Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Psychological and Physiological Acoustics
Session 1pPPb: Psychoacoustics and Perception (Poster Session)

1pPPb22. Reflection orders and auditory distance

Catarina Mendonça*, João Lamas, Tom Barker, Guilherme Campos, Paulo Dias, Ville Pulkki, Carlos Silva and Jorge A. Santos

***Corresponding author's address: Department of Signal Processing and Acoustics, Aalto University, Espoo, 02150, Finland, Finland, mendonca.catarina@gmail.com**

The perception of sound distance has been sparsely studied so far. It is assumed to depend on familiar loudness, reverberation, sound spectrum and parallax, but most of these factors have never been carefully addressed. Reverberation has been mostly analysed in terms of ratio between direct and indirect sound, and total duration. Here we were interested in assessing the impact of each reflection order on distance localization. We compared sound source discrimination at an intermediate and at a distant location with direct sound only, one, two, three, and four reflection orders in a 2AFC task. At the intermediate distances, normalized psychophysical curves reveal no differentiation between direct sound and up to three reflection orders, but sounds with four reflection orders have significantly lower thresholds. For the distant sources, sounds with four reflection orders yielded the best discrimination slopes, but there was also a clear benefit for sounds with three reflection orders. We conclude that at least three reflection orders are required so that reflection-related cues are accounted for in distance estimates. Also, these cues might interact differently with the direct sound pressure cues at different distances.

Published by the Acoustical Society of America through the American Institute of Physics

Received 22 Jan 2013; published 2 Jun 2013

INTRODUCTION

Distance perception is a complex process that requires multiple cue integration within and across sensory modalities. Both in visual and in auditory distance perception, distance cues might be classified as dynamic or static, absolute or relative. Static cues are those obtained by static listener/source, such as sound pressure level, reverberation level, sound frequency and interaural differences. Dynamic cues are available through motion. In auditory distance, such cues are the acoustic tau and motion parallax. Absolute and relative distance cues differ in the nature of the information they provide. While relative cues allow distance estimates by comparison with other landmarks or mental rules, absolute cues might yield direct estimates.

Most studies in auditory distance perception address absolute distance estimates. In absolute auditory distance perception, it is widely known that auditory space is often underestimated and subject to an auditory horizon effect (e.g. Brokhorst and Houtgast, 1999; Mershon and King, 1975; Mershon and Bowers; Nielsen, 1993). Some absolute distance cues are familiar pressure level, interaural difference, parallax, and direct-to-reverb energy ratio.

Familiar Pressure Level

For well-known ecological sounds, such as a footstep, a car, or a plane, loudness levels are familiar. Therefore, they can provide a gross estimation of distance based on pressure changes and expected pressure levels (Philbeck and Mershon, 2002).

Interaural Difference

For nearby sources, interaural level differences are greater when sounds are close than when sounds are distant. This cue is particularly important for sources less than one meter away from the listener, when such differences become perceptually available (e.g. Bronkorst and Adelbert, 2002).

Parallax

Moving objects that are far away will move less, in degrees, at the listener's perspective point, than objects moving a same distance but closer to the listener. The effect of this cue over distance estimates is more expressive for nearby sources, where parallax differences are most informative (e.g. Zuzuki, takane, Kim, and Sone, 1998). This dynamic cue provides good distance information, namely for familiar sources.

Direct-to-Reverb Ratio (D/R)

This cue alone is considered the most substantial element for distance perception, and has been shown to accurately predict distance estimates (Bronkhorst and Houtgast, 1999; Zahorik, 2002). Such cue provides absolute estimates because the level of the direct sound is inversely proportional to the square root of the source distance, while the level of the reverberation is roughly independent of source location (Mershon et al., 1989). Bronkhost and Houtgast (1999) proposed that the distance perception can be predicted by fitting a temporal window around the ratio of direct-to-reverberant sound energy. They assumed that the system makes an error in extracting the D/R ratio because it cannot perfectly separate the direct sound from the reverberant sound. Their model assumed a 6 ms fixed time window for the integration of direct sound. All energy arriving later would be used to estimate the reverberant sound. This would also explain the auditory horizon. This model has been subject to some criticism and reviews, as will be further described below.

Relative distance cues are known to affect auditory distance estimates, although their role has been sparsely analysed. Some of these cues are sound pressure level, frequency spectrum, and reverberant energy.

Sound Pressure Level

This cue alone can provide some relative information about a source position. In an anechoic environment, sound level for a source falls off by 6 dB for each doubling distance (Loomis, Klatzky, and Golledge, 1999).

Sound Frequency Spectrum

Due to the selective absorption of high frequencies with passage through air, sound source spectrum changes with distance. It has been demonstrated that spectral content provides relative distance cue to perceived auditory distance: a decrease in high-frequency content can lead to increases in perceived auditory distance, but only when compared to with similar sounds. Its effect is independent of changes in overall sound level (Little, Mershon, and Cox, 1992).

Reverberant Energy

Reverberation has been shown to degrade the perception of source direction, but enhance distance perception (Shinn-Cunningham, 2000). Room reflectance helps judge distance at least for far sources (Mershon, Ballenger, Little, McMurtry, and Buchanan, 1989).

The process of finding all auditory distance perception cues, how they operate, and how to predict distance estimates is far from over. Perhaps a most comprehensive approach, where all cues are taken into account, and weighted according to each environmental condition will one day be established. But such approach should be flexible enough to account for many factors. It is now known, for instance, that distance estimates strongly interact with vision. Having a memorized visual mental reference is enough to affect auditory distance judgement (Calcagno, Abregú, Eguia and Vergara, 2012). Also, the model from Bronkhorst and Houtgast (1999) based only on D/R ratio has revealed some problems. There is data supporting the idea that distance perception is based on binaural information from reflecting surfaces (Bronkhorst and Adelbert, 2002): auditory distance perception depends on the number of reflecting surfaces in the room, in particular, the number of reflecting walls. When two lateral walls are completely absorbent, the perceived distance of sound sources is close to the head, if absolute level and binaural cues are minimized. Their reviewed model integrates reflection directionality.

It has been demonstrated that how each cue is combined varies more than expected. The perceptual weight assigned to direct sound level, reverb level, and D/R ratio varies substantially as a function of both source type and angular positon (Zahorik, 2002).

In this work we address the effect of room reflections on relative distance perception. We intended to specifically assess the amount of reflected information that is needed in order to observe a benefit over the distance estimates.

METHODS

Participants

Five subjects, with ages comprised between 21 to 30 years old, took part of this experiment. All had normal hearing, verified through standard audiometric screening.

Stimuli

Stimuli consisted of 250 ms long white noise sounds convolved with dummy-head HRTF records from the CIPIC database (Algazi, Duda, Thompson, and Avendano, 2001) and auralized at several distances in a virtual room. The virtual room was 24 m wide, 50 m long and 11 m high. The listener was computed as being at half width, 12 m from each side, and only 5 m from the wall, facing the room in length. Sounds were auralized at 5, 10, 15, 20, 25, 30, and 35 m from the listener, always in the middle of the room, and therefore always at 0º azimuth. All sounds were auralized with direct sound only, one, two, three and four reflection orders. Stimuli were reproduced through s set of Etymotics ER-4S insert earphones for optimal external room noise cancellation.

Procedure

The general procedure for the distance perception experiments consisted of a 2-altenative forced choice task with a constant stimuli method. Two sets of stimulus pairs were created, one for intermediate and the other for far distances. The intermediate distance set had sounds at 15 m as the reference and sounds at 5, 10, 15, 20, 25 and 30 m as test stimuli. The far set had sounds at 25 m as reference and at 15, 20, 25, 30 and 35 m as test stimuli. In each pair, the reference stimulus always had the same amount of reflections as the test stimulus. All reflection orders were tested for each reference distance.

Sound duration was variable, according to the number of reflections. The inter-stimulus interval was fixed at 0.8 sec. Each stimulus pair was repeated 20 times.

Subjects sat in a darkened room and had to respond by clicking in a two-button mouse if the second sound was closer or farther than the first. Before starting the experience, subjects heard some tape recordings of people talking and ball playing in a room with exactly the same size as the synthesised room. They were instructed to imagine that they were in that room, not in the actual experiments room, while performing the task.

RESULTS

Figure 1 shows the average results for the intermediate distances set. In that graph, cumulative Gaussian curves were fit to the normalized distance estimates. From this figure, it is observed that overall distance judgments were accurate, “far” responses increasing with greater distances. It is also observed that distance estimates from direct sound-only, one, two, and three reflection orders did not present differences, while sounds with four reflection orders were clearly better discriminated in depth.

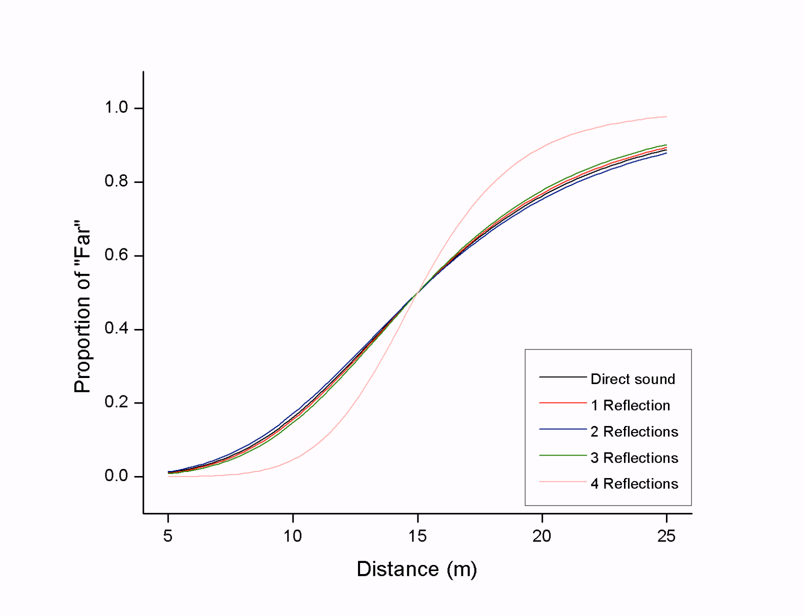

FIGURE 1. Normalized proportion of “Far” responses as a function of stimulus distance and number of reflection orders, fitted in cumulative Gaussian functions. Reference stimulus at 15 m.

All un-normalized pooled data were analyzed through curve fitting and the bootstrap technique. Through this method, cumulative Gaussian psychometric functions were fit to the binomial data and provided threshold, slope and spread information about the distance estimates. The bootstrapping computed standard deviations and confidence intervals of these functions (Effron, 1987). The confidence intervals were then used to determine if the different curve thresholds were significantly different from each other (Foster and Bischof, 1991). The method employed was that from Foster and Bischof (1997).

The psychometric function slopes provide us with a good measure of the quality of the stimuli discrimination. Steeper curves reveal better estimates and fewer errors. The slopes of the curves were .15, .12, .16, .16, and .26 for the direct sound, 1, 2, 3 and 4 reflection curves, respectively. The differences in slope between curves remain statistically insignificant, except for the 4 reflection order slope. This slope was statistically significant at the .01 level.

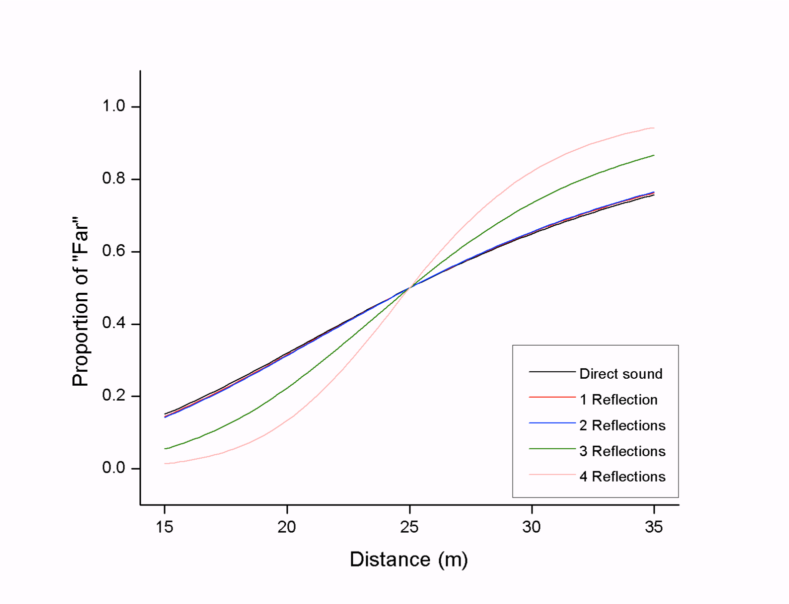

FIGURE 2. Proportion of "Far" responses as a function of stimulus distance and number of reflection orders, fitted in cumulative Gaussian functions. Reference stimulus at 25 m.

The results for the far distance estimates are presented in Figure 2. Overall, distance discrimination was poorer for these stimuli, with more errors and greater response variability. The normalized Gaussian curves reveal that sounds with no reflections, the sounds with one and two reflection orders had similar discrimination patterns. Sounds with three and four reflection orders had a clear benefit. The un-normalized psychometric fitting revealed as slopes .09, 0.9, .09, .14 and .18 for the direct sound, one, two, three, and four reflection order stimuli, respectively. Differences in slope between the direct sound, one and two reflection order curves remained insignificant. The slopes of the three and four reflection order curves were statistically different from the curves with fewer reflections, but not from each other, at the .01 level.

The fitted Gaussian distributions indicate the probability density functions for the distance estimates. From those functions the discrimination thresholds were obtained. The just noticeable difference (JND) was derived at the .84 point as JND=√2.σ. Results are shown in Figure 3.

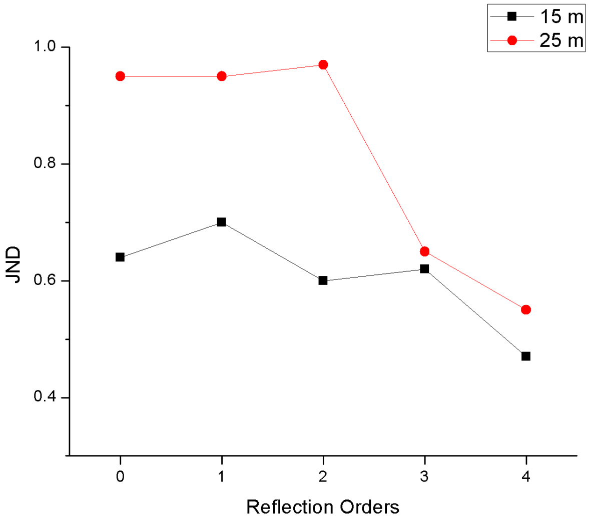

FIGURE 3. Distance discrimination thresholds, expressed in just noticeable difference (JND) values for both the intermediate (15 m) and far (25 m) stimuli pairs.

From the JND values, we observe that distance discrimination is poorer for far sounds than for intermediate ones, namely when none to two reflection orders are heard. This might reveal that listeners are using, for those sounds, mostly the direct sound level as a cue. This cue is less informative, as it varies less, between 15 and 35 m than between 5 and 25 m, easily explaining the poor discrimination for the far sounds with less reflected information. Interestingly, at the third reflection order discrimination thresholds become close for both stimuli sets. This might reveal that at this level reflection-related information starts to be accounted for in the distance estimates. With four reflection orders there are clearly lower thresholds in both stimuli sets, revealing a clear benefit obtained from the reflection cues.

CONCLUSION

In this work we addressed the effect of room reflections on relative distance perception. We intended to assess the amount of reflected information that is needed in order to observe a benefit over the distance estimates. Our results show that the amount of reflections necessary for a discrimination benefit varies with distance. At intermediate distances, between 5 and 25 m, the direct sound-only discrimination thresholds were similar to those for the 1, 2 and 3 reflection order sounds. This result reveals that at these distances estimates are based mostly on the direct sound pressure level. It is necessary to add four reflection orders to observe a benefit from this cue. With far stimuli, estimate accuracy is lower. Distance discriminability remains the same for sound from none to two reflection orders. It improves with three, and even more with four reflection orders.

In sum, to obtain reflection-related information with perceptual value for distance estimates, at least three reflection orders must be computed. The benefit of the sound reflections as a relative distance cue varies according to source position.

ACKNOWLEDGMENTS

This work was supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement No. 240453. It was also supported by the Portuguese Foundation for Science and Technology and FEDER, projects PTDC/EEA-ELC/112137/2009, FCOMP-01-0124-FEDER-007560, FCOMP-01-0124-FEDER-022674 and PEst-OE/ECI/UI4047/2011.

REFERENCES

Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. **(2001)**. "The CIPIC HRTF database." Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics. NY, 99-102.

Bronkhorst, A. W., Adelbert, W. **(2002)**. "Modeling auditory distance perception in rooms." Proc. EAA Forum Acusticum 2002, Sevilla.

Bronkhorst, A. W., and Houtgast, T. **(1999)**. "Auditory distance perception in rooms." Nature, **397**, 517-520.

Calcagno, E. R., Abregú, E. L., Eguisa, M. C., and Vergara, R. **(2012)**. "The role of vision in auditory distance perception. Percept. **41(2)**, 175-192.

Effron, B. **(1987)**. "Better bootstrap confidence intervals." J. Am. Stat. Assoc., **82**, 171-185.

Foster, D.H. and Bischof, W.F. **(1991).** "Thresholds from psychometric functions: superiority of bootstrap to incremental and probit variance estimators." Psychol. Bull. **109**, 152-159.

Foster, D.H. and Bischof, W.F. **(1997).** "Bootstrap estimates of the statistical accuracy of thresholds obtained from psychometric functions. Spatial Vis. **11**, 135-139.

Little, A. D., Mershom, D. H., and Cox, P. H. **(1992)**. "Spectral content as a cue to perceived auditory distance." Percept. **21(3)**, 405-416.

J. M. Loomis, R. L. Klatzky, and R. G. Golledge "Auditory distance perception in real, virtual and mixed environments." In *Mixed Reality: Merging Real and Virtual Worlds*, edited by Y. Ohta and H. Takamura (Ohmsha, Tokyo, 1999), Chap. 11.

Mershon, D. H., Ballenger, W. L., Little, A. D., McMurtry, P. L., and Buchanan, J. L. **(1989)**. "Effects of room reflectance and background noise on perceived auditory distance." Percept. **18(3)**, 403-416.

Mershon, D. H., and Bowers, J. N. **(1979)**. "Absolute and relative cues for the auditory perception of egocentric distance." **8**, 311-322.

Mershon, D. H., and King, L. E. **(1975)**. “Intensity and reverberation as factors in the auditory perception of egocentric distance.” Percept. Psychophys. **18**, 409-415.

Nielsen, S. H. **(1993)**. “Auditory distance perception in different rooms.” J. Audio Eng. Soc. **41**, 755-770.

Philbeck, J. W., and Mershon, D. H. **(2002)**. “Knowledge about typical source output influences perceived auditory distance. J. Acoust. Soc. Am. **111**, 1980-1983.

Shinn-Cunningham, B. **(2000)**. “Learning reverberation: Considerations for spatial auditory displays.” Proceedings of the International Conference on Auditory Display, 126-134.

Suzuki, Y., Takane, S., Kim, H-Y., and Sone, T. **(1998)**. “A modeling of distance perception based on auditory parallax model.” J. Acoust. Soc. Am. **103**, 3083.

Zahorik, P. **(2002)**. “Assessing auditory distance perception using virtual acoustics.” J. Acoust. Soc. Am. **111**, 1832-1846.