# EFFECTS OF SOURCE COLOURATION ON DISCRIMINATION OF FRONTWARD VERSUS REARWARD MOTION OF VIRTUAL AUDITORY IMAGES

*Sally-anne Kellaway 4300029204*

General Elective 1, ARCH9093, Semester 2 2013
Masters Program in Audio and Acoustics
Faculty of Architecture, Design and Planning, The University of Sydney

*HRTFs were recorded for 30 participants, and used to explore the impact of source spectra on the listener's ability to localise front versus rear virtual auditory images presented through the AKG KM1000 Earspeakers. With the aim of investigating the impact of formants from selected vocal vowel sounds on the HRTF based spectral cues of specifically selected front and rear locations, the first stage of this project was to explore through analysis in MATLAB, which was followed by the listening experiment as a second stage. The first stage uncovered parallels between the formants of the source spectra and directionally dependant HRTF based spectral cues, for which spectral interactions delivered evidence of enhancement and confusion of localisation from the listening experiment.*

## 1. INTRODUCTION

Hebrank and Wright (1974) touch on a consensus in their pioneering research that when studying the role of spectral cues in directional hearing, white noise is the best "vessel". From the earliest studies (e.g., (Butler, 1968), (Blauert, 1969)), and even up to current listening experiments, the practice of using short bursts of white noise has been the norm. Without explicitly discussing the reasons for these decisions in detail, the authors of these studies have implied something very specific - that the spectra of the source material will interact with and alter the spatial perception of virtual auditory sources in tests with human listeners.

Investigation of the spectral cues that determine source direction for virtual auditory display has a history rich with significant findings. The specific spectral cues for front-back directional distinctions are well documented (see, for example, Blauert, 1997; Wightman and Kistler, 1997; Zahorik, Bangayan, Sundareswaran, Wang & Tam, 2006; Sunder, et. al., 2012). These studies reach a consensus on the importance of source spectra in 3-7 kHz range. Furthermore, there is extensive consensus that there lies a significant challenge in achieving consistent and accurate listener identification of front and rear sources, when listening via headphones excluding the benefit achieved from active headtracking (Zahorik, Bangayan, Sundareswaran, Wang, & Tam, 2006) (Sunder, 2012), (Begault, 1990), (Kelly, 2003). Recognizing the unacceptably high rate of front-back reversals, each author presents solutions to this challenge; however, considering the impracticality of individualized Head Related Transfer Function (HRTF) measurement and head tracking (when considering the mass appeal of the application of such technology such as Virtual Reality Video Gaming), there exists the need to look beyond these high-requirement solutions, and to offer a solution separated from these spectra-based and individualized options.

This paper represents an investigation in to achieving consistent front-rear localisation, with specific consideration to the interaction between HRTF-based spectra and the spectra of the sources. This was achieved by exploring key features in the spectra of both recorded Head Related Impulse Responses (HRIRs) and selected source files.

This research project had two stages. The first stage was strictly exploratory, and makes use of the CIPIC HRIR database and individually recorded HRIRs from the University of Sydney (UoS) as tools to examine a large group of individual HRTF recordings for cues in both the frequency (spectra) and time (phase delay) domains for significant deviances from mean recorded results.

This exploration seeks to characterise cues that are both derived from and applicable to average listeners. Zahorik et al. discuss the vast range of variation between individual HRTF sets and go on to support the analysis of wider groups of individuals to arrive at the concept of an average listener (2006). Hence, all analysis is replicated across the UoS and CIPIC databases.

The analysis of the CIPIC database considers a 35 degree range of azimuth and 28 degree range of elevation in both frontal and rear hemispheres, while maintaining focus on the front-rear differences on each sagittal plane. This is to capture key magnitude and phase delay differences in truly lateralised sources. Sampling a range of azimuths between 35 and 60 degrees avoids the range near the median plane within which auditory spatial images typically have shown a “collapsed” or “inside the head” (non-externalised) result in previous binaural listening experiments (Toole, 1991). Characterising the properties of this phenomenon lies outside of the limits of this study.

The second stage of this research project was to implement a listening experiment to investigate the effect of source spectra on accurate front and rear localisation of virtual auditory images, and present the results of the listening experiment.

Martens (1987) describes the benefit of identifying the spectral cues for directional hearing as key to understanding how the auditory system evaluates the complex spectral profile presented to two ears. However, Martens wasn’t the first to see benefit in identifying spectral cues to the quest of understanding localisation for humans. There exists a so-called “cone of confusion” upon which Interaural Time Differences (ITD) are the same for sound sources located in many different directions, and yet the directions of those sources can be discriminated by human listeners, purportedly via the spectral differences that are exhibited in the HRTFs measured on that cone.

Blauert (1969) described this phenomenon more adequately as a hyperbola rather than a cone after initial research to attempt to explain the ability to localise without useful ITD cues. Blauert (1969) conducted a series of experiments using sources positioned within the median plane to identify spectral “directional bands” within which the binaural signals enabled the discrimination of the direction of Frontward, Rearward and Elevated (above ear-level) sources across a wide pool of participants. These results indicated that there is another layer of information that humans draw on to determine source direction when both azimuth and elevation varies, which triggered a wave of studies over the following decades that incrementally probed what cues humans could be drawing from within the “cone of confusion”. Comparison of Blauerts’ and Martens’ findings can start to paint a picture of where spectral cues may have a significant role.

The data that Blauert obtains is strongly echoed through Martens’ 1987 research. Even considering just the Front and Rear cues which Blauert reports, there is very specific agreement between the Front Directional Band cue of 3-6kHz, the Rear cue at 1kHz and Front between 10-13kHz. Within these bands, Martens HRTF energy plots (p. 276) also demonstrate the dominance of these cues, clearly showing that when these cues are active, the opposing coronal plane spectra is significantly lacking that specific spectral range.

Initial spectral analysis of the source files used for this research project indicated that there is the potential for significant interplay between the spectra of the selected source files and frontal and rear HRTF spectra. In particular, it was the formant frequencies of the three vowel sounds selected that were anticipated to interact with the HRTF spectra. In effect, these source-spectral features were hypothesized to exert an influence on a listener’s interpretations of spectral cues present in the signals presented to their ears via headphones.

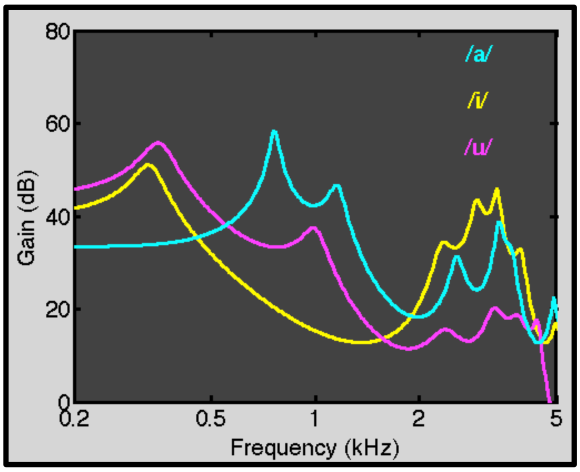


Figure 1 - Spectra of the three selected vowel source sounds graphed as Magnitude. Each peak forms a formant of the sound.

As demonstrated in Figure 1, magnitude differences of 20-40 dB that represent the formants of the selected source sounds are hypothesised to interact heavily with the HRTFs presented to the listeners. Depending on spectral range of the formants, it is predicted the interactions may cause a significant range of influences on localisation, ranging from a confusion or inability to distinguish frontal images from rearward images – this may occur when the two are opposed (a peak with a notch) – to reversals of localisation or (conversely) where cues are in agreement – to what is assumed to result in extremely high proportion of correct responses due to the exaggerated cue.

These predictions can be made when combining this knowledge of the source spectra with the Butler and Rogers’ proposition of Covert Peak Areas (1992). CPAs are specific spatial areas where narrow spectral bands (in practise) demonstrate recorded magnitude levels near maximum magnitude levels. Butler and Rogers delve deeply into extracting CPAs for the frontal hemisphere, while also identifying the elevation implied by the 1kHz wide noise bands used within their experiments. Regardless of the limitations represented by the limited areas for which CPAs are characterised, this effect can be extrapolated to apply to this project by considering that the extreme peaks that the selected sound sources incite will impart an inherent spatial nature to these unprocessed sounds.

Source sounds are limited to these three vowels as they demonstrate high directional tendencies from initial investigations. It is also apparent that the formats, which will be discussed in further detail, lie well within the normal frequency range of the AKG K1000, which are the Earspeakers that were selected for the listening experiments.

To ensure that the there results are attributable to either source spectra or HRTF spectra, the frequency range of the analysis is limited to the range between 200Hz and 5,000Hz. Including the range above 5,000Hz would introduce directional cues outside of the range of the main formants of the source and hence may complicate analysis. To ensure that full analysis of the HRTF could be executed, recordings of HRTFs were captured between 200Hz and 20,000kHz.

## 2. AIMS

At the conclusion of this exploration, this project hopes to explore and demonstrate the influence of the spectra of source sounds on front and rear virtual auditory image localisation. More specifically, using selected source sounds, this exploration will determine specific interactions that impart a directional tendency on source material. A listening experiment has been designed to expose the perceived motion between both the same and different source sounds, and the same HRTF-based cues and different HRTF-based cues. Analysing the results of this listening experiment and comparing where listeners are most able to correctly identify source movement from between front and rear locations with the spectral analysis of both HRTF spectra and source spectra will reveal the reasons for the performance, and will hopefully inform of specific directional cues that can be improved in future studies.

This project will also form as the collection of data, and provide direction for the commencement of analysis to determine directionally dependent phase delay features, such as amounts of phase delay in characterised spatial cue ranged, and the frequency of inflections in the phase-delay curves. It is assumed that such features potentially allow selected source material to be localized as either front or rear across a wide range of listeners, regardless of the spectral content of the source material. Developing an understanding of the role that phase delay plays in the auditory system's ability to localise sounds may provide the tools to negating the impact of source colouration on localisation on virtual auditory images.

## 3. METHOD

To gather the background data to derive the exploration and analysis for this project, two sets of HRIR measurements were processed and compared. The first set was recorded during a listening test at the UoS. The second set of data was obtained from the CIPIC database, recorded at the University of California (V. R. D. Algazi, R. O; Thompson, D. M; Avendano, C, 2001).

The data collection was conducted at the UoS in two sessions. The first was a collection of HRIRs for analysis. The second session was a listening experiment including the same participants of the data collection. Using the recorded HRIR data, the experiment was conducted to determine impact of the spectral content of source material on the listener's ability to localise front-to-rear source direction movement.

For the collection of the HRIRs, a chair was set for the listener (resulting listener height of 1.2 meters) in the anechoic room at the UoS (which has near-anechoic qualities). The participants were asked to block their ear canals with Anti-Nois™ earplug wax to support a pair of B&K Type 4101 miniature ear canal microphones, positioned in the centre of the ear canal. The Anti-Nois™ wax is used exclude the ear canal resonances. Exclusion of the ear canal from measurements is common, as they are constant even with changing source locations (Blauert, 1969). HRIRs were then recorded through a Fireface UCX interface. This particular style of listening test has been utilized previously for headphone listening tests in studies such as Møller et al. (1995).

To derive data for HRIR measurements, a swept logarithmic sine tone was played from two Yamaha MSP3s that were placed at front and rear positions with azimuth angles of approximately 40° and 140° degrees, and at a distance of approximately 1.5 meters from the centre of the subject's head. This style of HRIR capture is equivalent to the methods outlined in Farina's research and is widely used for similar applications (2000). At this stage, transfer functions for AKG K1000 Earspeakers were also recorded, with multiple re-seatings on each listener's head, so that a correction filter could be developed and used for individually-equalized reproduction in the listening experiment to follow.

The resulting data is then processed in the Mathworks software MATLAB. The HRIR data was subject to the same analysis as the CIPIC database, however, the Earspeaker Transfer Functions (ETF) were averaged across the 10 seatings recorded and stored for the future listening test. Recording and averaging the ETF (and through to the application of the transfer filter to "reverse" the impact of the Earspeakers themselves) is a method similar to Sunder et al (2012), and is an important element within the experimentation of spectra within the context virtual auditory display. Developing a transfer function to "cancel out" the headphones/Earspeakers used in the test ensures that colouration from the delivery method does not impact on the results.

The CIPIC database is a pre-recorded database of HRIRs for a wide range of vertical and horizontal locations for 44 individual subjects. Without the limitation of two fixed locations (as captured at the UoS), it was decided to make use of the wide range of azimuthal data to capture the influence of specific spectral and phase delay with changing azimuth. For each subject, elevation between -5° and 23° (for frontal sources) and for the elevation of 158° through 186° (for rear sources) was selected. This range was selected with the intention of capturing average spectral features for the wider source area. Spatial averaging was applied in the MATLAB software, to the range of elevations at each azimuth, to arrive at an average transfer function per point of azimuth, and then collated into a 4 Dimensional matrix for further processing.

To prepare for analysis, the data in both databases were subject to time alignment to extract the Head Shadowing component of the impulse. The head shadow is a fixed duration delay, which is not direction dependant on the median plane and within the cone of confusion (where all recordings were taken) (Hebrank, 1974). Including this delay in the analysis would disguise the phase delay values that form an element of this exploration. Once all HRIR data is time aligned, the data is then converted to frequency domain and the mean of the total subject pool response is calculated (at each azimuth and the meaned elevations).

This mean is then divided out to isolate variances from the average response, hence revealing direction-varying elements. At this stage, the Phase Delay data was extracted from the HRTFs in to separate matrices for further analysis.

This averaging of the means is modelled after the data preparation that Martens executed prior to the Principal Components Analysis (PCA) of data for the exploration in "Principal Components Analysis and Resynthesis of Spectral Cues to Perceived Direction", where the exploratory analysis of recorded HRTFs was carried out through coarse spectral analysis to allow for visual identification of "directionally-dependant changes in the overall shape of HRTFs" (Martens, 1987). This coarse analysis is then used to derive initial observations to inform a more in-depth Cluster analysis of the databases to determine any groups or patterns in the data.

Cluster analysis was carried out through MATLAB to identify coarse groups of similar behaviour across the participant range. Positive results from this style of analysis indicated that more detail could be achieved through Principal Components Analysis at a later stage.

For these datasets, extracting the data where the highest variance exists will reveal the key frequency and phase delay features which allow for the most definitive discrimination of front and rear virtual auditory sources. Most clearly, determining at which subject and azimuth the front to rear comparison shows the highest variance across the frequency spectrum will isolate the potentially most applicable phase delays to extrapolate and experiment with in future studies. On a secondary level, comparing all subjects on a global level will identify trending groups of frequency and phase delay responses, which can be used to predict the robustness of the frequency and phase delay values selected.

The secondary element of the UoS data collection was a listening test used to explore potential effects of the spectral colouration of source material on the perception of front to back source location. This test was carried out on the same pool of subjects as the individually recorded HRIR and ETF, in the same semi-anechoic room. The AKG KM1000 Earspeakers were used to reproduce 3 sound sources processed using HRIRs measured at 40 and 140 degrees of azimuth.

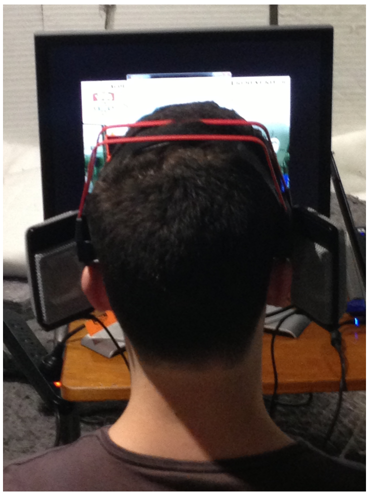

Figure 2 - Photograph showing the placement of the AKG KM1000 Earspeakers during the listening experiment. Permission attained for reproduction.

These locations were presented to the listener in a 2 Alternative Forced Choice test of 30 trials in 3 seatings. In each trial, the listener was presented with two sources – either an "A", "U" or "I" vowel sound convolved with either the frontal or rear HRTF. These combinations were randomised. The listener was then required to select whether the second sound moved forwards or rearwards in comparison to the first sound via a graphical user interface (GUI) developed and run through the MATLAB software.

The first seating was informed as being a practise seating. The second seating was presented in exactly the same manner as the first, in immediate succession. The participants were then offered a new GUI with each of the six source files attached to a button, which could be pressed to play each sound at will. This "training" session was then followed by a final seating of 30 trials.

## 4. RESULTS

Primary analysis of the results across all analyses and listening tests demonstrated consistent spectral features that assist in the discrimination between front and rear sources. It was also well acknowledged during the analysis of results that many features for future investigation were identified within phase delay as well.

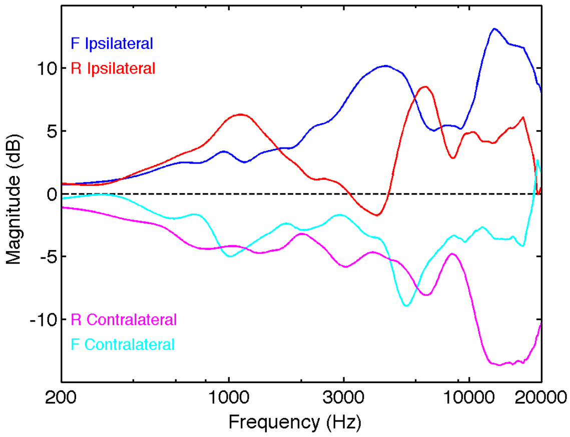


Figure 3 - Plot of Magnitude difference (from the mean response) for Front/Rear sources. Areas of greatest difference of Front from Rear describe major spatial cues.

Figure 3 presents the Magnitude deviance from the mean recorded result across 30 students that participated in the UoS listening experiment. Spectral cues that would most clearly distinguish front from rear are points where the Magnitude difference is at highest opposition between Front and Rear. This includes 1kHz (where rear dominates), 3-4kHz (where front dominates) and 8kHz (where rear dominates). Considering range beyond 10kHz, it can also be noted that 13kHz also plays an important role in discriminating frontal sources. This data has been raised previously as being standard across relevant literature.

We will now consider the findings from individual trial combinations, and across trial pairings to determine the impact to the source spectra. For the simplest trial combination within the listening experiment, the same source was convolved with both Front and Rear HRTFs and presented in succession. The results of this test demonstrate that the interplay between the spectra is detectable across the participant pool (and hence that the experiment is valid), and follows the trend as identified through analysis.

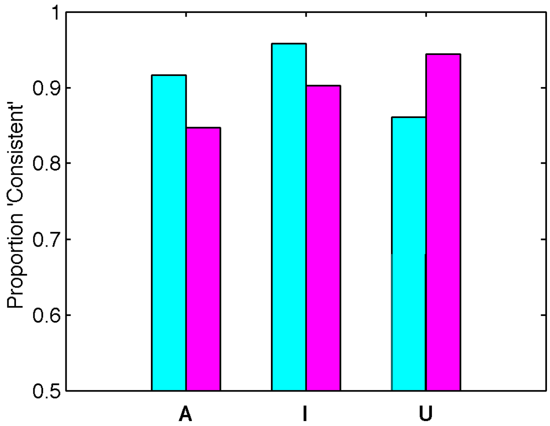


Figure 4 - For the same source presented with differing HRTFs in the same trial, the proportion of directional offset reports consistent with measured HRTF angles is shown for three vowel colourations. Cyan bars plot results for Forward movement and Magenta bars plot results for Rearward movement.

Figure 4 shows the result of trials where the same source sound was presented in succession with the HRTF applied as the varying factor. The three source sounds are charted with the proportion of consistent responses for "front" shown as the Cyan bar and the "rear" as the Magenta. A 'consistent' response is here defined as a response indicating frontward movement of the virtual source when the HRTF applied to the source in the first of two temporal intervals is that for incidence from the rear, while the HRTF applied to the source in the second temporal interval is that for incidence from the front. Isolating the HRTF as the changing feature identifies that participants are able to detect directional offset of the virtual auditory image with varying levels of consistency based on the interaction between source and HRTF spectra. Interaction between the source formants and spectral features of the HRTFs is clear in the more consistent identification of a greater proportion of frontward movement for "A" stimuli, and a greater proportion of rearward movement for "U" trials and rearward "U" trials. In particular, this can be understood as a demonstration that the alignment of formants and HRTF cues leads to a stronger bias for 'consistent' identification of source directional offsets.

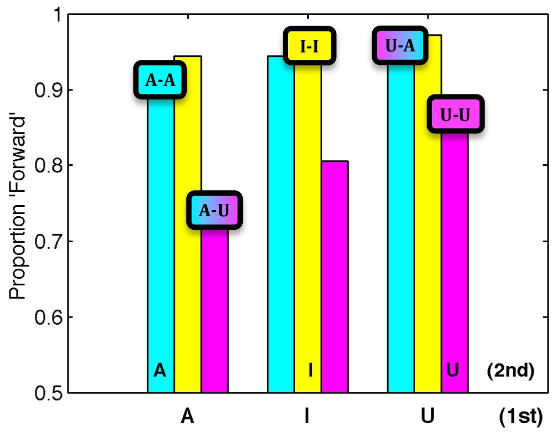


Figure 5 - Given sources processed to move from Rear to Front, across all three vowel sounds, proportion of 'Forward' responses is plotted. Consistent reports occurred at the highest proportion when the vowel sounds spectral shape was similar to the spectral shape of the HRTF used in processing that source for presentation to the listener.

Figure 5 adds another layer of complexity to the results achieved from the listening experiment. Showing the results of the specific combination of Rear-Front HRTF presentation across all source sound combinations, correct identification of a "forward" movement can be considered across the wider result pool of trial combinations.

The predicted outcome is truly demonstrated within the "A-U" or "I-U" combination. The spectra of the "U" favours the Rear spectral cues, where as the "A" and "I" are skewed more towards Frontal spectral cues. Hence, when the listeners were presented with a source that moved from Rear to Front while the spectra of the source was tending from Front to Rear, the accuracy of correct localization drops to approximately 76%.

## 5. DISCUSSION

As per Figure 5, this particular spatial movement has performed well for some vowel combinations ("U"-"I") while others have demonstrated significant room for improvement of correct identification ("A"-"U"). The spectra of the source sounds was introduced in Figure 1. Outlining the specific constructive and destructive interactions between the source spectra and the HRTF spectra will make clear why some combinations in the listening experiment achieved a higher correct response rate than others.

Noting the most significant alignment of spatial cues and formants, the shift between the second formant of "U" (987 Hz) to second formant of "I" (2.4 kHz) aligns with a key Rear cue (1 kHz) and then shifts towards a key Frontal key (4 kHz). This alignment demonstrates a tendency for these unprocessed sounds to have an inherent spatial nature, and to perceptually have a virtual auditory image without convolution with a HRIR.

To consider these interactions within the context of the listening experiment, Figure 6 presents the three sources convolved with the mean HRIRs for both the Front and Rear positions.

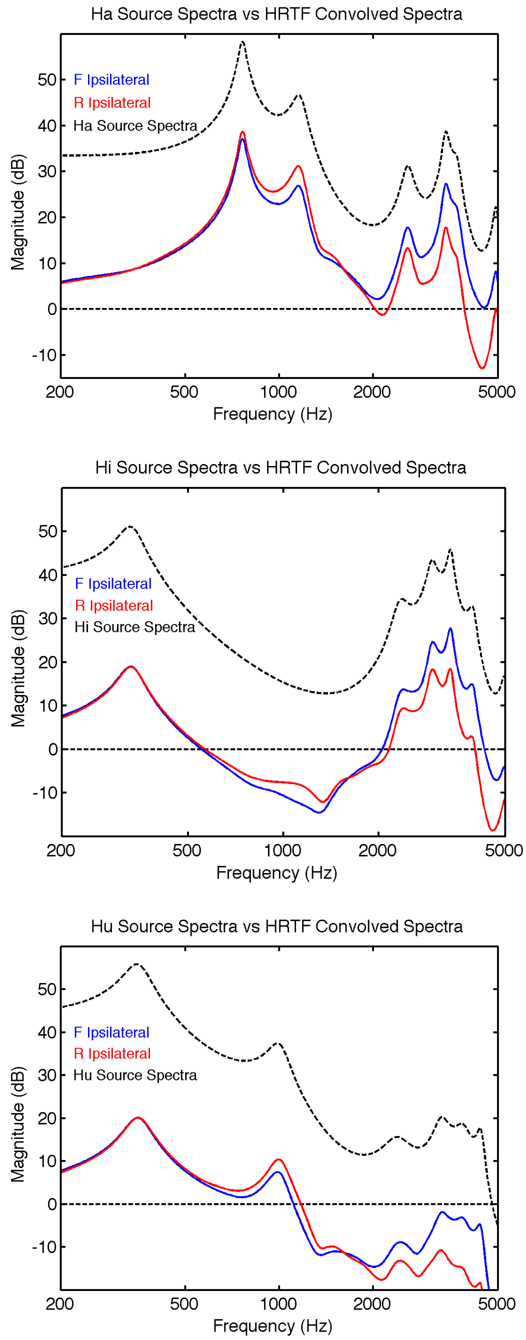


Figure 6 - Source spectra after convolution with mean Front (blue) and Rear (red) HRTFs

These graphs demonstrate the impact of the source spectra through the difference between the Front and Rear Responses charted. Comparing the raw source spectra (black dashed lines) to the resulting Front (blue) and Rear (red) spectra, differences between the source spectra and resulting spectra demonstrate constructive and destructive interactions that have contributed to the results of the listening experiment.

Immediately noticeable is the enhancement of specific spatial cues (eg Rear cues are dominant in the "U" spectra and Frontal cues in the "I" spectra), while other cues are counteracted by the source spectra (eg, Rear cues are near indistinguishable from front cues in the "I" spectra). This can be thought of as the source spectra "pushing" and "pulling" the spatial cues in the HRTF.

To delve in to further detail, particularly noticeable in the resulting "U" spectra is the impact of the second formant (creation of a peak at 934Hz in Rear). Similarly, the dominance of the Frontal cue range (major peak at 2.9kHz) in the resulting "I" spectra coupled with the absence of the Rear cues in the 900-1000Hz range describes on an introductory level why listeners in this listening experiment responded as they did.

For this figure, another consideration that is explored, is exactly which charted spectra (front or rear for that source sound) is dominant in which spectral range. This level of detail extends beyond a general comparison across all source sounds for the overall results and brings in to focus correlations between these results and the lowest scoring trials from Figure 5.

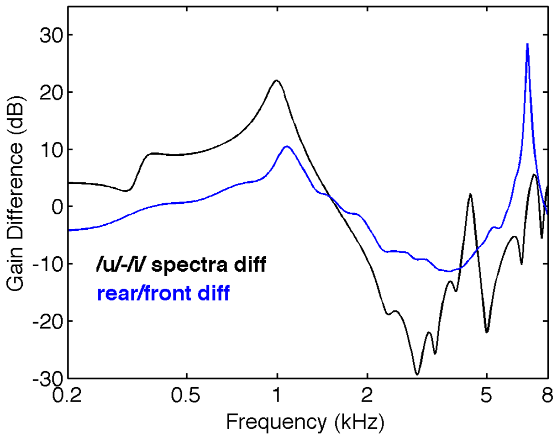


Figure 7 - Magnitude difference in spectra of "I" and "U" (blue) and Mean Magnitude difference in spectra of Rear and Front HRTF (black).

Figure 7 demonstrates that there are similar differences between the "I" and "U" spectra and Rear and Front spectra. Using the "I" and "U" spectra as specific examples of the potential extremes of manipulation of spatial image consistently leads us to consider the overall difference between the two spectra and compare this difference to the difference that occurs between the Front and Rear HRTF spectra. As has been alluded to throughout the report, the trends of both differences are somewhat consistent through the range, while observing the two key spectral cue ranges (appx 1kHz and 4kHz).

Specific consistency between the two differences in the Rearward cue range of 1-1.5kHz demonstrates a clear and overwhelming and exaggerated demonstration of the Rear spectral cue. A negative value in this range demonstrates that the "U" spectra has higher magnitude values for this cue range. The 30dB of difference between "U" and "I" spectra in the 3-4 kHz range follows the HRTF-based spectral cue.

Additionally, the frontal cue difference in the source spectra explains several results attained in the experiment. Again, referencing the specific case presented in Figure 5 where trials of Rear to Forward movement that ended on a Forward "U" sound show the most potential for an improvement of correct response rates. There are strong correlations from results in Figure 6 - where the forward cues for the "U" sound are not only lowest in magnitude out of the three source spectra, but are also proportionally lower in magnitude out of that source's individual spectra – with the clear 30dB in the frontal cues range that are missing from the "U" spectra (as per figure 7). Investigating a method to remedy this for a future project may help increase the proportion of correct response rates by helping the listener correctly distinguish the "U" source as Frontal when convolved with a Frontal HRTF.

### 5.1. Limitations of Results

As highlighted in Algazi and Duda's study (2002), the use of human participants in listening tests are problematic as postural changes caused by the constant need to shift cause variations in results. This observation is only specifically applicable in situations where head movement is allowed and not controlled in any way. For this experiment, while the listener's heads were not fixed to a specific location in the collection of HRTFs or listening experiment, the use of the Earspeakers negated the impact of head movement in identifying the location of the sources within the listening experiment. Additionally, the use of an averaged ETF (collected over ten seatings) allowed the researchers to average out any outlying results caused by misplacement of the Earspeakers.

Other elements of the UoS data collection and experiment that may have interfered with the results include hair interfering with the placement of the Earspeakers and placement of the mini microphones. The Earspeakers sit away from the head (as per Figure 2), not around the ears such as supra aural headphones, this allows for the hair to sit in any number of arrangements between the ear and Earspeaker.

## 6. CONCLUSION

This study has revealed that there is a significant amount of interaction between the spectra of the source sounds selected and the spectra of the Front and Rear HRTFs applied within the listening experiment at the University of Sydney. Through analysis of the HRTF spectra, Source sound spectra, results of the convolution of these elements, and the results of the listening experiment, it was demonstrated that sources that constructive and destructive interactions can manipulate the ability of the listener to correctly identify the front to back movement of virtual auditory sources.

More specifically, it was revealed that when the source and HRTF form constructive interactions (the source is not cancelling the cue in the HRTF) that correct source direction identification was the overwhelming outcome. Specifically considering the "U" Columns (Figure 6 references the vast support which the "U" F2 offer the Rear HRTF spectra), near 100% consistent identification occurred when a Front "A" or "I" followed the Rear "U". Conversely, when a destructive interaction occurred between the source spectra and the HRTF spectra (source spectra working against the HRTF spectra), listeners were less consistent in identifying the correct movement of the source.

It was revealed that the inherent directionality of these selected sources, and the extreme impact of these source sounds on the listeners ability to identify movement of the source, is due to the coherence of the formants of these sources to the spectra of the HRTFs at the selected locations. The parallels of magnitude difference between source spectra and HRTF spectra demonstrate specific spectral regions that (additionally) achieved the lowest scores of correct identification of spatial movement. These spectral ranges could greatly benefit from manipulations in phase delay to overcome such extreme spatial manipulations from the source spectra.

The applications of these findings are somewhat esoteric at this stage. It is clear that there is more to be learned about the interactions between source spectra and the spectra of HRTF filters for individual listeners – in both the impact of more complex vowel sounds, spoken word in general, and other sound sources (both highly resonant and non-resonant or varied in spectral quality). As such, this initial project is an excellent starting point to investigate the impact of source spectra on the perception of position in virtual binaural auditory display. The direct implications may be relevant to technologies such as virtual auditory display for Virtual Reality experiences and video games, however, to prove that simplified filters may improve front to back localisation would be a stretch and would require further study to determine whether there is an effectiveness in such a direct approach.

## 7. AKNOWLEDGMENTS

The author would like to acknowledge the assistance of William Martens, Luis Miranda and Manuj Yadav for organizing and setting up the listening experiment, and William Martens for providing the introductory MATLAB script from which the analysis of the results and generation of figures was derived.

## 8. REFERENCES

[1] Algazi, V. R; Duda, R. O; Duraiswami, R; Gumerov, N. A; Tang, Z. (2002). Approximating the Head-Related Transfer Function Using Simple Geometric Models of the Head and Torso. *Journal of the Acoustical Society of America, 112*(5), 11. doi: 10.1121

[2] Algazi, V. R; Duda, R. O; Thompson, D. M; Avendano, C. (2001, Oct 21-24). *CIPIC HRTF Database.* Paper presented at the IEEE Workshop on Applications of Signal Processing to Audio and Electronics, New York.

[3] Begault, Durand. (1990). *Challenges to the Successful Implementation of 3-D Sound.* Paper presented at the 89th AES Convention, Los Angeles USA.

[4] Blauert, J. (1969). Sound Localization in the Median Plane (Vol. 22, pp. 205-213). Acustica.

[5] Butler, Suzanne K. Roffler; Robert A. (1968). Factors that Influence the Localisation of Sound in the Vertical Plane. *The Journal of the Acoustical Society of America, 45*(6), 4.

[6] Farina, Angelo. (2000, February 19-22). *Simultaneous Measurement of Impulse Response and Distortion with a Swept Sine Technique.* Paper presented at the 108th AES Convention, Paris, France.

[7] Hebrank, J; Wright, D. (1974). Spectral Cues Used in the Localisation of a Sources on the Median Plane. *Journal of the Acoustical Society of America, 56*(6), 1829-1834.

[8] Kelly, Michael C.; Tew, Anthony I. (2003, 22-25 March). *A Novel Method for the Efficient Comparison of Spatialisation Conditions.* Paper presented at the 114th AES Convention, Amsterdam, DE.

[9] Martens, William L. (1987). Principal Components Analysis and Resynthesis of Spectral Cues to Percieved Direction. *ICMC Proceedings*, 274-281.

[10] Møller, H; Hammershøi, D; Jensen, C; & Sørensen, M. (1995). Transfer Chracteristics of Headphones Measured on Human Ears. *Journal of the Audio Engineering Society, 43*(4), 14.

[11] Rogers, M; R, Butler. (1992). The Linkage Between Stimulus Frequency and Covert Peak Areas as it Related to Monaural Localisation. *Perception & Psychophysics, 52*(2), 10.

[12] Sunder, Kaushik; Tan, Ee-Leng; Gan, Woon-Seng. (2012). *On the Study of Frontal-Emitter Headphone to Improve 3-D Audio Playback.* Paper presented at the 133rd AES Convention, San Francisco, USA.

[13] Toole, Floyd E. (1991). *Binaural Record/Reproduction Systems and Their Use in Psychoacoustic Investigations.* Paper presented at the AES Convention 91, New York.

[14] Zahorik, Pavel, Bangayan, Philbert, Sundareswaran, V., Wang, Kenneth, & Tam, Clement. (2006). Perceptual recalibration in human sound localization: Learning to remediate front-back reversals. *The Journal of the Acoustical Society of America, 120*(1), 343. doi: 10.1121/1.2208429

[15]

[16]