DRR & ILD CUE WEIGHTING IN AUDITORY DISTANCE PERCEPTION

Keerthi Doreswamy [1,2]; Norbert Kopčo [1,2]
[1]Institute of Computer Science, P. J. Safarik University, Kosice, 04001, Slovakia
Keerthi.doreswamy@upjs.sk
[2] Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Harvard Medical School/Massachusetts General Hospital, Charlestown, MA, 02129, USA
norbert.kopco@upjs.sk

Abstract

Perceiving the sound source distance is important in many everyday activities. The estimates of auditory distance are typically dominated by the overall received stimulus intensity. However, distance processing can also be guided by intensity-independent cues. Specifically, the interaural level differences (ILDs) provide distance information for lateral stimuli and, in reverberant space, the direct-to-reverberant energy ratio (DRR) provides distance information for sources from all directions. Here we examine which of these cues dominates in intensity-independent distance perception. We conducted a behavioral experiment in a virtual reverberant environment using broadband noise stimuli. Binaural room impulse responses (BRIRs) were manipulated to either eliminate one of the cues, or to make them incongruent. The average results show that the subjects are more sensitive to the variation in ILD than DRR, when only one of the cues is available. However, a quarter of the subjects were more sensitive to DRR. These individual differences were also observed when the cues were presented incongruently, confirming that while the ILD is dominant among normal-hearing listeners, some listeners prefer the DRR cue.

Keywords: Auditory Distance Perception, Virtualization, Computational Modelling, Psychoacoustics.

Acknowledgements

Spatial Audio Virtualization & Gamification for Hearing Assessment and Enhancement (European Research Executive Agency; HORIZON-MSCA-2022-SE-01), Computational assessment of plasticity in spatial hearing (Science Grant Agency of the Slovak Rep. VEGA, 2022-24).

INTRODUCTION

Localizing the objects of interest or threat is critical in daily life activities. Auditory modality is special in that it provides information even for objects that are occluded or behind the user (Kolarik et al., 2016; Zahorik et al., 2005). For example, consider a person reaching for a ringing phone in dark (Zahorik et al., 2005). Although auditory distance perception is critical in such scenarios, its functional mechanisms are not well understood (Kolarik et al., 2016; Zahorik et al., 2005). The estimates of auditory distance are typically dominated by the overall received stimulus intensity (Warren, 1999). However, in many situations, the emitted sound level is varying or unknown. In such cases, auditory distance perception relies on intensity-independent cues (Kopčo et al., 2012).

Recent results showed that robust intensity-independent distance perception is possible for nearby sources (up to 100 cm from the listener) in simulated reverberant environments. For such sources, two major intensity-independent distance cues exist, the inter-aural level difference (ILD) (Brungart, 1999; Shinn-Cunningham et al., 2005) and direct-to-reverberant energy ratio (DRR) which compares the sound energy received at the ears directly from the source to that reflected off the walls in reverberation (Mershon & King, 1975). ILDs provide distance information for stimuli off the midline, while DRRs provide distance information for sources from all directions. Although, we have started to understand which of the two cues the listeners use (N. Kopco et al., 2012; Kopčo & Shinn-Cunningham, 2011; Kopco et al., 2020), the question still remains which of these cues dominates in distance perception.

The current study focuses on understanding cue weighting in auditory distance perception. For this we examined auditory distance perception for lateral nearby sound sources located at various distances simulated along the interaural axis in a virtual reverberant environment (Fig. 1B). To assess the individual cue contribution and relative weighting, we manipulated the BRIRs (Kopčo et al., 2012; Shinn-Cunningham et al., 2005) to affect the availability and congruency of the DRR and ILD cues (Fig. 1A). Five types of stimuli were generated (Fig 1A). The unaltered BRIRs were used for the Congruent stimuli in which DRR and the ILD vary congruently with distance. In the ILD-only (or DRR-only) stimuli, the DRR (or ILD) was scaled so that it was fixed at the value of the 38-cm stimulus at all distances, while the other cue varied naturally. In the Incongruent(DRR) condition, the ILDs were unmodified while the DRRs were reversed at both ears by scaling the reverberant portion of the BRIRs so that the DRR of the stimulus at 15 cm corresponded to the DRR of the 100 cm unaltered BRIR, etc. Analogously, the Incongruent(ILD) condition was created by scaling the relative levels at the two ears. It was expected that the percepts would follow the cue that is weighted more when the cues were incongruent.

METHODS

SET UP & STIMULI

In a behavioral experiment, 15 human subjects (3 outliers) performed a distance discrimination task (Fig. 1C). The subjects were seated in a double-walled soundproof booth in front of an LCD display and a keyboard connected to a control computer which ran a Matlab (Mathworks) script controlling the experiment. The pre-generated stimuli were played through Fireface 800 sound processor (RME), and Etymotic Research ER-1 insert earphones.

A single set of non-individualized BRIRs was used, measured in a midsize classroom on a listener that did not participate in this study (Zahorik, 2002). Unless specified otherwise, all details of the measurement procedures, including the microphone, speaker, and the BRIR measurement technique used are identical to our previous studies (Kopco et al., 2020; Kopčo et al., 2012). Either the whole BRIR or only its reverberant tail were scaled to modify the ILD or DRR as shown in Fig. 1A.

On each trial, two 300-ms broadband noise stimuli were presented consecutively from two different randomly chosen distances at a roved level. The listener indicated which sound was closer. Each listener performed 81 trials per condition. Feedback was provided on each trial in all conditions except the incongruent ones.

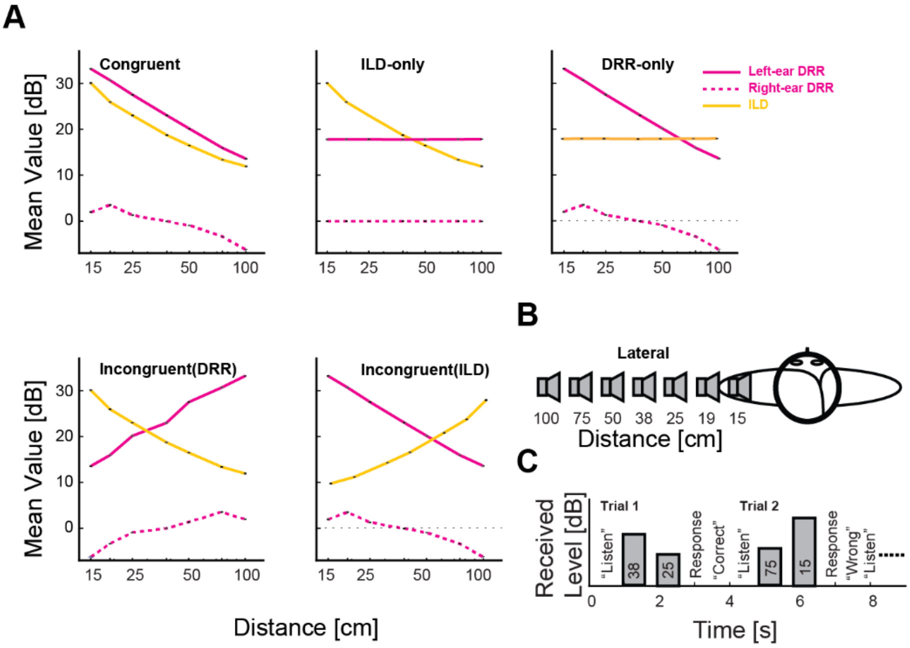

Figure 1. Experimental design. (A) Cue manipulation in different conditions. In the Congruent condition the cues ILD and DRR both were present and varied congruently with distance. In

Incongruent condition one of the cues was reversed as a function of distance. In the ILD-only condition the DRR cue values were fixed at the values corresponding to distance of 38 cm. In DRR-only condition ILD was fixed. (B) Simulated source locations. (C) Design of trials: The instruction "Listen" appeared on the screen, followed by presentation of two stimuli from different distances at random levels. Listeners responded by indicating whether the second stimulus was closer or more distant than the first stimulus. On-screen feedback was provided in all conditions except for the incongruent ones.

DATA ANALYSIS

The proportion of correct responses was computed for each stimulus condition, distance pair, and subject. A simple decision theory model based on psychophysical decision theory (Durlach & Braida, 1969) was used to compute nominal distance sensitivity d'_N across all distance pairs (Kopčo et al., 2012) for all conditions except the incongruent ones. d'_N represents the sensitivity for a nominal distance separation (for difference of log(distance) equal 1). A similar measure d_{ILD} was used for the incongruent conditions (d_{ILD} has the same meaning as d'_N except that positive values mean that responses follow the ILD more than the DRR while negative values mean responses following DRR). Statistical comparisons were done using repeated measures ANOVAs.

FINDINGS AND ARGUMENT

Fig. 2 shows the intensity-independent distance discrimination sensitivity for the three stimulus conditions of behavioral experiment. The sensitivity index d'_N was significantly higher for congruent condition than in the ILD-only and DRR-only conditions (3.36 vs. 2.59 and 1.27). These results are consistent with the hypothesis that both DRR and ILD are used as intensity-independent distance cues, contributing significantly to the distance percept, when both are available. Additionally, the value of d'_N for ILD-only condition was significantly higher than the DRR-only condition. This indicates that listeners were more sensitive to distance-dependent variations in ILD than to those in the DRR. A repeated measures ANOVA with the factors of sound types (3 levels) supports these observations, by finding a significant main effect of sound types ($F_{2,22} = 12.43$, $p = 0.0002$). Post-hoc pairwise comparisons with Bonferroni correction found significant differences for all three condition pairs ($p < 0.005$). However, the results also show that 1/4th of subjects (marked in black) was more sensitive to the DRR.

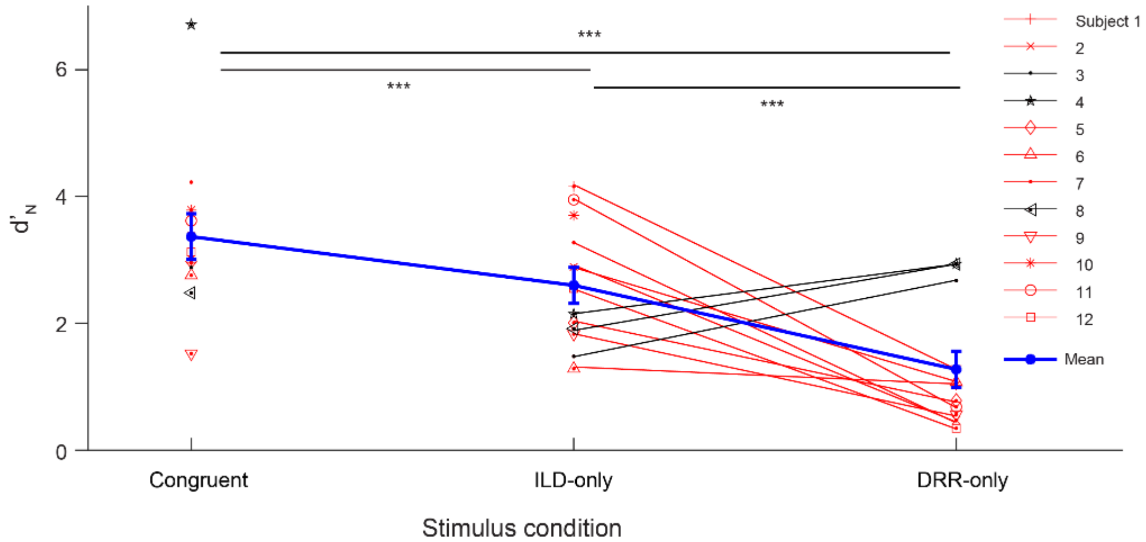

Figure 2. Discrimination sensitivity for the three stimulus conditions. The across-subject sensitivity index d'_N (+/-SEM) as well as individual values are shown. Horizontal lines and asterisks link conditions with performance significantly different in pairwise t-tests (*** for $p < 0.005$).

Fig. 3A shows the performance for the two incongruent conditions, this time as a scatter plot using the d_{ILD} measure. Results in Fig. 3A show that the two ways of creating the incongruent-cue stimuli are highly correlated, with a slight tendency for larger values in the Incongruent(DRR) stimuli. The black symbols in this figure indicate the 3 subjects whose d'_N was larger in DRR-only than ILD-only condition, showing that these three subjects were the only ones following the DRR cue in the incongruent conditions.

To directly test whether these differences in sensitivity to the individual cues predict how the cues are combined when available simultaneously, Fig. 3B shows a scatterplot of the across-incongruent-condition average d_{ILD} (from Fig. 3A) vs. the difference between d'_N in the ILD-only and DRR-only conditions (from Fig. 2). This figure clearly shows a separation between the two groups of subjects in both measures.

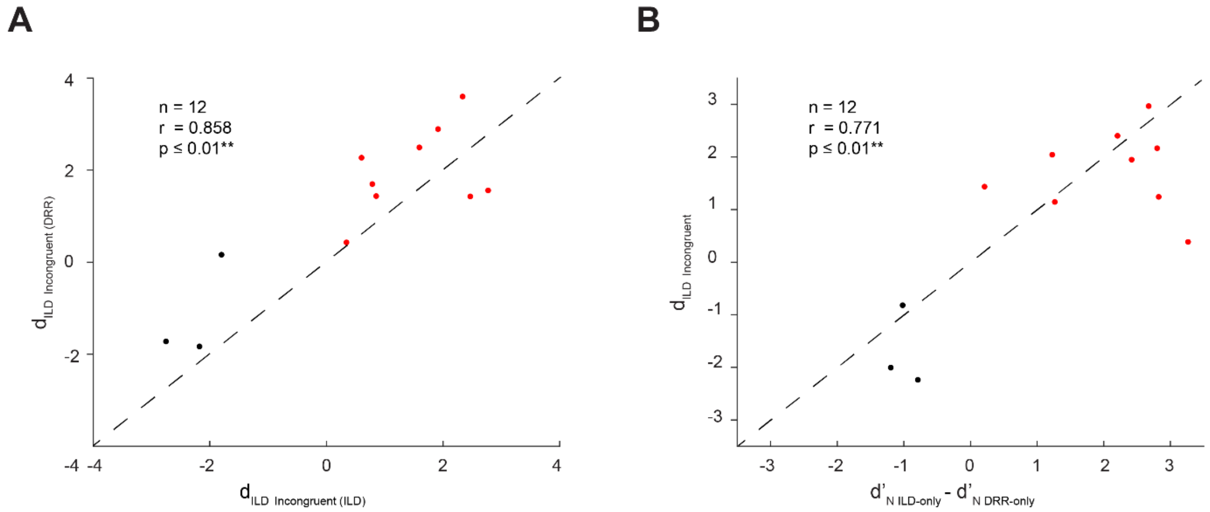

Figure 3. Correlation analysis. (A) Comparison of incongruent condition variants. (B) Comparison of incongruent and individual cue weighting.

CONCLUSIONS

The congruent condition had better performance than the ILD-only or DRR-only conditions, supporting the hypothesis that both cues contribute to the intensity-independent distance percepts. On average, the subjects were more sensitive to the distance-dependent variation in ILD than in DRR, suggesting that ILD is the dominant cue. However, there also were subjects more sensitive to the DRR cue. These between-subject differences were preserved even when the cues were pitched against each other in the incongruent condition, indicating that there is a large variation in the cue weighting even in in the normal-hearing subject population. These results are not consistent with those of (Kopčo & Shinn-Cunningham, 2011) who, based on modeling, suggested that monaural DRR is the primary cue used by the listeners when judging distance. It is possible that this difference is driven by the fact that only lateral stimuli were used in the current study, thus ILD was always available, while in (Kopčo & Shinn-Cunningham, 2011) also frontal stimuli with no ILDs were included. Thus, it is possible that listeners change the cue weighting dependent on the current context and cue availability.

REFERENCES

Brungart, D. S. (1999). Auditory localization of nearby sources. III. Stimulus effects. *The Journal of the Acoustical Society of America*, *106*(6), 3589–3602.

Durlach, N. I., & Braida, L. D. (1969). Intensity Perception. I. Preliminary Theory of Intensity Resolution. *The Journal of the Acoustical Society of America*, *46*(2B), 372–383.

Kolarik, A. J., Moore, B. C. J., Zahorik, P., Cirstea, S., & Pardhan, S. (2016). Auditory distance perception in humans: a review of cues, development, neuronal bases, and effects of sensory loss. *Attention, Perception & Psychophysics*, *78*(2), 373–395.

Kopco, N., Doreswamy, K. K., Huang, S., Rossi, S., & Ahveninen, J. (2020). Cortical auditory distance representation based on direct-to-reverberant energy ratio. *NeuroImage*, *208*, 116436.

Kopčo, N., Huang, S., Belliveau, J. W., Raij, T., Tengshe, C., & Ahveninen, J. (2012). Neuronal representations of distance in human auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(27), 11019–11024.

Kopčo, N., & Shinn-Cunningham, B. G. (2011). Effect of stimulus spectrum on distance perception for nearby sources. *The Journal of the Acoustical Society of America*, *130*(3), 1530–1541.

Mershon, D. H., & King, L. E. (1975). Intensity and reverberation as factors in the auditory perception of egocentric distance. *Perception & Psychophysics*, *18*(6), 409–415.

Shinn-Cunningham, B. G., Kopco, N., & Martin, T. J. (2005). Localizing nearby sound sources in a classroom: Binaural room impulse responses. *The Journal of the Acoustical Society of America*, *117*(5), 3100–3115.

Warren, R. M. (1999). *Auditory perception: A new analysis and synthesis*. Cambridge University Press.

Zahorik, P. (2002). Auditory display of sound source distance. *International Conference on Auditory Displays*.

Zahorik, P., Brungart, D. S., & Bronkhorst, A. W. (2005). Auditory Distance Perception in Humans: A Summary of Past and Present Research. *Acta Acustica United with Acustica*, 409–420.