Audio Engineering Society

Convention e-Brief 535

Presented at the 147th Convention
2019 October 16–19, New York, USA

Comparing Externalization Between the Neumann KU100 Versus Low Cost DIY Binaural Dummy Head

Kelley Jayne DiPasquale

SUNY Potsdam, Crane School of Music, Potsdam New York, 13676, USA

Correspondence should be addressed to Kelley Jayne DiPasquale (dipasqkj202@potsdam.edu)

ABSTRACT

Music is usually recorded using traditional microphone techniques. With technology continually advancing, binaural recording has become more popular, that is, a recording where two microphones are used to create a three-dimensional stereo image. Commercially available binaural heads are prohibitively expensive and not practical for use in typical educational environments or for casual use in a home studio. This experiment consisted of gathering recorded stimuli with a homemade binaural head and the Neumann KU 100. The recordings were played back for 34 subjects instructed to rate the level of externalization for each example. The study investigates whether a homemade binaural head made for under $500 can externalize sound as well as a commercially available binaural head the Neumann KU 100.

1 INTRODUCTION

Binaural sound recording techniques are used to create an immersive listening experience. When listening with headphones, typical stereo microphone techniques create a product that places the sound inside the listeners head, creating a one-dimensional experience. Binaural recording attempts to record sounds as they would be heard in real life and perceived as outside the head. This process is intended to create a three-dimensional experience.

Commercial binaural heads are prohibitively expensive for many music programs and home studios making the technology inaccessible for casual or educational use. The average cost of the most popular model, the Neumann KU100, is $8,500. By building a model with less expensive materials and maintaining comparable quality this technology may be more widely available. A homemade binaural head will be created for less than $500. By placing this monetary limit on materials, the goal is to explore the effectiveness and sound quality of an inexpensive model versus a professional model. The effectiveness and sound quality will be measured by evaluating the level of externalization in each model.

2 LITERATURE REVIEW

Although there has been a recent resurgence in the use of binaural sound and recording due to heavy use of headphones in recent years, this technology is approximately 140 years old. Binaural audio predates stereophonic sound and was used and researched until the creation of stereo recordings. Physically, as humans process sound binaurally the sound is directed by the pinna of the ears and funneled into the ear canal to the ear drum. Various high and low frequency signals are processed as they vibrate on the basilar membrane and are processed through the brain and nervous system. The physicality of sounds being impacted by the ears and face is called Head Related Transfer Function (HRTF) [1]

The ears and brain also interpret interaural time differences and interaural intensity differences. Both cues aide in the brain determining where a sound is coming from. Interaural time differences are measured by whether the sound reaches the left or

right ear first. Interaural intensity differences measure the volume of the sound when it reaches the ear. The face creates a sound shadow allowing the sound to reach one ear before the other and allowing humans to determine which direction the sound is coming from. [2] These two interaural cues, combined with the spectral cues created by the ears, head and torso, allow humans to decipher what direction and where the sound originates. [3], [4]

In a binaural recording head, the ears and microphones in the ears attempt to recreate this process. The ears in a binaural recording head direct and localize sounds into small diaphragm microphones that mimic the function of the ear drums. The physical features of the mannequin attempt to recreate the physical spectral cues. The distance between the ears on the mannequin head recreate the interaural time difference cues. This entire process aims to externalize the sound, which means that the sound is perceived to be outside the head as if listening live, while using headphones. [5]

3 METHODOLOGY

The methodology required for this study involved the building of a binaural head formally named Aurelio. (see alternate paper to be named), running tests at McGill University using a Neumann KU100 head and comparing those results to Aurelio in a listening test. Both heads were taken into an anechoic chamber and a series of prerecorded sounds were played out of speakers at 0, 45, 135 and 180 degrees. The prerecorded sounds included pink noise, a neighborhood ambiance, two examples of spoken word, a car driving by, a recording sample of La Traviata. One live recording was included, which was a rehearsal of Gustav Mahler's 5th Symphony performed by the McGill Symphony Orchestra. The goal was to analyze the externalization of each recording and attempt to correct any differences between Aurelio and Neumann KU100.

After this initial test it was clear that the high frequencies of Aurelio were very prominent in all the recordings. At that point a spectrogram was performed, and a new EQ was applied to Aurelio which involved emphasizing the bass frequencies, and the result was a comparable sound and spectrum to the Neumann KU100. The specifics of the equalizer are a low frequency shelf boost of 6.4dB at 492.2Hz and a high mid frequency cut of 9.4dB at 4kHz and a Q of 4.

The listening test involved compiling auditory stimuli from the preliminary experiment conducted at McGill University. The tests that were used are neighborhood ambiance at 0 and 135 degrees, pink noise at 0 degrees, drive by car at 180 degrees, and a clip from Mahler's 5th Symphony performed by the McGill Symphony Orchestra at 0 degrees. The test also included a completely internalized version of each sound example to measure the level of understanding in each participant. The internalized examples were created by panning each of the original sound files center. The object of using different angles of the stimuli was to see if direction also impacts the level of externalization. The order of which examples were played was determined by a coin flip, and completely random. All the sound examples were also loudness matched so that no stimuli had an advantage just due to the loudness level.

The participants were given verbal and written instructions as well as a written evaluation sheet, in which the participants rated the level of externalization on a scale from 1-5. 1 indicating complete internalization and 5 indicating complete externalization. Prior to the test the participants were given a briefing which explained and demonstrated how to listen for externalization versus internalization. Sound clips were played and explained to the participants which aided in the explanation. After the briefing, each trial was played two times, consecutively, and each participant rated each example. Each participant was given the same headphones for the experiment and could adjust the volume on the headphone amplifier, if needed.

The test was conducted in four different sessions. Two sessions were conducted in a classroom setting, one session in an open area that had some extraneous noise and one session was conducted in a closed studio setting. Each session used the same equipment and recordings, played in the same randomized order played twice each.

4 ANALYSIS

The data analyzed was compiled from paper forms completed by a total of 34 participants, which consisted of mainly undergraduate students and two university professors. This data was compiled and organized into a spreadsheet and several averaging formulas applied. Upon closer study of the data it was observed that many of the participants rated the mono example, which was completely internalized, as higher than both the Neumann KU 100 and Aurelio. After checking that there were not any issues with the listening examples, it was decided that these participants should be eliminated from the study. The justification being that these participants did not understand the assignment. From the original 34, 18 valid participants remained. To reiterate, the mono example had to be rated higher than both heads in order to be disqualified.

This data also suggested that students have not been exposed to immersive styles of listening and that the briefing may not have been enough for the participants to fully understand the listening task.

The following chart shows the externalization ratings from the 18 valid participants based on the type of stimuli played:

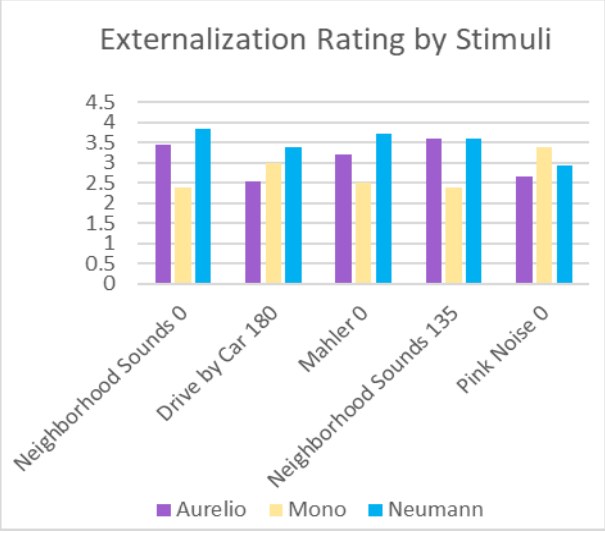

Figure 1. Externalization Rating by Stimuli of Aurelio, Mono and Neumann KU100

This figure illustrates that the closest ratings between Aurelio and the Neumann KU 100 were of the neighborhood sounds test. Both rated higher than the internalized example and Aurelio tying in externalization at 135 degrees and within .4 at 0 degrees. This graphic also shows that Aurelio did not rate well with the drive by car and Mahler example. The pink noise example is the outlier in that most participants rated the mono clip as externalized above both Aurelio and the Neumann KU100.

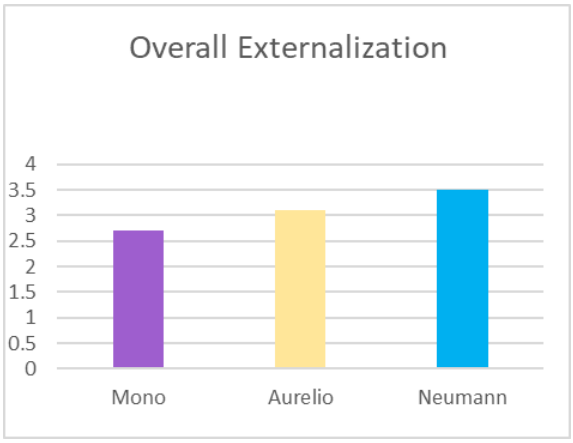

Figure 2. Overall Externalization across all stimuli of Aurelio, Mono and Neumann KU100

Illustrated in figure 2 is the overall externalization ratings from the 18 participants averaged among all the stimuli. The mono example rated the lowest and that was to be expected as those examples were completely internalized. The overall rating of Aurelio was slightly lower than the Neumann KU100. The Neumann KU100 was rated 3.5 by the 18 participants and Aurelio rated a 3.1. This data suggests that Aurelio is about halfway externalized with certain sound stimuli like neighborhood ambiances being more convincing than music recordings when compared to the Neumann KU100

5 DISCUSSION

There were some potential additional variables noticed when conducting this experiment. The listening environments had to be changed, therefore they were not perfectly controlled between each test. The level of noise in each environment potentially

distracted participants and could have skewed results. An example is, during the neighborhood ambiance example one of the listening environments had a lot of extraneous talking and noise. These external sounds may have been confused with the sounds being played back over the headphones. Therefore, the internalized examples may have seemed externalized and skewed the results.

The other variable that wasn't controlled for is the participants listening to the examples and changing their initial responses. Upon analyzing the participants sheets, many participants had crossed off and changed values. This potentially indicates that the listeners were comparing each example, and it favored one head over the other. The intention was that the participants compared each example to the briefing example, which was an example of optimal externalization. To correct this, the examples in each trial should have been isolated with the briefing example played between each example.

While Aurelio isn't completely equal to the Neumann KU100 it does achieve comparable results for significantly less money to build. It was anticipated that participants might have a difficult time adjusting to this type of listening, but it was surprising how much of the data pool had to be eliminated for that reason.

Many participants had many questions after the listening, and this was the most encouraging part of the experiment. Students, and faculty had a real interest in this field, and it shows that this kind of research is worthwhile.

These results can be applied to the world by making this technology more widely available for both educational and casual use in a home studio. There is value in building a binaural head, testing, and using a binaural head to educate musicians about the possibilities of spatial hearing and exposing young learners to the possibilities of three-dimensional sound. As listening technology moves toward immersive listening experiences it is valuable to expose students to these technologies at a younger age. By making things like binaural sound recording heads less expensive it allows for educational programs to experiment using this technology with their students, and for use in home studios.

This result fits into the pool of existing data. No new information was gathered regarding how humans localize sound or spatial hearing. The new information collected is that a binaural recording head can be created on a limited budget and it is effective in demonstrating externalization.

The next step in the research would be to continue exploring ways to easily and cost effectively make individualized binaural heads for each listener, so that they can be more widely used and explored. Individualized ears are the most important aspect of customization that should be and is currently being researched. Currently, molds of participant ears have been created to explore what the most easy and effective ways to create individualized binaural heads. Aurelio currently has been fitted with a set of audiology ears and a set of ears that have been specifically molded from an actual listeners ear.

It would also be worthwhile to conduct the experiment in a more controlled setting with more detailed explanations and examples of internalization versus externalization. This would lead to eventually incorporating and comparing recordings made with individualized ears versus neutral sets of ears.

6 CONCLUSION

To conclude, the research of creating and using binaural sound recording heads is a worthwhile endeavor. Although the homemade binaural head did not perform to the exact level of the Neumann KU100 it demonstrates the technology well, especially for educational or casual home use. Using the homemade binaural head to record ambient soundscapes, or scenes as demonstrated in the neighborhood sounds example of the experiment would be particularly relevant as it had the same score as the Neumann KU100. The hypothesis was proven as Aurelio does demonstrate externalization almost as well as the Neumann KU100 with a significantly lower price tag. Hopefully the impact of this research is that more people will bring this technology into their classrooms and home studios

and immersive sound technology, and binaural sound recording are more accessible to a wider audience.

7 ACKNOWLEDGEMENTS

This work was supported by generous funding from the Crane Institute for Music Business & Entrepreneurship (https://www.potsdam.edu/academics/crane-school-music/departments-programs/music-business-program-crane), as well as The Crane School of Music (https://www.potsdam.edu/academics/crane-school-music). Additional support and supervision provided by Dr. Michael Sitton, Dean, Crane School of Music; Professor Marc Ernesti, Executive Director, Crane Institute for Music Business & Entrepreneurship, Crane School of Music; Professor Denis Martin, Adjunct Instructor, Crane School of Music; and Dr. Douglas McKinnie, Audio & Visual Engineer, Crane School of Music

REFERENCES

[1] Moore, Brian C.J. "Controversies and Mysteries in Spatial Hearing." Audio Engineering Society, Audio Engineering Society, 1 Mar. 1999, www.aes.org/e-lib/browse.cfm?elib=8031.

[2] Rayleigh, John William Strutt. The Theory of Sound in 2 Vol. Dover Publ., 1945.

[3] "Chapter 5." The Science of Sound, by Thomas D. Rossing et al., Addison-Wesley, 2003, pp. 80–90.

[4] Blauert, Jens. "Chapter 2." Spatial Hearing: the Psychophysics of Human Sound Localization, by Jens Blauert, MIT Press, 1983, pp. 50–67.

[5] Stuart, Robert J., and Peter J Craven. "A Hierarchical Approach for Audio Capture, Archive, and Distribution." Journal of the Audio Engineering Society, Audio Engineering Society, 3 May 2019, www.aes.org/e-lib/browse.cfm?elib=20456.