Published in final edited form as:
J Acoust Soc Am. 2007 January ; 121(1): 411–419.

Individual differences in the masking level difference with a narrowband masker at 500 or 2000 Hz

Emily Buss[a)], **Joseph W. Hall III**, and **John H. Grose**
Department of Otolaryngology/Head and Neck Surgery, University of North Carolina School of Medicine, Chapel Hill, North Carolina 27599

Abstract

The masking level difference (MLD) for a narrowband noise masker is associated with marked individual differences. This pair of studies examines factors that might account for these individual differences. Experiment 1 estimated the MLD for a 50 Hz wide band of masking noise centered at 500 or 2000 Hz, gated on for 400 ms. Tonal signals were either brief (15 ms) or long (200 ms), and brief signals were coincident with either a dip or peak in the masker envelope. Experiment 2 estimated the MLD for both signal and masker consisting of a 50 Hz wide bandpass noise centered on 500 Hz. Signals were generated to provide only interaural phase cues, only interaural level cues, or both. The pattern of individual differences was dominated by variability in NoSπ thresholds, and NoSπ thresholds were highly correlated across all conditions. Results suggest that the individual differences observed in Experiment 1 were not primarily driven by differences in the use of binaural fine structure cues or in binaural temporal resolution. The range of thresholds obtained for a brief NoSπ tonal signal at 500 Hz was consistent with a model based on normalized interaural correlation. This model was not consistent for analogous conditions at 2000 Hz.

I. INTRODUCTION

The masking level difference (MLD) is the detection advantage obtained for some stimuli when the signal and the masker are presented to the two ears with different interaural characteristics (Hirsh, 1948). In the typical example of an MLD, thresholds for a diotic stimulus, with the signal and masker in phase at the two ears (NoSo), are compared to thresholds for a diotic masker and a signal that is out of phase at the two ears (NoSπ). Thresholds in the NoSπ condition can be substantially lower than those in the NoSo condition, particularly at low signal frequencies. Large individual differences have been reported in the MLD when the maskers are spectrally narrow (e.g., Bernstein *et al.*, 1998).

There are several lines of evidence suggesting that inherent fluctuations of narrowband maskers may affect the magnitude of the MLD. Grose and Hall (1998) estimated temporal weighting functions which showed that binaural cues impacted performance differently depending on the masker envelope: Portions of the signal coincident with masker envelope dips were given greater perceptual weight in the NoSπ detection task than other portions of the signal coincident with peaks or intermediate values. No such preferential dip weighting was observed for the NoSo condition. A follow-up study measured detection thresholds for a brief pure tone signal, with 15 ms $\cos^2$ ramps, and a 50 Hz wide band of Gaussian masking noise, both centered on 500 Hz. The brief signal was coincident with either a masker envelope dip or peak. Thresholds in the NoSo condition did not vary reliably with signal placement, while those in the NoSπ condition were on the order of 10 dB lower in a dip than a peak, a result that was referred to as a “dip advantage.” In other words, the MLD was larger for a signal presented coincident

---

[a)] Author to whom correspondence should be addressed. Telephone: (919) 966-8019. Electronic mail: ebuss@med.unc.edu.

with a masker envelope dip (Buss *et al.*, 2003; Hall *et al.*, 2004). These results are consistent with the conclusion that detection of a long duration Sπ signal is based on the interaural cues coincident with modulation dips in an No masker.

As is typical in experiments on the MLD, the studies described above used low-frequency narrowband stimuli. Therefore, stimulus fine structure almost certainly played an important role in these results (e.g., Zurek and Durlach, 1987). While the MLD is typically larger at low than high frequencies, substantial MLDs can also be obtained at frequencies above those for which interaural changes in fine structure are thought to be useful in binaural hearing, a finding that is attributed to the use of interaural envelope difference cues (McFadden and Pasanen, 1978). It has been suggested (Bernstein and Trahiotis, 1996) that the utilization of interaural fine structure and interaural envelope differences may rely on the same underlying cue—detection of a change in interaural correlation. The finding of a larger MLD with a narrowband masker at low as compared to high frequencies has been argued to reflect the large contribution of fine structure cues at low frequencies (van de Par and Kohlrausch, 1997).

Addition of a brief Sπ signal to an No narrowband masker dip or peak can introduce binaural fine structure and envelope differences, depending on the interactions between the signal and masker stimulus components. This is illustrated in Fig. 1 with stimuli from Buss *et al.* 2003. In the top row of panels, solid grey waveforms show representative masker samples for conditions in which a brief signal is co-incident with a masker envelope dip (A1) and a masker envelope peak (B1). The dotted lines indicate the temporal placement (but not the amplitude) of the brief tonal signal. For a 72 dB SPL masker, thresholds in the NoSπ condition are on the order of 60 dB SPL in the *dip* condition and 70 dB SPL in the *peak* condition (Buss *et al.*, 2003). Panels A2 and B2 show a magnified view of the stimulus presented to the left and right ear in the NoSπ condition (overlaid traces), with a signal at the associated threshold level. The dark lines indicate the temporal envelopes of the left and right ear stimuli, which have been raised for visual clarity.

Two features of these stimuli are noteworthy. First, the fine structure of the composite signal-plus-masker waveforms presented to the left and right ears is approximately out-of-phase in the center of the masker dip (panel A2) and only slightly phase delayed in the masker peak (panel B2). Generating 100 random samples of the masker and adding a signal at threshold (60 dB for dips or 70 dB for peaks), interaural phase approaches 180° in the middle of the masker dip for all 100 samples, while the maximum interaural phase difference was approximately 30° in the *peak* condition. Summarized in this way, the interaural fine structure cue at threshold could be characterized as less salient in the *peak* than the *dip* condition. Recall that the detection of a long-duration low-frequency Sπ signal in a narrowband No masker is thought to be based largely on fine structure cues (van de Par and Kohlrausch, 1997), and that NoSπ detection for such a long duration tone is argued to be based on cues in the masker dips (Buss *et al.*, 2003). This implies that NoSπ thresholds for a brief signal in the *dip* condition should likewise rely on interaural fine structure cues. If fine structure cues in a peak are not as salient and informative as those in a dip, this opens up the possibility that interaural envelope cues may play an increased or dominant role in the NoSπ, *peak* condition. If this speculation is correct, then repeating the experiment at a higher frequency, where fine structure is no longer available, should eliminate or greatly reduce the dip advantage by virtue of eliminating the interaural fine structure cues in the dip, while leaving unchanged the envelope cues in the peak. Experiment 1 tested this prediction; in addition, performance was compared across signal conditions and frequencies in an attempt to better characterize individual differences obtained in the narrowband masker MLD.

A second feature of note in the stimuli shown in Fig. 1 is the fact that the temporal epoch containing optimal information about a brief signal is by its nature brief, suggesting that fine temporal resolution in deriving a cue to the presence of a signal could result in better performance. There is a wide range of estimates of binaural temporal resolution in the literature. Some data suggest that the binaural system is sluggish and integrates information over hundreds of milliseconds (e.g., Grantham and Wightman, 1979), while other data suggest more precise resolution (e.g., Bernstein *et al.*, 2001). Much of this variability is likely due to the different cues characterizing the signal interval across experimental paradigms (Holube, Kinkel, and Kollmeier, 1998). Grantham (1984) suggested that binaural temporal resolution may be different for interaural time and interaural level cues, further complicating the issue. If observers differ in the temporal resolution with which they are able to process brief interaural cues, then this could affect their ability to make use of brief cues and this factor could underlie the individual differences observed for the narrowband MLD. Experiment 2 estimated narrowband MLD for fluctuating and stable binaural cues to test this possibility.

II. EXPERIMENT 1

Experiment 1 tested the hypothesis that the dip advantage previously demonstrated at 500 Hz would not be found at 2000 Hz. Further, it was hypothesized that the individual differences obtained with long-duration 500 Hz stimuli would be related to those found for brief signals coincident with masker dips. This result would support previous claims that detection with a long duration signal is dominated by cues coincident with dips at 500 Hz (Buss *et al.*, 2003). Further, if the underlying detection cues associated with interaural fine structure and interaural envelope differences are the same, then the pattern of individual differences should be consistent across the 500 and 2000 Hz signal frequencies.

A. Methods

1. Observers—Observers were 28 adults, 18–49 years of age, with thresholds 15 dB HL or better at octave frequencies 250–8000 Hz, and no reported history of ear disease. Of this group, 14 had previously participated in psychoacoustic studies. Five observers had experience listening to MLD stimuli. These five observers spanned the range of individual differences in NoSπ performance. For example, in the NoSπ condition for a brief 500 Hz signal positioned in a masker envelope dip, one of these five had the lowest threshold and another had the second highest threshold with respect to the entire group of 28 observers.

An additional three observers were omitted from the study because of excessive threshold variability, defined as 10 dB or more variation across thresholds in a single condition. It should be noted that these cases of threshold variability did not resemble simple practice effects, in that no clear trends towards improvement were evident. All other data are reported below.

2. Stimuli—The signal was either a 500 Hz or a 2000 Hz pure tone, with random starting phase. In the *peak* and *dip* conditions, the signal was of brief duration, ramped on and off with 15 ms $\cos^2$ ramps (no steady state), and temporally centered in the masker. In the *long* condition, the signal was 200 ms in duration and was gated on with 15 ms $\cos^2$ ramps, the onset occurring 100 ms after masker onset.

Maskers were 50 Hz wide bands of noise centered on the signal frequency (either 500 or 2000 Hz), with a level of 72 dB SPL and duration of 409 ms, including 15 ms $\cos^2$ ramps. Bands of noise were generated digitally in the frequency domain as an array of 2^{14} points, with independent Gaussian draws assigned to the real and imaginary components at points within the 50 Hz passband. This array was transformed into the time domain via IFFT. At the DAC rate of 24.4 kHz used here, the resulting waveform was approximately 670 ms in duration and could be repeated seamlessly (no discontinuity between the ending and starting points of the array). Three copies of the noise array were concatenated. For the *peak* condition, the maximum envelope value of the middle sample was identified, and the extended array was truncated in such a way as to place that point in the temporal center of a 409 ms stimulus. Similarly, for the *dip* condition the minimum was identified and placed in the temporal center of a 409 ms stimulus. A new masker, based on independent samples of Gaussian noise, was generated prior to every stimulus presentation, both within and across trials. Thresholds were estimated for NoSo and NoSπ in all three conditions. These stimuli are functionally identical to those used by Buss *et al.* 2003 and those discussed in the introduction.

3. Procedure—Stimuli were presented using deeply inserted earphones (Etymotic: ER-2). Detection thresholds were estimated using a three-alternative forced-choice, three-down one-up track estimating the 79% point on the psychometric function (Levitt, 1971), with feedback provided visually. In this procedure, three masker intervals were presented with a temporal separation of 500 ms. The signal was added to the masker in one randomly chosen interval, and the observer's task was to indicate which interval contained the signal. The signal level was adjusted in steps of 4 dB until two track reversals were obtained and then in steps of 2 dB for the remaining six reversals. Threshold estimates were computed as the average signal level at the last six track reversals. Conditions were run in blocks, with the order of blocks randomized across observers. A minimum of three threshold estimates were obtained in each condition, and a fourth was obtained if the first three spanned a range of 3 dB or more. The average of 3 to 4 estimates is reported.

B. Results

The distribution of MLDs is shown in Fig. 2. Panels indicate data for the 500 Hz (left) and 2000 Hz (right) conditions. Symbols indicate the median, boxes indicate the 25th-to-75th percentile span, bars indicate the 10th-to-90th percentile span, and stars show the maximum and minimum MLDs. Conditions are indicated on the abscissa. As suggested in this figure, there were substantial individual differences in some of the conditions. For example, the median MLD for a long duration signal at 500 Hz was 18.1 dB, with values for individuals ranging from 7.5 to 26.5 dB. The median MLD for a long duration signal at 2000 Hz was 6.9 dB, with values for individuals ranging from 1.6 to 18.3 dB.

The MLD data (NoSo-NoSπ) were submitted to a repeated measures ANOVA, with two levels of FREQUENCY (500 and 2000 Hz) and three levels of CONDITION (*long*, *dip*, and *peak*). There was a main effect of FREQUENCY ($F_{1,27}=277.9$, $p < 0.0001$), reflecting the fact that MLDs were larger at 500 than 2000 Hz. There was also a significant main effect of CONDITION ($F_{1,27}=5.42$, $P < 0.05$) and a significant interaction of FREQUENCY and CONDITION ($F_{1,27}=41.53$, $p < 0.001$). *Post hoc* contrasts indicated that at 500 Hz the MLDs associated with the *long* and the *dip* conditions were not different ($p=0.17$), but that the MLDs in both of these conditions were significantly greater than the MLD in the *peak* condition ($p < 0.005$). At 2000 Hz, the MLDs were significantly different in all three conditions (*ong* > *peak* > *dip;* $p < 0.005$) Two-tailed t-tests with Bonferroni adjustment for multiple tests confirmed that the MLD was greater than zero for all six conditions (2 freq × 3 cond) at $\alpha =0.01$. In addition to these within-subjects effects, there were substantial individual differences ($F_{1,27}=223.8$, $p < 0.0001$).

Closer examination of these MLD data in terms of the constituent NoSo and NoSπ thresholds suggests that individual differences are dominated by across-observer variation in NoSπ (as opposed to NoSo) thresholds. Figure 3 shows constituent thresholds in the 500 Hz, *long* condition, plotted as a function of the 500 Hz, *long* MLD. Individual differences in the MLD for a long duration, 500 Hz tone were highly correlated with NoSπ thresholds ($R=-0.95$, $p < 0.0001$) but not with NoSo thresholds ($R=-0.06$, $p =0.75$). This indicates that individual differences in the MLD are driven by variability in the NoSπ condition. Across all six conditions of Experiment 1, correlations between NoSo thresholds and associated values of the MLD ranged from R=−0.02 to R=−0.44. In contrast, correlations between NoSπ thresholds and associated values of the MLD were much higher, ranging from R=−0.84 (2000 Hz, *peak*) to R =−0.96 (500 Hz, *dip*). The one exception was the 2000 Hz, *dip* condition, where the correlation was only R=−0.59. The relatively weak correlation in this condition may be due to the reduced inter-observer variance and small values of the MLD.

Figure 4 shows the thresholds for brief NoSπ signals as a function of the NoSπ thresholds for the long duration signal. At 500 Hz (left panel), those observers with the highest (worst) thresholds in the *long* condition showed comparable thresholds in the *dip* and *peak* conditions. In contrast, those with the lowest (best) thresholds in the *long* condition tended to show the biggest dip advantage. The relationship between performance in the brief and long conditions is best illustrated in terms of the slope of the regression line that best fits each set of data. At 500 Hz the slope of the best-fitting line was significantly greater than one for the *dip* condition and less than one for the *peak* condition (at α =0.005). In contrast, at 2000 Hz there is a “peak advantage” for better-performing observers. Slopes of the best-fitting lines were less than one for both *dip* and *peak* conditions (at α=0.005). The slopes were compared by performing a regression analysis on the difference between thresholds in the *dip* and *peak* condition for each observer, as a function of threshold in the *long* condition for that observer. The slope of the regression line for this difference variable was significantly less than zero, indicating that the dip/peak difference decreased significantly with increasing thresholds in the *long* condition ($F_{1,26}$ =4.39, $p < 0.05$).

As suggested by Fig. 4, NoSπ thresholds in the brief signal conditions are positively correlated with those in the associated *long* condition. For the 500 Hz data, those correlations are R=0.92 and R=0.77 for *dip* and *peak* conditions, respectively, and the correlation between thresholds in the *dip* and *peak* conditions is R=0.80. Corresponding correlations for 2000 Hz data are R=0.75, R=0.73, and R=0.83. This suggests some degree of uniformity in individual differences observed across conditions within frequency. Further, NoSπ thresholds in the *long* duration condition are correlated across the 500 and 2000 Hz signal frequencies, with R=0.78, suggesting that the underlying source of individual differences is not frequency-specific.

C. Discussion

The MLD in the *long* condition was highly variable across observers, with mean (and standard deviations) of 17.2 (5.8) and 8.1 (4.5) dB at 500 and 2000 Hz, respectively. These results are comparable to those of Bernstein *et al.* 1998, who reported means of 15.8 (4.7) and 5.8 (3.1) at 500 and 4000 Hz, respectively. As noted by Bernstein *et al.*, individual differences in the MLD for a long duration signal are more strongly associated with NoSπ (as compared to NoSo) thresholds.

The pattern of MLDs in the 500 Hz *dip* and *peak* conditions replicates the results of Buss *et al.* 2003, showing the dip advantage. The pattern of MLDs was quite different at 2000 Hz, where results indicate a slight peak advantage. As in the previous study, MLDs in the 500 Hz *dip* condition were comparable to those in the *long* condition, suggesting that much of the MLD in the *long* condition could be based on use of binaural cues present in the dip. In contrast, at the 2000 Hz frequency the MLD in the *long* condition exceeded that in either the *dip* or the *peak* condition. This result was not predicted. It is possible that neither the *dip* nor the *peak* condition represents the optimal placement for a signal at 2000 Hz. If the optimal signal placement were at an intermediate point in the masker envelope, associated with more pronounced interaural envelope cues, then binaural release would be larger in the *long* condition where such cues were available (as opposed to the *dip* or *peak* condition). This result may also be associated with the finding of greater temporal integration of binaural information at high frequencies where temporal fine-structure is not available as a cue (Bernstein and Trahiotis, 1999).

As expected, there were extensive individual differences at both frequencies. At 500 Hz, the ‘dip advantage’ for NoSπ thresholds was largest for those observers with lowest thresholds in the 500 Hz *long* condition. Analogously, the peak advantage for NoSπ thresholds at 2000 Hz was largest for those observers who attained the lowest thresholds in the 2000 Hz *long* condition. Whereas it is difficult to assess the relative importance of fine structure versus envelope cues in the peak condition at 500 Hz, the pattern of results obtained here is consistent with an interpretation that good performance in the *dip* condition at 500 Hz is based on differences in sensitivity to interaural fine structure cues. In contrast, good performance in the *peak* condition at 2000 Hz must be associated with cues related to interaural envelope differences. The fact that better-performing observers demonstrate a more pronounced dip/peak difference at both frequencies suggests a range of performance in utilization of both cues.

III. EXPERIMENT 2

One possible interpretation of the results of Experiment 1 is that the individual differences in the MLD can be explained in terms of differences in binaural temporal resolution. If the binaural cue coincident with a masker envelope dip is superior to the binaural cue coincident with a masker envelope peak, then thresholds for a long duration tone should be determined in part by the extent to which the detection is based on cues derived from those brief epochs. Degradation in the temporal specificity with which those cues are derived would be associated with degradation in those cues. Alternatively, the individual differences in the pure-tone MLD may be driven by a factor that is largely independent of the binaural temporal resolution.

Experiment 2 assessed the role of temporal resolution in the use of interaural phase and level cues, and compared the pattern of individual differences with that observed in Experiment 1. Narrowband stimuli were constructed in such a way as to minimize the possible effects of binaural temporal resolution on the detection process. For some of the stimuli used here, the magnitude of the interaural cue resulting from signal-masker interaction was constant across the entire duration of a 200 ms signal presentation. With consistency of the binaural cue across masker dips and peaks, binaural temporal resolution should not be a relevant factor for signal detection. Thus, if binaural temporal resolution were not an important factor contributing to the individual differences observed in Experiment 1, we would expect to see a similar pattern of individual differences in the present experiment.

A. Methods

1. Observers—A subset of 12 observers from Experiment 1 participated. Effort was made to ensure that the observers recruited for this study represented a range of binaural performance, as estimated in Experiment 1. Represented in this group were the observers with the lowest and the highest MLD in the 2000 Hz *long* condition and the observers with lowest and the second highest MLD in the 500 Hz *long* condition.

2. Stimuli—As in Experiment 1, maskers were 50 Hz wide bands of noise centered on the signal frequency (500 Hz), presented at 72 dB SPL. In contrast to the stimulus generation procedures of Experiment 1, maskers in this experiment were generated via multiplication in quadrature in the time domain (for discussion of this method in binaural psychoacoustics, see Amenta *et al.*, 1987). Two independent samples of 25 Hz, lowpass Gaussian noise were generated, $n_A(t)$ and $n_B(t)$. Each noise was multiplied by a 500 Hz tone, where one tone was 90° phase-advanced relative to the other. The two products were summed to produce a bandpass Gaussian noise masker sample, M

$$M(t) = n_A(t) \cdot \cos(\omega t) + n_B(t) \cdot \sin(\omega t).$$

Like the masker, the signal (S) was also a band of Gaussian noise, 50 Hz wide, centered on 500 Hz and generated by multiplication in quadrature. In the ILD condition, the signal was identical to the masker, $S(t) = M(t)$. In the IPD condition, the noise samples used to generate the masker were also used to generate the signal, but the tones multiplied by those samples to generate the signal were 90° phase-shifted relative to those used in generation of the masker

$$S(t) = n_A(t) \cdot \cos\left(\omega t + \frac{\pi}{2}\right) + n_B(t) \cdot \sin\left(\omega t + \frac{\pi}{2}\right).$$

In the RAN condition the signal was generated based on independent samples of lowpass noise. That is

$$S(t) = n_C(t) \cdot \cos(\omega t) + n_D(t) \cdot \sin(\omega t),$$

where $n_A \neq n_C$ and $n_B \neq n_D$. Thus, in the RAN condition the signal and masker were independent samples of bandpass Gaussian noise. In all three signal conditions, stimuli in the NoSo condition were diotic, with the left and right ear defined as $M(t) + \alpha S(t)$, where α controls the amplitude of the signal. In the NoSπ condition, the stimulus delivered to the left ear was defined as $M(t) + \alpha S(t)$, while the stimulus delivered to the right ear was defined as $M(t) - \alpha S(t)$.

As in the *long* conditions of Experiment 1, the signal was 200 ms in duration and temporally centered in the masker. In contrast to Experiment 1, the signal was ramped on and off with 25 ms (as compared to 15 ms) $\cos^2$ ramps, a change that was instituted to guard against spectral artifacts.[1]

For NoSπ signal presentation, these conditions are associated with an interaural level cue only (ILD), an interaural phase cue only (IPD) or both types of cues (RAN). The approach of isolating IPD and ILD cues is similar to that taken by Jeffress and McFadden (1970). It should also be noted that IPD and ILD cues were constant for the duration of the signal in the associated conditions, but cues were fluctuating over the course of the signal presentation in the RAN condition.

Thresholds for both NoSo and NoSπ were estimated in all three conditions. In the NoSo ILD condition, the signal and the masker were identical and added in phase. In contrast, the signal and masker added in random phase or orthogonal phase in the NoSo RAN and IPD conditions, respectively. As a consequence, a given signal level resulted in a larger increase in energy in the NoSo ILD condition as compared to the RAN and IPD conditions.

3. Procedures—Procedures were similar to those used in Experiment 1. The task was a three-alternative forced-choice with feedback provided visually. The signal level was adjusted in a three-down one-up track estimating 79% correct. Conditions were run in blocks, with the order of blocks randomized across observers. A minimum of three threshold estimates was obtained in each condition, and a fourth was obtained if the first three spanned a range of 3 dB or more.

B. Results

The MLDs from Experiment 2 are shown in Fig. 5, plotted following the same convention as used in Fig. 2. The median MLD for the RAN condition was 19.8 dB. This closely resembles MLD in the 500 Hz *long* condition of Experiment 1, supporting the implicit assumption that

---

[1] Previous work suggests that spectral cues do not play a significant role in determining thresholds for the brief tonal signal in the paradigm used in Experiment 1 (Buss *et al.*, 2003). While this is likely to also be the case for the 50 Hz bandpass signal in Experiment 2, temporal specificity was not an issue in this paradigm and 25 ms ramps were adopted as a conservative measure to guard against the influence of spectral cues.

the cues underlying the MLD for the tonal signal of Experiment 1 can be characterized using a narrowband signal. The median MLD for the IPD condition was 16.3 dB, and that for the ILD condition was 4.5 dB.

Thresholds in the NoSo condition were relatively consistent across the three conditions when expressed in terms of ΔL. Median thresholds in the RAN and IPD conditions were 73.1 and 73.6 dB SPL, respectively. Because masker and signal added in random or orthogonal phase, the level of a masker-plus-signal at threshold was approximately 3.3 dB higher than the 72 dB level of a masker alone. The median threshold in the ILD condition was 64.9 dB SPL. Because masker and signal added in phase, the level of a masker-plus-signal at threshold was approximately 3.1 dB higher than a masker alone. These results are generally in agreement with those of Bos and de Boer (1966) for a similar stimulus bandwidth.

Because NoSo thresholds were not constant across conditions when expressed in terms of the level of the signal, differences in the MLD across conditions do not solely reflect differences in NoSπ thresholds. The median NoSπ threshold was 55.8 dB in the RAN condition, 60.7 dB in the IPD condition, and 62.6 dB in the ILD condition. A repeated measures ANOVA resulted in a significant effect of condition ($F_{2,22}$=47.18, $p < 0.0001$). *Post hoc* contrasts revealed that NoSπ thresholds in the RAN condition were lower than those in the IPD condition ($p < 0.001$), and thresholds in the IPD condition were lower than those in the ILD condition ($p < 0.05$).

As in conditions with pure tone signals from Experiment 1, MLDs with narrowband signals are associated with marked individual differences, and the variability in the MLD was more highly correlated with NoSπ than NoSo thresholds (mean R=−0.91 and R=−0.26, respectively). Individual differences in NoSπ thresholds were highly correlated with those in Experiment 1. Figure 6 shows NoSπ threshold in the narrowband noise conditions of Experiment 2 as a function of associated thresholds in the 500 Hz *long* condition of Experiment 1. Correlations between NoSπ thresholds in the three conditions tested here and the six conditions of Experiment 1 were all significant ($p < 0.005$, one-tailed) and ranged between R=0.72 and R=0.95, with an average correlation of approximately R=0.85 for each of the three narrowband signal conditions. These correlations are comparable to those among NoSπ thresholds from Experiment 1.

C. Discussion

The pattern of individual differences obtained with narrowband signals was quite similar to that measured in Experiment 1 with a pure tone signal, both in the *long* condition and the two conditions utilizing a brief signal (the *dip* and *peak* conditions). A subset of the stimuli in Experiment 2 was designed to minimize the possibility that binaural temporal resolution could have contributed to the detection of the signal. Under these conditions, the pattern of individual differences was quite similar to that found in Experiment 1, with NoSπ thresholds correlating highly across the experiments. Considered together, these results suggest that it is unlikely that binaural temporal resolution plays a material role in the individual differences that characterize the MLD for narrow-band masking noise. Instead, the results suggest that there are “good performers” and “poor performers,” a categorization that applied in a similar way in all binaural (NoSπ) conditions of Experiments 1 and 2, irrespective of the temporal fluctuation of the binaural cue. Comparable results in MLD, interaural time discrimination and interaural intensity discrimination conditions have been reported previously (e.g., Koehnke *et al.*, 1986).

The approach taken here bears some resemblance to that of Bernstein *et al.* 1998. In that study, individual differences in sensitivity were compared for narrowband MLD and for interaural time and level differences. Individual differences across conditions were weakly correlated ($R \approx 0.4$) when compared to the current data ($R \approx 0.8$). One possible reason for this discrepancy is that Bernstein *et al.* used an interaural difference paradigm that featured a 400 Hz wide stimulus and compared it to NoSπ thresholds for a 50 Hz wide masker, whereas the current experiments used a 50 Hz bandwidth for both paradigms. Given the closer correspondence of the conditions used in the current study, it seems likely that individual differences in NoSπ signal detection are closely related to those in the discrimination of interaural phase (or time) and level under comparable conditions.

IV. GENERAL DISCUSSION

The results of Experiment 1 support the hypothesis that the dip advantage observed previously with 500 Hz stimuli does not occur for comparable stimuli at 2000 Hz, where temporal fine-structure information is no longer useful for binaural hearing. Rather, a peak advantage at 2000 Hz was observed, presumably due to interaural envelope cues in masker peaks. Similar individual differences occurred in both the long duration and short duration NoSπ conditions. Results of Experiment 2 suggest that the individual differences obtained in Experiment 1 are not related to differences in the temporal acuity with which fluctuating IPD and ILD cues are utilized. These results are consistent with an interpretation that limitations in binaural temporal resolution are not responsible for the individual differences in NoSπ thresholds for a narrowband noise masker.

One very common way to characterize the binaural cue underlying the MLD is in terms of the reduction in interaural correlation associated with addition of a signal (e.g., Durlach *et al.*, 1986; for discussion see Bernstein and Trahiotis, 1996). Previous data for a brief 500 Hz tonal signal coincident with masker envelope dips and peaks were well characterized in terms of correlation at the output of a simplified model of the auditory periphery (Buss *et al.*, 2003). That model closely followed the approach developed by Bernstein, Trahiotis and colleagues (e.g., Bernstein *et al.*, 1999; Bernstein and Trahiotis, 1996), but also included a rectangular temporal window centered on the signal. A similar approach was pursued here to see if the individual differences observed could be characterized in terms of differential sensitivity to changes in interaural correlation. In contrast to the previous modeling efforts, a double exponential temporal window was adopted here, similar to the one that has been fitted to binaural temporal window data (e.g., Kollmeier and Gilkey, 1990).

Threshold estimates as a function of differential criterion levels of correlation were obtained with the following procedures, executed in MATLAB. Stimuli were generated using the same algorithm as used in Experiment 1. The brief *Sπ* signal was added to the diotic masker at a range of *S*/*N* ratios, corresponding to signal levels of 50–95 dB SPL, in either the *dip* or the *peak* condition. These stimuli were then submitted to a binaural normalized correlation model which included the following stages: (1) exponential compression of the envelope, with an exponent of 0.23;[2] (2) half-wave, square-law rectification; (3) low-pass filtering;[3] (4) application of the double exponential window centered on the signal;[4] and (5) calculation of the normalized correlation (Bernstein and Trahiotis, 1996). In order to capture possible effects of off-time listening, results were computed for a family of nine temporal window placements, equally spaced between −15 and 15 ms relative to the temporal center of the signal; the correlation estimate associated with each sample was the minimum across these nine possibilities. This procedure was repeated for 1000 masker samples. The first 25 correlation functions in each condition are shown in Fig. 7.

---

[2] Results were not very sensitive to the value of this exponent. Lopez-Poveda *et al.* 2003 have reported different compressive functions at 500 and 2000 Hz—values of 0.23 and 0.28, respectively. Using these different exponents did not materially change the results reported.

[3] Filter parameters are those described by Bernstein and Trahiotis (1996).

[4] This window was taken from Kollmeier and Gilkey (1990). The mean of the time constant measured in that study (τ =25 ms) was adopted.

The signal level associated with a criterion correlation spanning 0.45–0.95 (the range associated with observer's thresholds in the 500 Hz *dip* condition) was then determined based on the correlation functions for *dip* and *peak* conditions at both 500 and 2000 Hz. At 500 Hz, results were quite consistent with the pattern of results obtained experimentally: There was a dip advantage overall and the dip/peak differences grew with increases in the criterion level of correlation (with larger values thought to characterize performance of the better-performing subjects).

This approach was less consistent with 2000 Hz data. Consistent with experimental results, the signal level associated with a given criterion level of correlation was higher for 2000 Hz than 500 Hz conditions. However, for the 2000 Hz peak condition there was no indication of a binaural advantage (i.e., estimated signal levels exceeded NoSo thresholds) in all but two cases. In contrast, thresholds were lower in the NoSπ than the NoSo condition for 27 of 28 observers. Perhaps most inconsistent with empirical data, at a correlation of 0.98, corresponding to a signal level at threshold in the 500 Hz *dip* condition for the two most sensitive observers, thresholds in the 2000 Hz *dip* condition were underestimated by over 10 dB. The reason for this underestimation can be seen in Fig. 7. A criterion correlation of 0.98 corresponds to a signal level of approximately 60 dB SPL in the 2000 Hz *dip* condition and over 70 dB SPL in the 2000 Hz *peak* condition, and thus a dip advantage. This failure to predict a peak advantage at 2000 Hz is quite consistent and remains relatively unchanged with the use of a broader temporal window or a window with less pronounced central emphasis (e.g., a rectangular or a Hanning window).

Thus, constant criterion correlation at threshold appears to capture some trends in the data, particularly at 500 Hz, but appears to be inconsistent with the modest MLD and peak advantage obtained at 2000 Hz. This result is in contrast to those of Bernstein and Trahiotis (1999), where a range of binaural data with long-duration signals was successfully modeled based on constant criterion correlation, including data collected with a 50 Hz wide band of Gaussian noise centered on 500 Hz. This discrepancy suggests that some aspect of the model may not capture the transient high-frequency cues examined in the current study.

The current approach assumes that any difference in binaural hearing between 500 and 2000 Hz is due to the elimination of cues based on fine-structure at the higher frequency. This is a common assumption in the binaural literature (e.g., Zurek and Durlach, 1987), and one that has received empirical support. For example, the MLD for a narrowband masker is comparable at low and high frequencies when the low-frequency fine-structure is preserved in the envelope at high frequencies (van de Par and Kohlrausch, 1997). There is some evidence that temporal processing of binaural information at low and high frequencies may differ in detail, however. For example, sensitivity to dynamic changes in interaural intensity may be better at high than at low frequencies (Grantham, 1984). Temporal integration of binaural information has also been argued to differ fundamentally for frequencies where fine-structure information is available and frequencies where binaural differences are envelope-based (Bernstein and Trahiotis, 1999).

Another feature of the present approach that might be problematic is the assumption that correlation in the dips and peaks are equally useful as detection cues. If, instead, forward masking or some other level-dependent factor rendered correlation in the dips a less effective cue, the result would be an underestimate of thresholds in the *dip* condition. Using wider stimulus bandwidths than those employed here, forward masking has been shown to play a role in the MLD, with NoSπ thresholds following the same decline as a function of temporal separation for masker offset as NoSo thresholds (Deatherage and Evans, 1969; Yama, 1992). Further investigations are currently underway to explore the possible role of forward masking on the ability to make use of binaural information in the dips of narrowband maskers.

Regardless of whether the NoSπ data of Experiment 1 can ultimately be modeled in terms of correlation at threshold, results of Experiment 2 suggest that the predominant source of individual differences does not lie in differential ability to resolve brief interaural correlation cues or in differential sensitivity to IPD and ILD. Across the 12 listeners who participated in both Experiments 1 and 2, correlations between NoSπ thresholds in the 500 Hz *long* condition and those in the narrowband signal conditions associated with steady interaural cues (ILD and IPD) were R=0.83. This is comparable to the correlation of R=0.95 for the fluctuating-cue RAN condition. Because the signal in the ILD condition of Experiment 2 did not introduce interaural correlation cues other than those at the signal onset and offset, this consistency of individual differences may reflect a more basic aspect of binaural auditory processing. This is consistent with the suggestion that there are good binaural listeners and poor binaural listeners (Bernstein *et al.*, 1998), and that the factor underlying performance is common to both IPD and ILD (Koehnke *et al.*, 1986).

V. CONCLUSIONS

1. The MLD for a brief 500 Hz pure tone signal in a 50 Hz wide band of noise is larger if that signal is coincident with a masker envelope dip as opposed to an envelope peak. This result has been attributed to a dip advantage for utilizing the binaural cue in the NoSπ condition. The dip/peak difference is reversed at 2000 Hz, where there is a slight peak advantage.
2. A similar pattern of individual differences in NoSπ thresholds was obtained for brief and for long duration tones, as well as for 500 and 2000 Hz signal frequencies. This is consistent with the idea that there are good binaural listeners and poor binaural listeners.
3. Thresholds for detection of a narrowband signal associated with nonfluctuating interaural phase or intensity difference cues demonstrated the same pattern of individual differences as seen with pure tones (and fluctuating cues). This result suggests that temporal resolution is not a dominant factor contributing to individual differences for an *S*π tone in No narrowband noise.
4. Individual differences in criterion levels of interaural correlation are consistent with 500 Hz NoSπ thresholds for pure tones presented coincident with masker dips and peaks. The 2000 Hz pure tone data are less consistent with this approach; a cue based on interaural correlation, at least as computed here, is consistent with an elevation in NoSπ thresholds at 2000 Hz, but not with a significant release from masking (MLD) and not with the peak advantage observed empirically.

Acknowledgements

This work was supported by NIH Grant No. R01 DC00397. A partial report of these data was presented at the 149th meeting of the Acoustical Society of America (Buss, Hall, and Grose , **117**, 2563). The authors are grateful to Associate Editor Armin Kohlrausch, Bill Hartman, and an anonymous reviewer for helpful comments on earlier versions of this manuscript.

References

Amenta CA III, Trahiotis C, Bernstein LR, Nuetzel JM. Some physical and psychological effects produced by selective delays of the envelope of narrow bands of noise. Hear Res 1987;29:147–161. [PubMed: 3624080]

Bernstein LR, Trahiotis C. On the use of the normalized correlation as an index of interaural envelope correlation. J Acoust Soc Am 1996;100:1754–1763. [PubMed: 8817901]

Bernstein LR, Trahiotis C. The effects of signal duration on NoSo and NoSπ thresholds at 500 Hz and 4 kHz. J Acoust Soc Am 1999;105:1776–1783. [PubMed: 10089601]

Bernstein LR, Trahiotis C, Akeryod MA, Hartung K. Sensitivity to brief changes of interaural time and interaural intensity. J Acoust Soc Am 2001;109:1604–1615. [PubMed: 11325131]

Bernstein LR, Trahiotis C, Hyde EL. Inter-individual differences in binaural detection of low-frequency or high-frequency tonal signals masked by narrowband or broadband noise. J Acoust Soc Am 1998;103:2069–2078. [PubMed: 9566329]

Bos CE, de Boer E. Masking and discrimination. J Acoust Soc Am 1966;39:708–715.

Buss E, Hall JW III, Grose JH. The masking level difference for signals placed in masker envelope minima and maxima. J Acoust Soc Am 2003;114:1557–1564. [PubMed: 14514209]

Deatherage BH, Evans TR. Binaural masking: Backward, forward, and simultaneous effects. J Acoust Soc Am 1969;46:362–371. [PubMed: 5804106]

Durlach NI, Gabriel KJ, Colburn HS, Trahiotis C. Interaural correlation discrimination: II. Relation to binaural unmasking. J Acoust Soc Am 1986;79:1548–1557. [PubMed: 3711454]

Grantham DW, Wightman FL. Detectability of a pulsed tone in the presence of a masker with time-varying interaural correlation. J Acoust Soc Am 1979;65:1509–1517. [PubMed: 489821]

Grantham DW. Discrimination of dynamic interaural intensity differences. J Acoust Soc Am 1984;76:71–76. [PubMed: 6747114]

Grose JH, Hall JW III. Masker fluctuation and the masking-level difference. J Acoust Soc Am 1998;103:2590–2594. [PubMed: 9604353]

Hall JW III, Buss E, Grose JH, Dev MB. Developmental effects in temporal resolution of binaural hearing. J Speech Lang Hear Res 2004;47:13–20. [PubMed: 15072524]

Hirsh IJ. The influence of interaural phase on interaural summation and inhibition. J Acoust Soc Am 1948;20:536–544.

Holube I, Kinkel M, Kollmeier B. Binaural and monaural auditory filter bandwidths and time constants in probe tone detection experiments. J Acoust Soc Am 1998;104:2412–2425. [PubMed: 10491703]

Jeffress LA, McFadden D. Differences of interaural phase and level in detection and lateralization. J Acoust Soc Am 1970;4:1169–1179.

Koehnke J, Colburn HS, Durlach NI. Performance in several binaural-interaction experiments. J Acoust Soc Am 1986;79:1558–1562. [PubMed: 3711455]

Kollmeier B, Gilkey RH. Binaural forward and backward masking: Evidence for sluggishness in binaural detection. J Acoust Soc Am 1990;87:1709–1719. [PubMed: 2341675]

Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am 1971;49:467–477. [PubMed: 5541744]

Lopez-Poveda EA, Plack CJ, Meddis R. Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing. J Acoust Soc Am 2003;113:951–960. [PubMed: 12597188]

McFadden D, Pasanen EG. Binaural detection at high frequencies with time-delayed waveforms. J Acoust Soc Am 1978;63:1120–1131. [PubMed: 649871]

van de Par S, Kohlrausch A. A new approach to comparing binaural masking level difference at low and high frequencies. J Acoust Soc Am 1997;101:1671–1680. [PubMed: 9069634]

Yama M. Effects of temporal separation and masker level on binaural analysis in forward masking. J Acoust Soc Am 1992;91:327–335. [PubMed: 1737881]

Zurek PM, Durlach NI. Masker-bandwidth dependence in homophasic and antiphasic tone detection. J Acoust Soc Am 1987;81:459–464. [PubMed: 3558963]

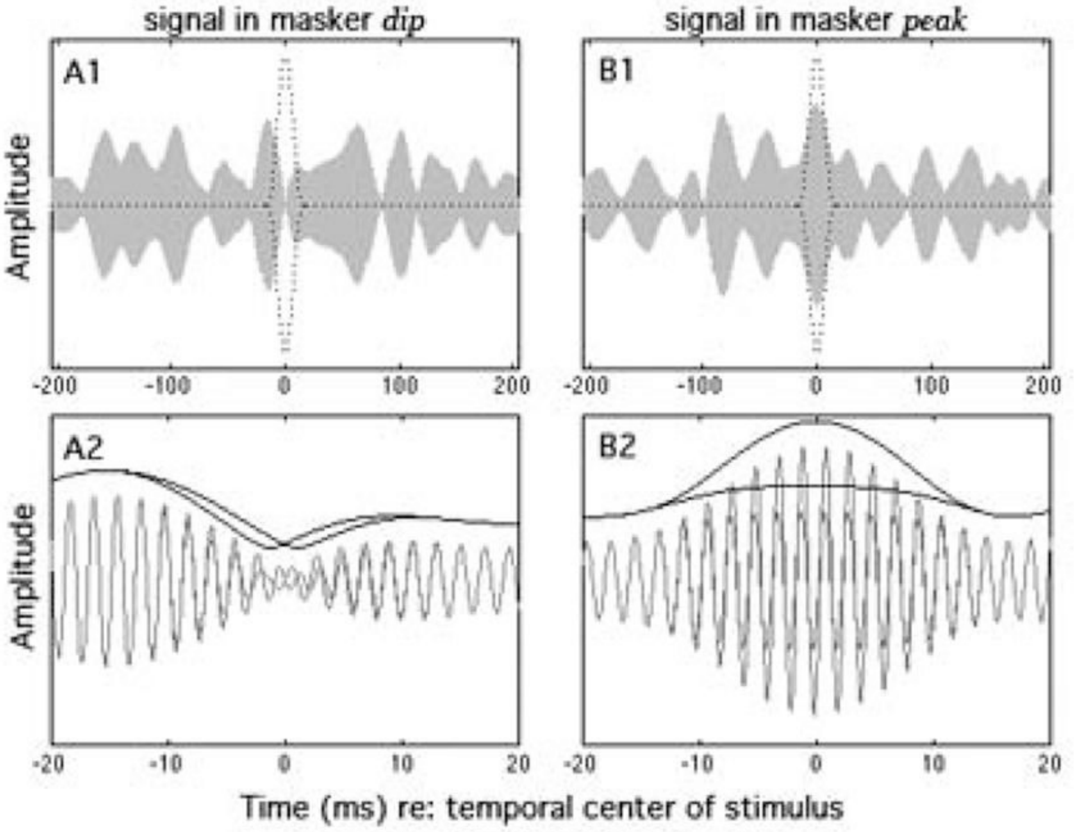

FIG. 1.

Illustration of interaural cues available at threshold for a brief signal in a narrowband masker in the NoSπ condition, based on stimuli and thresholds from Buss *et al.* 2003. The left column of panels corresponds to a signal presented in a masker envelope dip and the right column to a signal presented in a masker envelope peak. The top row shows amplitude of a representative masker as a function of time, with dotted lines indicating the temporal position (but not the level) of an added signal. The bottom row of panels is a magnified view of the stimuli presented to the left and right ears (overlaid traces) with a diotic masker (as shown in the top row) and an *S*π signal at threshold (60 dB for the dip condition and 70 dB SPL for the peak condition). Grey lines indicate the temporal fine-structure and dark lines show the raised envelope.

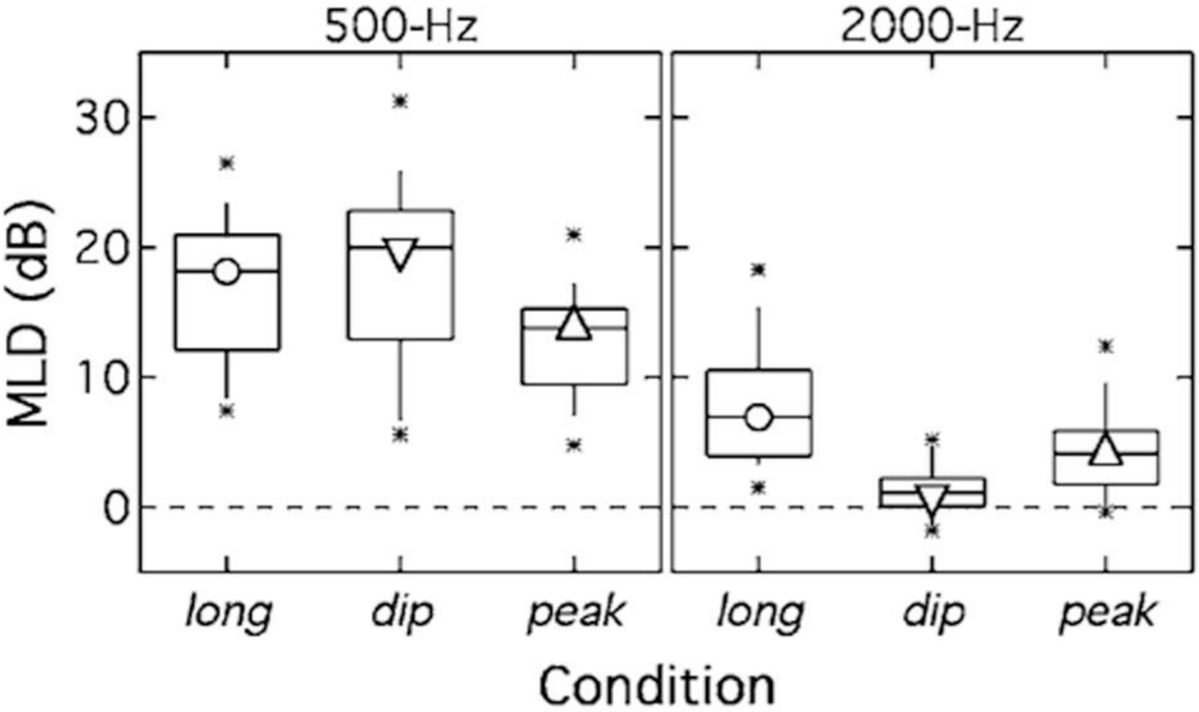

FIG. 2.
The distribution of MLDs is plotted for 500 Hz (left panel) and 2000 Hz (right panel) conditions. Abscissa labels and symbols indicate condition. The median of each distribution is indicated with a symbol, the span between the 25th and the 75th percentiles is indicated with a box, the 10–90th percentile with bars, and the maximum and minimum of each distribution with stars.

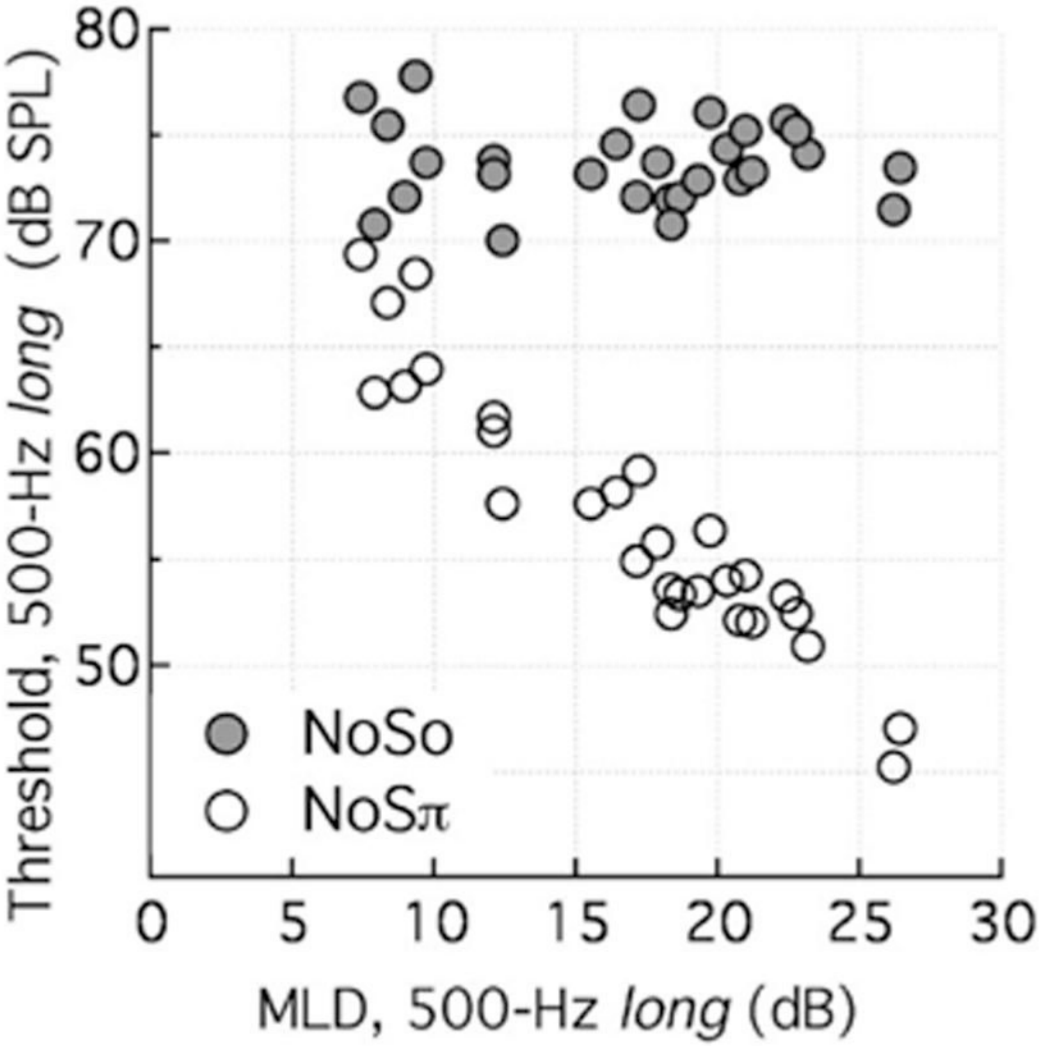

FIG. 3.

For each observer, NoSo and NoSπ thresholds in the 500 Hz *long* condition are plotted in dB SPL as a function of the MLD in the 500 Hz *long* condition, in dB.

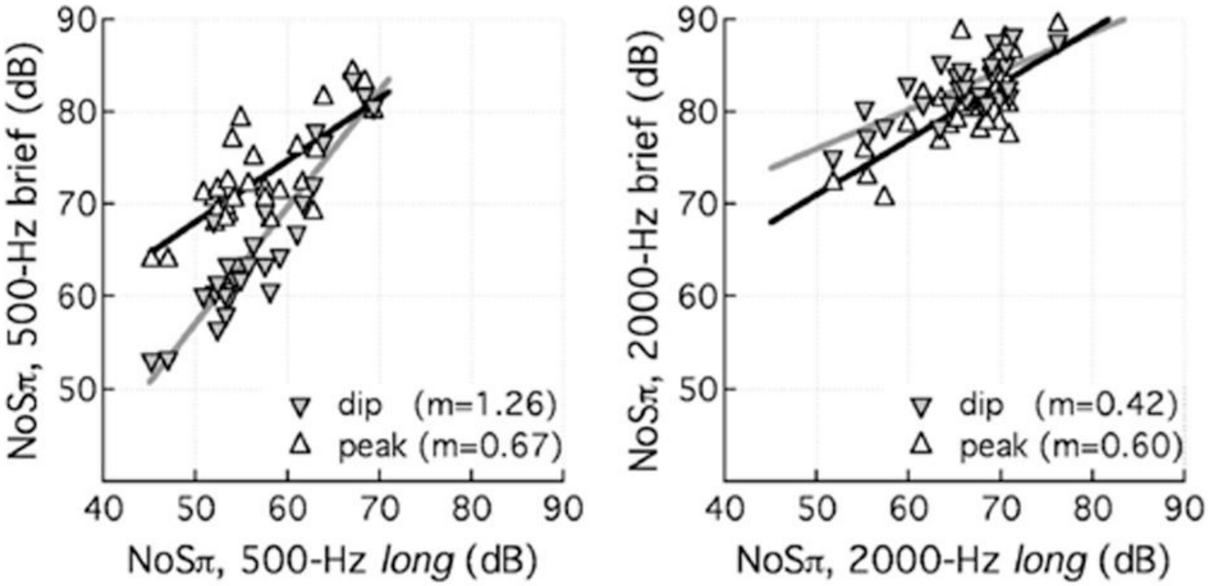

FIG. 4.

NoSπ thresholds for brief signals are plotted as a function of NoSπ thresholds in the *long* condition. The left panel indicates data for the 500 Hz conditions, and the right for the 2000 Hz conditions. Upward-pointing triangles indicate *peak* data, downward-pointing triangles indicate *dip* data, and lines indicate the best linear fit to the data (grey=dip; black=peak). The *m*-values indicated in the lower right corner of each panel show the slopes of these linear fits.

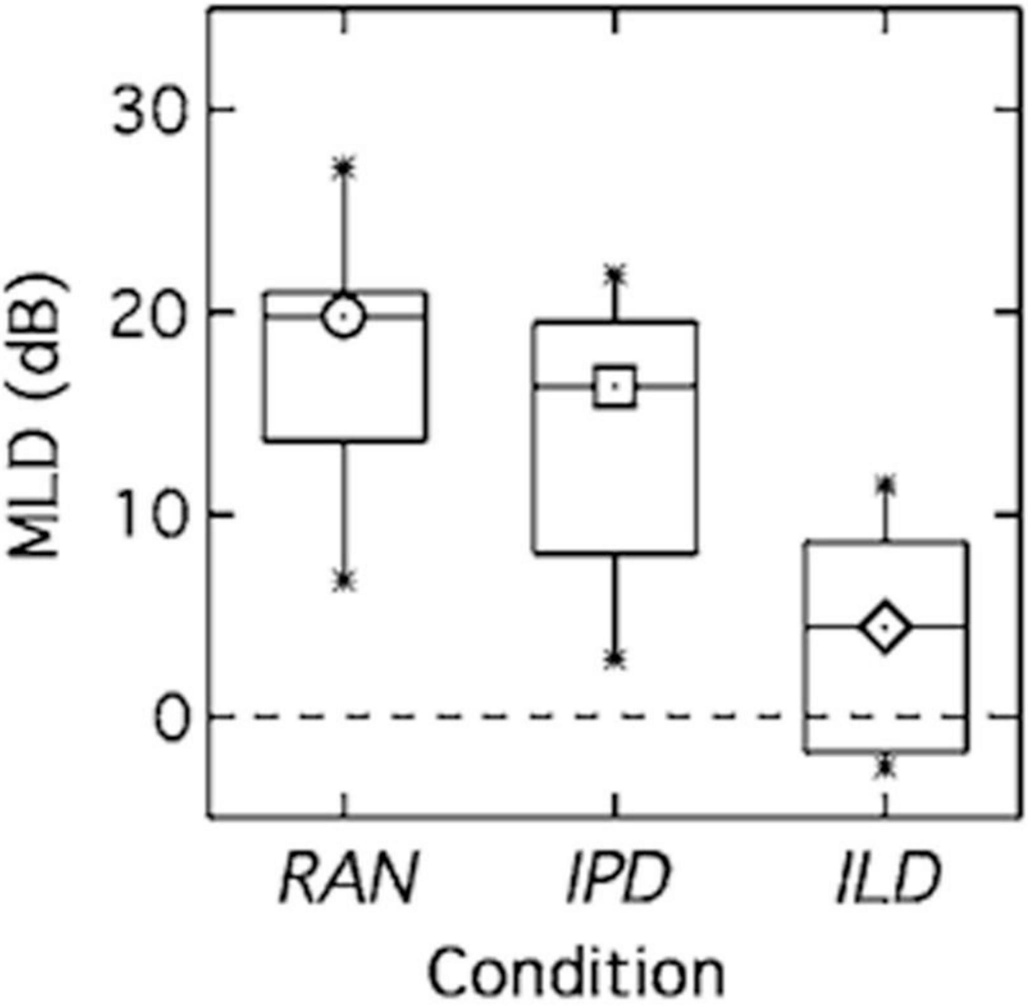

FIG. 5.

The distributions of MLDs in Experiment 2 using narrowband noise signals are plotted following the conventions of Fig. 2. Due to differences in signal stimuli, NoSo thresholds were not constant across conditions. As a result, these values of MLD do not solely reflect differences in the NoSπ threshold.

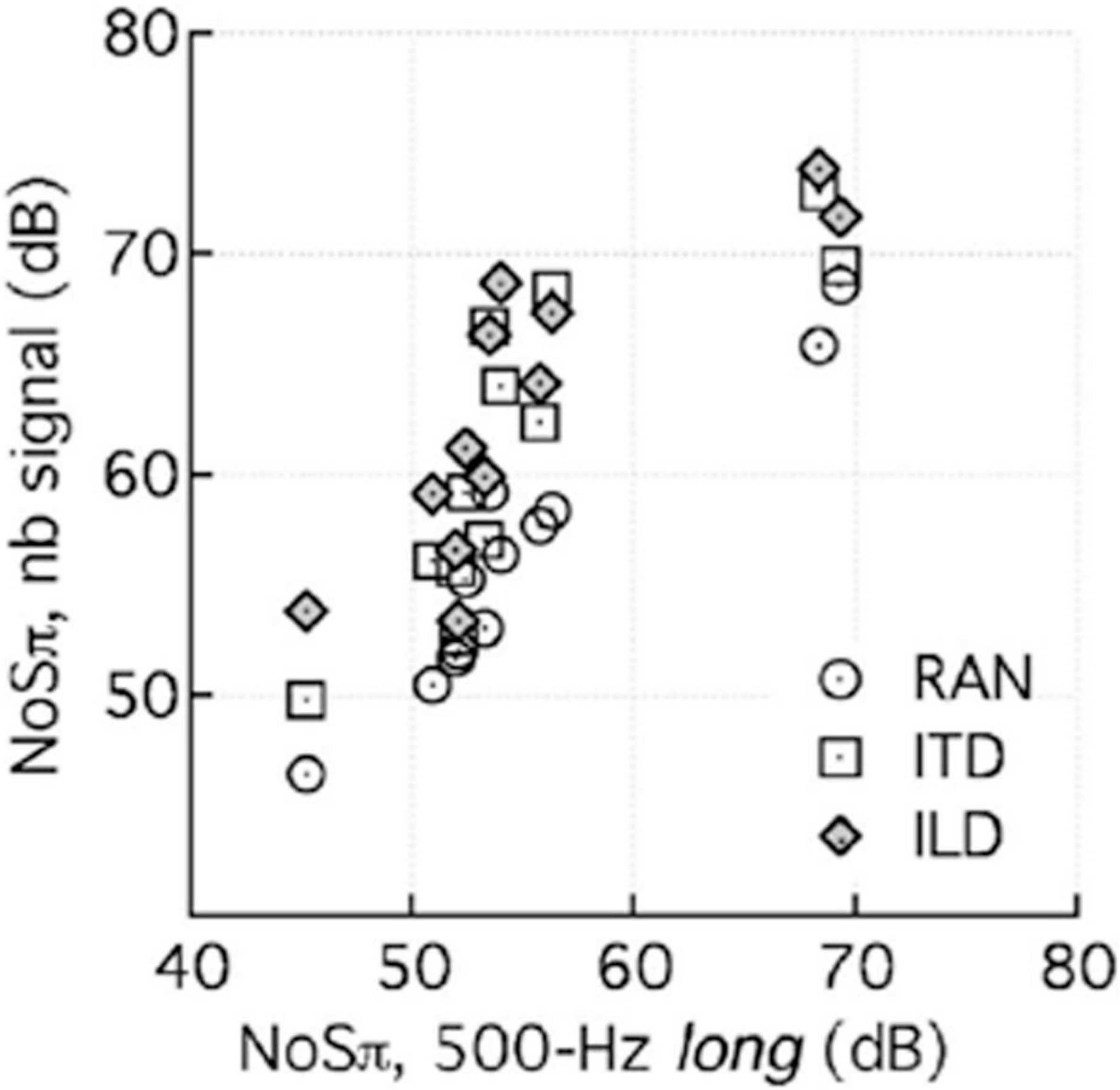

FIG. 6.

NoSπ thresholds for narrowband signals of Experiment 2 are plotted as a function of NoSπ thresholds in the 500 Hz *long* condition of Experiment 1. Dotted circles indicate RAN thresholds, squares indicate ITD thresholds, and diamonds indicate ILD thresholds.

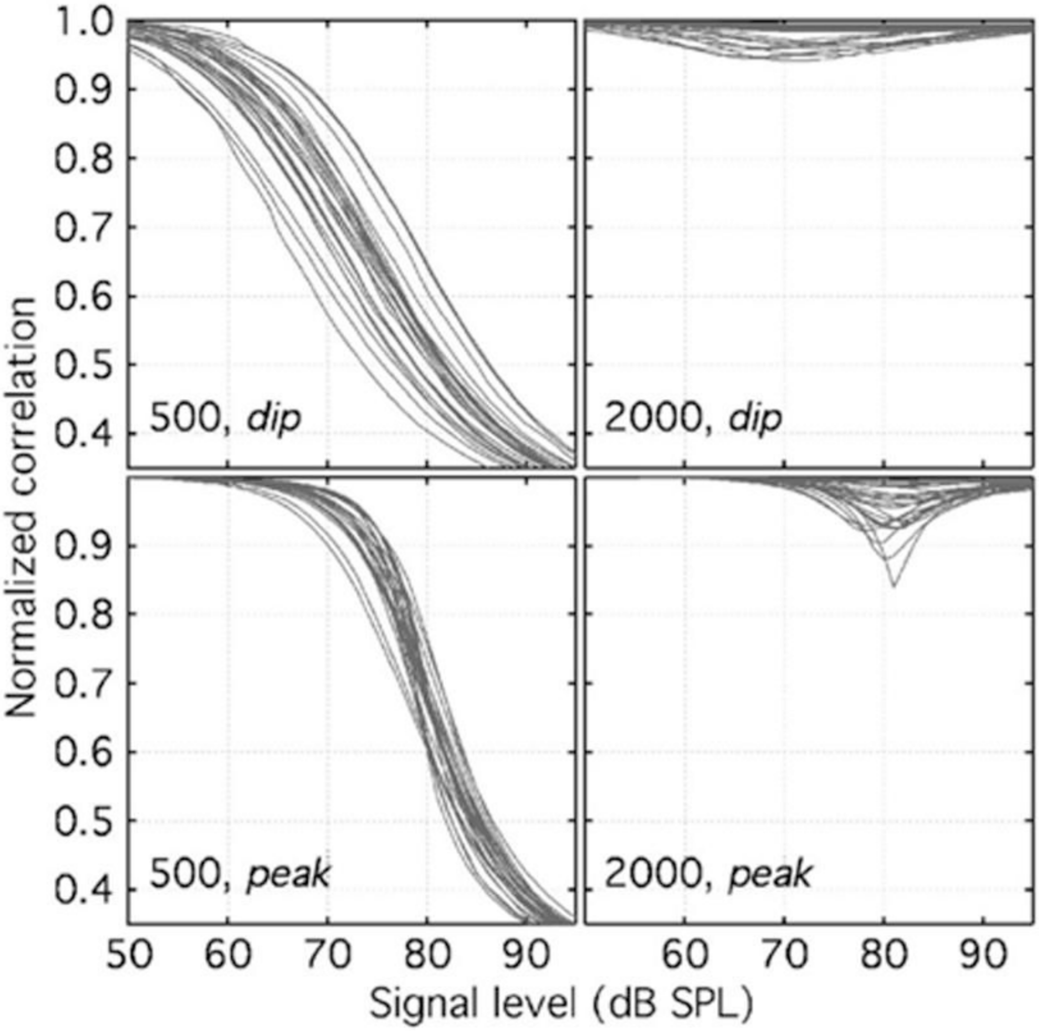

FIG. 7.

Normalized correlation for a range of signal levels is plotted for the first 25 masker+signal tokens used to predict thresholds. Data for the 500 Hz stimuli appear in the left column and those for 2000 Hz data in the right; the top row shows results for the *dip* conditions, and the bottom for the *peak* conditions.