891	6.8
E	2156	338.1	*F*	550	124.6
	2536	105.4		631	59.6
	2663	131.6		663	278.4
	2849	336.8		1925	334.4
	3066	206.3		2201	277.7
	3133	284.7		3439	117.9
	3321	159.2		3609	0.0
	3546	321.8		3651	243.0
	3906	25.1		3789	197.9
	4689	181.6		4420	8.7
G	255	357.1	*H*	355	293.5
	434	171.0		1887	198.6
	797	245.3		1980	237.9
	2009	55.9		2382	185.9
	2399	226.8		2957	179.4
	3246	57.8		2967	249.9
	3397	236.9		3218	117.2
	3828	100.5		3375	126.8
	4015	167.4		3445	345.8
	4202	126.5		4304	307.5
I	709	53.2	*J*	208	275.3
	1227	208.7		308	217.6
	1436	214.6		1268	27.2
	1478	161.5		1293	150.0
	1556	178.5		1343	142.4
	2027	111.6		2167	267.2
	2514	162.3		2350	255.1
	3119	355.4		3031	109.3
	4005	193.7		3641	214.8
	4661	31.6		4678	304.0

4. Conditions

There were three listening conditions. In the *fixed condition*, the same masker sample was presented on every observation interval throughout an entire block of 60 trials. The ten different masker samples were presented in random order across blocks.

In the *random-by-trial condition*, the same masker

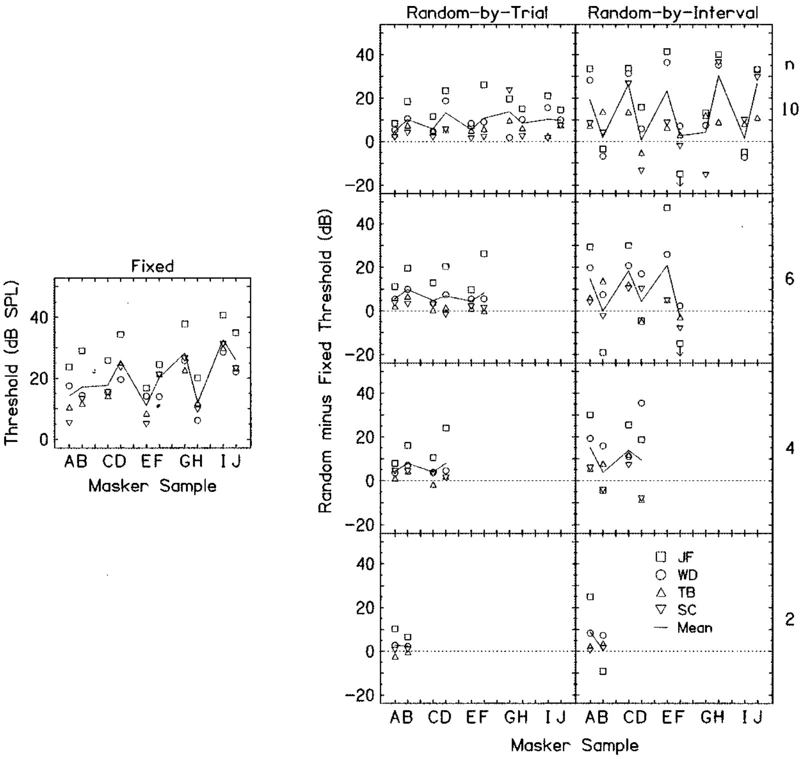

FIG. 1. Individual (symbols) and mean (lines) adaptive results of the four listeners for the ten masker samples (A–J). The floating panel shows the signal thresholds for each masker sample in the fixed condition. The remaining panels show the threshold differences for each masker sample between the random-by-trial and fixed conditions (left column) and between the random-by-interval and fixed conditions (right column) for the four masker-set sizes (rows).

sample was presented on both observation intervals of a single trial, but the particular masker sample was chosen quasi-randomly across trials from *n* possibilities, where *n* equalled 2, 4, 6, or 10 in different blocks of trials. The particular samples used when $n<10$ were chosen arbitrarily but always included the same samples as in the tests using fewer samples. The same sets of maskers were used throughout the experiment. Performance for each masker sample was monitored with a separate adaptive track. Thus there were *n* interleaved tracks in each test. To obtain the 60 responses necessary to complete the track for each masker sample, the particular sample was chosen at random, without replacement, from a pool of 60 instances of each of the *n* samples. Thus there were 60*n* trials in each block. Breaks were provided after every 60 trials within a block to make the experimental session as similar to that in the fixed condition as possible.

Finally, in the *random-by-interval condition*, different masker samples were presented on the two observation intervals of a trial, and the particular masker samples were chosen quasi-randomly across trials from *n* possibilities ($n=2$, 4, 6, or 10). To simultaneously measure performance for each masker sample and monitor the influence of the masker sample in the nonsignal interval on the response pattern, the masker samples were presented in fixed pairs (A and B, C and D, E and F, G and H, I and J). The pairings were made arbitrarily and were maintained throughout the experiment. Both members of a sample pair were always presented within the same trial, but the order of their presentation within the trial was random. A separate adaptive track monitored performance for each masker sample, yielding *n* interleaved tracks. Data in the random-by-interval condition were collected in the same manner as in the random-by-trial condition.

The initial signal levels were 25, 35, and 45 dB SPL in the fixed, random-by-trial, and random-by-interval conditions, respectively. Example trials were provided at the beginning of each block. In those trials the signal level was 15 dB higher than the initial signal level used in the actual trials. Each listener was generally tested on all three conditions, presented in random order, during each 2-h listening session.

B. Results and discussion

Figure 1 presents the individual (symbols) and mean (lines) results of the four listeners for the ten masker samples (A–J). Plotted in the floating panel are the signal thresholds in the fixed condition. The remaining panels show the threshold differences between the fixed and each of the two random conditions (columns) for the four masker-set sizes (rows). The mean threshold differences between the random-by-trial and fixed condition were consistently positive (left column), indicating that uncertainty about which sample was to be presented on a trial made the signal harder to detect. In contrast, the threshold differences between the random-by-interval and fixed condition fluctuated markedly (right column). When the set size was ten, the difference was around 30 dB for one sample and near 0 dB for the other sample of each masker pair. Every listener had at least one negative difference score, indicating a lower threshold in the random-by-interval than the fixed condition. It seems unlikely that this oscillating pattern of difference scores resulted from peculiarities in the influence of uncertainty on particular masker samples. A more plausible scenario is that the listeners treated one masker sample of a pair in the random-by-interval condition as being more likely to contain the signal. This possibility is explored in the second experiment.

		stimulus				
		(1) x′y	(2) yx′	(3) y′x	(4) xy′	
RESPONSE	1	28	65	77	31	
	2	62	25	13	59	
		90	90	90	90	360

		stimulus		
		(1) + (3) sig in int 1	(2) + (4) sig in int 2	
RESPONSE	1	105	96	
	2	75	84	
		180	180	360

$d'_s = 0.11$

		stimulus		
		(1) + (4) x in int 1	(2) + (3) x in int 2	
RESPONSE	1	59	142	
	2	121	38	
		180	180	360

$d'_m = -0.86$

FIG. 2. *Top*: A representative stimulus-response matrix from a single fixed pair of masker samples in the random-by-interval condition. *Middle*: The submatrix used to derive the listener's sensitivity to the signal, d'_s, independent of any bias toward or against the masker sample. *Bottom*: The submatrix used to derive the listener's preference to select an interval based on the presence of a particular masker sample, d'_m, independent of the presence of the signal. A negative d' resulting from this calculation merely indicated that the listener favored masker sample *y* rather than masker sample *x*. All values of d'_m are reported as positive values. See text for details on all three matrices.

II. EXPERIMENT 2: CONSTANT STIMULI

A. Introduction

Experiment 2 introduces a technique for simultaneously measuring both signal detectability and masker bias in the random-by-interval condition. This technique employs the method of constant stimuli, from which information about the slopes of the psychometric functions in fixed- and random-masker conditions may be obtained. Furthermore, the signal thresholds derived in this experiment are free of the influence of masker bias. This allows the evaluation of the contributions of the size of the masker set, of the particular masker sample, and of the individual listener to the magnitudes of both the random-by-trial and random-by-interval uncertainty effects.

B. Method

The listeners, stimuli, and conditions were the same as in experiment 1. The differences between the two experiments were that, here, (1) the signal level was fixed throughout a block of trials rather than being adjusted adaptively, and (2) both the observation interval in which the signal was presented and the masker sample that contained the signal were recorded with the response on each trial. Using this method, it was possible to estimate both the listener's sensitivity to the signal and the extent to which the listener was biased toward intervals containing particular masker samples.

The top of Fig. 2 illustrates a representative stimulus-response matrix from a single fixed pair of masker samples in the random-by-interval condition. The stimuli are listed along the top of the matrix as $\langle x'y\rangle$, $\langle yx'\rangle$, $\langle y'x\rangle$, and $\langle xy'\rangle$, where the letter indicates the particular masker sample, the letter sequence indicates the presentation order between the two observation intervals, and the prime indicates the signal interval. The responses are listed along the side of the matrix as $\langle 1\rangle$ for ''signal in the first interval'' or $\langle 2\rangle$ for ''signal in the second interval.''

To determine the listener's sensitivity to the signal in a particular masker sample independent of bias toward or against that sample, the responses were pooled over columns in which the signal was presented in the same observation interval. Thus the responses were added across the $x'y$ and $y'x$ [columns (1) and (3)], and across the yx' and xy' [columns (2) and (4)] stimulus patterns. This pooling resulted in the two-by-two matrix shown at the middle of Fig. 2. The index of detectability for the signal (d'_s) was calculated using this new matrix by taking the total proportion of correct responses [in the example, $(105+84)/360=0.53$] and looking up the d' value corresponding to that proportion correct in a forced-choice table (Green and Swets, 1964; Swets, 1964).

Similarly, to determine the listener's preference to select an interval based on the presence of a particular masker sample, independent of the presence of the signal, the responses were combined over columns in which a particular masker sample (chosen arbitrarily as sample *x*) was presented in the same observation interval. Thus, the responses were added across the $x'y$ and xy' [columns (1) and (4)], and across the yx' and $y'x$ [columns (2) and (3)] stimulus patterns. This yielded another two-by-two matrix, shown at the bottom of Fig. 2, from which the masker bias (d'_m) was calculated. In some cases, a negative d' resulted from this calculation. That merely indicated that the listener favored masker sample *y*, rather than masker sample *x*. All values of d'_m are reported as positive values.

For the random-by-interval condition, each stimulus pattern ($x'y$,yx',$y'x$,xy') was presented a total of 180 times. To obtain measures of d'_s and d'_m that were based on different data sets, the data for each pattern were divided into two sets of 90 trials each. The 360 responses (4 patterns$\times$90 trials) from the even-numbered blocks formed one set and those from the odd-numbered blocks formed the other. Each set, such as the one in Fig. 2, was used to calculate both d'_s and d'_m. Thus d'_s obtained from the even-numbered trials could be compared to d'_m obtained from the odd-numbered trials, and the reverse. Note that Fig. 2 shows the data from only one of the two sets of responses; a complete figure for one masker pair (*x* and *y*) would include three parallel matrices containing the other data set. When there were ten masker samples, five such pairs of matrices were produced.

For the random-by-trial and fixed conditions, the two possible stimulus patterns ($x'x$ and xx') were each presented a total of 180 times. Thus d'_s was calculated from a total of 360 trials. No measure of d'_m was necessary or possible.

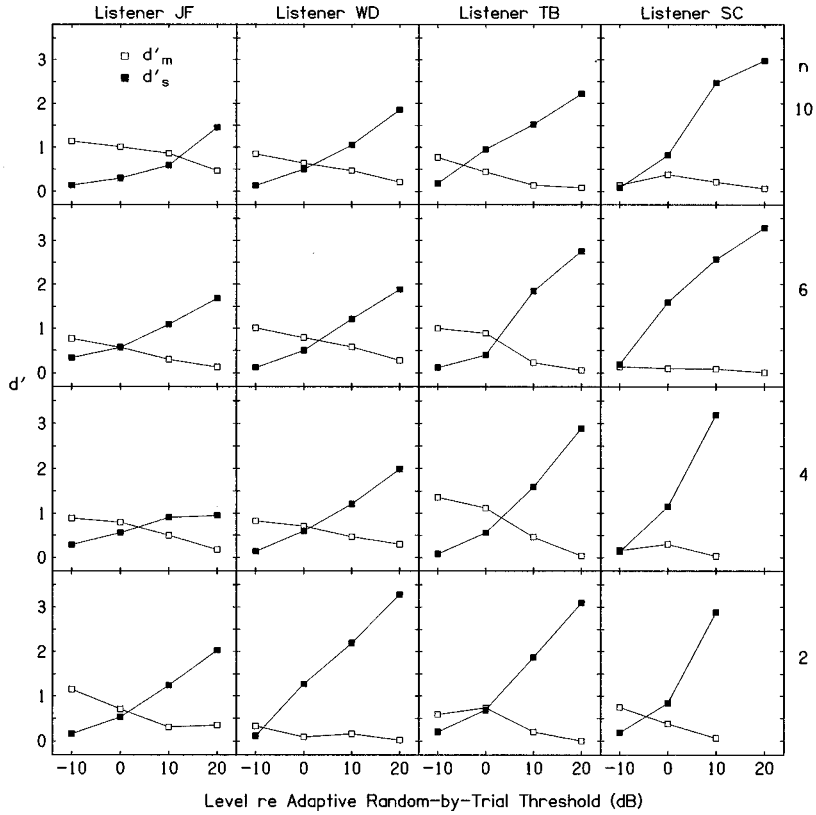

FIG. 3. Individual psychometric functions (columns) for d_s' (filled squares) and d_m' (open squares) averaged across the sample pairs tested for each set size (rows). The abscissa shows the signal level relative to the adaptively measured random-by-trial threshold.

For the two random conditions, the signal was presented at four different levels denoted −10, 0, +10, and +20 dB. For the fixed condition, the signal levels were chosen specifically for each listener, but typically included the values 0, −5, −10, and −20 dB. These levels are all expressed relative to the signal level at threshold obtained for each listener, number of masker samples, and particular sample in the random-by-trial condition in experiment 1. The signal levels were defined relative to the adaptive random-by-trial thresholds so that the levels used would produce a reasonable range of d' values for each masker sample in each condition, despite the large differences in the amounts of masking produced by each masker individually. Breaks were provided after every 60 trials. Data were collected in three-block sets in each condition in the order: (1) random-by-interval, (2) random-by-trial, (3) random-by-interval, (4) fixed. The second phase of data collection in the random-by-interval condition was added when it was realized that more trials were needed to have independent estimates of d_s' and d_m'.

C. Results and discussion

1. Masker bias

The d' values measured in the random-by-interval condition are plotted in Fig. 3 for each listener (columns) and each masker-set size (rows). The even- and odd-numbered trials produced similar estimates of d_s', and of d_m', and therefore were averaged, yielding one estimate of d_s' and one of d_m' at each signal level for each pair of masker samples. Each panel shows the values of d_s' and d_m' averaged over the sample pairs tested for each set size.[1] For example, each point in the top row of panels represents the mean of five estimates of d', one estimate for each of the five pairs of masker samples, and each point in the bottom row represents one estimate of d' from the single pair of masker samples tested.

For all listeners, as expected, d_s' (filled squares) increased from near 0 to above 1 as the signal level was increased from −10 to +20 dB. In contrast, for three of the four listeners (all but SC), d_m' (open squares) frequently decreased from around 1 to 0 as the signal level was increased over the same range. Thus, at low signal levels, the majority of listeners made their responses based more on the presence of a particular masker sample than on the presence of the signal. Individual listeners consistently favored the same sample of a pair regardless of the masker-set size, but the favored sample often differed across listeners. When only one sample of the masker pair had masking components lower than the signal frequency, listeners JF, WD, and TB all favored that sample (B over A, F over E), but those same three listeners differed in their preferences for the remaining samples. The magnitude of d_m' at the signal level of −10 dB did not differ systematically across the masker pairs. For listener SC, d_m' was essentially constant at about 0 at all signal levels, indicating that his responses were influenced almost entirely by the presence of the signal.

In 77% of the cases where there was a bias (24 of 31 instances), threshold in the fixed condition was higher for the favored than for the unfavored masker sample. Likewise, in 78% of the cases where the adaptive random-by-interval threshold was lower than the adaptive random-by-trial or fixed threshold (25 of 32 instances in Fig. 1), threshold in the fixed condition was higher for that masker sample than for the other sample in the pair. Thus listeners tended to favor the more effective masker when they had difficulty hearing

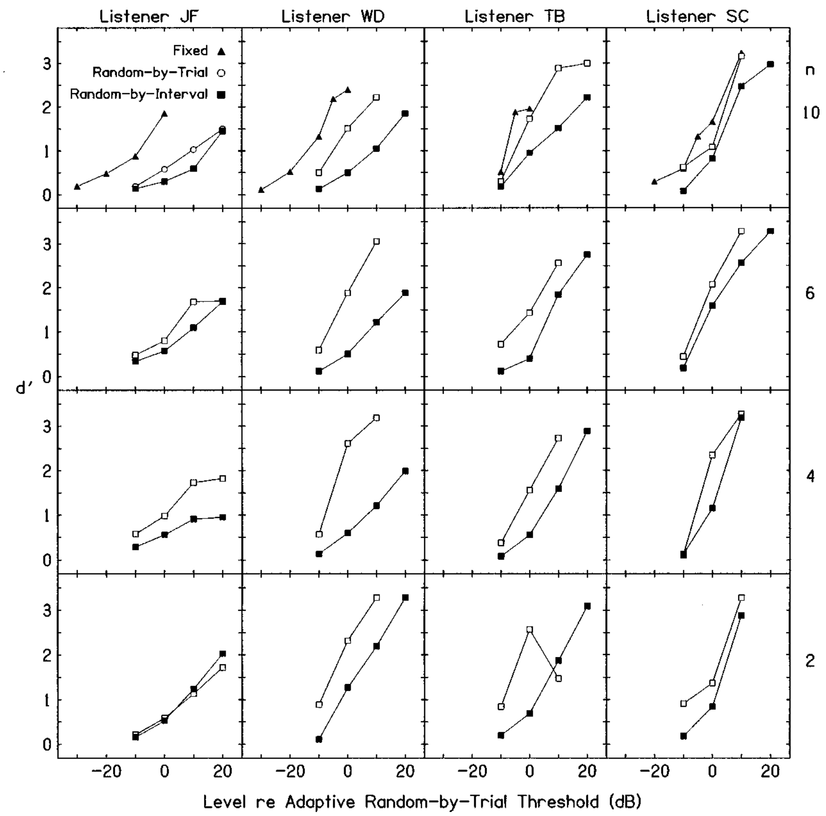

FIG. 4. As in Fig. 3, but for d_s' in the fixed (filled triangles), random-by-trial (open circles), and random-by-interval (filled squares) conditions.

the signal. Threshold in fixed conditions is traditionally thought to be proportional to the amount of masker energy contained in the peripheral auditory filter centered on the signal frequency. Perhaps the listeners who were susceptible to masker bias tended to favor the maskers with the higher fixed thresholds, because those maskers, with their extra energy near the signal frequency, sounded more like they contained the signal. It is possible that the three listeners who showed masker bias made their judgments on such a timbral aspect of the stimulus, while the remaining unbiased listener focused on the signal frequency. This interpretation is in accord with the reports of Neff and her colleagues that individual listeners adopt either holistic or analytic listening strategies when faced with masker uncertainty (Neff *et al.*, 1993; Neff and Dethlefs, 1995).

These masker-bias results show that when the detection task is difficult due to uncertainty, listeners essentially change the task. Here, rather than responding randomly when the signal was difficult to detect, listeners instead responded systematically to particular masker samples. While that result might be attributed to the signal-like qualities of some masker samples, there are data showing response bias even when the randomized variable had no signal-like characteristics. For example, Lee (1994) asked listeners to indicate which of two observation intervals contained the higher rate of sinusoidal amplitude modulation of a tonal carrier. When she randomly presented two different carrier frequencies on the two intervals of a trial, one of her two listeners consistently selected as signal the interval with the higher-frequency carrier. We also have preliminary results from listeners who were asked to detect the longer of two temporal intervals each marked by two brief tones. When we randomly marked the temporal intervals by low-frequency tones in one observation interval and high-frequency tones in the other, listeners responded systematically to the tone frequency rather than to the longer temporal interval. Particular tonal frequencies do not inherently sound more signal-like when the signal is a higher modulation rate or longer temporal interval. Thus listeners appear to adopt the task of imposing order on the randomized variable when the assigned task becomes difficult.

2. Psychometric functions

Figure 4 is plotted in the same manner as Fig. 3, but shows the mean psychometric functions for d_s' for the fixed (filled triangles; top row only), random-by-trial (open circles), and random-by-interval (filled squares) conditions.[1] To derive the signal levels that corresponded to 79% correct detections from each of the psychometric functions that went into the mean functions in Fig. 4, the data for each function (d_s' versus signal level in dB) were fitted with a least-squares straight line from which the slope and intercept of the function was determined. Inspection of individual functions showed that a linear fit was acceptable for estimating the relative steepness of the functions.

Table II lists the mean slopes of the psychometric functions in the three conditions calculated across the masker samples in each masker set. Results are shown for the individual listeners and their mean. The slope of the psychometric functions was generally steepest in the random-by-trial condition, intermediate in the fixed condition, and shallowest in the random-by-interval condition. The mean psychometric slope was (1) steeper in the random-by-trial than the random-

TABLE II. Mean slopes of the psychometric functions in the three conditions calculated across all of the masker samples (n) in each masker set. The individual listener means, and the grand mean and ± one standard error of the mean, are listed.

n	Listener	Condition: Fixed	Random-by-trial	Random-by-interval
10	JF	0.048	0.053	0.042
	WD	0.062	0.086	0.057
	TB	0.146	0.101	0.067
	SC	0.109	0.141	0.108
	Mean	0.091	0.095	0.069
	se	(0.004)	(0.004)	(0.004)
6	JF	0.050	0.045	0.045
	WD	0.062	0.131	0.060
	TB	0.148	0.092	0.093
	SC	0.115	0.165	0.117
	Mean	0.094	0.108	0.079
	se	(0.006)	(0.006)	(0.003)
4	JF	0.048	0.045	0.043
	WD	0.068	0.131	0.062
	TB	0.155	0.118	0.095
	SC	0.117	0.177	0.153
	Mean	0.097	0.117	0.088
	se	(0.008)	(0.004)	(0.002)
2	JF	0.048	0.050	0.063
	WD	0.068	0.156	0.104
	TB	0.129	0.090	0.099
	SC	0.095	0.119	0.135
	Mean	0.085	0.104	0.100
	se	(0.006)	(0.000)	(0.000)

by-interval condition for every set size, and in 11 of the 16 individual cases, (2) steeper in the random-by-trial than the fixed condition for every set size and for 10 of the 16 individual cases, and (3) steeper in the fixed than the random-by-interval condition for all but the set size of 2, and in 11 of the 16 individual cases. Using molar psychophysics, Neff and Callaghan (1987) measured the slopes of psychometric functions in d'/dB units with stimuli and conditions similar to those tested here. The mean slopes of their four listeners were in the same general range as the present report at 0.088 in the random-by-trial condition (termed ''fixed'' by them) and 0.095 in the random-by-interval condition (termed ''random'' by them). Two of their four listeners had steeper slopes in the random-by-trial than the random-by-interval condition. They did not report results for the fixed condition. Differences in the psychometric slope across conditions are usually thought to indicate that different internal nonlinear transformations of the stimulus scale occur either at peripheral or central sites in each condition (Egan, 1965; Laming, 1986; Saberi and Green, 1997). Such nonlinearities may be modeled as a power-law transformation of the stimulus scale. On a logarithmic scale, the exponent becomes the constant of proportionality, i.e., the slope. Laming has described psychometric slopes from one to eight for visual and auditory tasks (see also Egan, 1965; Dai, 1995; Saberi and Green, 1997).

The actual signal SPL estimated to yield 79% correct detections was calculated by determining the relative signal level that corresponded to a d_s' of 1.15 (79% correct) from the fitted functions and then adding that value to the appropriate random-by-trial threshold measured adaptively. This was a straightforward process for the fixed and random-by-trial conditions, for which there were separate adaptive thresholds for each masker sample. For the random-by-interval condition, however, the calculation of the actual signal SPL was complicated by the fact that each pair of masker samples produced only one d_s' psychometric function. The actual signal SPL used in that condition was calculated by adding the single 79% correct value from the d_s' function separately to each of the two adaptive threshold values of the two maskers in the pair. This calculation assumes that the signal threshold in each of the maskers in a pair would have been increased by the same amount due to random-by-interval uncertainty had it been possible to measure each one independently.[2] It is doubtful that this assumption is strictly true. However, the results obtained with it are orderly and seem more reasonable than the adaptive measurements of the influence of random-by-interval uncertainty shown in Fig. 1, for which no correction for bias was applied.

Figure 5 is plotted in the same manner as Fig. 1, but shows the results obtained from the actual signal SPLs for 79% correct detections derived from the d_s' psychometric functions.[1] These results are free of the influence of masker bias. In contrast to the mean difference scores calculated from the adaptive thresholds shown in Fig. 1, the mean differences between the random-by-interval and fixed condition (right column) were consistently larger than those between the random-by-trial and fixed condition (left column). Thus uncertainty about which masker was to be presented impaired performance, and the magnitude of that impairment was greater in the random-by-interval than the random-by-trial condition.

3. Size of the masker set

Figure 6 shows the mean effect of random-by-trial (open squares) and random-by-interval (filled squares) uncertainty as a function of the number of different masker samples presented in each test. Each point represents the mean difference score calculated across all of the samples tested for each set size in Fig. 5.[1] Thus the left-most points are based on difference scores from two samples and the right-most points from ten samples. The error bars indicate ± one standard error of the mean. Note that because the psychometric slopes sometimes differed among the fixed and random conditions (Table II), the calculated magnitudes of the uncertainty effects depends on the percent-correct level chosen for examination.

Two aspects of these data deserve notice. First, performance was clearly degraded with only two samples for random-by-interval uncertainty, and four samples for random-by-trial uncertainty. The demonstration of clear effects of uncertainty with only a few masker samples shows that listeners are very susceptible to the influence of masker uncertainty, and presumably to other forms of uncertainty as well. This susceptibility provides additional support for the idea that everyday situations are better represented by conditions with than without uncertainty. It also shows the feasibility of quasimolecular investigations of the stimulus characteristics that lead to large or small uncertainty effects.

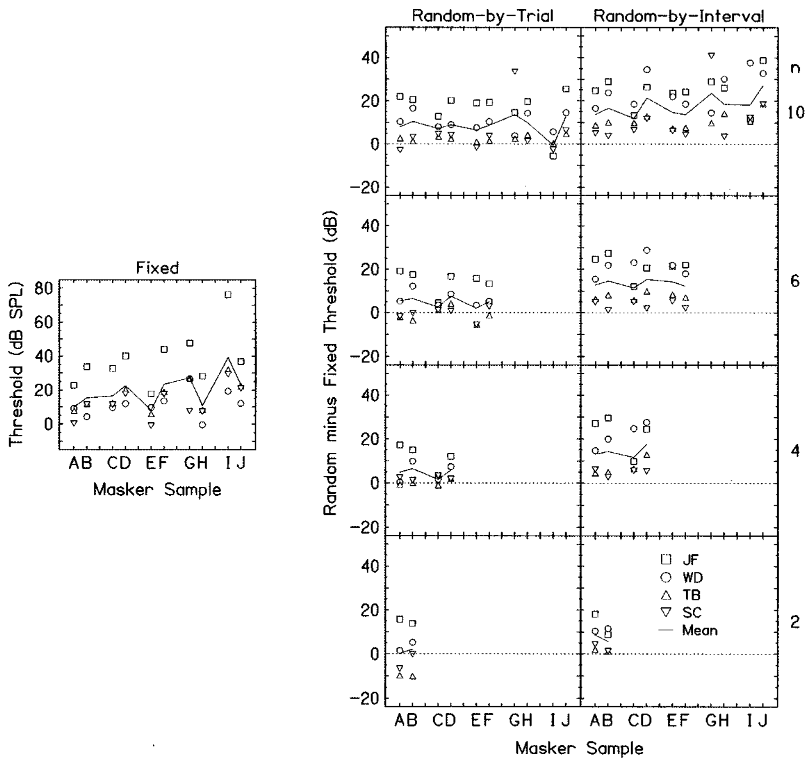

FIG. 5. Parallel to Fig. 1, but the values were derived from the d_s' psychometric functions, thereby removing the influence of masker bias from the threshold estimates in the random-by-interval condition. Note that the ordinate range in the floating panel is greater than in Fig. 1.

Second, the mean detrimental influence of random-by-trial uncertainty grew by about 7 dB, and that of random-by-interval uncertainty by about 11 dB, as the size of the masker set was increased from two to ten. The difference in the rate of growth between the two types of uncertainty was due almost exclusively to performance with a set size of two. For set sizes of four and greater, random-by-interval uncertainty was consistently about 9 dB greater than random-by-trial uncertainty. For both types of uncertainty, the mean growth reached a plateau for set sizes of four and six, but increased again for the set size of ten. The magnitude and form of this growth should be interpreted with caution, because only samples A and B were tested with the smallest set size. When only those samples were included in the analysis, both types of uncertainty grew by about 8 dB, with the same pattern as for the whole data set, as the set size was increased from two to ten. Though based on a limited set of masker samples, the roughly parallel increase in the two types of uncertainty with set size suggests that listeners may use the same detection strategy less successfully, or less efficiently, in the random-by-interval than the random-by-trial condition.

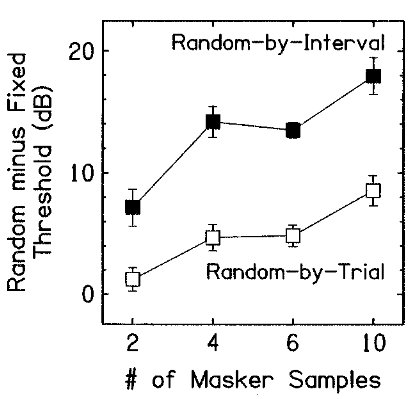

FIG. 6. The mean effect of random-by-trial (open squares) and random-by-interval (filled squares) uncertainty as a function of the number of different masker samples presented in each test. Each point represents the mean difference score calculated across all of the samples tested for each set size in Fig. 5. The error bars indicate ± one standard error of the mean.

For comparison, the mean adaptive estimates of random-by-trial uncertainty were on average 1 dB larger than the constant-stimuli estimates. For set sizes of two, four, six, and ten, the adaptive values were 3, 6, 7, and 9 dB, and the constant-stimuli values 1, 5, 5, and 9 dB. The mean adaptive estimates of random-by-interval uncertainty were on average 4 dB smaller than the constant-stimuli estimates. For set sizes of two, four, six, and ten, the adaptive values were 5, 10, 8, and 13 dB, and the constant-stimuli values 7, 14, 13, and 18 dB. To the extent that the mean of quasimolecular adaptive estimates corresponds to molar measurements in other experiments, the present results indicate that molar and molecular estimates are similar for random-by-trial uncertainty, but that molar analyses may underestimate the influence of random-by-interval uncertainty.

4. Sample-specific influence of masker uncertainty

The influence of uncertainty differed among the masker samples. As shown in Fig. 5, when the set size was ten, the mean influence of random-by-trial uncertainty ranged from −1 dB (sample I) to 14 dB (sample G) and the mean influence of random-by-interval uncertainty ranged from 12 dB (sample C) to 27 dB (sample J). The mean magnitude of the uncertainty effect for each masker sample showed a Spearman rank correlation of $r_s = 0.65$ $(p < 0.05)$ between the two types of uncertainty. Thus the effects of the two types of uncertainty tended to be both large, or both small, for the same masker samples.

There were significant Spearman rank correlations between the mean magnitude of the random-by-interval uncertainty effect and the frequency difference between the nearest masker components below and above the signal frequency ($r_s = 0.65$, $p < 0.05$, two-tailed test), and between the adaptive measure of random-by-trial uncertainty and that same frequency difference ($r_s = 0.64$, $p < 0.05$):[3] The smaller the frequency difference, the greater the uncertainty effect. Both random-by-interval ($r_s = 0.55$, $p < 0.10$) and adaptive random-by-trial ($r_s = 0.81$, $p < 0.05$) uncertainty were also significantly correlated with the frequency distance to the masking component nearest to the signal frequency (but not with the frequency distance to the nearest component exclusively above or below the signal frequency): The closer the nearest component, the greater the uncertainty effect. Other reports have also associated masking components close to the signal frequency with the production of uncertainty effects (Neff and Callaghan, 1987; Lutfi and Doherty, 1994). One interpretation of these relationships is that listeners use a narrow attentional filter matched to a single peripheral filter in the fixed condition, but use a wider attentional filter whose output represents the sum of the outputs of multiple separate peripheral filters in the random conditions (Lutfi, 1993; Neff *et al.*, 1993). The increase in the bandwidth of the attentional filter would lead to an increase in signal threshold. As observed, this threshold increase would be greatest when the frequency separation of the masking components closest to the signal frequency was small.

5. Individual differences

Finally, there were systematic differences in the performance patterns across individual listeners. For example, listener JF, who had the largest uncertainty effect in 70% of the cases in Fig. 5 (squares), also showed the highest threshold in every fixed condition in Fig. 5 (squares), the shallowest mean psychometric function in every case (Table II), and the most masker bias (her average d'_m across set sizes at -10 dB from Fig. 3 was 0.99, compared to 0.76 for WD, 0.93 for TB, and 0.30 for SC). In contrast, listener SC, who had or tied for the smallest uncertainty effect in 59% of the cases in Fig. 5 (inverted triangles), also showed the steepest mean psychometric functions in 58% of all the cases and 88% of the random cases in Table II, and the least masker bias (right column of Fig. 3).

Using molar psychophysics, Neff and her colleagues have reported remarkably large threshold differences across listeners in random-masker conditions (Neff *et al.*, 1993; Neff and Dethlefs, 1995). Using quasimolecular psychophysics, the present listeners show similarly marked threshold differences in the random-masker conditions, but also show large threshold differences in the fixed condition (see floating panels in Figs. 1 and 5). The threshold differences in the fixed condition are important because with the quasimolecular approach the uncertainty effect is measured by subtracting threshold in the fixed- from that in the random-masker condition. This analysis reveals that the actual uncertainty effect is sometimes different from that apparently indicated by the threshold in the random-masker condition. For example, listener WD, who had the largest or second largest uncertainty effect in 89% of the cases in Fig. 5 (circles) had the lowest threshold in 70% of the fixed conditions (Fig. 5; circles). Therefore, her actual thresholds in the random-masker conditions (fixed threshold plus the difference threshold plotted in Fig. 5) were much lower, on average by 26 dB, than those of listener JF. The actual thresholds of WD in the random-masker conditions thus give the mistaken impression that she was not influenced much by uncertainty.

III. SUMMARY

A quasimolecular psychophysical approach was used to investigate how uncertainty about the frequency content of a ten-tone masker sample affected the ability to detect a 1000-Hz signal. The masker sample was either fixed throughout a block of two-interval forced-choice trials (fixed condition) or was randomized across (random-by-trial condition) or within (random-by-interval condition) trials. The results showed the following.

(1) *Masker bias:* When the signal level was low and different masker samples were presented on the two observation intervals of a trial, listeners often based their responses more on the presence of a particular masker sample than on the presence of the signal (Fig. 3). Signal threshold in the fixed condition was higher in the favored than than the unfavored sample in the majority of cases. Adaptively measured signal thresholds in the random-by-interval condition were clearly skewed by this bias (Fig. 1). A method was described by which masker bias can be identified and separated from sensitivity to the signal (Fig. 2).

(2) *Psychometric functions:* Psychometric functions measured with the same masker sample were generally steepest in the random-by-trial condition, intermediate in the fixed condition, and shallowest in the random-by-interval condition (Fig. 4, Table II).

(3) *Size of the masker set:* Performance was clearly degraded with only two masker samples for random-by-interval uncertainty and four samples for random-by-trial uncertainty. As the size of the masker set was increased from two to ten, the mean magnitude of the effect of both types of uncertainty grew in parallel by an average of 9 dB (Fig. 6).

(4) *Sample-specific influence of masker uncertainty:* The magnitudes of both types of uncertainty effects varied by 15 dB across individual masker samples. Signal detection was most affected by both types of masker uncertainty in samples that had frequency components close to the signal frequency.

(5) *Individual differences:* There were systematic differences in the performance patterns across individual listeners. The two listeners who were the most and least affected by both types of masker uncertainty also showed, respectively, the shallowest and steepest psychometric functions, and the most and least masker bias. Threshold in the random-masker conditions did not reliably reflect which listeners were most influenced by uncertainty.

ACKNOWLEDGMENTS

These data were collected in the Psychology Department at the University of Florida. We wish to thank Dr. David M. Green for helpful discussions that led to experiment 2. Dr. Donna L. Neff, Dr. Robert A. Lutfi, Wilbur J. Wright, and two anonymous reviewers provided helpful comments on earlier drafts of this paper. This work was supported by NIDCD Grant No. 1 R29 DC02997-03 awarded to BAW and No. 1 R29 DC03648-01 awarded to KS.

[1]The data of listener JF in the random-by-interval condition for masker samples E and F collected with a set size of six are omitted from the mean values reported for her in Figs. 3 and 4. That is because the psychometric function fitted to her data for those samples did not yield a signal level for 79% correct detections that was between -100 and $+100$ dB. To include her data in the grand means in Figs. 5 and 6 and Table II, the average difference across samples between the magnitudes of her random-by-interval uncertainty effects (Figs. 5 and 6) and her psychometric slopes (Table II) when $n=10$ and $n=6$ was calculated. That value was then subtracted from her uncertainty effect for samples E and F when $n=10$ and used as the estimate of her performance for those samples when $n=6$.

[2]A single-interval method would not necessarily solve this problem. At first glance it may appear that the hit rate in a single-interval task could be considered the sensitivity to the signal and the false-alarm rate the masker bias. However, sensitivity to the signal (d'_s) in a single-interval task is calculated from both the hit and false-alarm rates. Furthermore, it would be difficult to determine the proportion of hits made due to the presence of the signal versus those made due to a bias toward the masker, or the proportion of false alarms made due to the tendency to say ''signal'' versus those made due to a bias toward the masker.

[3]For masker samples A and E that had no components below the signal frequency, the frequency difference was calculated as the frequency of the nearest component above the signal frequency (minus zero).

Dai, H. (**1995**). ''On measuring psychometric functions: A comparison of the constant-stimulus and adaptive up-down methods,'' J. Acoust. Soc. Am. **98**, 3135–3139.

Dai, H., and Wright, B. A. (**1995**). ''Detecting signals of unexpected or uncertain durations,'' J. Acoust. Soc. Am. **98**, 798–806.

Egan, J. P. (**1965**). ''Masking-level differences as a function of interaural disparities in intensity of signal and of noise,'' J. Acoust. Soc. Am. **38**, 1043–1049.

Egan, J. P., Greenberg, G. Z., and Schulman, A. I. (**1961**). ''Interval of time uncertainty in auditory detection,'' J. Acoust. Soc. Am. **33**, 771–778.

Gilkey, R. H., Robinson, D. E., and Hanna, T. E. (**1985**). ''Effects of masker waveform and signal-to-masker phase relation on diotic and dichotic masking by reproducible noise,'' J. Acoust. Soc. Am. **78**, 1207–1219.

Green, D. M. (**1961**). ''Detection of auditory sinusoids of uncertain frequency,'' J. Acoust. Soc. Am. **33**, 897–903.

Green, D. M. (**1964**). ''Consistency of auditory detection judgements,'' Psychol. Rev. **71**, 392–407.

Green, D. M., and Swets, J. A. (**1964**). *Signal Detection Theory and Psychophysics* (Wiley, New York).

Kidd, G. Jr., Mason, C. R., and Rohtla, T. L. (**1995a**). ''Binaural advantage for sound pattern identification,'' J. Acoust. Soc. Am. **98**, 1977–1986.

Kidd, G. Jr., Mason, C. R., and Dai, H. (**1995b**). ''Discriminating coherence in spectro-temporal patterns,'' J. Acoust. Soc. Am. **97**, 3782–3790.

Kidd, G. Jr., Mason, C. R., Rohtla, T. L., and Deliwali, P. S. (**1998**). ''Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns,'' J. Acoust. Soc. Am. **104**, 422–431.

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.

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

Lee, J. (**1994**). ''Amplitude modulation rate discrimination with sinusoidal carriers,'' J. Acoust. Soc. Am. **96**, 2140–2147.

Leek, M. R. (**1987**). ''Directed attention in complex sound perception,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 278–288.

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. (**1994**). ''Discrimination of random, time-varying spectra with statistical constraints,'' J. Acoust. Soc. Am. **95**, 1490–1500.

Lutfi, R. A., and Doherty, K. A. (**1994**). ''Effect of component-relative-entropy on the discrimination of simultaneous tone complexes,'' J. Acoust. Soc. Am. **96**, 3443–3450.

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 (A).

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., and Green, D. M. (**1987**). ''Masking produced by spectral uncertainty with multicomponent maskers,'' Percept. Psychophys. **41**, 408–415.

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

Pfafflin, S. M. (**1968**). ''Detection of auditory signal in restricted sets of reproducible noise,'' J. Acoust. Soc. Am. **43**, 487–490.

Pfafflin, S. M., and Mathews, M. V. (**1966**). ''Detection of auditory signals in reproducible noise,'' J. Acoust. Soc. Am. **39**, 340–345.

Saberi, K., and Green, D. M. (**1997**). ''Evaluation of maximum-likelihood estimators in nonintensive auditory psychophysics,'' Percept. Psychophys. **59**, 867–876.

Spiegel, M. F., and Watson, C. S. (**1981**). ''Factors in the discrimination of tonal patterns. III. Frequency discrimination with components of well-learned patterns,'' J. Acoust. Soc. Am. **69**, 223–230.

Spiegel, M. F., and Green, D. M. (**1982**). ''Signal and masker uncertainty with noise maskers of varying duration, bandwidth, and center frequency,'' J. Acoust. Soc. Am. **71**, 1204–1210.

Swets, J. A. (**1964**). *Signal Detection and Recognition by Human Observers* (Wiley, New York), pp. 682–683.

Veniar, F. A. (**1958a**). ''Signal detection as a function of frequency ensemble. I,'' J. Acoust. Soc. Am. **30**, 1020–1024.

Veniar, F. A. (**1958b**). ''Signal detection as a function of frequency ensemble. II,'' J. Acoust. Soc. Am. **30**, 1075–1078.

Watson, C. S. (**1987**). ''Uncertainty, informational masking, and the capacitity of immediate auditory memory,'' in *Auditory Processing of Complex Sounds*, edited by W. A. Yost and C. S. Watson (Erlbaum, Hillsdale, NJ), pp. 267–277.

Watson, C. S., and Kelly, W. J. (**1981**). ''The role of stimulus uncertainty in the discrimination of auditory patterns,'' in *Auditory and Visual Pattern Recognition*, edited by D. J. Getty and J. H. Howard (Erlbaum, Hillsdale, NJ), pp. 37–59.

Watson, C. S., Wroton, H. W., Kelly, W. J., and Benbassat, C. A. (**1976**). ''Factors in the discrimination of tonal patterns. I. Component frequency, temporal position, and silent intervals,'' J. Acoust. Soc. Am. **57**, 1175–1185.

Wright, B. A., and McFadden, D. (**1990**). ''Uncertainty about the correlation among temporal envelopes in two comodulation tasks,'' J. Acoust. Soc. Am. **88**, 1339–1350.