SPATIAL UNMASKING OF NEARBY PURE-TONE SOURCES IN A SIMULATED ANECHOIC ENVIRONMENT

3aPP1

Norbert Kopčo[1,2] and Barbara Shinn-Cunningham[1,2,3]

[1]Hearing Research Center, [2]Departments of Cognitive and Neural Systems and [3]Biomedical Engineering, Boston University

1. BACKGROUND

Spatial separation of target from noise improves target detectability and intelligibility.

Previous studies of spatial unmasking have examined:
- detection of tones and complex sounds
- speech intelligibility

No studies have examined these effects as a function of source distance

Past studies results can be explained by a combination of monaural, binaural, and informational (un)masking

Data from past headphone experiments:
- is plentiful (especially for 500-Hz tones)
- predict spatial unmasking effects
- is well predicted from models of auditory processing

Localization cues and localization behavior of nearby sources has been examined recently by Brungart and Rabinowitz (1999) and Shinn-Cunningham, Santarelli, and Kopčo (2000)
- unique, extra large ILDs accross all frequencies
- small positional changes cause large monaural and binaural cue changes
- ILDs vary with both direction and distance

2. MOTIVATION

Current state:
- No studies of spatial unmasking for nearby sources
- No quantitative modeling of spatial unmasking
- Available headphone data
  - cannot be used o predict unmasking at all spatial combinations of tone and masker
  - no complete set of data for single subjects available
- Collected data can be used for future studies of unmasking of complex sounds/speech and reverberation

GOALS

- Measure spatial unmasking of pure tones for nearby sources
- Compare relative importance of monaural and binaural processing for unmasking
- Study significance of the distance dimension for spatial unmasking (only monaural effects as in far field, or also binaural effects?)
- Compare behavioral data with predictions of available models

3. METHODS

SUBJECTS
- 1F, 3M students
- normal hearing

SIMULATION AND SPATIAL REGION
- simulated near-field anechoic auditory space
- all sounds in frontal horizontal plane (FIG 1)
- tone positions: distances 15 cm, 1 m
  azimuths -90, -45, 0, 45, and 90°
- masker positions: distances 15 cm, 1 m
  azimuths 0, 45, and 90°
- total of 60 target/masker configurations

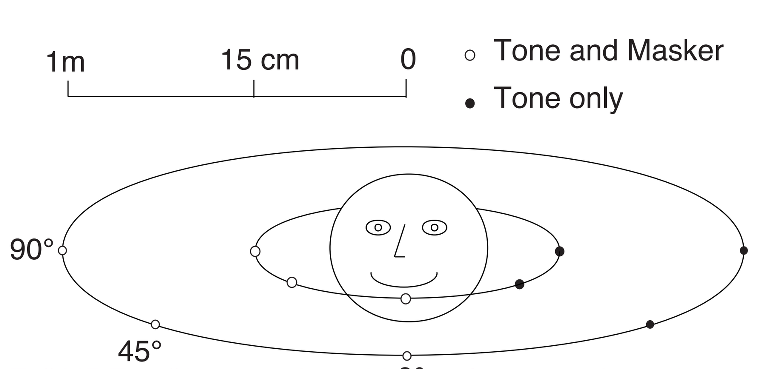


FIGURE 1 Simulated positions of tone and masker

- HRTFs used in simulation:
  - individually measured in an echoic room
  - MLS technique
  - cos²-windowed to exclude reverberation
  - assume symmetric head

STIMULI

Tones:
- f_0 either 500 or 1000 Hz
- 165-ms tone burst gated by 30-ms $\cos^2$ ramps
- temporally centered within the noise bursts

Maskers:
- 250-ms white noise bursts
- lowpass-filtered at 5000 Hz
- equalized so the better-ear rms energy in f_0-centered ERB filter fixed at 64 dB SPL

EQUIPMENT
- stimuli generated using TDT PD1, PA4, SM3, HB6
- played back through Etymotic Research ER-1 insert ear-phones
- response and feed-back provided via handheld terminal (QTERM)

THRESHOLD DETERMINATION
- 3-down-1-up adaptive procedure (tracking 79.4% correct)
- three-interval, two-alternative forced choice task
- each threshold measured 3 times in separate runs
- additional runs until std error of mean less than 1dB

OVERALL PROCEDURE
For both f_0 500 and 1000 Hz:
- Each configuration tested at least 3x (180 runs)
- Runs blocked so that each block contained all configs
- 3 Blocks organized into 6 sessions
- Each 1-hour long session contained 10 runs
  - masker location fixed in session
  - target locations in random order in session
- order of sessions random within blocks (1st practice)

4. ANALYSIS / MODELS

Spherical head model (Brungart and Rabinowitz, 1999; Shinn-Cunningham, Santarelli, and Kopčo 2000)
- compared with individual HRTFs and with HRTFs of the KEMAR manikin
- Colburn (1977) model
  - models processing in auditory nerve and brainstem processing of binaural information
  - implemented in simplified form defined by equations 10, 14, 17 and 19 in Colburn (1977)
  - for 1000-Hz data, the model extended as in Stern and Shear (1996) to incorporate dependence of binaural coincidence counter distribution on signal frequency.

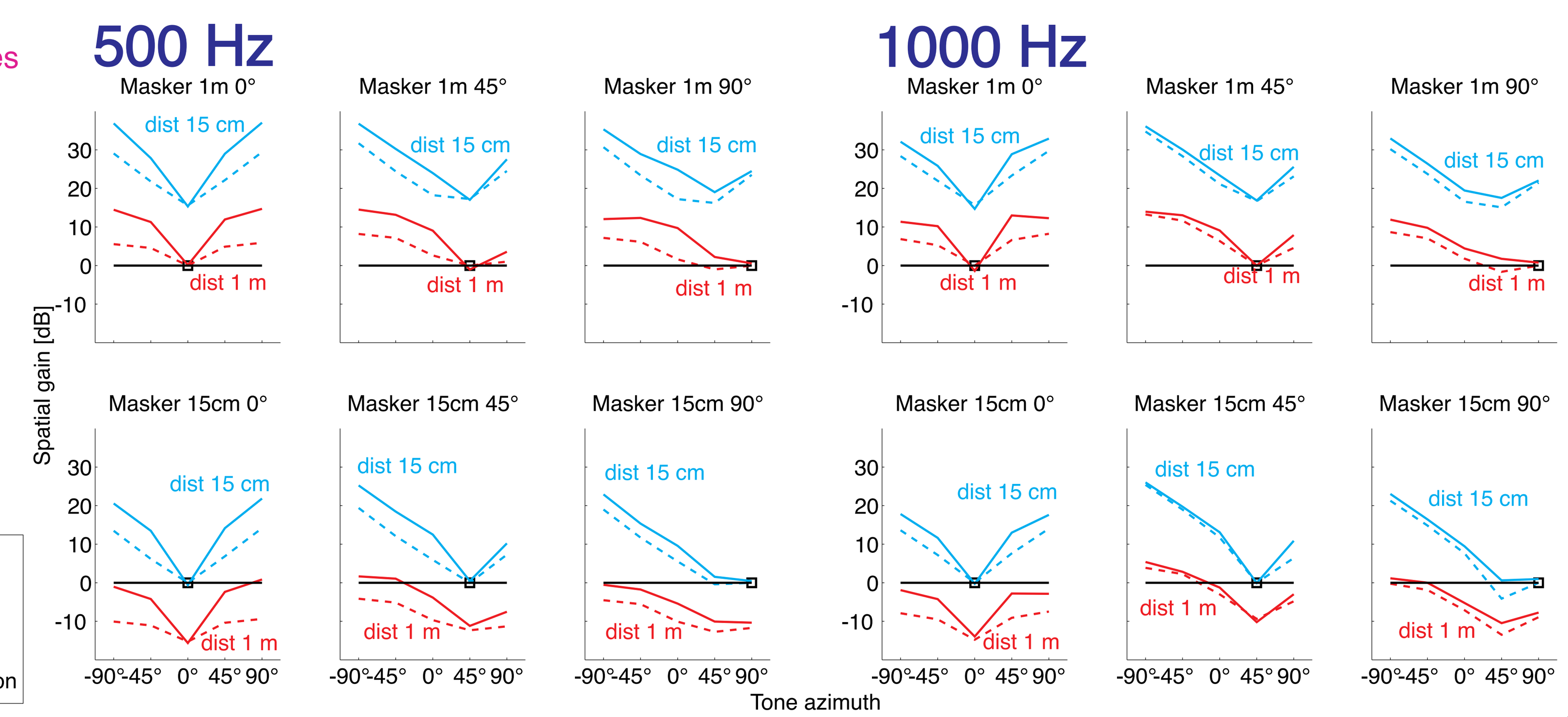
5. SPATIAL UNMASKING

Figure 2 shows average (cross-subject) energy that target must emit to be detectable, relative to target and masker co-located. Overall spatial gain given by solid lines. Better-ear contribution given by dashed lines.

Results:
- Change in spatial configuration leads to huge changes in tone detectability for nearby sources (loss of 15 dB to gain of 40 dB)
- Overall, monaural effects (change in tone level at better ear) are stronger than binaural effects.
- Monaural effects dominate at 1000 Hz for maskers off the midline.
- Binaural processing is significant for masker at the midline, especially for 500 Hz tones.
- At 500 Hz, binaural processing can compensate for loss in detectability due to increased distance of the tone (bottom left graph)

FIGURE 2 Spatial unmasking of pure tones as a function of tone frequency and spatial position of tone and masker

Left: 500 Hz tone
Right: 1000 Hz tone
Upper row: masker fixed at 1 m
Lower row: masker at 15 cm

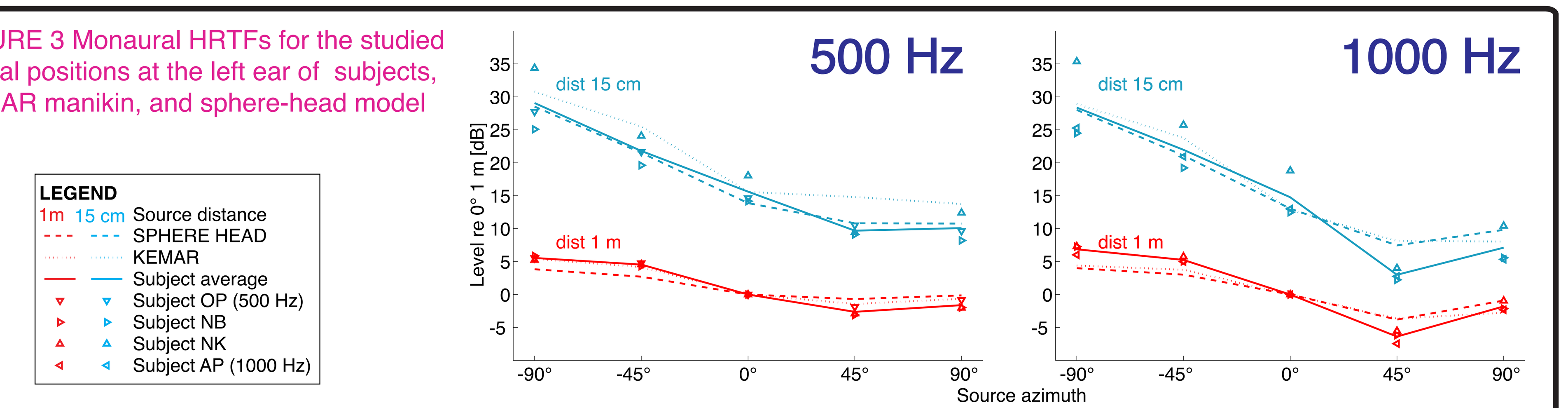


6. MONAURAL FACTORS

Monaural performance is determined by the monaural HRTF from a given spatial position to the ear. Figure 3 compares individual HRTFs for the subjects used in the study, KEMAR, and the spherical head model.

- Spherical head predictions closer to subject average than KEMAR HRTFs, especially at 500 Hz
- Large inter-subject variance at 15 cm (measurements sensitive to exact source placement)

FIGURE 3 Monaural HRTFs for the studied spatial positions at the left ear of subjects, KEMAR manikin, and sphere-head model



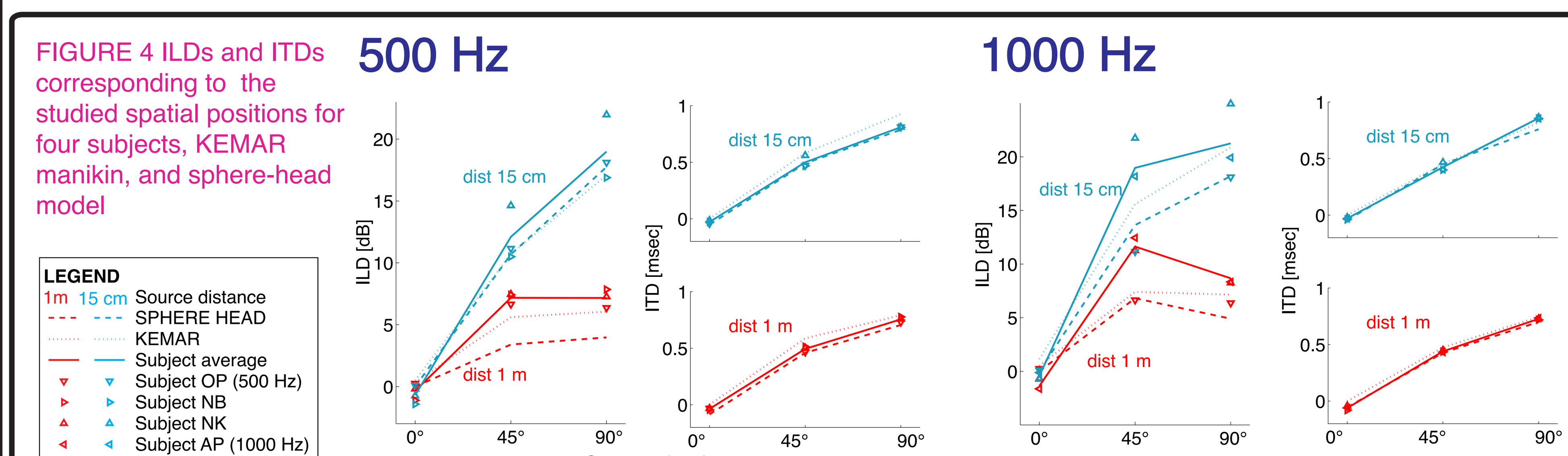
7. BINAURAL FACTORS

Binaural contribution to detection performance is determined by:
- ITD and ILD corresponding to the tone and masker position
- the subject's binaural sensitivity,
- overall noise level (constant in the present study).

Figure 4 shows the ITD and ILD observed at spatial positions of interest. Table 1 gives the measured binaural sensitivities.

Similar ITDs for 500 and 1000 Hz lead to dramatically different IPDs.

FIGURE 4 ILDs and ITDs corresponding to the studied spatial positions for four subjects, KEMAR manikin, and sphere-head model



Binaural sensitivity is defined as

$$\text{BMLD} = T(\text{NoSo}) - T(\text{NoS}\pi)$$

The average values in Table 1 were used in modeling.

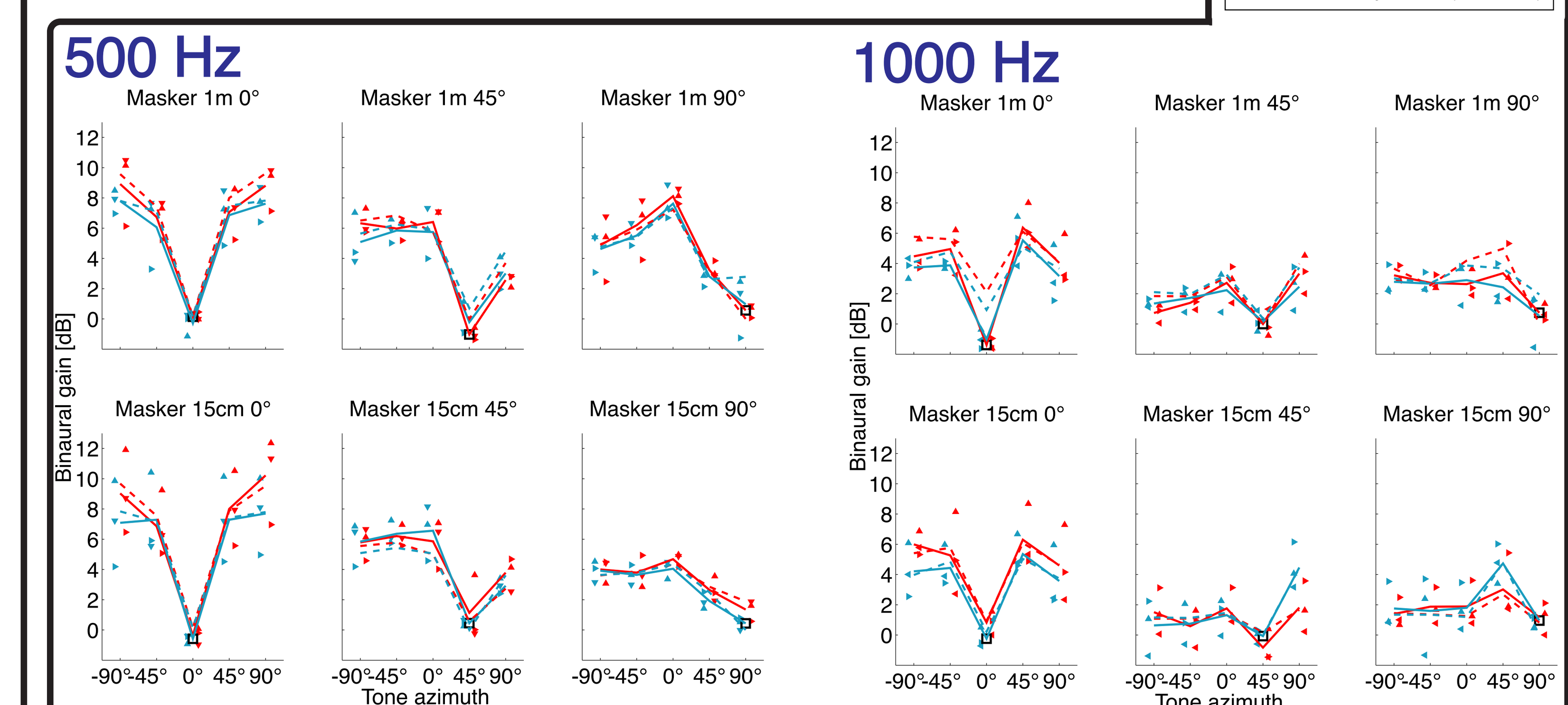
TABLE 1 Measured binaural sensitivity of the subjects

	Measured BMLD [dB]				
	OP	NB	NK	AP	Average
500 Hz	14.5	11	15.6		13.7
1000 Hz		7.5	13.1	8.7	9.8

8. BINAURAL UNMASKING

Binaural contribution to spatial unmasking at 500 and 1000 Hz:
- large amount of unmasking at both frequencies
- very good match between model prediction and behavioral data
- observable differences between model and data at 1000 Hz
- binaural unmasking due to change only in tone distance observed
- comparable amount of unmasking for close and distant masker positions
- large inter-subject differences

FIGURE 5 Predicted and measured binaural contribution to spatial unmasking



9. DISCUSSION OF MODELING

Biggest challenge: fitting dependence on noise ILD. The fit depends on two factors:
- distribution of binaural coincidence counters as a function of noise ITD and frequency
  - Colburn function $p(\tau, f_C)$
- number of stimulated coincidence counters as a function of noise ILD and frequency
  - Colburn function $q_B(\alpha, f_C)$

Above results obtained with:
- Colburn (1977) definition of $p(\tau, f_C)$ for 500-Hz data
- Stern and Shear's (1996) $p(\tau, f_C)$ for 1000-Hz data
- Frequency-independent definition of $q_B(\alpha, f_C)$ given by equation 35 from Colburn (1977)
- small mismatch observable at 1000 Hz

Results show need for unified definition of $p(\tau, f_C)$

10. CONCLUSIONS

MEASURED DATA
- Large effects of spatial unmasking for nearby sources (-15 to 40 dB)
- Monaural (better-ear) effects prevail (25 dB)
- Binaural processing important at low frequencies and for masker in midline (10 dB)
- Binaural processing influences distance unmasking for nearby sources
- Binaural unmasking comparable for near and far sources

MODELING
- Monaural effects, as well as ITD and ILD, can be modeled accurately using the spherical head model
- Binaural unmasking can be modeled using Colburn (1977) model
- Binaural modeling very sensitive to assumptions about coincidence counters distribution

11. REFERENCES

Brungart, DS & WM Rabinowitz (1999). "Auditory localization of nearby sources I: Head-related transfer functions," J Acoust Soc Am, 106, 1465-1479.

Colburn, HS (1977). "Theory of binaural interaction based on auditory-nerve data. II: Detection of tones in noise," J Acoust Soc Am., 64: 525-533 and supplementary material, AIP document no. PAPS and JASMA-61-525-98.

Shinn-Cunningham, BG, SG Santarelli, and N Kopčo (2000). "Tori of confusion: Binaural localization cues for sources within reach of a listener," J Acoust Soc Am., 107: 1627-1636.

Stern RM & GD Shear (1996). "Lateralization and detection of low-frequency binaural stimuli: Effects of distribution of internal delay," J Acoust Soc Am., 100: 2278-2288.

Supported by AFOSR grant F49620-98-1-0108