

A triangulation method for determining the perceptual center of the head for auditory stimuli

PACS REFERENCE: 43.66.Qp

Brungart, Douglas¹; Neelon, Michael²; Kordik, Alexander³; Simpson, Brian⁴

¹AFRL/HECB, Air Force Research Laboratory, WPAFB, Ohio USA 45433

²University of Wisconsin, 423 Psych Bldg, W J Brogden, Madison, Wisconsin USA 53706

³Sytronics Inc., 4433 Dayton-Xenia Road, Dayton, Ohio USA 45432

⁴Veridian, 5200 Springfield Pike, Ohio USA 45431

Tel: 937-255-3660 ext. 414

Fax: 937-656-7680

E-mail:douglas.brungart@wpafb.af.mil

ABSTRACT

Although sound source locations are traditionally expressed in a coordinate system with its origin at the midpoint of the listener's interaural axis, there is little evidence that listeners actually use this coordinate system to judge the relative locations of sounds. In this experiment, location pairs where nearby and distant sound sources appeared to be at the same angle in azimuth were used to triangulate the location of the perceptual center of the head. The results show that an auditory parallax effect generally shifts the perceptual center of the head several centimeters in front of the listener's interaural axis.

1. INTRODUCTION

Over the past century, dozens of experiments have been conducted to examine how accurately listeners are able to judge the locations of sound sources and to identify the auditory cues that listeners use to make these localization judgments. A common requirement of all of these auditory localization experiments has been the selection of a coordinate system to represent the actual and perceived locations of sound sources. Although no single coordinate system has been adopted as a standard, most of the coordinate systems that have been used in auditory research have been similar in two important ways. 1) They have been based on polar coordinates (probably because the directional auditory cues used for localization depend only on the direction of the sound source at distances greater than one or two meters); and 2) they have used an origin that was located at the midpoint of the interaural axis (Blauert (1983), for example, explicitly defined the origin of his coordinate system as the point "halfway between the upper margins of the entrances to the two ear canals).

Although it is difficult to argue with the practical utility of this anthropometrical definition of the "center of the head", there is little evidence to suggest that it accurately represents the "perceptual" center of the head. In this context, we refer to the "perceptual" center of the head as the origin of the internal coordinate system that listeners use to encode the apparent locations of sounds. The origin of this coordinate system is the point where a sound would appear to originate from a location "exactly" in the center of the head. Judgments about the absolute locations and relative directions of sound sources are presumably also made relative to this origin. Thus, one would expect that sound sources at different distances that are perceived to originate from the same direction will be in line with the "perceptual" center of the head. In this regard, the auditory center of the head is analogous to the direct visual "egocenter", which has been defined as the location in the head towards which rods point when they are judged to be pointing directly to the self (Howard and Templeton, 1966). Thus, it seems appropriate to refer to the auditory "center of the head" as the auditory egocenter.

Although we know of no studies that have specifically examined the location of the auditory egocenter, there is evidence to suggest that it is located somewhere on the median sagittal plane. The best evidence for this comes from auditory lateralization studies, which have shown that listeners consistently report that sounds that are more intense at the left ear and/or arrive first at the left ear appear to be located on the left side of the head, sounds that are more

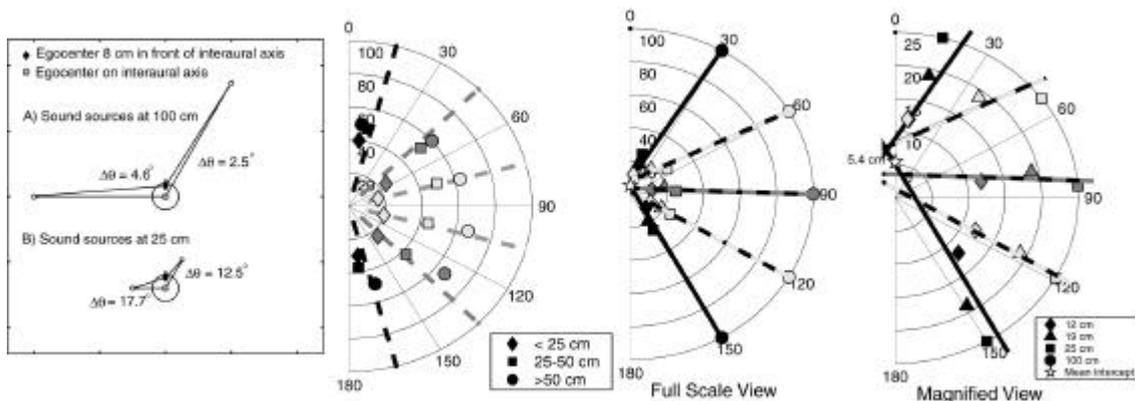


Figure 1: Auditory Parallax Effect for Near and Far sources

Figure 2: Median response biases for free-field localization

Figure 3: Isoazimuth lines for virtual near-far matching experiment

intense at the right ear and/or arrive first at the right ear appear to be located on the right side of the head, and sounds that have no interaural level or time differences appear to be located in the center of the head. In the free field, the only sound locations that produce binaural signals with no interaural time and intensity differences are in the median sagittal plane. If these points are assumed to be in line with the auditory egocenter, then it follows that the auditory egocenter must lie somewhere in the median sagittal plane.

The real question, then, is where on the median sagittal plane the auditory egocenter is located. To this point, little effort has been made to address this question. We believe the reason for this oversight is that the actual location of the auditory egocenter in the median plane is essentially irrelevant when the sound source is located 1 meter or more away from the listener. This is illustrated in the top panel of Figure 1, which shows the effect of a mismatch between the auditory egocenter and the geometric center of the head for 1 m sound sources at 90° and 300° in azimuth. In general, a discrepancy in the locations of the auditory egocenter and the geometric center of the head will lead to a difference between the angles of the sources relative to these two locations: we refer to this difference in angle as an auditory parallax effect and measure its magnitude by the difference between the two angles ($\Delta\theta$). When the source is at 90° , the azimuth θ of the source relative to an egocenter located 8 cm in front of the interaural axis (94.5°) is only about 4.5° greater than the azimuth of the source relative to the geometric center of the head (90°). When the source is located at 300° , the difference in θ between the two egocenters is only about 2.5° . Both of these $\Delta\theta$ values are smaller than the minimum audible change in the angle of a sound source (Mills, 1958). Thus there is no reason to believe that parallax effects due to the location of the auditory egocenter within the head has any meaningful effect on the perception of the relatively distant sound sources that have been used in the vast majority of auditory localization experiments.

When the sound source is near the listener, however, the location of the auditory egocenter within the head may produce much larger parallax effects. This is illustrated in the bottom panel of Figure 1, which shows that the relative angles of sound sources located 25 cm from the head can be shifted more than 10° by a change in the location of the egocenter. It is this region where the location of the auditory egocenter can have an important influence on the spatial perception of sound sources.

Although there is no direct way to determine the location of the auditory egocenter, it should be possible to measure its location indirectly by taking advantage of the auditory parallax effect that occurs for nearby sources. This indirect measurement technique requires the measurement of "isoazimuth" lines defined by loci of points where sound sources at different distances appear to be located in the same direction relative to the listener. By definition, the auditory egocenter should lie at the intersection of all the isoazimuth lines that occur in auditory space. The remainder of this paper describes a series of experiments that have used isoazimuth lines to

triangulate the auditory egocenter. The next section briefly reviews an earlier experiment that examined how well listeners are able to point to the locations of nearby sound sources in the free field. Section 3 describes two experiments that used virtual sound sources to measure “isoazimuth” lines for nearby sources. Section 4 describes an experiment that measured “isoazimuth” lines for nearby sources in the free field. Finally, Section 5 reviews the results of these experiments and attempts to estimate the location of the auditory egocenter within the head.

2. AUDITORY LOCALIZATION OF NEARBY SOURCES

In general, listeners appear to be able to localize the directions of nearby sound sources as well as they can localize the directions of far-field sources. This was shown in an earlier experiment that measured how well listeners could identify the locations of nearby random-amplitude noise bursts by moving a pointer to the perceived location of the sound (Brungart et al., 2000). One interesting outcome of this experiment was that listeners were generally able to distinguish between the isolated increases in interaural level difference (ILD) that occur when a nearby sound source at a fixed azimuth approaches the head and the correlated increases in interaural time delay (ITD) and ILD that occur when a sound source at a fixed distance moves from 0° to 90° in azimuth. This is illustrated in Figure 2, which shows the median azimuth errors for the randomly located sound sources near the horizontal plane (-20° to +20° elevation). The data have been divided into six non-overlapping azimuth bins (centered every 15° from 15° to 75° on the right side of the listener) and three non-overlapping distance bins (<25 cm, 25-50 cm, and >50 cm). Note that the median azimuth errors are shown relative to the midpoints of each bin (shown by the dashed lines) and that the data have been corrected for front-back reversals. Also note that the symbols have been plotted at the median distance of the responses within each bin. The most important feature of this figure is that the median response errors show no signs of the kind of systematic auditory parallax shown in Figure 1- the median azimuth errors were generally no larger for the close stimuli than they were for the far stimuli. This is true despite the large ILDs that occur for nearby stimuli. Apparently listeners were able to interpret the large ILDs associated with nearby sources as a distance cue and avoid becoming confused about the azimuth location of the stimulus.

Although the absence of an auditory parallax in Figure 2 seems to suggest that the auditory egocenter is located near the geometric center of the head, there is really no evidence to support this hypothesis. What the absence of parallax really indicates is that the listeners were able to accurately translate locations from the internal coordinate system they used to encode the apparent positions of nearby sounds into the coordinate system they needed to move a pointer to that apparent location. The near and far points shown in each azimuth bin in Figure 2 were not necessarily perceived in the same direction, but the listeners were able to point to the locations where they heard the sounds without any systematic biases in their responses.

3. VIRTUAL TRIANGULATION EXPERIMENTS

The free-field localization data cannot be used to triangulate the auditory egocenter because it provides no information about the apparent azimuthal locations of the near and far sound sources. Triangulation cannot be achieved without identifying two or more isoazimuth lines comprised of sound sources at different distances that appear to originate from the same angle relative to the listener. This section describes two experiments that used virtual sound sources to identify isoazimuth lines and triangulate the auditory egocenter. In both cases, the stimuli were synthesized from HRTFs that were measured in an anechoic chamber with an acoustic point source. The HRTFs were measured every 1° in azimuth for source locations 12 cm, 19 cm, 25 cm, and 100 cm from the center of the head (Brungart and Rabinowitz, 1999), corrected for the response of the headphones (Sennheiser HD540), and used to generate 251-point linear-phase digital filters at a 44.1 kHz sample rate.

Isoazimuth Lines for Virtual Sound Sources

In the first experiment, listeners were asked to move the location of a nearby virtual sound to match the azimuthal position of a more distant virtual sound. Prior to each trial, the first sound

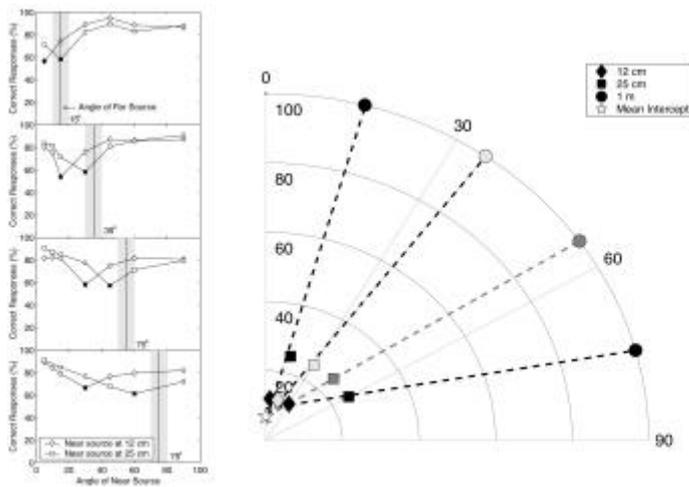


Figure 4: Isoazimuth lines for a virtual cocktail party listening task with a 1 m talker at a fixed azimuth location (shaded regions in the left panel of the figure) and 12 cm or 25 cm talkers at varying azimuth locations. The blackened symbols in the left panel represent near-far source configurations where overall performance was lowest, presumably because the target and masking sources were perceived at the same angle in azimuth. These "isoazimuth" points are plotted by the corresponding symbols in the right panel. The lines in the right panel show the isoazimuth lines that best fit these points, as determined from the 1st principal components. The white star shows the mean intercept of these lines, 6.5 cm in front of the interaural axis.

source (the reference) was set at a distance of 1 m and an angle of 30°, 60°, 90°, 120°, or 150°. The second sound source (the probe) was set at a distance of 12 cm, 19 cm, or 25 cm at the same angle as the reference. Then the listeners were presented with a series of six 250 ms burst of filtered gaussian noise: one burst at 0° and 1 m; two bursts of the reference sound at 1 m; one burst at 0° at the probe distance; and two bursts of the probe sound. They were then asked to either accept the target and probe sounds as matched in azimuth or to move the probe sound left or right by 2° or 10° and listen the stimulus again. This process was repeated until the listeners accepted the pair of target and probe locations as "matched in azimuth". Then the target and probe locations were recorded and the next trial was started.

The task was difficult and time consuming, and it required a great deal of concentration and motivation to perform properly. Participation in the experiment was therefore limited to three investigators from our laboratory who had substantial experience with virtual audio displays and some knowledge about the hypothesis under test. Each of these listeners participated in 100-200 trials collected at random reference angles and probe locations.

The results of this experiment are shown in the polar plots in Figure 3. The symbols in the figure represent the median matched probe locations (diamonds, triangles, and squares) associated with the 1 m reference locations (circles) at each reference angle. The lines show a best linear fit to the data that was calculated by extracting the first principle component for all the matching reference and probe locations at each reference angle (Kistler and Wightman, 1992). Thus each line represents a linear estimate of the locus of points that would be perceived at the same location in azimuth, i.e. an "isoazimuth" line. These isoazimuth lines are clearly influenced by an auditory parallax effect that causes the lines to intersect the median plane roughly 5.4 cm in front of the interaural axis (illustrated by the white star in the figure). Thus, the results of this experiment suggest that the auditory egocenter is located somewhere between the midpoint of the listener's interaural axis and the front of the head.

Isoazimuth Lines That Maximize Speech Interference

Our experiences with the virtual matching experiment indicated that matching the apparent azimuth locations of virtual sound sources by incremental shifts in the azimuth of a probe sound was too onerous a task to be conducted by any but the most intrinsically motivated subjects. However, in the course of conducting a separate, unrelated experiment we discovered a different experimental technique that makes it possible to indirectly measure the relative apparent azimuth locations of two sounds at different distances without requiring any direct judgments about the apparent locations of the stimuli. This technique is based on the Coordinate Response Measure (Brungart, 2001), a speech perception task that presents listeners with stimuli containing two phrases of the form "Ready (call sign) go to (color) (number) now" and requires them to identify the color (red, blue, green, or white) and the number (1-8) contained in the phrase addressed to the target call sign (Baron). In this particular experiment, one of the competing talkers (the far source) was presented within a fixed 10° range of azimuth

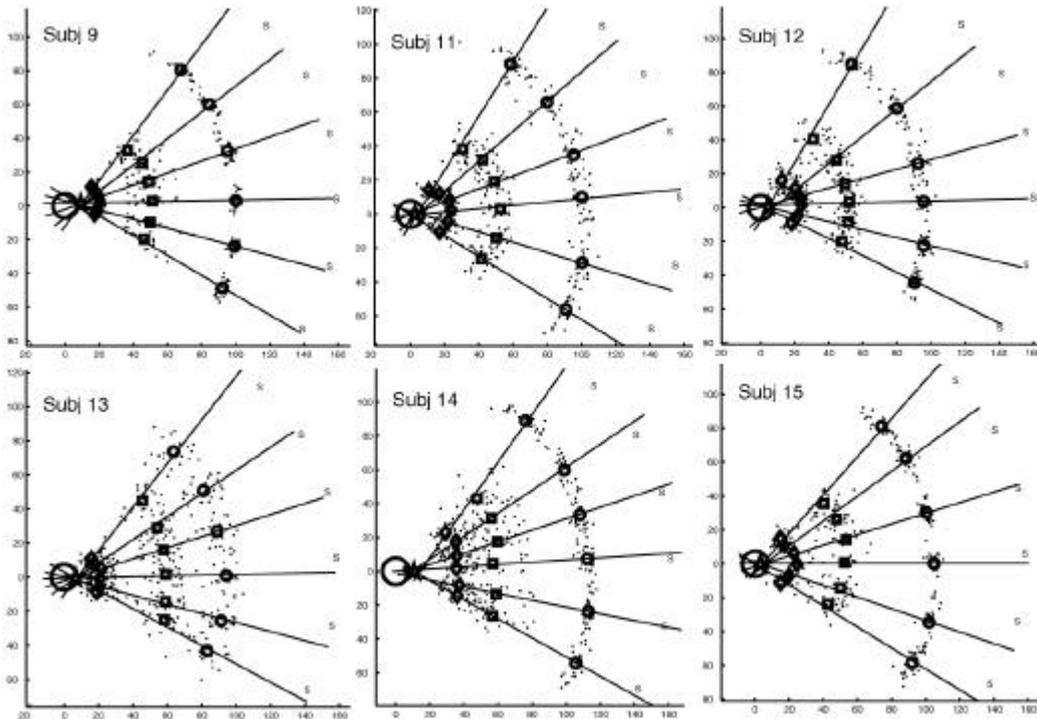


Figure 5: Isoazimuth lines measured for six listeners in the free field

locations at a distance of 1 m (illustrated by the shaded regions in the left panel of Figure 4), and the second competing talker (the near source) was presented at a distance of either 12 cm or 25 cm and at azimuth values ranging from 0° to 90° . Overall performance was measured in terms of the percentages of correct color and number identifications for each near-far source configuration (left panel of Figure 4). The results of this experiment were analyzed under the assumption that the listeners were using differences in apparent direction to help segregate the near and far talkers, and that performance in the task was minimized when the near and far sources were perceived in the same direction relative to the listener. Thus, the isoazimuth lines shown in the right panel of Figure 4 were determined by taking the 12 cm and 25 cm source angles that minimized performance in the speech perception task for each far source angle (shown in black in the left panel of the figure), plotting these 12 cm, 25 cm, and 1 m source locations in polar coordinates, and using principle components analysis to determine the best linear fit for each set of isoazimuth points. The resulting isoazimuth lines show a strong auditory parallax effect and a triangulated auditory egocenter location 6.5 cm in front of the listener's interaural axis. Thus, the results of the second experiment show virtually the same auditory egocenter as the first experiment, despite the use of a different experimental technique, a completely different set of listeners, and a much larger number of trials (at least 280 trials for each data point shown in the Figure 4). It is important to note, however, that both of the virtual experiments were conducted with the same set of KEMAR HRTFs. Thus it is still conceivably possible that the parallax effects seen in Figures 3 and 4 are the result of some artifact in the HRTFs or a mismatch between the KEMAR HRTFs and the individual HRTFs of the listeners. In order to address this issue, a third experiment was conducted that required the listeners to match the azimuth locations of near and far sources in the free field.

4. ISOAZIMUTH LINES FOR FREE-FIELD SOURCES

The third experiment was conducted in a sound-treated listening room. The listeners were blindfolded and seated on a bench where they were asked to immobilize their heads by biting down on a custom-molded bite bar. An arc of six small fixed loudspeakers was located roughly 1.5 m from the listener's head with approximately 15° spacing from -45° to the left to 30° to the right of the median plane. The listeners participated in the experiment while holding a point source that was equipped with an electromagnetic tracking device (Polhemus FasTrak). Prior to running the experiment, the tracking device was used to record the locations of the listener's left and right ear canal openings and the tips of their noses. These locations were used to define a

coordinate system centered at the midpoint of the interaural axis that was used to record the response locations in the experiment (Brungart et al., 2000).

The stimulus in each trial of the experiment consisted of a continuous series of noise tokens that alternated between the hand-held point source and one of the six fixed loudspeakers. The alternating sequence consisted of two 100 ms noise bursts from the far speaker, followed by one 200 ms burst of noise from the near speaker, with 100 ms of silence between each noise burst. The listener's task was to move the hand held point source to a close (4-5 inches from face), intermediate, or far (arm's length) location where it appeared to match the azimuth location of the fixed sound source, and then respond by pressing a footswitch. Each block of trials consisted of 5 repetitions of each of the 6 fixed speaker locations. A total of six paid volunteer subjects participated in the experiment (4 males and 2 females).

The results of the experiment are shown in Figure 5. The symbols show the mean response locations of the close, intermediate, and far matching conditions of the six speaker locations (S's in the figure). The lines represent the linear estimates of the isoazimuth lines extracted from the first principal component of the data for each of the six speaker locations. The white stars represent the mean intersection of the isoazimuth lines.

These results again show evidence of a strong auditory parallax effect. All six of the subjects systematically responded at more medial locations in the "near" matching condition than in the "intermediate" or "far" matching conditions. This caused the isoazimuth lines to consistently converge at a location in front of the listener's interaural axis. The distance in front of the interaural axis ranged from 2.7 cm to 10.6 cm across the six listeners, with an average intercept 6.12 cm in front of the interaural axis. Thus, as in Experiments 1 and 2, the results of Experiment 3 suggest that the auditory egocenter falls somewhere in front of the interaural axis.

5. CONCLUSIONS

This paper has presented the results of three experiments that have attempted to use the triangulation of isoazimuth lines to determine the location of the auditory egocenter within the listener's head. Although there were substantial differences in the techniques used in these three experiments, all three of the experiments suggest the same general conclusion about the location of the auditory egocenter: it is located not at the midpoint of the interaural axis, but roughly 6 cm in front of the interaural axis on the median sagittal plane.

At this point, it is worthwhile to comment on why the auditory egocenter might be at this location. A likely explanation is that the auditory egocenter is related in some way to the location of the eyes. For example, it is possible that people learn to judge the relative locations of near and far sound sources from their previous experiences where they were able to see sound sources at different distances. This might cause them to learn to associate equality in perceived *auditory* azimuth with configurations where the near and far sources are also lined up *visually*. Further research is needed to determine whether the visual modality influences the location of the auditory egocenter for sounds sources located outside the normal field of vision.

REFERENCES

- Blauert, J. (1983). *Spatial Hearing*. Cambridge, MA: MIT Press.
- Brungart, D. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *Journal of the Acoustical Society of America*, **109**, 1101-1109.
- Brungart, D., Durlach, N., & Rabinowitz, W. (1999). Auditory localization of nearby sources. II: Localization of a Broadband Source. *Journal of the Acoustical Society of America*, **106**, 1956-1968.
- Brungart, D. & Rabinowitz, W. (1999). Auditory localization of nearby sources. I: Head-related transfer functions. *Journal of the Acoustical Society of America*, **106**, 1465-1479.
- Brungart, D., Rabinowitz, W., & Durlach, N. (2000). Evaluation of Response Methods for the Localization of Nearby Objects. *Perception and Psychophysics*, **62**, 48-65.
- Brungart, D. & Simpson, B. (2001). Auditory Localization of Nearby Sources in a Virtual Audio Display. In *Proceedings of 2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, October 21-24, 2001, pp. 107-110.
- Howard, I. & Templeton, W. (1966). *Human spatial orientation*. London: John Wiley & Sons.
- Kistler, D. & Wightman, F. (1992). A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction. *Journal of the Acoustical Society of America*, **91**, 1637-1647.
- Mills, A. (1958). On the minimum audible angle. *Journal of the Acoustical Society of America*, **30**, 237-246.