Spatial unmasking of nearby pure-tone targets in a simulated anechoic environment

Norbert Kopčo and Barbara G. Shinn-Cunningham[a)]
Hearing Research Center, Boston University, Boston, Massachusetts 02215

(Received 31 December 2002; revised 12 August 2003; accepted 15 August 2003)

Detection thresholds were measured for different spatial configurations of 500- and 1000-Hz
et al.

can lead to large changes in the energy of the target and masker reaching the two ears. A few previous studies hint that, in some conditions, binaural performance can be worse than monaural performance using the better ear, particularly when there are large ILDs in the stimuli (e.g., see Bronkhorst and Plomp, 1988; Shinn-Cunningham *et al.*, 2001). Given that large ILDs can arise when sources are within reach of the listener, studies of binaural unmasking for nearby sound sources may shed light on these reports.

The current study examined spatial unmasking of pure tone sources within reach of a listener in a simulated anechoic environment. Individually measured head-related transfer functions (HRTFs) were used to simulate sources. This approach allowed realistic spatial acoustic cues to be presented to the subjects while still allowing detailed analyses of the stimuli reaching the subjects during the experiment. The main goals of the study were to (1) measure how target threshold depends on target and masker azimuth and distance for nearby sources, (2) characterize better-ear effects by analyzing how the TMR varies with the spatial configurations tested, (3) evaluate the binaural contribution to spatial unmasking, particularly for spatial configurations in which large ILDs arise, and (4) investigate the degree to which results can be accounted for by a model of binaural interaction.

II. SPATIAL UNMASKING OF NEARBY PURE TONE TARGETS

A. Methods

1. Subjects

Four graduate students with prior experience in psychoacoustic experiments (including author NK) participated in the study. One subject was female and three were male. Subject ages ranged from 25 to 28 years. All subjects had normal hearing as confirmed by an audiometric screening.

2. HRTF measurement

Individualized HRTF measurements were made with subjects seated in the center of a quiet classroom (rough dimensions of $5\times9\times3.5$ m; broadband T_{60} of approximately 700 ms). Subjects were seated with their heads in a headrest so that their ears were approximately 1.5-m above the floor. Measurements were taken for sources in the right front horizontal plane (at ear height) for all six combinations of azimuths (0°, 45°, 90°) and distances (0.15 m, 1 m) relative to the center of the head (defined as the intersection of the interaural axis and the median plane) as shown in Fig. 1.

The Maximum-Length-Sequence (MLS) technique (e.g.,

simulated in the left hemifield, this approximation should introduce no significant perceptual artifacts in the simulated stimuli).

The measured HRTFs reflect the radiation characteristics of the loudspeaker used, which is not a uniformly radiating point source. For sources relatively far from the head, any differences in the measurement caused by the directivity of the source should be minor. For sources 15-cm from the center of the head, the effect of the source directivity may be significant. Therefore, the current study focuses on how distance influenced the signals reaching the ears for the particular source used (the Bose loudspeaker in question). The issue of how well the current results may generalize to other nearby sources is considered further in Sec. III, where empirical HRTF measurements are compared with theoretical predictions from a spherical head model that assumes a perfect point source.

In a similar vein, HRTFs measured for sources close to the head are much more sensitive to small displacements in the source (re: the intended source location) than more distant sources. However, given that all acoustic analyses and predictions of performance were made using the same measured HRTFs used to simulate the headphone-presented stimuli, any conclusions regarding which acoustic factors influence performance are justified, even if other measurement techniques might yield slightly different estimates of near-source HRTFs for the positions reported here.

3. Stimulus generation

Target stimuli consisted of 165-ms-long pure tones of 500 or 1000 Hz gated on and off by 30-ms cos-squared ramps. The 500-Hz target frequency was chosen so that results could be compared with previous studies of binaural masking level differences (BMLDs) and spatial unmasking of tones, most of which include a 500-Hz target condition. The 1000-Hz target was included in order to examine what happens for a higher target frequency where target and masker ITDs are still likely to have a large impact on detection but ILDs are larger than at 500 Hz. The target was temporally centered within a broadband, 250-ms-long masker. On each trial, the masker token was randomly chosen from a set of 100 pregenerated samples of broadband noise that were digitally low-pass filtered with a 5000 Hz cutoff frequency (ninth-order Butterworth filter, as implemented in the signal-processing toolbox in Matlab, the Mathworks, Natick, MA).

In most cases, target and masker were simulated as arising from different locations in anechoic space by convolving the stimuli with appropriate individualized head-related impulse responses (time-domain representation of the HRTFs). The simulated spatial configurations included all combinations of target at azimuths ($-90°$, $-45°$, 0°, 45°, 90°) and distances (0.15 m, 1 m) and masker at azimuths (0°, 45°, 90°) and distances (0.15 m, 1 m). A total of 60 spatial configurations was tested (10 target locations $\times$ 6 masker locations; see Fig. 1). In a subset of trials, traditional BMLDs were measured using the same stimuli without HRTF processing.

For nearby sources, keeping the masker presentation level constant would result in the received level (at the subject's ears) varying widely with masker position. In order to keep the received level of masker relatively constant, the levels of the HRTF-processed masker stimuli were normalized to keep constant the rms energy falling within the equivalent rectangular band (ERB; Moore, 1997) centered on the target frequency at the ear receiving the more intense masker signal (the right ear for all of the tested configurations). In other words, the virtual stimuli actually simulated a masker whose distal energy level was adjusted up or down (depending on the masker spatial location) until the proximal stimulus level was constant at the more intense ear. In our analysis, the amounts by which the distal masker was adjusted were added back to the raw thresholds to predict the amount of spatial unmasking that would have occurred if the distal masker level had been constant.[2]

For the 500-Hz center frequency, the rms levels were adjusted using a 100-Hz-wide ERB. For the 1000-Hz target, the ERB width was set to 136 Hz. The masker signals were preprocessed in Matlab so that the right- (more-intense-) ear rms masker level in the ERB would be 64 dB SPL when played via headphones. BMLDs were measured with the low-pass-filtered noise spectral level fixed at 64 dB SPL.

At least three separate runs were performed for each subject in each condition. Final threshold estimates were computed by taking the average threshold across the repeated adaptive threshold estimates. Additional adaptive runs were performed as needed for every subject and condition to ensure that the standard error in this final threshold estimate was less than or equal to 1 dB for each condition and spatial configuration tested.

The study was divided into two parts, one measuring thresholds for the 500-Hz target and one for the 1000-Hz target. Three subjects performed each part (two of the four subjects performed both). For each target, subjects performed multiple sessions consisting of ten runs. Subjects were allowed to take short breaks between runs within one session, with a minimum 4-h break required between sessions. Each subject performed one initial practice session consisting of four practice runs and six runs measuring detection thresholds for NoSo and NoSπ conditions (where NoSo represents a sinusoidal diotic signal, i.e., with zero interaural phase difference, in the presence of a diotic noise; NoSπ represents a sinusoidal signal with interaural phase difference equal to π in the presence of a diotic noise). Subjects then performed 18 additional sessions (180 runs; 3 runs each of every combination for 6 target positions and 10 masker positions). In each of these sessions, a full set of thresholds was determined for one masker position (the order of the ten target positions was randomized within each session). These sessions were grouped into three blocks of six with each block containing a full set of thresholds. The order of masker positions was separately randomized for each block and subject. Any additional runs were performed after completion of the initial 19 sessions. Each subject performed approximately 20 h of testing per target frequency.

B. Results

1. Binaural masking level difference

Table I shows the BMLD (see Durlach NoSoke2TD [0 Ts6.e2TD [1(egmc(ta2TD [bromere2TD [1)3 [1) 8 ocks427283.6ks427283c

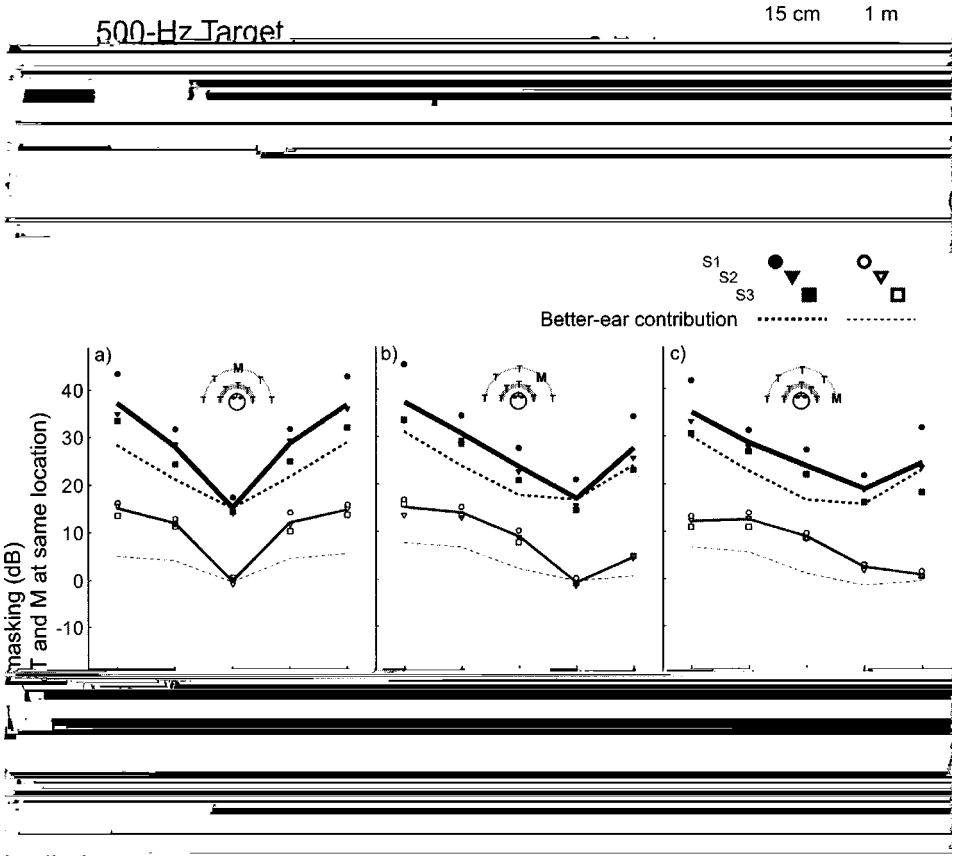

FIG. 2. Spatial unmasking for the 500-Hz target. Each panel plots spatial unmasking (the difference between target detection threshold when target and masker are at the same spatial location and when target and masker are in the spatial configuration denoted in the plot) as a function of target azimuth for a fixed masker location. Across-subject averages are plotted for target distances of 15-cm (thick solid lines) and 1-m (thin solid lines). Individual subject results are plotted as symbols. Dashed lines show the estimated better-ear contribution to spatial unmasking. The spatial configurations of target and masker represented in each panel are denoted in the panel legend.

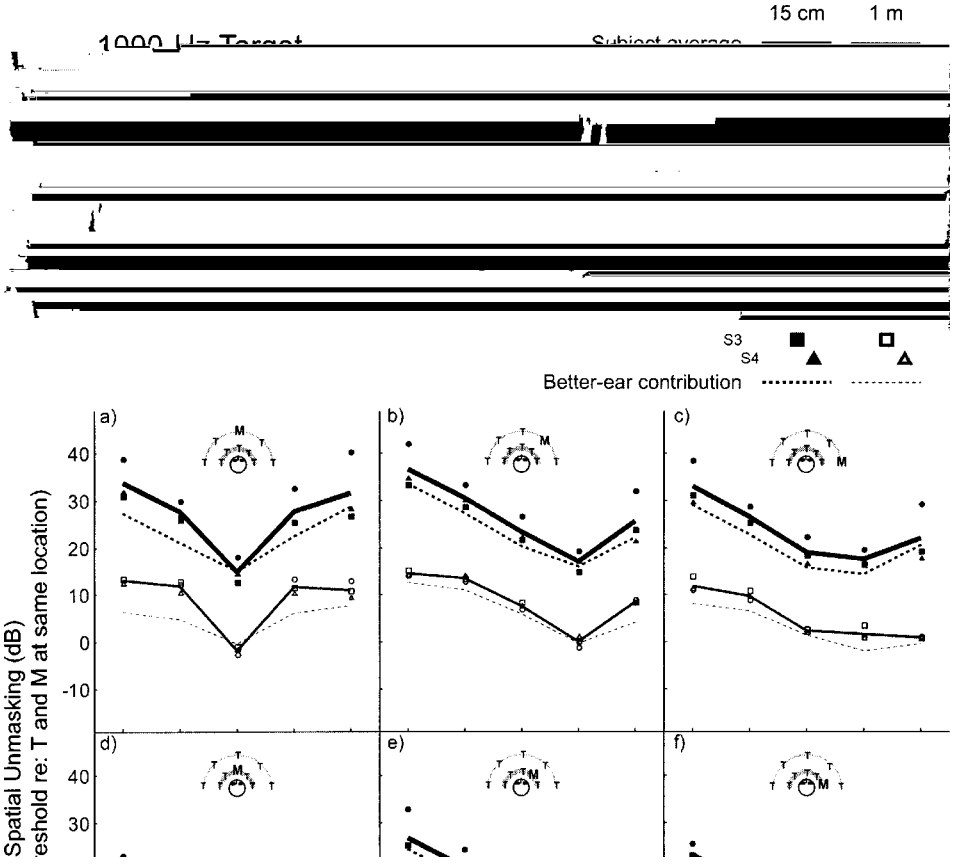

FIG. 3. Spatial unmasking for the 1000-Hz target. See caption for Fig. 2.

for distant targets. For example, in Fig. 3e, the difference between thresholds for the −90° and 45° targets is more than 25 dB for nearby targets (thick line) but less than 20 dB for distant targets (thin line).

Similarly, spatial unmasking resulting from a fixed angular separation of target and masker is larger for nearby maskers than for distant maskers. For example, as discussed above, for a 1000-Hz target when the masker is at (45°, 15 cm) [Fig. 3(e)], spatial unmasking for a 15-cm target (thick line) decreases by more than 25 dB when the target azimuth changes from −90° to +45°. However, when the masker is at (45°, 1 m) [Fig. 3(b)], this same angular displacement of the 15-cm target (thick line) produces a change in spatial unmasking of roughly 20 dB (compare the leftmost point and the point producing the least spatial unmasking, where the target is at 45°).

Angular separation of target and masker can actually make performance worse when target distance differs from masker distance. Usually, separating target and masker in azimuth improves target detectability compared to when the target and masker are in the same direction, but not in every case. When the masker is at 0° (panels a and d in both Figs. 2 and 3) the least amount of spatial unmasking occurs (thresholds are highest) when the target is at 0° (the same direction as the masker); when the masker is at 45° (panels b and e in Figs. 2 and 3) the least unmasking arises when the target is in the 45° masker direction. However, when the masker is at 90° (panels c and f in Figs. 2 and 3), angular separation of target and masker does not always increase the amount of unmasking. Specifically, for a masker at (90°, 1 m) [Figs. 2(c) and 3(c)] [illegible] produces less unmasking of

whether the models capture the acoustic effects that are important for predicting the amount of spatial unmasking as a function of nearby target and masker locations. As noted in Sec. II, the current measurements do not try to compensate for the radiation characteristics of the loudspeaker used; as such, any consistent discrepancies between predictions from a spherical-head model and measured results (from KEMAR and the human subjects) may reflect influences of the radiation characteristics of the loudspeaker used (which is not a point source) or other differences between the assumptions of the spherical-head model and properties of the physical sources and heads measured.

A. Methods

KEMAR HRTFs were measured using a procedure identical to that used for the human listeners (see description in Sec. II). HRTF predictions for a spherical head model (Brungart and Rabinowitz, 1999; Shinn-Cunningham *et al.*, 2000) were computed using a head with radius of 9-cm and diametrically opposed ears. These results are compared to the HRTFs measured for the four subjects who participated in the spatial unmasking experiment.

For all of the HRTFs, the magnitude spectra, ILD, and ITD were determined for the equivalent rectangular band (ERB) centered at a given frequency. Magnitude spectra were calculated as the rms energy in the HRTF falling within each ERB filter (100-Hz width centered at 500 Hz and 136-Hz width centered at 1000 Hz)compaap TD (!)0 Tin

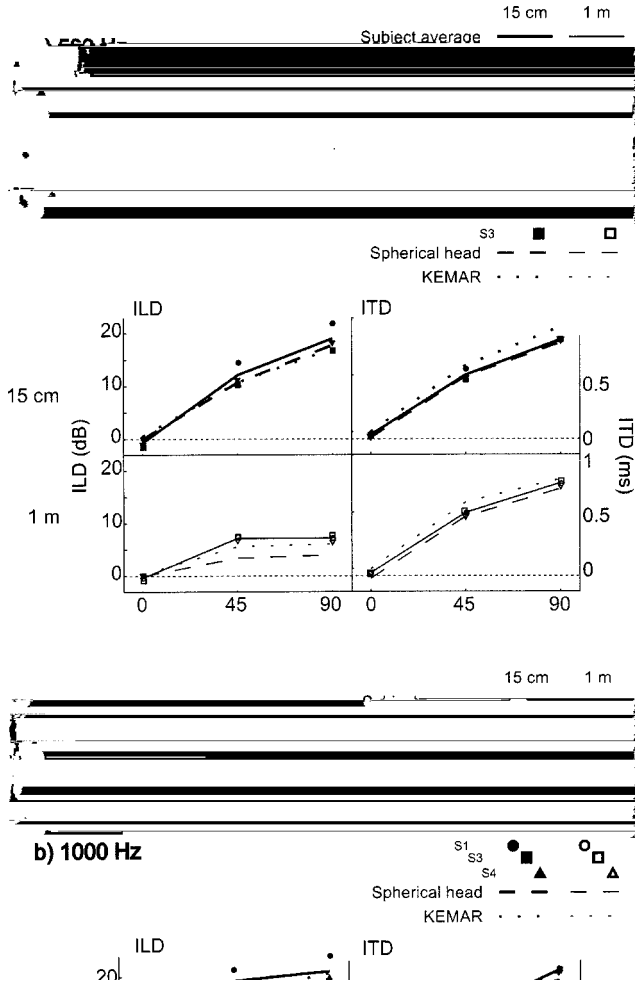

FIG. 5. ILDs and ITDs in HRTFs for individual subjects, KEMAR manikin, and the spherical head model. (a) 500 Hz. (b) 1000 Hz.

reaching the ears to vary monotonically with lateral angle of the source, human HRTF measurements show that this is not strictly true. In particular, the 1000-Hz human measurements [symbols and solid lines in Fig. 4(b)] show that less energy reaches the contralateral ear when a source is at 45° than when it is at 90° for both source distances (thick and thin lines are nonmonotonic with azimuth) Similarly, at 500 Hz [Fig. 4(a)] the gain to the contralateral ear is comparable for 45° and 90° sources rather than decreasing for the 90° source (thick and thin lines). This nonmonotonicity [which may in part be a consequence of the acoustic ''bright spot;'' e.g., see Brungart and Rabinowitz (1999)] is underestimated in both the spherical-head model (dashed lines) and KEMAR (dotted lines) HRTFs, especially at 1000 Hz [compare lines to human subject results for sources at 45°, especially in Fig. 4(b)].

2. Interaural differences

Figure 5 shows the ILDs and ITDs in the measured HRTFs at 500 and 1000 Hz [Figs. 5(a) and (b), respectively] for the spatial positions used in the study. As in Fig. 4, results for individual subjects (symbols), the across-human-subject average (full lines), KEMAR (dotted lines), and a spherical head model (dashed lines) are shown as a function of target azimuth. Results for near sources are shown in the top of each subplot with heavy lines and filled symbols. Thin lines and open symbols plot results for far sources (bottom row of each half of the figure). The left column shows ILD results and the right column shows ITD results.

ILDs were calculated directly from the measurements plotted in Fig. 4. As a result, there are large intersubject differences in the ILDs (left panels in Fig. 5) that are directly related to the intersubject differences in the monaural spectral gains. For instance, subject S1 has much larger ILDs at both 500 and 1000 Hz for the 15-cm source [filled circles in the left columns of Figs. 5(a) and (b)] than any of the other subjects (other filled symbols).

As expected, for both frequencies [Figs. 5(a) and (b)] ILDs are much larger for sources at 15-cm (thick lines in top left panels) compared to 1-m (thin lines in the bottom left panels) with ILDs at 500 and 1000 Hz approaching 20 dB for the nearby sources at 90° (rightmost point in the top left panels). The spherical-head (dashed lines) and KEMAR (dotted lines) results tend to underestimate ILDs for lateral sources, although for the 500-Hz, 15-cm sources [Fig. 5(a), top left panel], both spherical-head and KEMAR results are within the range of human observations. Discrepancies between human and model results are most pronounced for a 1000-Hz source at a distance of 1-m [Fig. 5(b), bottom left panel] and are greater for the spherical-head predictions (dashed lines) than KEMAR measurements (dotted lines).

ITDs [the right panels in Figs. 5(a) and (b)] vary primarily with source angle and change only slightly with distance and frequency. For most of the measured locations, both spherical-head and KEMAR results are in close agreement with human measurements.

C. Discussion

Both spherical-head and KEMAR HRTFs provide reasonable approximations to how acoustic parameters in human HRTFs vary with source location. In general, both KEMAR and the spherical head measurements fall within the range spanned by the individual subject measurements. However, both spherical-head predictions and KEMAR measurements slightly overestimate the gain at the contralateral ear when a source is at 45° (especially at 1000 Hz) and tend to modestly underestimate the ILD for sources off midline, particularly at the 1-m distance. These small differences cannot be attributed to loudspeaker characteristics, given that (1) the discrepancies are similar for both KEMAR measurements (using the same loudspeaker) and spherical-head predictions (assuming a perfect point source) and (2) the differences are, if anything, larger for the more distant, 1-m source (where the loudspeaker directivity is less influential) than the nearby source. Thus, we conclude that generic HRTF models capture the important features of the HRTFs measured in human subjects and that the effects of the source transmission characteristics do not strongly influence the signals reaching the ears even for nearby sources, at least for the frequencies considered in the current study.

Intersubject differences in the HRTFs are large, especially for nearby sources. Of the four subjects, one subject showed consistently larger spectral gains and consistently

larger ILDs than the other subjects when the source was at 15-cm. While it is possible that some of the intersubject differences arise from inaccuracies in HRTF measurement (e.g., from hand-positioning the loudspeaker), the fact that one subject has consistently larger gains and ILDs for all nearby source locations suggests that real anatomical differences rather than measurement errors are responsible for the observed effects. It is also interesting to note that the observed intersubject differences are much smaller for the 1-m source, suggesting that intersubject differences in HRTFs are especially important when considering sources very close to the listener.

IV. BETTER-EAR AND BINAURAL CONTRIBUTIONS TO SPATIAL UNMASKING

A. Analysis

For each subject, estimates of the better-ear and binaural contributions to spatial unmasking were derived from the acoustic parameters of the HRTFs and the behavioral thresholds.

The better-ear contribution to spatial unmasking was estimated by calculating the TMR in the ERB filter centered on the target frequency at the better ear for each spatial configuration when target and masker emit the same level (and thus would yield a TMR of zero when at the same location). The resulting TMR predicts the amount by which target thresholds decrease or increase simply because of acoustic effects at the better ear (i.e., if the calculated TMR is +2 dB, it implies that at detection threshold, the intensity of the target at the better ear was 2 dB more for the given spatial configuration than if the target and masker were at the same spatial location; thus, the better-ear contribution for such a configuration is +2 dB). The subject-specific binaural contribution to spatial unmasking was estimated by subtracting the estimated better-ear contribution to spatial unmasking (derived from individually-measured HRTFs) from the individual behavioral estimates of spatial unmasking.

B. Results

angle (i.e., the modulation of binaural gain with target azimuth) is smaller when the masker is laterally displaced (right panels) than when the masker is at 0° (left panels), particularly for the 1000-Hz target (Fig. 7). For instance, looking at the bottom left panel of Fig. 7(a), when the masker is at (0°, 15 cm) the binaural contributions to spatial unmasking for the 1000-Hz target for subject S1 range from 0 to 8 dB depending on the target azimuth. However, when the masker is at (90°, 15 cm) [bottom right panel in Fig. 7(a)], binaural unmasking is roughly constant, independent of target angle (roughly 0–2 dB).

conditions where the model fails to account for behavioral data, parallel lines plot a range of ± 1 dB around the actual model predictions. Predictions for the nearby target are shown as dashed black lines; predictions for the far target are shown as solid gray lines.

Model predictions of binaural unmasking are non-negative for all spatial configurations. Predictions are exactly zero whenever the target and masker are at the same spatial location and positive whenever the target and masker have differences in either their IPDs or ILDs at the target frequency. Thus, in theory, predictions of binaural unmasking are positive whenever the target and masker are at different distances but in the same direction off the median plane because of differences in ILDs in target and masker. However, in practice, predictions are near zero for all configurations when the target and masker are in the same direction for subjects S2, S3, and S4 [Figs. 6(b), 6(c), 7(b), and 7(c)]. Predictions for subject S1 [who has the largest ILDs for 15-cm sources and the largest BMLDs at both frequencies; Figs. 6(a) and 7(a)] are greater than zero for both target frequencies when the target and masker are at different distances but the same (off-median-plane) direction. For instance, in the top center and top right panels of Figs. 6(a) and 7(a) [masker at (45°, 1 m) and (90°, 1 m)], the black dotted lines (predictions for the target at 15 cm) are above zero for all target azimuths, including the target at 90°; in the bottom center and right panels of Figs. 6(a) and 7(a) [masker at (45°, 15 cm) and (90°, 15 cm)], the gray solid lines (predictions for the target at 1 m) are positive for all azimuths.

essentially diotic). For most of the configurations with the masker at 0°, model predictions agree well with observed results. In contrast, larger discrepancies between the modeled and measured results arise when the masker is at 45° and 90° (conditions in which there are significant ILDs in the masker).

While there are some conditions in which the model predictions consistently over- or underestimate binaural unmasking [e.g., results for subject S1 at 1000 Hz in Fig. 7(a) or for subject S3 at 1000 Hz in Fig. 7(b)], there are other conditions for which changing the single subject-specific ''binaural sensitivity'' of the model cannot account for discrepancies between the model predictions and the measurements [e.g., results for subject S2 at 500 Hz in Fig. 6(b) or for subject S4 at 1000 Hz in Fig. 7(c)].

The current results suggest that subjects differ not only in their overall sensitivity to binaural differences, but also in the dependence of binaural sensitivity on the interaural parameters in masker and/or target. In particular, binaural sensitivity appears to depend on the interaural level difference in the masker differently for different subjects. As a result, individualized model prediction errors are generally larger when there are large ILDs in the masker than when the masker has near-zero ILD. While the Colburn model has been tested (and shown to predict results relatively well) in many studies in which target and masker vary in their interaural phase parameters, there are few studies that manipulate the target and masker ILD. These results suggest the need for additional behavioral and theoretical studies of the effects of ILD in binaural detection tasks.

Even though there are specific conditions for which predictions fail to account for the results for a particular subject, the model captures many of the general patterns in results, including the tendency for binaural unmasking to decrease as the ILD in the masker increases and how the amount of binaural unmasking depends on the angular separation of target and masker and the frequency of the target.

VI. GENERAL DISCUSSION

The current study is unique in measuring how tone detection thresholds are affected by target and masker location when sources are very close to the listener. Results show that for sources very close to the listener, small changes in source location can lead to large changes in detection threshold. These large changes arise from changes in both the TMR (affecting the better-ear contribution to spatial unmasking) and ILDs (affecting the binaural contribution to spatial unmasking).

The current results demonstrate how the relative importance of better-ear and binaural contributions to spatial unmasking change with target and masker location, including source distance (in contrast to previous studies that considered only angular separation of relatively distant sources). The relative importance of better-ear contributions to spatial unmasking increases as masker distance decreases, probably because of increases in the ILD in the masker, which reduce the amount of binaural unmasking. The better-ear contribution also increases as target distance decreases, primarily because the TMR changes more rapidly with target angle when

vary not only in overall magnitude but as a function of the interaural differences in the masker.

While predictions from the Colburn model (taking into account differences in the stimuli presented to the individual subjects as well as individual differences in binaural sensitivity) cannot account for some small but significant intersubject differences in spatial unmasking, rough predictions of the amount of spatial unmasking capture most of the observed changes in detection threshold with spatial configuration. For instance, generic acoustic models of HRTFs (e.g., KEMAR measurements or spherical-head model predictions) combined with predictions of binaural unmasking using ''average'' model parameters should produce predictions that fall within the range of behavior observed across a population of subjects.

VII. CONCLUSIONS

(1) Acoustic cues (particularly TMR and ILD) vary dramatically with source distance and direction for nearby sources. Therefore, when source distance varies, the effect of source location on both the better-ear and binaural contributions to spatial unmasking is complex.
(2) For nearby sources, the better-ear contribution to pure-tone spatial unmasking can be very large (as much as 25 dB) compared to conditions where sources are relatively far from the listener.
(3) The binaural contribution to spatial unmasking decreases with increasing masker ILD. As a result, the binaural contribution to spatial unmasking is smaller for lateral sources very near the head than for more distant sources at the same lateral angle relative to the listener.
(4) Intersubject differences in spatial unmasking are larger for nearby sources than for far sources, in part because there are larger acoustic differences in HRTFs for nearby sources compared to more distant sources. However, there also are subject-specific differences both in binaural sensitivity and on how ILDs influence binaural sensitivity.
(5) Predictions based on Colburn's analysis (1977b) show the correct general trends in binaural detection for both near and far sources, but cannot account for small, but consistent, subject-specific differences in performance, particularly when large ILDs are present in the masker.

ACKNOWLEDGMENTS

This work was supported in part by AFOSR Grant No. F49620-98-1-0108 and the Alfred P. Sloan Foundation. Portions of this work were presented at the Spring 2000 meeting of the Acoustical Society of America. H. Steven Colburn provided valuable input throughout this work, including help in putting the results in appropriate context. Les Bernstein, Adelbert Bronkhorst, and an anonymous reviewer provided valuable criticism and comments on a previous draft of this paper.

$$p(\tau,f_0)=\begin{cases} C(e^{-2\pi k_l|0.2|}-e^{-2\pi k_h|0.2|})/0.2, & |\tau|<0.2 \text{ ms},\\ C(e^{-2\pi k_l|0.2|}-e^{-2\pi k_h|0.2|})/|\tau|, & |\tau|\geqslant 0.2 \text{ ms},\end{cases}$$

$$k_h=3\times 10^6,$$

$$k_l\begin{cases} 0.1(f_0 10^{-3})^{1.1}, & f_0\leqslant 1200 \text{ Hz},\\ 0.1(1200\times 10^{-3})^{1.1}, & f_0>1200 \text{ Hz},\end{cases} \tag{A3}$$

$$C=\begin{cases} 0.1534, & f_0=500 \text{ Hz},\\ 0.2000, & f_0=1000 \text{ Hz}.\end{cases}$$

$G(f)$ is given by

$$G(f_0)=\begin{cases} \sqrt{10}, & f_0\leqslant 800 \text{ Hz},\\ \sqrt{10}\dfrac{800}{f_0}, & f_0>800 \text{ Hz}.\end{cases} \tag{A4}$$

$\gamma(\tau)$ is given by

$$\gamma(\tau)=\begin{cases} 2.359\times 10^{-4}+1.5207\times 10^{8}\tau^4-1.764\times 10^{4}\tau^2 & \\ \qquad +0.993, & |\tau|\leqslant 0.006,\\ -97.3236|\tau|+1.139, & |\tau|>0.006,\end{cases} \tag{A5}$$

where τ is in milliseconds.

Finally, function $R(\alpha,K)$ characterizes the decrease in the number of activated auditory nerve fibers in the ear receiving the less intense signal as a function of masker ILD. The current implementation uses a modified version of Eq. (35) from Colburn (1977b):

$$R(\alpha_n)=\begin{cases} \left(\dfrac{10\log_{10}\alpha_n^{-2}K}{40}\right)^2, & \alpha_n^{-2}K\leqslant 10^4,\\ 1, & \alpha_n^{-2}K>10^4,\end{cases} \tag{A6}$$

where K is the ratio of the spectrum level at the more intense ear to the detection threshold in quiet. This implementation of the model assumes that the auditory nerve fibers at each target frequency have thresholds uniformly distributed (on a dB scale) over a 40-dB range above the absolute detection threshold for that frequency.

[1]System identification using a MLS depends on circular convolution techniques. Theoretically, the approach requires the MLS to be concatenated with itself and presented an infinite number of times to ensure that the system is in its steady-state response prior to measuring the response (see Vanderkooy, 1994). The resulting estimated system response is a time-aliased version of the true system response. In the current measures, the MLS was presented twice and the response to the second repetition was recorded. Given the length of the MLS used, the room characteristics of and ambient noise in the environment in which we were measuring, and the noise in our measurement system, the steady-state response can be approximated with only two repetitions of the MLS and no significant time aliasing is present in our measurements.

[2]Note that this analysis assumes that detection performance depends only on the target-to-masker ratio or TMR and is independent of the overall masker level, an assumption that is not valid if the masker is near absolute threshold or at very high presentation levels. For instance, imagine two masker locations so distant from the listener that the masker is inaudible. These masker locations would produce identical signal detection thresholds if the experiment were performed with the distal stimulus intensity fixed; however, our technique might adjust the masker by different amounts for these two masker locations in order to achieve a fixed proximal stimulus level at the ear of the listener, producing two different estimates of spatial unmasking. While holding the distal masker intensity fixed may seem more natural and intuitive than holding the proximal stimulus level constant, the overall presentation level of the masker would span an extraordinarily large range in the current experiments because the masker distance varied between 15 cm and 1 m in addition to varying in direction. Therefore, we elected to fix the proximal masker intensity.

Bronkhorst, A. W., and Plomp, R. (**1988**). ''The effect of head-induced interaural time and level differences on speech intelligibility in noise,'' J. Acoust. Soc. Am. **83**, 1508–1516.

Brungart, D. S., and Rabinowitz, W. M. (**1999**). ''Auditory localization of nearby sources. I. Head-related transfer functions,'' J. Acoust. Soc. Am. **106**, 1465–1479.

Colburn, H. S. (**1977a**). ''Theory of binaural interaction based on auditory-nerve data. II: Detection of tones in noise,'' J. Acoust. Soc. Am. **61**, 525–533.

Colburn, H. S. (**1977b**). ''Theory of binaural interaction based on auditory-nerve data. II: Detection of tones in noise. Supplementary material,'' J. Acoust. Soc. Am. AIP document no. PAPS JASMA-61-525-98.

Doll, T. J., and Hanna, T. E. (**1995**). ''Spatial and spectral release from masking in three-dimensional auditory displays,'' Hum. Factors **37**, 341–355.

Duda, R. O., and Martens, W. L. (**1998**). ''Range dependence of the response of a spherical head model,'' J. Acoust. Soc. Am. **104**, 3048–3058.

Durlach, N. I., and Colburn, H. S. (**1978**). ''Binaural phenomena,'' in *Handbook of Perception*, edited by E. C. Carterette and M. P. Friedman (Academic, New York), pp. 365–466.