34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Influence of sound source width on human sound localization

Nathaniel T. Greene and Gary D. Paige

***Abstract*— Free-field sound localization experiments generally assume that a loudspeaker can be approximated by a point-source; however, a large loudspeaker may extend beyond the width that two sources can be discriminated. Humans can accurately discriminate sound source locations within a few degrees, thus one might expect localization precision to decrease as a function of sound source diameter, much as precision is lower for localizing the center of a wide, blurry light source. In order to test the degree to which humans differentially localize small and large sound sources, auditory targets were presented using a single 25.4 cm by 10.2 cm elliptical loudspeaker with the primary axis oriented both horizontally and vertically in different sessions. Subjects were seated with their heads fixed by a bite bar in a darkened, echo-attenuating room facing a cylindrical, acoustically transparent screen at a distance of 2 meters. Auditory targets consisted of repeating bursts (5 Hz) of low frequency band-pass noise (0.2 – 1 kHz, 75 dB SPL). Subjects were instructed to quickly and accurately guide a laser pointer mounted on a cylindrical joystick towards targets, presented randomly within a field ± 40° in azimuth by ± 10° in elevation, with oversampled points located every ten degrees along the primary meridians. Localization accuracy and precision (mean and standard deviation of localization error at oversampled locations) were not significantly different between speaker orientations, and were comparable to baseline measurements recorded using a 7.6 cm circular speaker. We conclude that low frequency sound localization performance is not dependent upon the size of the sound source as predicted theoretically, and is well approximated by a point source.**

I. Introduction

A. Sound Localization

Mammalian sound localization is performed principally based upon three cues generated by the head: the width and mass of the head introduce differences in time and level between the two ears (interaural time and level differences; ITD and ILD), and the shape of the pinna induce notches in the spectrum of incoming sound. ITDs and ILDs provide localization information in the horizontal plane of the head containing the interaural axis, i.e. in azimuth, and spectral notches provide information regarding vertical localization. ITDs predominate below ~1 kHz, ILDs predominate above ~3 kHz, and spectral notches are present above ~4 kHz. [1]

Experiments conducted to assess sound localization ability typically utilize loudspeakers located a small distance from the listener, and assume that the sound source can be approximated by a virtual point source located in the center of the real loudspeaker. Under these conditions, frequencies targeting ITD discrimination produce the most accurate and precise localization, with minimum audible angles on the order of a few degrees [2]. However, the angle subtended by a large loudspeaker can be non-trivial, extending beyond the minimum audible angle.

The output of a loudspeaker can be approximated using an array of point sources equal in area to the loudspeaker. Intuitively one may suspect that each source could provide separate ILD and ITD cues which are combined to form the perception of a single broad object, just as many points of light can combine to form a single large visual object. Further, it seems obvious that while the center of such an object may be localized with great accuracy on average, the precision of this identification could be worse than for a single small source. Is a large auditory source localized with lower precision than a small source? Evaluation of the waveform [3] at the ears may aid in predicting such an effect.

B. Acoustics

A listening point is in the farfield (i.e. the point beyond which the radiated sound can be described by spreading spherical, as opposed to plane waves) when the radius of the vibrating piston (i.e. loudspeaker), a, is much smaller than the distance, r, to the surface, and can be approximated as the Rayleigh distance

$$R_0 = ka^2/2, \tag{1}$$

where the wave number, k, is the ratio of 2π to the wavelength. The pressure waveform for an arbitrary point in the farfield, L (Fig. 1A), which is r away from the center of a baffled circular piston that lies on the x-y plane, at an angle θ away from the z-axis, is found by integrating the output of each point (B) on the surface of the piston, and is given by the Rayleigh integral [4]:

$$p(r,\theta;t) = \frac{jkP_0}{\pi} e^{j\omega t} \int_0^{\pi} d\psi \int_0^{a} \sigma \frac{e^{-jkR}}{R} d\sigma \ , \tag{2}$$

where P_0 is an idealized pressure amplitude, and

$$R = \sqrt{r^2 + \sigma^2 - 2r\sigma \sin\theta \cos\psi} \ . \tag{3}$$

Evaluating these integrals yields a solution dependent upon the first order Bessel function of the first kind, $J_1(x)$:

$$p(r,\theta;t) = \frac{jaP_0}{r} \frac{J_1(ka\sin\theta)}{\sin\theta} e^{j(\omega t - kr)} \ . \tag{4}$$

Research supported by the Center for Navigation and Communication Sciences at the University of Rochester (NIDCD P30-DC005409).

N. T. Greene is with the Biomedical Engineering Department, University of Rochester, Rochester, NY 14642 USA (phone: 585-276-0281; fax: 585-756-5334; e-mail: Nate_Greene@urmc.rochester.edu).

G. D. Paige is with the Department of Neurobiology and Anatomy, University of Rochester, Rochester, NY 14642 USA (e-mail: Gary_Paige@urmc.rochester.edu).

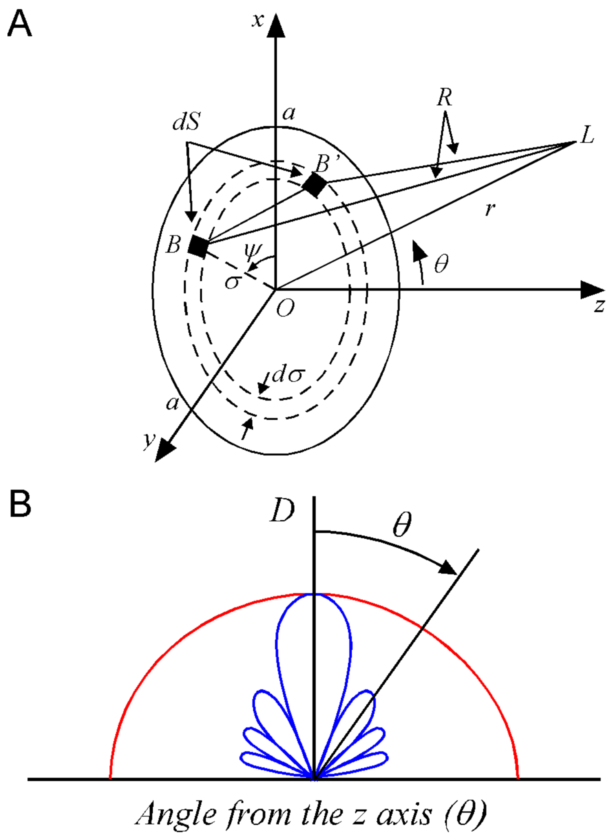

Fig. 1. A. Calculating the radiation from a circular piston of radius a. **B.** The directivity pattern of a pure tone produced by example small (red) and large (blue) circular pistons.

Since the piston can be modeled as an array of point sources, we must take into account how the resulting radiation from each source interacts with each other. The radiation from these sources will interact both constructively and destructively, dependent upon the angle (θ). This interference pattern is given by the directivity, which is the ratio of the pressure, p(r,θ;t), at a given angle (θ) to the pressure amplitude on the z-axis, p(r,0;t):

$$D(\theta) = \frac{2J_1(ka\sin\theta)}{ka\sin\theta}\,. \tag{5}$$

The principal effect of increasing the radius of a circular, baffled piston, therefore, is to increase the directivity (or decrease the beam width) of the sound produced (Fig. 1B).

Directionality also depends on the frequency of the radiated sound (via the wave number, k), such that high frequencies will be radiated in a more directional pattern than low frequencies. This suggests that we should expect the pressure amplitude to approach zero, producing an intensity null, when kasinθ is equal to a zero of the Bessel function (α), or rewritten:

$$\theta=\sin^{-1}(\alpha/ka) \tag{6}$$

We estimate the directivity our elliptical speaker in each orientation (major axis horizontal or vertical) as if it was being produced by circular speakers whose diameters match the dimensions of the major and minor axes of the elliptical speaker. It follows, therefore, that the lowest frequency to produce a null (at ± 90°) for a 10 inch diameter circular piston is at approximately 3 kHz, and the lowest frequency to produce a null for a 4 inch diameter piston is over 5 kHz.

C. Binaural hearing

The preceding treatment assessed the directivity of sound with respect to the angle deviated from directly in front of the sound source. Both the pressure waveform (4) and directivity (5) depend upon the distance (r) and angle from center (θ), thus the waveform present at two symmetrically positioned ears will be identical when the sound source is on the midline of the head and pointed directly at the center of the head. Further insight into localization ability can be gleaned by assessing deviation from this case.

Rotating the sound source around the head, while keeping the source pointed at the center of the head (i.e. for a head located at the apex of the curve in Fig. 1B), causes the distance (r) to decrease to one ear, and increase to the other. The result is a change in amplitude (by 1/r) and phase (by $e^{j(\omega t-kr)}$) of the sound arriving at the ears, producing the ILD and ITD cues used to localize sounds in azimuth. Importantly, these differences in amplitude and phase between the two ears are independent of speaker width (a), so long as the ears remain symmetrically oriented (i.e. maintain equal |θ|).

For the current experiment, therefore, theory predicts that localization performance will be identical for any loudspeaker in the far field, $r > R_0$ (e.g. a < 47 cm for 1 kHz). However, additional and unrecognized cues may affect perception, and evaluating these effects can provide insight into the mechanisms by which humans localize sound.

II. Methods

A. Subjects

Seven healthy young adults (3 males, 4 females; 18-30 yrs old) participated in this experiment. Subjects were screened for hearing and vestibular disorders prior to participating in the study. Experiments were conducted in accordance with the 1964 Declaration of Helsinki, and the protocol was approved by the University of Rochester Research Subjects Review Board. Informed consent was obtained prior to data collection, and subjects were reimbursed for their participation. All subjects were familiar with the apparatus and task, but were naïve to the experimental conditions.

B. Stimulus and apparatus

The experimental environment has been described previously [5-7]. Briefly, subjects were seated in a darkened, echo attenuating room. Head position was leveled, centered, and fixed with a personalized bite-bar two meters (2m) from a cylindrical section of acoustically transparent black cloth, behind which auditory and visual stimuli were presented. Subjects were instructed to quickly but accurately orient a laser pointer mounted on a cylindrical joystick towards the perceived source of the sound (Fig. 2A).

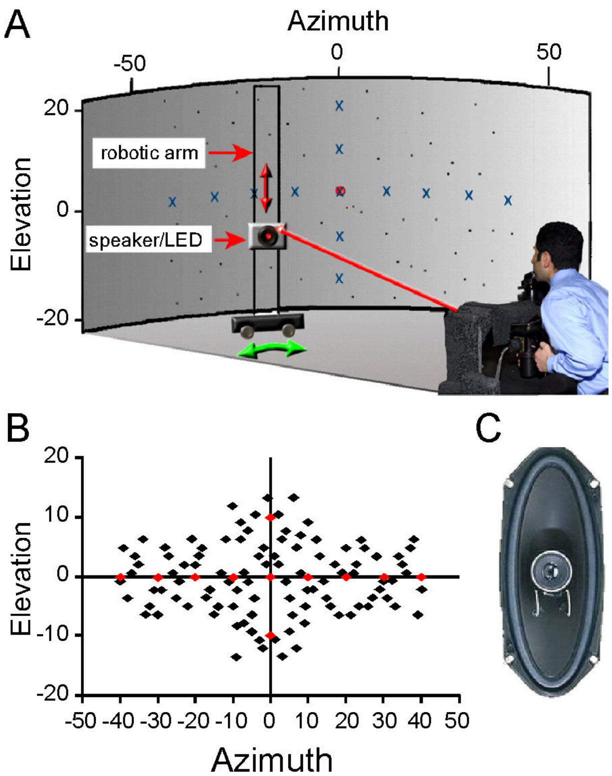

Fig 2. A. Schematic of the test chamber, subject orientation, and laser pointing apparatus. **B.** The target field (in degrees from center) included semi-random distracter (black) and oversampled (red) target locations. **C.** The sound source was a 2-way 10 x 4 inch (25.4 x 10.2 cm) elliptical speaker (Pioneer TS-A4103).

Sound localization was evaluated using trains of calibrated band-pass (0.2 - 1 kHz) noise bursts (150 ms duration, 10 ms rise/fall time), presented at 75 dB SPL (RMS; N_0 = 46 dB SPL), and repeated continuously at 5 Hz. Low frequencies were selected in order to selectively radiate sound from the large cone in a 2-way 25.4 cm by 10.2 cm elliptical speaker (Fig. 2C; Pioneer TS-A4103, crossover frequency > 7 kHz), and to limit the possible localization mechanisms to ITD comparisons [1]. The speaker was mounted on a robotic arm capable of translating the speaker anywhere within our experimental range of ± 40° in azimuth, and ± 10° in elevation and hidden behind the cylindrical screen of speaker cloth (Fig. 2A). A white noise masker was presented by separate bilaterally located speakers during robot movements in order to mask changes in target location. The auditory target and laser pointer were turned on concurrently at the beginning of each trial, and remained on until the subject localized the target, signaled by a button press. The target array presented during experiments is shown in Fig. 2B. Oversampled targets (red points) along the X and Y meridians were presented 10 times each, in order to provide an estimate of the mean and variance for each condition. Distracters were interspersed with the oversampled points to minimize predictability.

III. Results

Only localization in azimuth is considered here, due to poor subject localization ability to low frequency sounds in elevation. *Accuracy* of the subjects' localization performance is quantified as the mean localization error at each oversampled speaker location. The population mean and standard deviation (SD) is shown in Fig. 3A. In general, subjects tend to overshoot targets as a function of eccentricity [5].

Qualitatively, localization performance in azimuth appears essentially identical between vertical (red) and horizontal (blue) speaker orientations. To quantify the dependence of localization *accuracy* on speaker size the difference in mean localization error between vertical and horizontal speaker orientations is shown in Fig. 3B. As suggested above, performance is similar across orientations and target locations. Subjects localized eccentric targets with slightly lower error when the speaker was vertical; however, a one-way analysis of variance reveals no significant differences between groups, and Bonferroni corrected t-tests reveal no significant deviations from zero.

The *precision* of each subject's response was assessed by calculating the SD of localization error at each oversampled location. The mean and SD of precision across the population is shown in Fig. 3C. Average precision is similar for the two speaker orientations across stimulus locations, and is comparable to results using a 7.6 cm speaker [6,7].

The dependence of localization *precision* on speaker size is quantified as the difference in SD between vertical and horizontal orientations for each subject, and the mean difference is shown in Fig. 3D. Unexpectedly, the difference is slightly positive at most locations, suggesting that localization precision is slightly worse in the vertical than horizontal speaker orientation; however, a one-way ANOVA and a Bonferroni corrected Levene test for equality of variance reveal no significant differences across the population.

Examination of localization precision reveals that, counterintuitively, subjects are capable of localizing a large speaker with precision comparable to that of a narrow speaker (S.D. ~ 3°) and with a lower SD than the width of that speaker (~ 7° subtended for a 25.4 cm loudspeaker). These results suggest that loudspeaker diameter does not affect localization performance in the far field.

IV. Conclusion

In the first part of this report we reviewed the acoustics describing how source radius affects the sound radiated by a circular, baffled piston, and predicted that a listener will not be capable of discriminating two loudspeakers on the basis of speaker width using two symmetrically located ears. However, just as with any complex physical system, many assumptions are made when performing the calculations necessary to describe the acoustics. Similarly, while the basic mechanisms used for sound localization are generally well understood, there may exist additional and unrecognized cues that are used by the brain in forming a coherent auditory percept.

A direct test of localization ability when sound sources of various widths are presented is necessary in order to determine whether the brain is actually capable of

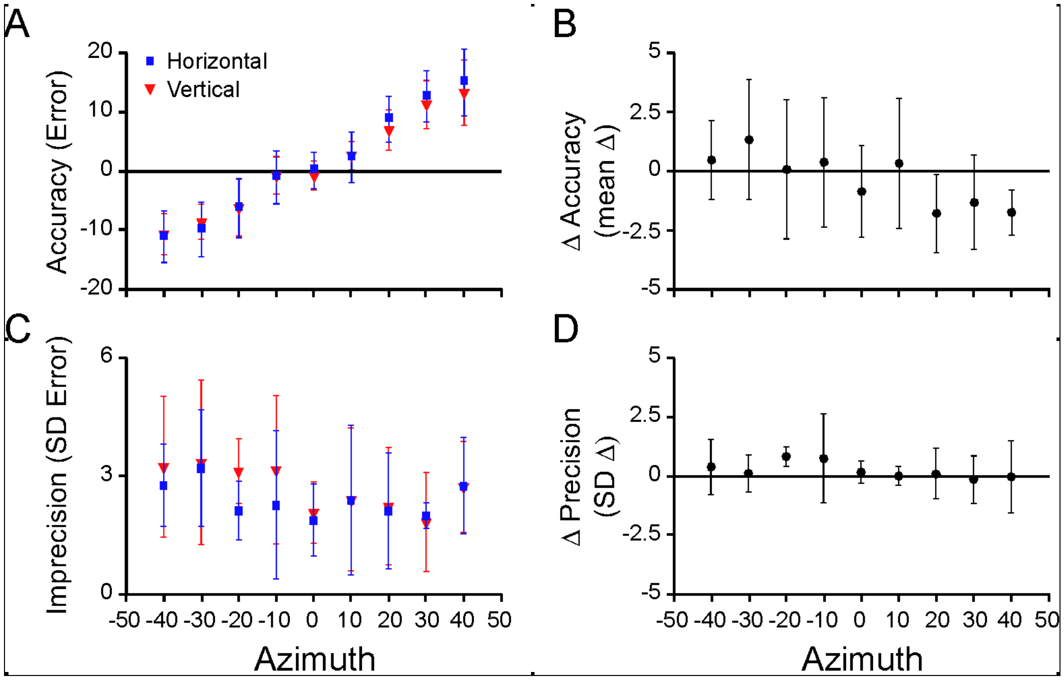

Fig 3. A. Mean localization error across subjects when the major axis of the elliptical speaker is vertical (red) and horizontal (blue). Error bars represent the SD for the population of localization means. **B.** Difference between the means in the vertical and horizontal orientations at each oversampled location. **C.** Mean of individual subject's localization error SD, which are listed as *imprecision* since large values indicate poor performance. **D.** Mean of differences in individual subject SD across speaker orientations for each oversampled location.

performing such discrimination. For this reason we conducted a study examining the accuracy and precision of human sound localization as a function of azimuthal sound source width. Under laboratory listening conditions subjects localized the sound sources with equal accuracy and precision, suggesting no dependence upon target size. However, the current results are limited by several factors.

First, a single elliptical loudspeaker was used to present stimuli. Use of this device ensured that no spectral differences were presented, thus ensuring any effects observed were due to source width, but limited the possible sizes available. Our current results suggest that in the far field, loudspeakers of different widths are indistinguishable on the basis of localization performance.

Second, localization comparisons were limited to discrimination of low frequency sounds in order to selectively radiate sound from the large cone of the loudspeaker. Directionality is more strongly modulated by, and ILD discrimination predominates, at high frequencies. Thus, localization of high frequencies could show significant effects on localization accuracy by speaker size, which is not visible for low frequency sounds.

Finally, sounds are not often presented under laboratory conditions. In particular, most environments include objects and walls that produce reverberations that the listener must discriminate in order to accurately localize the sound. While these effects contribute to the perceived spaciousness of the auditory object [8,9], the effects on localization ability are not clear. Overall, our results indicate that, unlike vision, hearing is independent of source object size in isolation, within the limits of our experimental parameters.

ACKNOWLEDGMENT

We thank the Human Subjects Core of the Center for Navigation and Communication Sciences at the University of Rochester for providing auditory and vestibular screening of subjects, K. Davis, W. O'Neill, and P. Allen for providing advice and discussion, and M. Dobreva and E. Clark for technical assistance.

REFERENCES

[1] J. Blauert, *Spatial Hearing: The Psychophysics of human sound localization*, Cambridge, MA:MIT Press,1997.

[2] Mills, "Auditory Localization," ed: Tobias, *Foundations of Modern Auditory Theory*, Academic Press, p 310, 1972.

[3] D.T. Blackstock, *Fundamentals of Physical Acoustics*, New York:J. Wiley, pp. 440-464, 2000.

[4] J. W. S. Rayleigh, *The Theory of Sound*, 2nd ed., vol. 2, London:Macmillan, 1896, secs. 278, 302.

[5] Razavi, B., W. E. O'Neill, and G. D. Paige, "Auditory spatial perception dynamically realigns with changing eye position." *J Neurosci*, vol. 27 no. 38, pp. 10249-58, 2007.

[6] Q. N. Cui, W. E. O'Neill, and G. D. Paige, "Perception of auditory, visual, and egocentric spatial alignment adapts differently to changes in eye position." *J. Neurophysiol.* vol. 103 no. 2, pp. 1020-35, 2010.

[7] M. S. Dobreva, W. E. O'Neill, and G. D. Paige, "Influence of aging on human sound localization." *J. Neurophysiol.* vol 105 no 5, pp 2471-86, 2011.

[8] L. L. Baranek, *Concert halls and opera houses: music, acoustics, and architecture*, Springer, 2004.

[9] F. Rumsey, *Spatial Audio*, Oxford, UK: Focal Press, 2001.