

Effects of cuing on perceived location of auditory sources

Norbert Kopčo and Barbara Shinn-Cunningham

Hearing Research Center, Department of Cognitive and Neural Systems
Boston University, 677 Beacon St., Boston, MA 02215
{kopco, shinn}@bu.edu

This study explores auditory processing mechanisms underlying the computation of sound source location. A previous study examined whether a cue sound that indicated which left-right spatial hemifield to attend could improve localization accuracy (Kopčo, Ler and Shinn-Cunningham, 2001) in the same way that it decreases response latency (Spence and Driver, 1994). Results show that, in an ordinary room, an informative preceding auditory cue does not improve localization accuracy; however, the presence of a preceding cue causes consistent localization bias of the target stimulus for cue-target delays as long as 300 ms. The current paper presents acoustic analysis that examines the extent to which localization bias can be explained by purely acoustic effects (including room reverberation) as opposed to neural processing effects (e.g., see Carlile, Hyams and Delaney, 2001).

Keywords: auditory computation, auditory perception, sound localization, psychoacoustics, auditory attention, head-related transfer functions, reverberation

1. Introduction

One of the goals of computational intelligence is to determine how biological systems solve difficult problems in order to apply these approaches to technical applications. The way in which biological systems compute sound source location from the acoustic signals reaching the ears can be instructive in this way, because biological systems do a remarkably good job of determining sound locations from what are often ambiguous, noisy acoustic spatial cues.

Figure 1 shows a diagram that outlines the neural pathway involved in the computation of the sound source position. As shown in Figure 1, the neural pathway for extracting source location is very hierarchical, involving many stages of processing even before the level of the cortex. Basic spatial cues are first extracted in the brainstem before being integrated in the midbrain. The resulting signals then reach the cortex, which governs high-level spatial behaviors. In addition to the feed-forward paths shown in Figure 1 (see also Shinn-Cunningham,

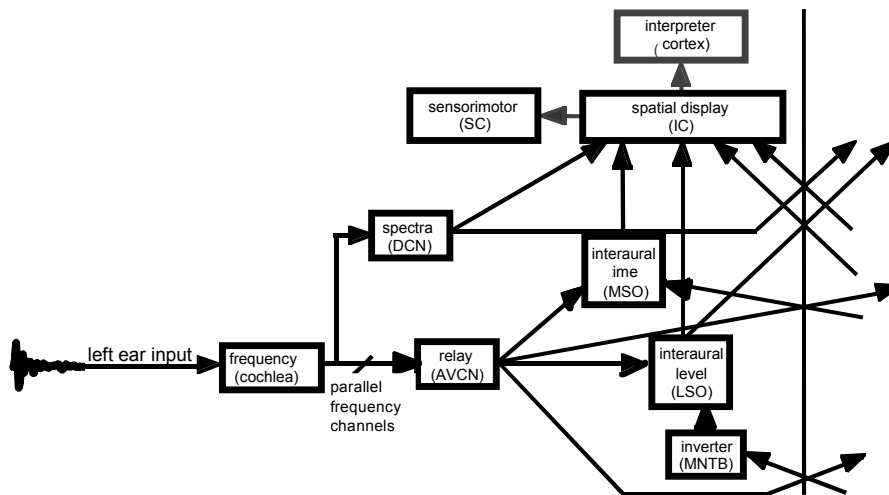


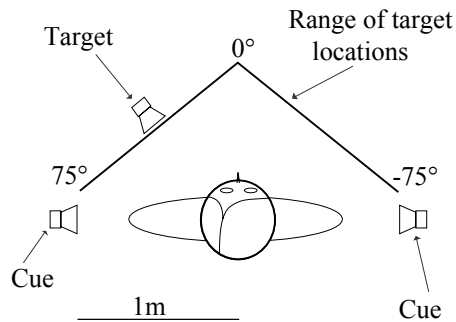
Figure 1. Simplified, block diagram of the forward neural pathway for computing source position (left half of symmetrical model; taken from Shinn-Cunningham, 2001). Abbreviations of physiological nuclei implicated in spatial auditory processing: DCN - dorsal cochlear nucleus; AVCN - anteroventral cochlear nucleus; MNTB - medial nucleus of the trapezoid body; LSO - lateral superior olive; MSO - medial superior olive; IC - inferior colliculus; SC - superior colliculus. Computations change over time (on many time scales) in order to maintain accurate spatial perception and include complex temporal dynamics, particularly at and above the level of the IC.

2001), there are complex temporal interactions and feedback in the spatial auditory pathway (which are just beginning to be characterized and are not included in Figure 1) that appear to aid the system in computing robust estimates of sound source location in the face of noise and uncertainty.

The goal of the present paper is to examine the extent to which simple acoustic interactions, rather than neural processing mechanisms (particularly temporal interactions), explain behavioral data obtained in a previous experiment studying cued auditory localization in a reverberant room (Kopčo et al., 2001).

Kopčo and colleagues investigated whether automatic and strategic auditory attention (Spence and Driver, 1994) influence human sound localization. The study tested whether the presence of an informative cue activates automatic and/or strategic auditory attention and improves sound localization accuracy (especially given that it reduces reaction times; Spence and Driver, 1994; and increases sensitivity in detection tasks; Sach, Hill and Bailey, 2000). Listeners indicated the azimuth angle of a target sound in the presence of a preceding cue stimulus that

Figure 2. Experimental setup from Kopčo et al. (2001). In every trial, a cue sound from either $+90^\circ$ or -90° azimuth preceded a target sound positioned between $+75^\circ$ and -75° . Four experimental conditions were explored, varying the percentage of trials in which the cue correctly indicated the target hemifield (0, 50, 75, or 100%). The experiment was conducted with the listener located in the center of a quiet room (5 m x 9 m, $T_{60} = 650\text{ms}$).



provided different amounts of information about the target location from block to block (see Figure 2). Results, which are summarized in Figure 3, show that the preceding cue generally degraded localization accuracy, independent of the amount of information provided by the cue, even when the stimulus onset asynchrony (SOA) between the cue and the target stimulus was 300 ms.

There are two possible explanations for the observed localization bias due to the preceding cue. First, in a reverberant room, acoustical interactions between the cue and the target stimuli could cause systematic distortions of the acoustic spatial cues due to the target stimulus (such as interaural temporal differences or ITDs between the left and right ears). Such physical effects represent cue-target interactions that are external to the listener (i.e., that arise before the processing shown in Figure 1). Alternatively, the preceding cue may influence the internal, neural processing of the subsequent target (Figure 1). Such neural interactions could come about through temporal dynamics in the midbrain, or at higher stages of processing such as the cortex. The present study analyzes whether simple acoustic interactions can explain the effects of a preceding cue on localization of a target in order to rule out simple, parsimonious explanations for the observed behavior.

2. Methods

Acoustical interactions of the cue and target are investigated by examining the total signals reaching the ears of the listener for spatial configurations like those used in the behavioral study. This analysis focuses on how the energy due to the preceding cue source distorts the ITD in the onset of the target sound, which is assumed to be the dominant acoustical cue for horizontal auditory localization (Middlebrooks and Green, 1991). In order to simulate the total stimuli received at the ears of the listeners, the signals that would reach the listener from cue and target stimuli at appropriate locations are simulated using reverberant head-related

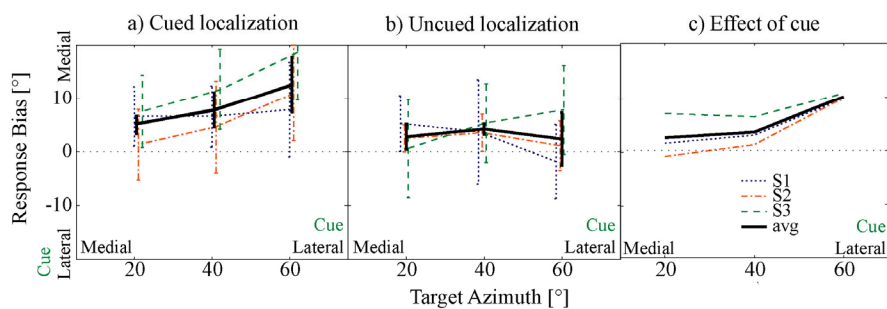


Figure 3. Effect of the cue on perceived target azimuth (from Kopčo et al., 2001). Each panel plots the mean and standard deviation in results for three individual subjects as well as an across-subject average. **a) Localization bias (mean signed localization error) of the target in the presence of a cue. b) Localization bias of the target in isolation. c) Effect of cue on localization bias (difference in bias for cued and uncued conditions).**

transfer functions (HRTFs) measured in the room used in the study. Such HRTFs, which can be used to generate the total signal reaching the ears of the listener for an arbitrary source at a particular location in the room (e.g., see Shinn-Cunningham, Desloge and Kopčo, 2001), embody the acoustical interaction of the sound with the room and with the listener's body and pinna.

The HRTFs used in this study were measured on a Knowles Electronic Manikin for Acoustic Research (KEMAR) positioned in the same room as in the behavioral study. A 20-ms-noise-burst cue from $+90^\circ$ was simulated 50 ms before a 2-ms-click target stimulus from $+75^\circ$. This particular condition was chosen for analysis because the acoustic effects of the cue should be greatest for this condition compared to the others used in the behavioral study. In particular, for this condition, the cue and the target stimulus are in closer spatial and temporal proximity than in other conditions and the cue energy is relatively large compared to the target stimulus. If acoustic interactions are too small to explain localization bias for this configuration, they will be too small to explain bias in other configurations, suggesting that the preceding cue influences how subsequent signals are processed in the spatial auditory pathway.

Left and right ear stimuli were bandpass filtered between 0.2 and 2 kHz and windowed through a 20-ms-wide Hanning window. The resulting 20-ms-long signal snippets were then cross-correlated to show how ITD varied during the course of the total cue-target stimulus. In order to get a better picture of how the ITD due to the target is affected by the presence of the cue, ITDs were examined in detail at the onset of the target. In this analysis, interaural phase differences (IPDs) between the left and right ear signals were computed as a function of frequency in a 40-ms-long time slice containing the end of the cue stimulus and the onset of the target stimulus. This computation was also performed for the

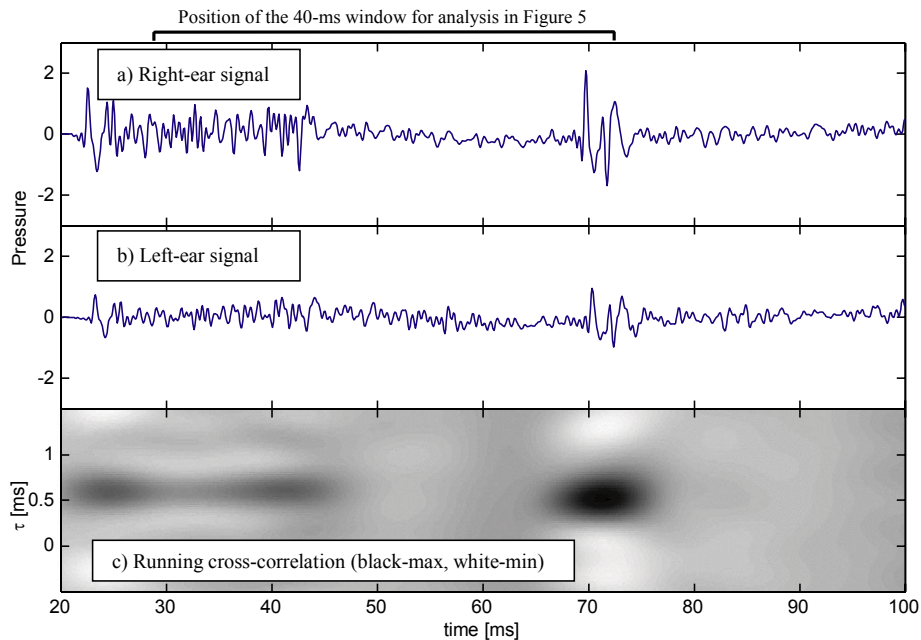


Figure 4. Signals received at the ears of the listener for a click-target at 75° and a noise-burst cue at 90° (50 ms later). a) The acoustic pressure signal received at the right ear. b) The acoustic pressure signal received at the left ear. c) The running, broadband cross-correlation of the left and right ear signals.

target signal in isolation. In both cases, the raw IPDs were calculated by dividing the left and right ear phase spectra; these IPDs were then converted into the corresponding ITD by dividing by frequency.

3. Results

Figure 4 shows both the left and right ear pressure waveforms at the ears (panels 4a and 4b, respectively) and the running cross-correlation of the left- and right-ear signals (panel 4c). In Figure 4c, the cross-correlation energy is plotted in grayscale; the black elongated region on the left corresponds to ITD in the cue and the second energy peak at approximately 70 ms corresponds primarily to the ITD in the target stimulus. These two cross-correlation peaks are clearly separated, despite the fact that the reverberant energy due to the cue is still relatively strong at the onset of the target stimulus (e.g., see panels 4a and 4b). The peak in the cross-correlation function corresponding to the target stimulus occurs at a slightly smaller ITD than that of the cue (lower value of τ along the vertical dimension),

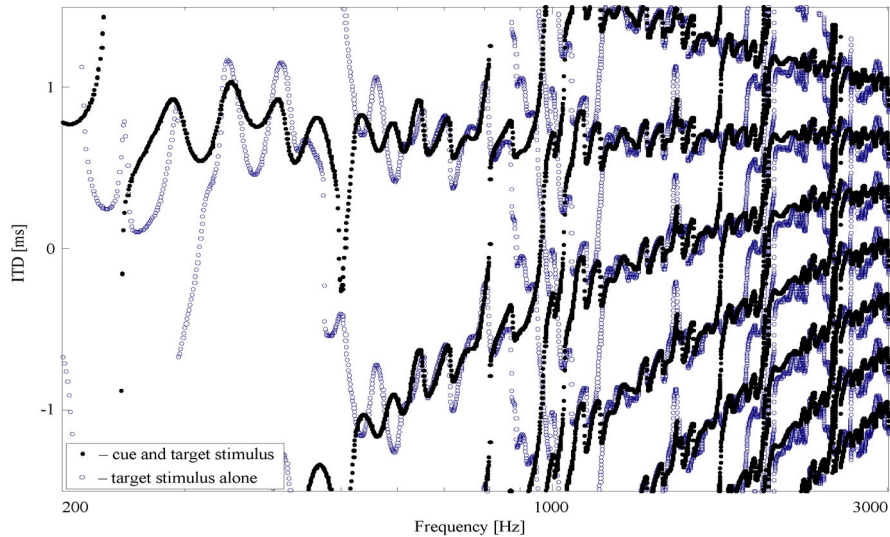


Figure 5. Effect of the cue on the ITD between left and right ear signals as a function of frequency in a 40-ms window centered on the onset of the target stimulus.

corresponding to the smaller ITD associated with target stimulus at 75° compared to the cue at 90° . In addition, the cross-correlation peak corresponding to the target stimulus is broader in the vertical dimension, due to the relatively short duration of the target stimulus compared to the width of the window used in the cross-correlation.

Figure 5 compares the ITD at the onset of the target in the cued and uncued conditions, where the interaction between the cue and the target stimulus should be greatest. The main effect of the cue shown in Figure 5 is an increase frequency-to-frequency variability in the ITD. However, this variation is also large for the uncued stimulus. More importantly, the ITD computed for the cued condition is not systematically biased towards 0, a result which could have explained the observed localization bias.

The results of the analysis shown in Figures 4 and 5 suggest that simple acoustic interactions of the cue and target stimuli cannot explain changes in the perceived target location due to the cue.

4. Conclusions and discussion

Given that 1) the information value of the cue has no effect on behavior, and 2) the localization bias due to the cue increases with the spatial proximity of the cue

and target, attentional modulation cannot explain the observed effects of the cue on localization of the target. The current results also show that a simple bottom-up acoustic analysis cannot explain the behavioral data. This suggests that the observed effects (Kopčo et al., 2001) are due to a neural interaction between the representations of the cue stimulus and the subsequent target stimulus. Such effects may be related to short-term adaptation effects (e.g., see Kashino, 1998; Kashino and Nishida, 1998; Carlile et al., 2001). However, more central explanations may also play a role (e.g., see Sach, Hill and Bailey, 2000).

More thorough analysis of the acoustic interaction is needed to determine where in the auditory processing pathway interactions of cue and target arise. Specifically, analysis that considers both acoustical interactions and temporal interactions at low levels of processing (e.g., in the brainstem) may be capable of explaining these results. Such analysis is currently being undertaken to begin to pinpoint where in the human auditory pathway the response to a preceding cue stimulus influences the perception of a later-arriving target source.

5. References

- Carlile S, Hyams S, Delaney S (2001) Systematic distortions of auditory space perception following prolonged exposure to broadband noise. *Journal of the Acoustical Society of America* 110:416-424.
- Kashino M (1998) Adaptation in sound localization revealed by auditory after-effects. In: 11th International Symposium on Hearing.
- Kashino M, Nishida Sy (1998) Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect. *Journal of the Acoustical Society of America* 103:3597-3604.
- Kopčo N, Ler A, Shinn-Cunningham BG (2001) Effect of auditory cuing on azimuthal localization accuracy. *Journal of the Acoustical Society of America* 109:2377.
- Middlebrooks JC, Green DM (1991) Sound localization by human listeners. *Annual Review of Psychology* 42:135-159.
- Sach AJ, Hill NI, Bailey PJ (2000) Auditory spatial attention using interaural time differences. *Journal of Experimental Psychology: Human Perception and Performance* 26:717-729.
- Shinn-Cunningham BG (2001) Models of plasticity in spatial auditory processing. *Audiology and Neuro-otology* 6:187-191.
- Shinn-Cunningham BG, Desloge JG, Kopčo N (2001) Empirical and modeled acoustic transfer functions in a simple room: Effects of distance and direction. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. New Pfalz, New York.
- Spence CJ, Driver J (1994) Covert spatial orienting in audition: Exogenous and endogenous mechanisms. *Journal of Experimental Psychology: Human Perception and Performance* 20:555-574.