Changes in Spatial Perception Through Headphones

Andria Poiarkoff
MA Audio and Media Technologies
Sonic Arts Research Centre
November 17, 2008

Abstract

Auditory spatial perception is both a physical and psychological phenomenon, varying from person to person. How and where sound is perceived is based on a number of factors including but not limited to loudness, pitch, distance and duration. When played through headphones, what was once a three dimensional sound is said to shift into a lateralized, internalized image. This study examines the change in lateralization and overall spatial perception in five types of headphones ranging from in-ear to open circumaural. The findings show as the headphones progress from internal and closed to external and open lateralization is lost, externalization is gained, and individual sounds change from a point in space to a perceived size, shape, and depth.

Introduction

It is a common belief that “there are large errors in sound position perception associated with headphones, especially for the most important visual direction, out in front…it is very difficult to externalize sounds and avoid the inside-the-head sensation.” (Kyriakakis, 1998)

When discussing spatial perception of a headphone listener it is said that “the position of the image is located to the left or right as expected…but the image seems to be within the listener’s head – it is not perceived to be in the real external world.” (Hartmann, 1999)

However, to state these ideas as fact requires one to assume that only one type of headphone exists, or that all headphone types create the same spatial image for an individual listener, all of which are lateralized and internalized. As there are many different types of headphones on the market, this study was designed to explore whether or not spatial image, lateralization, and internalization changes with different types of headphones.

In order to determine if listeners’ spatial perception changes with different headphones, an experiment was devised to record individuals’ images on both horizontal and vertical planes. These images will be combined into composite images for both planes on each set of headphones. The final composite images should provide evidence for or against consistent lateralization and internalization across different headphone types.

Methodology

Design

Participants were tested individually.

For the purpose of this experiment, headphones are the independent variable.
Participants were provided with 5 different headphones:
In-ear
Ear buds
Supra-aural
Circumaural (closed)
Circumaural (open)

The dependent variable is the listeners' individual spatial images. This is subjective, and different for each listener.

Participants were given two audio tracks as stimuli.

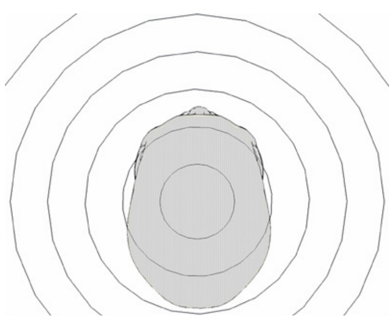
Participants

Participants were aged 21-30. The study group consisted of 6 subjects, 4 male and 2 female. All participants have a background in audio and music. Participation was voluntary. No other characteristics were found that would have an impact on the results.

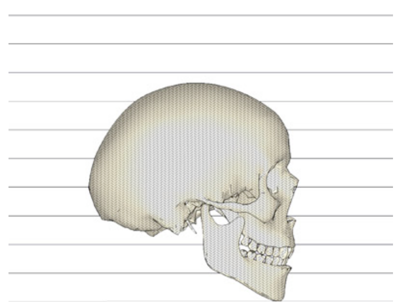
Materials

Participants were given a packet of diagrams on which they were to draw their perceived spatial image. Each image was labeled as to what headphones and track were to be used. Two different diagrams were used, one for horizontal images and one for vertical images.

Horizontal



Vertical



A color-coded key was provided along with corresponding drawing utensils. For the horizontal diagram color-coding was done according to instrument, and for the vertical diagram according to frequency.

Instrument	Color
Kick drum	Violet
Snare drum	Blue
Cymbals	Orange
Vocals	Pink
Chimes	Red
Left Guitar	Green
Right Guitar	Aqua

Frequency	Color
High	Red
Mid	Yellow
Low	Green

Two 24-bit audio tracks were used as stimuli.
The first track was section of music chosen based on a wide horizontal spread. The 40-second track was cut from a multitrack recording created by the author for the specific reason that the actual level and degree of panning for each instrument is readily available.
The second track was a 1 minute section of an a cappella song by the artist Imogen Heap. This track was used to judge vertical spread, which is in itself psychological rather than physical. Because of this, the creation of a track specifically for the purpose of judging vertical spread is not a practical science. Instead, this track was chosen by the consensus of multiple listeners as being vertically perceptible.

Participants were given five different headphones.
To limit speculation on engineering differences between brands, all headphones were sourced from the same manufacturer.

Type	Model
In-ear	Sennheiser CX300
Ear bud	Sennheiser MX460
Supra-aural	Sennheiser HD435
Circumaural (closed)	Sennheiser HD280
Circumaural (open)	Sennheiser HD595

All tests utilized an iMac running OS X 10.4. Audio files were played using iTunes and were run through a Digidesign Mbox2.

Procedure

Participants were taken individually to a workstation arranged for the experiment. They were given a packet that included written instructions and 10 diagrams: 5 horizontal and 5 vertical. Each diagram included instructions as to which headphones to use and which track to play.

The headphones were made available on a table next to the testing area.

Examples of spatial images were provided when necessary for clarification. Participants were asked to indicate not only the location of the sounds, but also the perceived size and shape (if such a thing was perceived). These were to be drawn on the provided diagrams.

<u>Example run:</u>

Instructions are read
Examples are given if needed
Subject turns page to first diagram
Diagram reads: In-ear headphones, Track 1
Subject listens to Track 1 using In-ear headphones
Subject uses the color code provided to draw their perceived image
Subject progresses to the next diagram
Diagram reads: In-ear headphones, Track 2
Subject listens to Track 2 using In-ear headphones
Subject uses the color code provided to draw their perceived image

This is repeated for each set of headphones for a total of 10 diagrams per participant. Subjects were permitted to repeat tracks as necessary, and to adjust the volume if needed.

Results

The composite diagrams of perceived horizontal image can be found in Figure 1, a-e on the next page.

In Figure 1a (in-ear), the instruments were predominantly perceived as being located inside the head, spanning from ear-to-ear across the forehead. A few less internalized sounds were heard directly to the left and right of the respective ears. The majority of the sounds are focused in the area of the temples. In Figure 1b (ear buds) one can see that the sounds have become slightly less internalized, but still lateralized. The sounds are perceived once again around the forehead, however the images are slightly forward of the forehead rather than encompassed by it. The sound is no longer focused in the temples, but rather is pushed to the front of the head and next to the ears, creating a gap. Figure 1c (supra-aural) shows a transition into predominately externalized sound. While all subjects reported hearing the kick drum at some point between the front and back of the head, most other sounds were heard outside the skull. This image also distinctly shows that the sound is now heard on all sides of the head rather than being focused toward the front. Figure 1d (circumaural closed) provides an image that is not so different spatially from Figure 1c, however now instead of drawing simple ellipses to represent the sounds, as was done by the majority of subjects up to this point, this Figure shows more asymmetrical distinct shapes. Figure 1e (circumaural open) deals with this same phenomenon, only spread wider and more evenly around the skull.

Horizontal

Vertical

Figure 1

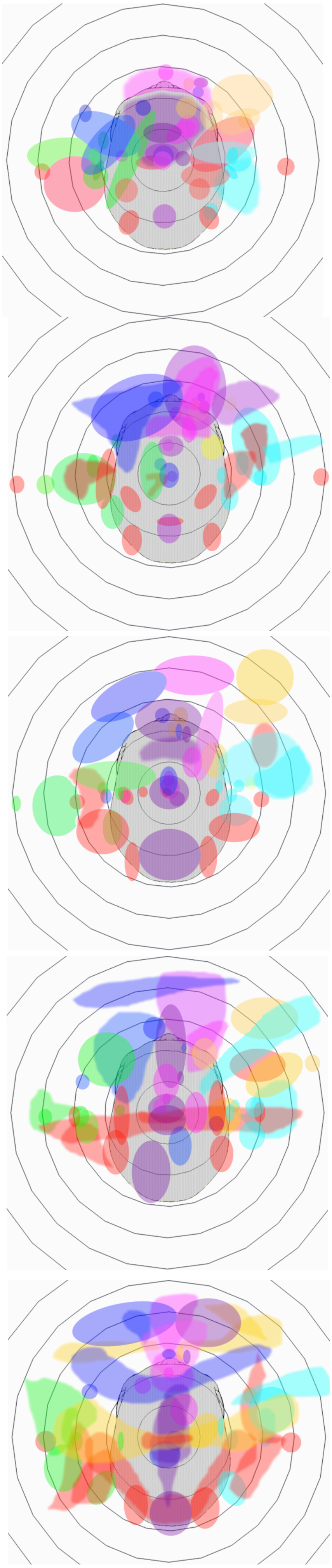
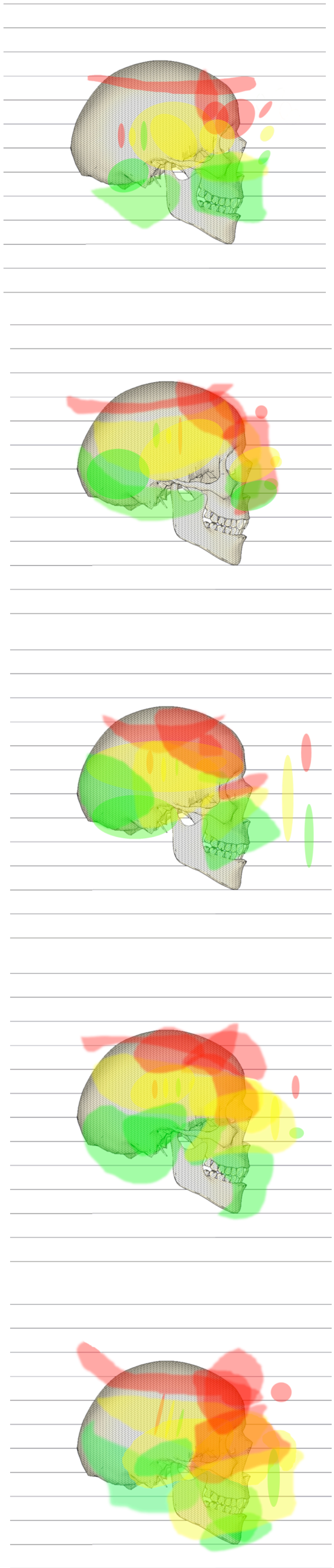


Figure 2



a) In-Ear *Sennheiser CX300*

b) Ear Bud *Sennheiser MX460*

c) Supra-Aural *Sennheiser HD435*

d) Circumaural Closed *Sennheiser HD280*

e) Circumaural Open *Sennheiser HD595*

The composite diagrams of perceived vertical image can be found in Figure 2, a-e.

In every case but one, the subjects perceived high, mid, and low frequencies as a vertical image. These subjects all represented the frequencies as having the high frequencies above the low frequencies with the mid frequencies somewhere in the middle, often overlapping.

Figure 2a (in ear) shows the frequencies spanning from above the eyes to the back of the neck and down to the edge of the upper jaw. Figure 2b (ear buds) shows a similar image, but raised above the jaw in the front and ending at the base of the skull in the rear. In Figure 2c (supra-aural), the frequencies have expanded to fill the skull almost perfectly, with the highs perceived as radiating from behind the forehead and lows from either the bottom rear of the skull or the lower jaw. The mid frequencies fill the space between. Figure 2d (circumaural closed) provides a similar image, except the high and low frequencies have now started to cross to the outside of the skull, extending above the front and below the rear of the head. In Figure 2e (circumaural open) these frequencies extend beyond 2d to cover the entire face and base of the skull.

Discussion

The results of this study serve to support the idea that spatial perception in different types of headphones is **not** uniform, lateral, or internal. As the headphones progress from internal and closed to external and open, lateralization is lost, externalization is gained, and individual sounds change from a point in space to a perceived size, shape, and depth.

As can be seen from the diagrams, listening to the same tracks through different headphones will dramatically alter the listeners' perceived spatial image. It is also clear that even with in-ear headphones most listeners will perceive sound to be outside the head, especially on the horizontal plane. The vertical spread, though entirely psychological, still supports the theory: by the time the listener has progressed through the series of headphones, the spatial image has extended through and beyond the skull in every direction, changing greatly along the way. The assumption that headphones in general create a lateralized, internalized sound is unfounded. It would be more correct to say that headphones, rather than representing space as it is, create a sphere of perception extending through and beyond the skull; and for each listener, stimuli, and headphone type the size and position of that sphere will vary.

An exciting and unexpected result is the perceived shift from shapeless orbs of sound into defined spatial images when listening on circumaural headphones. In Figure 1, it can be seen that the supra-aural headphones provide some sense of shape, however comparing 1c to 1d and 1e shows that the subjects perceived a much greater expanse of shape and depth with the circumaural headphones. Subjects reported these headphones as having similar horizontal spread but a fuller, richer sound that the preceding headphones. This would correlate to the large asymmetric shapes and sweeping colors in the last two diagrams of Figure 1.

These findings may be attributed to the fact that different headphones involve the ear in different ways. In-ear headphones are only influenced by the ear canal and inner ear, while circumaural headphones are involving the mechanics of the pinnae and even a small area of the head surrounding the ear. This may cause the sound to be perceived as being located outside of the skull rather than reflecting from the inside.

Developing this theory further could involve experiments to determine perception of specific frequencies and levels for each headphone design as well as any differences within the mechanics of the headphones themselves. Different stimuli would present with different results, and so experimenting with other forms of controlled stimuli would be beneficial. It may also be worthwhile to test subjects with no prior audio or music background, as they may be less accustomed to focused listening and may have less perception of a spatial image.

This study has only started to explore the phenomenon of spatial perception in headphones, but serves to support the idea that most common assumptions of perception in headphones is unfounded, and that more research needs to be done in this area. Currently there are numerous studies theorizing solutions to fix the problem of lateralization and internalization in headphones. However, after the results of this study, this does not seem as if it should be a matter of fixing, but a matter of exploring and enhancing what perception and externalization is already present.

Annotated Bibliography

Begault, D.R., 1990, The Composition of Auditory Space: Recent Developments in Headphone Music, *Leonardo*, 23(1), pp. 45-52.

This article focuses primarily on a program developed to enhance spatial perception in headphones. Within the details of this program are excellent descriptions of how current headphones try to create space, as well as how that space is interpreted by the human auditory system. In addition, there are detailed descriptions of the perceptual anomalies created by the use of headphones, such as front-image distortion and spatial inversion.

Carlile, S., 2006, LISTENING TO THE WORLD AROUND US, *SPECIAL ISSUE: MECHANISMS OF HEARING DAMAGE: Spatial perception and masking* 34(3-1).

This article looks at the physical/biological aspects of spatial perception. It contains well-worded explanations about how the human auditory system processes sound, and what characteristics of sound are factored into localization. There is a section explaining how cues and focus can contribute to the individual interpretation of sound location. One of the best features of this work is a paragraph dedicated to explaining the principle dimensions of sound, including direction, distance, and spaciousness. There are also numerous useful diagrams.

Hartmann, W.M., 1999, How we localize sound, *Physics Today*, 52(11), pp. 24-9.

This article focuses on human auditory localization. It details how our perception works biologically and proceeds to discuss how technology is used to test this perception, as well as the limitations of both our bodies and our technology. There are detailed descriptions of spatial imaging within headphones as well as how they created the phenomenon of lateralization, and how this relates to overall localization. In addition, this particular article contains excellent diagrams.

Kyriakakis, C., 1998, Fundamental and technological limitations of immersive audio systems, *Proceedings of the IEEE*, 86(5), pp. 941-51.

This article explores different methods of creating 3D audio, as well as ways to potentially improve these systems. While a majority of the writing focuses on loudspeakers, it also goes into great detail pertaining to 3D audio in headphones and the limitations in current technology. There is great emphasis on the fact that 3D sound, no matter how technologically advanced, will always be at the mercy of individual perception.

Liitola T. *Headphone Sound Externalization* (2006).

This thesis focuses on improving localization in headphones. This work is interesting as it focuses very specifically on the aspect of out-of-head localization (externalization). Lateralization is seen as something not to be improved upon, but to be completely eradicated. The author uses different techniques in an attempt to externalize headphone sound, thereby enhancing localization.

Okamoto, M., Kinoshita, I., Aoki, S. & Matsui, H., 1997, Sound image rendering system for headphones, *Consumer Electronics, IEEE Transactions on*, 43(3), pp. 689-93.

This document details a system for creating realistic 3D sound through headphones. There are a few excellent discussions on the flaws in current headphone systems, as well as ideas as to how to remedy these problems. The most useful aspect of this article is the detail pertaining to the experimental design. The methods of executing the experiments as well as the methods of recording information on individual perception are very helpful, as they take complex tasks and form them into simple, reliable processes. This is an excellent reference for a basis in headphone-based perceptive experiments.

Wang, Z. & Ben-Arie, J., 1996, Conveying visual information with spatial auditory patterns, *Speech and Audio Processing, IEEE Transactions on*, 4(6), pp. 446-55.

This paper details an experiment in spatial perception. The experiment involves taking a visual image, representing it in audio, then having the listeners attempt to spatially re-create the image. There are very interesting results dealing with bandwidth and sound color in relation to perception of elevation. All tests were done through headphones in an isolated environment. Discussions of relationships between visual and auditory spatial perception, as well as azimuth vs. elevation localization, are useful in exploring auditory localization via headphones.