3aPPb13. The role of amplitude modulation in auditory distance perception

Pavel Zahorik[1,2,3] & Paul W. Anderson[2]

[1]Division of Communicative Disorders, University of Louisville School of Medicine, Louisville, KY 40292; [2]Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY 40292; [3]Heuser Hearing Institute, Louisville, KY 40203

ABSTRACT

The ratio of direct to reverberant sound energy (D/R) has been shown to be a primary acoustic cue to perceived sound source distance. Because it is unclear exactly how D/R might be encoded in the auditory system, a variety of more physiologically plausible correlates to D/R have been identified, including: spectral variance, interaural correlation, and temporal cues. Here, following recent neural work by Kuwada and Kim [ARO, 2014], we describe a new correlate to D/R and perceived distance related to the amplitude modulation (AM) depth of the signal at the listener's location. This cue is caused by the change in the modulation transfer characteristics of the room as a function of source distance. Results from an apparent distance estimation task confirm the efficacy of this AM depth cue in a reverberant soundfield (approximate broadband T_{60} = 3 s), when level cues are made ineffective. Distance estimates were found to be more accurate when the source signal (1-octave band of noise centered at 4 kHz) had AM (32 Hz, 100% depth), and this facilitation was only observed in reverberation. The facilitation was most evident for monaural input, indicating that the AM depth cue is likely processed monaurally.

PACS: 43.66.Qp

INTRODUCTION

It has long been realized that the ratio of direct to reverberant sound energy (D/R) varies systematically with distance in reverberant environments (Békésy, 1938). Unlike level cues to distance, D/R is not confounded with the acoustic power of the sound source (Zahorik and Wightman, 2001). It is therefore considered to be a primary distance cue (Mershon *et al.*, 1989), and can provide absolute distance information (Mershon and King, 1975; Mershon and Bowers, 1979). Particularly unsatisfying about this account, however, is that there is no evidence of D/R being coded directly in the auditory system. For this reason, other correlates to D/R have been explored psychophysically, including spectral variance (Jetzt, 1979), interaural correlation (Bronkhorst and Houtgast, 1999; Bronkhorst, 2001; Larsen *et al.*, 2008), spectral centroid (Larsen *et al.*, 2008), and temporal cues (Zahorik, 2002b). Here, following recent work by Kim and colleagues (Kim *et al.*, 2014), we add amplitude modulation (AM) depth to this list of potential D/R correlates (see Fig. 1). It is a particularly appealing correlate, given the prevalence of AM sounds in natural environments, and the known acoustical transformations imposed by those environments on AM reaching the ears.

METHODS

Subjects

Ten subjects (7 female, 3 male) participated in the experiment. All had audiolometrically-verified normal hearing (pure tone, air-conductive thresholds ≤ 25 dB HL from 125 to 8000 Hz). Subject age ranged from 18 to 27.3 years (median age: 21.1 years). Subjects received course credit for participation in the experiment. All procedures involving human subjects were approved by the University of Louisville Institutional Review Board.

Stimuli and Procedure

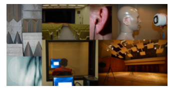
Listeners were presented with sounds at different simulated source distances and asked to estimate the egocentric distance of each sound, using methods broadly similar to those described in Zahorik (2002a). Virtual auditory space techniques were used to simulate sound field listening to sources at distances ranging from 0.35 to 5.6 m. The simulations were based on techniques described in previous work (Zahorik, 2009), and used non-individualized head-related transfer functions measured from a fixed distance of 1.4 m in anechoic space. Because of this, the near-field binaural cues to distance were unavailable. Two types of sound fields were simulated: anechoic and reverberant (room volume: 500 m^3, approx. broadband T_{60} = 3 s). As shown in Fig. 2, the ratio of direct to reverberant sound energy varied systematically with distances in the reverberant environment, and therefore was likely a primary acoustic cue to source distance. The source incidence angle was 90-degrees to the listener's right, at ear level.

The source signal was a 1-octave band of noise, centered at 4 kHz, 2 s in duration (500 ms rise/fall raised-cosine gate). In certain conditions, this signal was sinusoidally amplitude modulated (100% modulation depth) at a frequency of 32 Hz. This frequency was chosen because room reverberation is known to cause significant attenuation to amplitude modulation at this frequency and above (Houtgast and Steeneken, 1985; Zahorik *et al.*, 2011; Zahorik *et al.*, 2012). Fig. 3 demonstrates how AM depth (gain) varies systematically with distance in the reverberant environment, but not in anechoic space.

In order to limit listener's use of level cues in performing the distance estimation tasks, two types of level controls were implemented. First, level was equalized for distance by adjusting the gain of the simulated source in order to compensate for the 6 dB loss per distance doubling observed in anechoic space. Second, level was randomly varied (roved) over ±6 dB from trial to trial in the experiment. Listeners were also explicitly instructed to ignore any loudness differences between trials. Testing was conducted both monaurally (ipsilateral ear) and binaurally. Listeners provided 10 estimates for each target distance (presentation order randomized).

---

www.zaplab.org



RESULTS

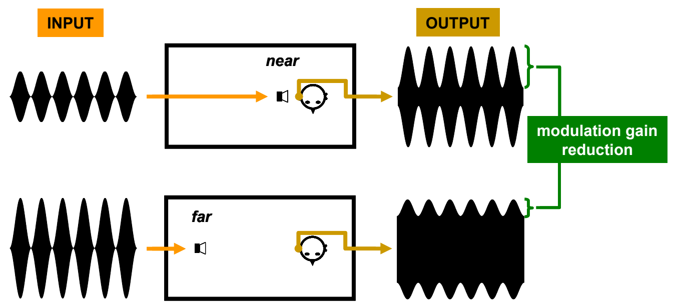


Fig. 1. Stylized example of modulation loss with increasing distance in reverberant space.

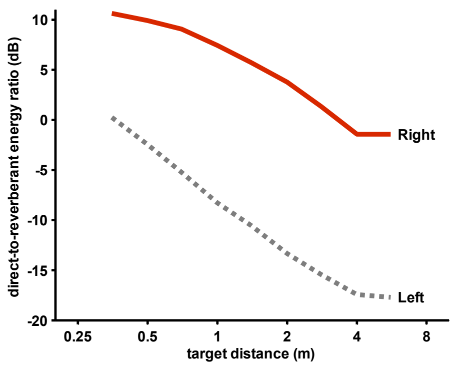


Fig. 2. Direct-to-reverberant energy ratio (D/R) in the 4 kHz octave-band as a function of target distance for the simulated reverberant room sound field, with source at 90 degrees. Values for both left (contra-lateral) and right (ipsi-lateral) ears are shown.

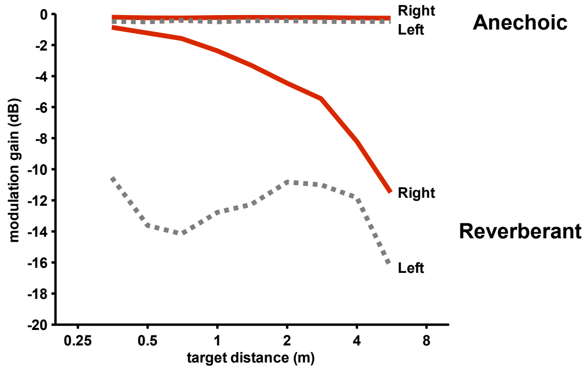
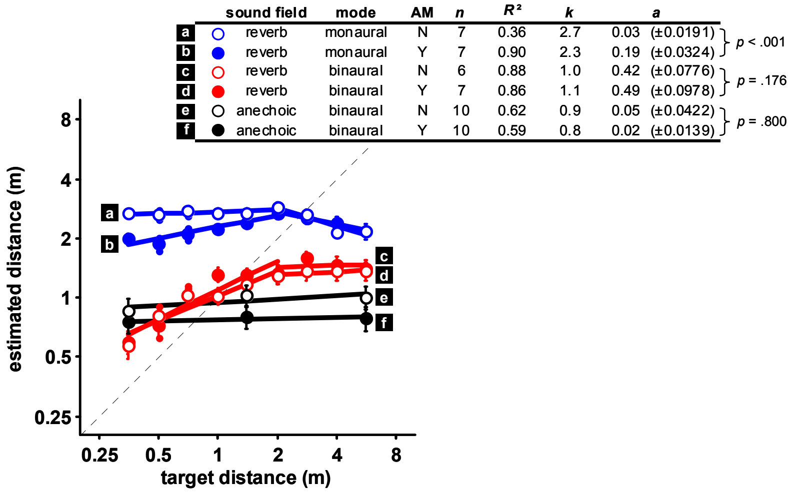


Fig. 3. Amplitude modulation gain as a function of target distance for the simulated reverberant and anechoic sound fields, with source at 90 degrees. Values for both left (contra-lateral) and right (ipsi-lateral) ears are shown. The source signal was the same as used for subsequent apparent distance testing: a 1-octave wide noise carrier centered at 4 kHz, sinusoidally amplitude-modulated at 32 Hz, 100% depth.



		sound field	mode	AM	*n*	R^2	*k*	*a*	
a	○	reverb	monaural	N	7	0.36	2.7	0.03 (±0.0191)	$p < .001$
b	●	reverb	monaural	Y	7	0.90	2.3	0.19 (±0.0324)	
c	○	reverb	binaural	N	6	0.88	1.0	0.42 (±0.0776)	$p = .176$
d	●	reverb	binaural	Y	7	0.86	1.1	0.49 (±0.0978)	
e	○	anechoic	binaural	N	10	0.62	0.9	0.05 (±0.0422)	$p = .800$
f	●	anechoic	binaural	Y	10	0.59	0.8	0.02 (±0.0139)	

Fig. 4. Estimated distance as a function of simulated target distance for six different listening conditions (a-f). Symbols indicate the (geometric) mean distance estimate pooled across listeners, and bars display the 95% confidence interval. Symbol coding for the conditions is shown in the accompanying table (Sound field: reverb or anechoic; Listening mode: monaural or binaural; Source signal AM: yes or no), along with the parameters of the best-fitting power function of the form $d' = kd^a$, where d' is the estimated distance and d is the simulated distance. Additional table entries include: the number of listeners contributing to each function, n; the proportion of variance accounted for by each fit, R^2; and ±1 standard error about the exponent parameter, a. For the reverberant room sound fields, separate power function fits were performed for near (≤ 2m) and far (≥ 2m) sources. Because there was relatively little change in the far functions, only the parameters for the near functions are shown.

Following data analysis strategies from past work (Zahorik, 2002a; Zahorik *et al.*, 2005), power functions were fit to the data to describe the relationship between estimated distance and target (simulated) distance. Because distance estimates in this setting were found to plateau after 2 m, perhaps due to the effect of the auditory horizon (Békésy, 1949), analysis was focused on distances 2 m and closer, and excluded subjects whose distance estimates had a significant inverse relationship with target distance. Results of this analysis are shown in Fig. 4. These results demonstrate that:

1. Distance estimates are significantly more accurate (greater slope in the log-log plots) in a reverberant room when source signals have AM relative to those that do not. In this setting, the improvement was limited to monaural listening, however. The strong binaural information available at 90 degrees may have trumped the distance information resulting from the AM cue in this case. Overall, these results are consistent with the notion that at modulation frequency of 32 Hz, distant-dependent modulation loss may serve as a cue to distance in reverberant sound fields. To our knowledge, this is the first psychophysical demonstration of the effectiveness of this cue.

2. There can be a substantial binaural advantage for distance estimation accuracy. The improvement in fitted function slope was at least 0.23, $p < .001$. This result is consistent with previous fact that additional information is available to the listener under binaural listening conditions, and that this additional binaural information is known to improve distance estimation accuracy, particularly for lateral sources (Gardner, 1969; Holt and Thurlow, 1969). Further study will be needed to more fully characterize this apparent interaction between listening mode (monaural/binaural) and the AM cue.

3. Distance accuracy is highly limited in anechoic space, when no near-field or level cues are made available to the listener. This result has been known since at least the work of Gardner (1969). The fact that AM does not significantly improve distance accuracy in anechoic space supports the hypothesis that the use of AM as a distance cue is specific to reverberant sound fields.

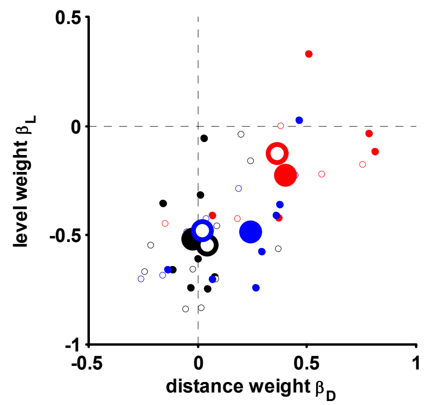


Fig. 5. Scatterplot of estimated perceptual weights assigned to level and non-level distance cues for all listeners and listening conditions in the experiment. Small symbols indicate weights for individual listeners. Listening conditions are indicated by the coding scheme shown in Fig. 4. Large symbols indicate weight centroids for each condition.

To more fully examine listeners' interpretation of multiple acoustic cues to distance, we computed the relative weights listeners placed on the (random) level cue versus all other cues to distance, using multiple regression techniques analogous to those used and described in (Zahorik, 2009). Figure 5 is a scatterplot of the distance and level weights for all listeners in the experiment, grouped by condition (same symbol scheme is in Fig. 4). Results from this analysis indicate that:

1. Most listeners were unable to ignore loudness/level cues, even though they were explicitly instructed to do so. This was evidenced by the non-zero level weights in nearly all cases. Negative level weight indicates an inverse relationship between level and estimated distance, which, of course, is what occurs most typically under natural listening conditions. This result may suggest that listeners attended to other unintended level cues in the reverberant room. One likely strategy involves attention to the level of the reverberant tail, which on average would have increased with increasing distance, given our procedure to equalize overall level.

2. There is a clear progression in the relative weight of the level cue. As additional distance information was made available to listeners, primarily through the addition of binaural cues, less weight was generally placed on level cues. This result is consistent with the notion that perceptual weighting strategies for auditory distance perception are plastic depending on the distance cue availability (Zahorik, 2002a).

CONCLUSION

Amplitude modulation gain can provide (monaural) distance information.

REFERENCES

Békésy, G. v. (**1938**). "Über die Entstehung der Entefernungsempfindung beim Hören (On the origin of the sensation of distance in hearing)," Akustiche Zeitschrift **3**, 21-31.

Békésy, G. v. (**1949**). "The moon illusion and similar auditory phenomena," Am. J. Psychol. **62**, 540-552.

Bronkhorst, A. W. (**2001**). "Effect of stimulus properties on auditory distance perception in rooms," in *Physiological and Psychological Bases of Auditory Function*, edited by D. J. Breebaart, A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs, and R. Schoonhoven (Shaker, Maastricht, The Netherlands), pp. 184–191.

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

Gardner, M. B. (**1969**). "Distance estimation of 0 degrees or apparent 0 degree-oriented speech signals in anechnoic space," J. Acout. Soc. Am. **45**, 47-53.

Holt, R. E., and Thurlow, W. R. (**1969**). "Subject orientation and judgment of distance of a sound source," J. Acout. Soc. Am. **46**, 1584-1585.

Houtgast, T., and Steeneken, H. J. M. (**1985**). "A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria," J. Acout. Soc. Am. **77**, 1069-1077.

Jetzt, J. J. (**1979**). "Critical distance measurement of rooms from the sound energy spectral response," J. Acout. Soc. Am. **65**, 1204-1211.

Kim, D., Carney, L., Bishop, B., and Kuwada, S. (**2014**). "A mechanism for neural coding of sound-source distance: experiment and model," Association for Research in Otolaryngology Abstracts **37**, 92-93.

Larsen, E., Iyer, N., Lansing, C. R., and Feng, A. S. (**2008**). "On the minimum audible difference in direct-to-reverberant energy ratio," J. Acout. Soc. Am. **124**, 450-461.

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," Perception **18**, 403-416.

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

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

Zahorik, P. (**2002a**). "Assessing auditory distance perception using virtual acoustics," J. Acout. Soc. Am. **111**, 1832-1846.

Zahorik, P. (**2002b**). "Direct-to-reverberant energy ratio sensitivity," J. Acout. Soc. Am. **112**, 2110-2117.

Zahorik, P. (**2009**). "Perceptually relevant parameters for virtual listening simulation of small room acoustics," J. Acout. Soc. Am. **126**, 776-791. PMCID: PMC2730711.

Zahorik, P., Brungart, D. S., and Bronkhorst, A. W. (**2005**). "Auditory distance perception in humans: A summary of past and present research," Acta Acustica **91**, 409-420.

Zahorik, P., Kim, D. O., Kuwada, S., Anderson, P. W., Brandewie, E., Collecchia, R., and Srinivasan, N. K. (**2012**). "Modulation detection by human listeners in reverberant sound fields: carrier bandwidth effects and binaural versus monaural comparison," Proceedings of Meetings on Acoustics **15**, 050002. PMCID: PMC3579647.

Zahorik, P., Kim, D. O., Kuwada, S., Anderson, P. W., Brandewie, E., and Srinivasan, N. K. (**2011**). "Amplitude modulation detection by human listeners in sound fields," Proceedings of Meetings on Acoustics **12**, 050005. PMCID: PMC3400464.

Zahorik, P., and Wightman, F. L. (**2001**). "Loudness constancy with varying sound source distance," Nat. Neurosci. **4**, 78-83.

Created: 28-APR-2014