DOI: 10.22032/dbt.39956
DOI (proceedings): 10.22032/dbt.39936

Verband Deutscher
Tonmeister e.V.

Full Reviewed Paper at ICSA 2019

Presented * by VDT.

Towards the Perception of Sound Source Directivity Inside Six-Degrees-of-Freedom Virtual Reality

Thomas Robotham[1], Olli S. Rummukainen[1], Emanuël A. P. Habets[1]
International Audio Laboratories Erlangen[1], Germany, email: thomas.robotham@iis.fraunhofer.de

Abstract

Sound source directivity is a measure of the distribution of sound, propagating from a source object. It is an essential component of how we perceive acoustic environments, interactions and events. For six-degrees-of-freedom (6-DoF) virtual reality (VR), the combination of binaural audio and complete freedom of movement introduces new influencing elements into how we perceive source directivity. This preliminary study aims to explore if factors attributed to 6-DoF VR have an impact on the way we perceive changes of simple sound source directivity. The study is divided into two parts. Part I comprises of a control experiment in a non-VR monaural listening environment. The task is to ascertain difference limen between reference and test signals using a method of adjustment test. Based on the findings in Part I, Part II implements maximum attenuation thresholds on the same sound source directivity patterns using the same stimuli in 6-DoF VR. Results indicate that for critical steady-state signals, factors introduced by 6-DoF VR potentially mask our ability to detect loudness differences. Further analysis of the behavioral data acquired during Part II provides more insight into how subjects assess sound source directivity in 6-DoF VR.

1. Introduction

The directivity of a source is a measure of the distribution of sound when propagated, dictated by its shape, size and material properties [12]. When sound is emitted into a diffuse-field environment, what arrives at our ears is a summation of direct and reflected components, all with individual characteristcs, initially determined by the source's directivity pattern. For auralizations inside VR, realistic sound effects play an import role in our sense of presence [1] and for sound sources, the directivity is a key component [11]. Altering this directivity on the same audio source could potentially lead to manipulation of perceptual aspects such as localization [2], specifically distance [24], or even increased presence [25].

Some research regarding auralizations suggest that alterations in source directivity *can* be perceived by subjects [4], even in complex free-field/outdoor environments [8]. However, a comparison between objective and subjective evaluations of source directivity suggests that whilst a significant difference in objective parameters is present between omni-directional and 'realistic' instrument sources, results of subjective testing showed no perceived significant difference [22]. Although not stated as significant, subjects were however able to perceive differences between omni-directional and a highly narrow-beamed source, with 1/16th of its surface area set to 10 dB louder than the remaining area. As the evaluation was auralized at various static positions, these results raise the question if such differences could be audible within a 6-DoF VR environment. A recent study conducted by Sloma and Neidhardt [18] explores such effects. Using two characteristics of directivity, omni-directional and loudspeaker data, various testing phases were conducted involving a static position with head movements, and full movement with guided paths. The aim was to state, for all phases, which source directivity was

---

[1] A joint institution of the Friedrich-Alexander-University Erlangen-Nürnberg (FAU) and Fraunhofer Institute for Integrated Circuits (IIS).

present. Subjects should stick to pre-determined trajectories, which were: a straight line (provoking fluctuations in distance attenuation) and a portion of the circumference (maintaining a constant distance). The selected stimulus was music, additional room acoustics were present. The authors conclude by stating that subjects were unable to distinguish between sources with different directivity respones at static positions with head rotations, that room acoustics only have a high influence during static listening positions, and that listener movement itself has a considerable contribution to our perception of source directivities. However, one could speculate that the simultaneous presence of distance attenuation, non-steady-state signal type, and inclusion of room acoustics may heavily influenced the results.

Considering a source inside an ideal anechoic environment, we hear only the direct sound. If this source has a frequency-dependent directional response, walking around the sound source alters timbre and loudness due to magnitude changes in frequency spectrum. At discrete angular azimuth (θ) positions A and B ($A \neq B$), the two frequency repsonses possess a difference. If the frequency response remains the same at points A and B but only overall gain changes (i.e., frequency independent directivity), we are essentially comparing the loudness at two different positions of the same stimuli. Difference limen (DL), that is the just noticable difference (JND) of a given attribute, for loudness/intensity have been extensively researched. Depending on method, aural presentation (monaural, binaural etc.,), measure, presentation level, and stimuli, the results may differ. However, some concensus within literature suggests for intensity discrimination using signals presented monotically (see Figure 1) at 40 dB SL, thresholds lie around $\approx$ 0.5 dB [9, 10], with subjects being less sensitive at levels below 40 dB SL [7, 15]. The unit dB SL refers to **S**ensation **L**evel. This is the level in dB above individual listeners auditory threshold [23]. The effect remains similar when stimuli are presented diotically with subjects being slightly more sensitive [3] (i.e., we are able to detect smaller differences). For a detailed review of intensity DLs, we refer the reader to [17]. The problem becomes even more complicated with the introduction of varying interaural differences (IADs) or with non-steady-state signals [13].

The main aim of this preliminary study is to validate if the inclusion of 6-DoF elements directly influences our ability to perceive a loudness difference of $A \neq B$ inside VR. This includes (but not limited to) two main aspects. One, the inclusion of consistent movement means an instant A/B comparison is not possible between discrete points. Therefore, the speed of movement, or rate of change may influence our ability to notice such differences. Two, as mentioned, the orientation of the listener's head, also influences noticable level differences. Again, this may be further affected by consistent fluctuations in IAD states due to consistent head movements. The variation of loudness within this 6-DoF context is directly mapped to the subjects' position relative to the sound source and thus, its directivity. Further information may also be gained from behavioural analysis of tracking data, regarding how subjects conducted the listening tests. The scope is not to ascertain new DL for 6-DoF VR, but to validate if deviations from traditional DL levels are present in the context of source directivity.

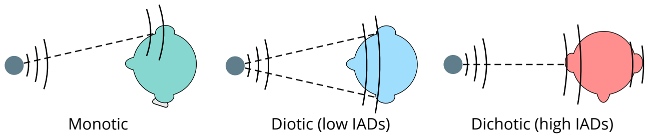


Fig. 1: Presentation of various head orientations in determining specific DL.

2. Study Overview

The study is divided into two experiments. Experiment I ascertains JND data from a control experiment with signals presented in a dioctic manner (Figure 1). Experiment II explores if any factors associated with 6-DoF VR have any effect on detecting changes in loudness in the context of source directivity, based on data from Experiment I. Throughout Experiments I and II, both the directivity functions and stimuli remain constant.

2.1. Directivity Patterns

Two first-order directivity patterns were chosen for investigation: A) cardioid and B) dipole. Both patterns are two of the most commonly used for source directivity, as they are simple to describe using a parametric function. So called zero- and first-order directivity patterns (such as omnidirectional, cardioid, bi-directional, super-cardioid, etc.) can be constructed by computing a weighted sum of an omnidirectional and a dipole pattern. As such, they are integrated into most 6-DoF audio renderers that include sound source directivity [6,21]. Both cardioid and dipole patterns provide a smooth curve which subjects can easily understand as louder and quieter when walking in a specific direction around the sound source (Figure 2). The correlation between off-axis source and listener angle (θ) attenuation for the two patterns, are taken from [14,19] and in terms of dB sound pressure level (SPL) are given as:

$$y \text{ dB (SPL)} = 20\log_{10}\left(\frac{1}{2}\cos\left(\frac{\theta\pi}{180}\right)+\frac{1}{2}\right) \tag{1}$$

$$y \text{ dB (SPL)} = 20\log_{10}\left(\cos\left(\frac{\theta\pi}{180}\right)\right), \tag{2}$$

where θ is the angle in degrees. The relationship between the two patterns also provides further information. As θ increases from 0°, the angular distance which the user must 'move' around a dipole pattern to reach equal levels of attenuation from a cardioid pattern is halved (illustrated in Figure 2). Therefore, a comparison of results between both known rates of attenuation may provide further insight into subjects ability to detect changes, without them having to physically 'move' twice as fast.

The scope of this study limits the source directivity patterns to be frequency independent. However, due to the equal loudness contour of the human ear, it is possible that the results for a broadband attenuation are influenced by certain frequency bands more than others.

2.2. Stimuli

Depending on the psychophsyical methodology employed, steady-state signals are often presented to subjects to eliminate the effects of level fluctuation [13]. As such, pink noise was selected as a steady-state signal allowing subjects to be highly critical with no temporal variations. However, in 6-DoF VR, there are seldom situations in which steady-state signals are present, either due to the signals themselves (music, speech, environmental, etc.) or the influence of distance attenuation provided by the subjects' movements. Therefore, two non-steady-state signals were chosen, viz. anechoic male speech and anechoic cello. Using a CORTEX head and torso simulator and Bayerdynamic DT770 closed headphones, playback was calibrated to an absolute level of 67 dB SPL using pink noise. Any louder than this, playback would become uncomfortable for louder portions of the non-steady-state signals. Normally, subjects' individual auditory thresholds should also be added to the playback level however, due to limited availability, individual auditory thresholds were not measured and accounted for, and the reproduction system was calibrated to an absolute playback level. It should be noted that this procedure can result in larger deviations across subjects.

3. Experiment I

The goal of Experiment I is to ascertain JND values of loudness for signals presented monaurally in terms of dB SPL using two attenuation functions.

3.1. Method

The Method of Adjustment (MoA) was selected as the method to identify difference limen between a *reference* signal and a *test* signal [20]. The objective is to vary the level of the test signal, such that a JND in loudness can be heard when compared to the reference signal. Subjects may alternate between the reference and test signals at any time, with no limitation on the number of times they may compare. Both *reference* and *test* signals are presented monaurally over headphones where only one signal is played at a time. Whilst other psychophysical methods such as method of constant stimuli may be more accurate [26], MoA puts the test signal under the subjects control. The benefit of this approach is that the subject is an active contributor to finding the criteria, thus paying more attention, as opposed to being presented signals in a passive manner and asked to make a forced choice between two signals. This methodology is also more comparable to a 6-DoF VR scenario where, given a static source, the subject is able to indirectly influence loudness via their body movements (e.g., via distance attenuation or directivity) and, the stimulus presentation is continuous.

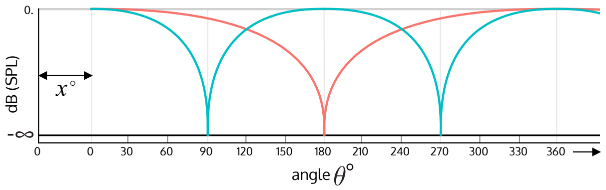


Fig. 2: Directivity patterns cardioid (orange) and dipole (blue) as a function of the angle θ, where x represents the initial randomized angular distance.

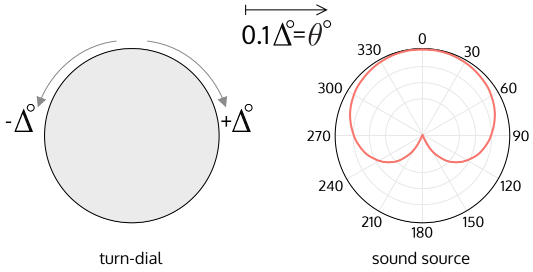


Fig. 3: Mapping of the turn dial increments and decrements (left) to angular position around cardiod pattern (right).

For the test, subjects controlled the test signal via a turn-dial button connected to a MaxMSP patch. The starting loudness of the both signals corresponds to the on-axis playback level at 0°. Rotating the dial resulted in a change in angular distance from the on-axis position; clockwise increases the angle, anti-clockwise decreases the angle. To eliminate any learning effects, the amount the dial must be initially rotated before the angular distance starts to increment is randomized across all items between 0° and 30° (see Figure 2). As the rotation of the dial emulates the movement around a sound source, continued clockwise movement results in the directivity pattern repeating itself after 360°. Finally, 1° rotation of the turn dial being equal to 1° angular increase along the directivity function was considered too sensitive. Therefore, it was remapped such that 10° rotation of the turn dial resulted in 1° angular change around the source (see Figure 3).

3.2. Subjects and Procedure

Twenty subjects participated in the control test, fifteen male and five female ranging from ages 20 - 42. All subjects were a mixture of trained and expert listeners at Fraunhofer IIS, none of which reported any hearing impairments. The test was conducted in a soundproof listening booth and instructions for subjects were presented both verbally and through text. Operation of the test was via keyboard used for switching between test and reference signals, and a turn-dial for the loudness control. After finding a JND value, subjects would press the turn-dial to move onto the next item. Each stimuli was repeated three times, resulting in ($2_{Directivities} \times 3_{Stimuli} \times 3_{Rep}$) 18 items. Average test time was around 20 minutes.

3.3. Results

The mean and 95% confidence interval (CI) of subjects' results for the control test can be seen in Figure 4. Statistical analysis was performed using an analysis of variance. Mean responses per sample show no significant difference between

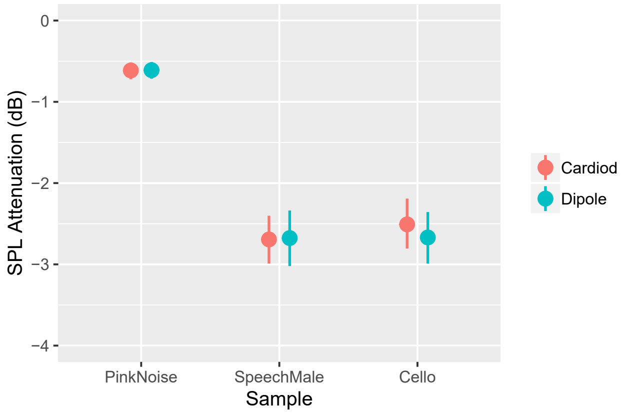


Fig. 4: Mean JND values as a physical measure in dB (SPL) between reference and test signals. Whiskers denote the bootstrapped 95% confidence intervals.

directivity patterns. It is therefore reasonable to argue the rate of attenuation per angular turn of the dial, provided no bias into subjects response. A significant effect was found under the sample type used ($F_{(2,38)} = 91.534, p < 0.001$, $\eta^2_G = 0.532$). For the steady-state pink noise signal, mean values are ≈-0.6 dB SPL in accordance with literature [17]. For non-steady-state signals, mean values are ≈ -2.6 dB SPL. Additionally, whilst the task proved harder for temporally fluctuating signals, CIs are still small and consistent across all samples and directivity patterns, suggesting that absolute level calibration at 67 dB SPL was sufficient for this test.

4. Experiment II

The goal of Experiment II, is to investigate if the introduction of self-motion and binaurally presented signals have an influence on subjects' ability to detect the thresholds found in Experiment I when incorporated into two sound source directivity patterns.

4.1. Method

Inside 6-DoF VR, continuous movement allows the subject to fully explore around a sound source. As such, an instant A/B comparison with the subjects' current position, against the on-axis position is not possible without by-passing the effect of self-movement. Therefore, to investigate whether continuous self-movement and constantly fluctuating binaural cues affect subjects' ability to detect loudness change, a different test methodology is used. Values found in Experiment I for steady and non-steady state signals are employed as thresholds, limiting the attenuation of the directivity pattern at a certain level. As the subject walks around the sound source, the level changes according to the cardioid or dipole directivity pattern as expected, until the attenuation is equal to that of JND values found in Experiment I. At this point, the level is maintained until returning above the threshold (see Figure 5). If during the assessment the subjects are unable to detect a JND in signal level, it may be hypothesized that influencing factors involved in 6-DoF VR impede, or mask, the ability to detect the changes in loudness heard in Experiment I. The limitation of this methodology is that it only provides a binary 'Yes/No' response, and further investigation would be required to ascertain if these JND values are either higher or lower. However, it is possible to extract data by monitoring the subjetct's behavior within the 6-DoF VR environment. By additionally recording user behavior during each test item, factors such as; rate of change with respect to subject movements, and speed of head movements may be analyzed and cross-referenced with subjects 'Yes/No' responses.

To conduct Experiment II, a 6-DoF VR environment was created in Unity. The overall system architecture is described in more detail in [16]. For indicating the position of the sound source, a sphere was placed in the center of the VR world. The height of the sphere was tethered to the HMD height restricting the users movements explicitly around the lateral plane of the source. A path was rendered at a 1 meter radius around the source as a guide for subjects to walk along. The angle θ between on-axis position (0°) and the users' position was directly mapped to the cardioid and dipole directivity patterns. For binaural audio, a parametric renderer was integrated with interaural time and level differences modelled from a spherical head model [5]. No distance attenuation was modeled to ensure that level fluctuations were induced purely by the directivity pattern. Using a CORTEX head and torso simulator and Bayerdynamic DT770 closed headphones, the playback level was calibrated such that the pink noise, when presented diotically at the on-axis position, was 67 dB SPL (consistent with Experiment I). In addition to the thresholds ascertained in Experiment I (see Figure 5), two sanity check conditions were also added. One included a threshold at -10 dB SPL, and the other with no attenuation at all. If the same changes always are audible by subjects, it is highly likely this leads to listener fatigue and reduced concentration levels, thus these sanity check conditions provide noticeable random variation and post-test subject screening. Finally, to remove any learning effects, an initial angular distance randomized between 0° and +30° degrees must be walked before the attenuation curve begins.

4.2. Subjects and Procedure

The same twenty subjects who participated in Experiment I also participated in Experiment II. The test was conducted in a virtual reality lab at Fraunhofer IIS using the HTC Vive Pro system, of which the VR space was calibrated to a size $2.3\ m \times 2.0\ m$. Instructions for the subjects were to walk to a starting position and click thumbpad on the Vive controller. Then, a test item began playing and subjects should walk along a circular path exploring a 180° area of interest around the sphere (see Figure 5). The task was to answer if they could hear a JND in absolute playback level of the signal when exploring around the sound source. To answer, rotating the hand-held Vive controller $\geq 30°$ to the *left* and pressing the trigger means yes, $\geq 30°$ to the *right* and clicking means no. For visual feedback, the controller would turn green and orange, indicating the respective choices. After an answer was given, subjects would go back to the start place to repeat the process. Once the test was completed, a text prompt would appear informing subjects they had finished. Instructions for

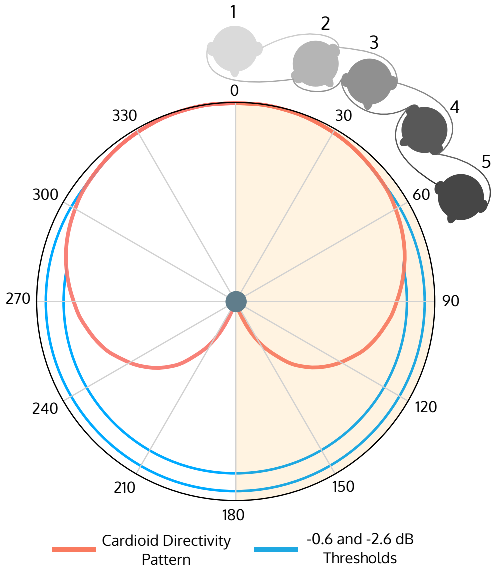


Fig. 5: Representation of -10 dB SPL threshold for cardioid directivity pattern of sound source, with example of user walking around the area of interest (*light orange*).

the test were presented both verbally and also written inside the VR world for reference. Here, test items were only repeated twice, plus -10 and 0 dB SPL sanity checks per stimuli resulting in a total of 21 items.

4.3. Results

The results of subjects' 'Yes/No' responses for Experiment II are shown in Figure 6. Green bars indicate the number of times subjects could hear a just noticeable difference, and orange if no difference in loudness could be heard. All subjects responded correctly to the sanity check thresholds (0 dB = *could not* hear difference, -10 dB = *could* hear a difference) and therefore, these responses are not plotted. Respective thresholds for the different stimuli are shown below each item. Binomial distribution tests were conducted for statistical analysis regarding the 'Yes/No' pairing of each item, and are overlayed on top of all 'No' responses. If the number of 'Yes' responses occupies the same range as the binomial test confidence intervals, the null hypothesis may be accepted (i.e., that no significant effect is present and subjects are equally likely to respond 'Yes' or 'No'.). For non-steady-state signals Cello and Speech, no significant difference can be observed between subjects responding 'Yes/No'. However, for the noise signal, in almost all observations, subjects could *not* detect a difference in loudness. The number of 'Yes' responses does not overlap with the binomial test intervals therefore, a significant effect is present that results in subjects not being able to hear any difference. Both observations regarding the signal type can be made also for cardioid and dipole directivity patterns.

Comparing these results with the data from Experiment I, it is reasonable to conclude that for the critical noise signal, the inclusion of 6-DoF elements impacted subjects ability to notice any change in loudness. Mean values in Figure 4 indicate that 50% of subjects *could* detect a JND smaller than -0.6 dB.

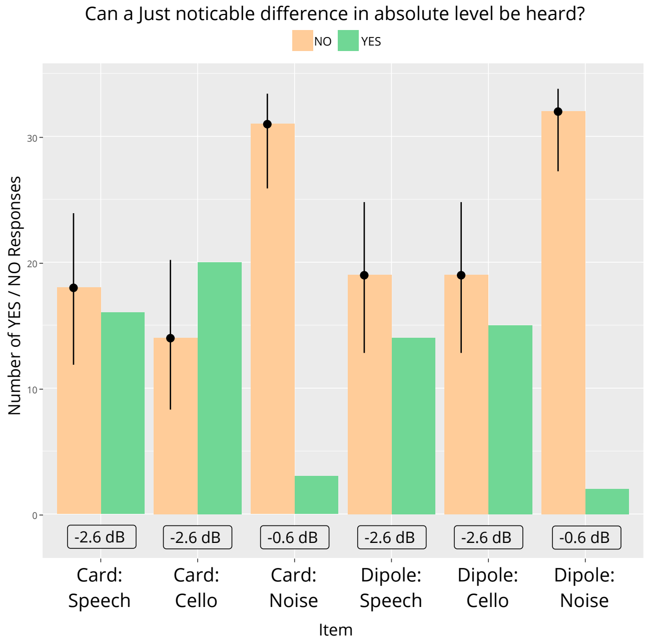


Fig. 6: Frequency of subjects' responses for Experiment II as bar plots. Binomial test data provided over the top showing probability of success for selected answer within 95% bootstrapped confidence intervals.

However, the results in Figure 6 show that a JND in loudness was detected only five times across all presentations. As this signal is steady state and no inherent temporal fluctuations are present, subjects' ability to recognize loudness changes may be hindered due to the shifting inter-aural time and level differences induced by head movements. The result is that this level difference of -0.6 dB is no longer audible inside 6-DoF. Conversely, 'Yes/No' responses in Experiment II for the Cello and Speech signals remained statistically equal. As the results from Experiment I show that 50% of subjects could hear this difference, this may indicate that JND value of -2.6 dB for non-steady-state signals has remained the same. As the methodology of Experiment II was designed to confirm or deny if subjects could hear previous findings, further testing would be needed to confirm this hypothesis. However, if such a threshold exists for non-steady state signals inside 6-DoF VR, this information could prove useful in perceptually optimizing directivity data for such sound sources.

5. Behavioral Analysis

5.1. Position, Orientation and Time

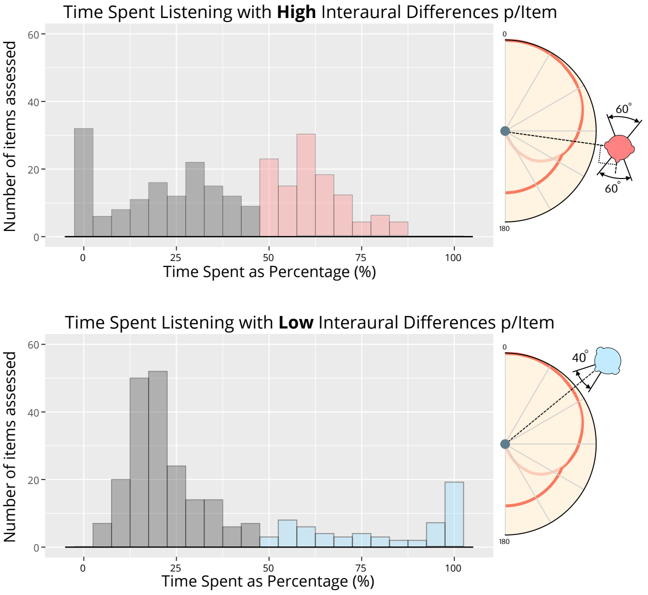
To further assess subjective results, tracking data of head movements and user position over time was recorded for each subject per test item. Analysis was conducted on total average: distance walked, head movements (pitch, yaw, roll and source relative yaw), and time taken per test item. For 6-DoF VR, exploring if freedom of movement combined with subjects' head rotations have any affect on our perception of source directivity is of particular interest. All analysis showed no significant difference between either the signal type used or the directivity pattern employed. In comparison to the ‘Yes/No’ answers in Figure 6, where a large difference can be seen between steady and non-steady state signals, no such difference can be observed in the tracking data. This indicates ‘Yes/No’ responses provided by subjects were not due to uncertainty, which would be reflected in subjects needing significantly more time, or moving greater distances. Considering the relative yaw movement data, further insight may be gained by analyzing *how* subjects listened throughout the test.

5.2. Interaural Differences

To assess if subjects listening with ‘high’ or ‘low’ IADs (orientations shown in Figure 1) had an effect on the results, source relative yaw tracking data was divided into ‘low’ and ‘high’ angular categories. For items that were assessed with relative head rotations (Yaw) of 0°, with ±20° variation were labeled as having ‘low’ IADs. Conversely, items that were assessed with 90° or 270° with ±30° variation were labeled as having ‘high’ IADs. The angular ranges (±20° and ±30°) for these categories was based on visual inspection of all raw tracking data. A smaller angular distribution was chosen for ‘low’ IADs as it appears easier for subjects to maintain a more accurate head position when looking at the source. For each subject, the percentage of time spent listening to each item with both IAD categories was calculated. The number of items where subjects spent over 50% of their time evaluating with ‘high’ or ‘low’ IADs was then counted and cross-referenced with their ‘Yes/No’ response. A visual representation of this IAD categorization is shown in Figure 7 and the results of cross-referencing responses with IADs are shown in Figure 8. This analysis aims to provide a preliminary insight into possible behaviors and not to establish new methods of analysis. By initial observation, the results in Figure 7 would indicate that *if* subjects were to listen for any JND whilst moving around a sound source in an intentional manner, the most frequent method would be with higher levels of IADs (as indicated in ‘Red’). Far fewer evaluations were conducted with subjects spending over 50% of their time with lower IADs.

Tab. 1: Table of logit regression analysis showing significant predictive variables on the binary outcome of subjects being able to hear a JND in loudness.

Coefficients	Estimate	St. Err	Z-Val	*P-value*
*Intercept**	-1.846	0.853	-2.164	0.031
Dipole × Cardioid	0.477	0.324	1.469	0.142
Dipole × *Null****	-2.621	0.769	-3.406	0.001
Cello × Speech	-0.442	0.336	-1.257	0.209
Cello × *Noise****	-2.798	0.527	-5.306	1.12e-07
High IADs	0.022	0.012	1.779	0.075
*Low IADs**	0.025	0.010	2.431	0.015



Time Spent Listening with **High** Interaural Differences p/Item

Time Spent Listening with **Low** Interaural Differences p/Item

Fig. 7: Number of answers given against the time spent listening with specific IADs. Responses given where over 50% of time was spent listening with high IADs is shown in Red, and low IADs in Blue.

For statistical analysis of data, a logistic regression model was used to determine if the categorized ‘high’ and ‘low’ IADs had a significant influence on subject responses. Generally, this method is used to determine if predictor variables (which may be both continuous or binary) can be used to model the log odds of a certain binary outcome (‘Yes/NO’). For this analysis, the predictor variables; ‘*Directivity*’, ‘*Stimuli*’, ‘*Percentage of time spent with high IADs*’ and ‘*Percentage of time spent with low IADs*’, were used to identify any significant effects on the binary ‘Yes/No’ outcome. Results in Table 1 show which of the variables have a significant effect. For categorical variables ‘*Directivity*’ and ‘*Signal*’, ‘Estimate’ shows the log odds of one variable over another changing the binary outcome. For continuous variables of IAD, every unit increase results in the log odds increasing by the estimate amount (i.e., the estimate equates to a single unit, hence why these are smaller values). From this, we can see the most significant effects on binary outcomes are the ‘Noise’ stimuli (already apparent in Figure 6), and when no attenuation curve (‘Null’) is used. However, a significant effect is also found if subjects spent more than 50% of their time evaluating with low IADs. This result is also reflected in Figure 8 where *all* responses with ‘low’ IADs (blue) are in the ‘Yes’ column, suggesting we are more likely to hear a loudness difference when maintaining low IADs.

6. Discussion

For Experiment II, subjects were free to move around the sound source, no instructions were given advising subjects to maintain a specific head orientation. Given this, it is interesting to see that whilst conducting the evaluation with lower IADs made a significant impact on the outcome, it was far more prevalent to conduct the experiment with higher IADs. Furthermore, as previously mentioned, all responses given where subjects spent more than 50% of time listening with lower IADs, were *all* positive. This would imply that even though critical listening is more acute with consistent lower IADs, in 6-DoF VR where, subjects must move around the sound source, maintaining persistent head orientation towards the sound source is either too unnatural, or most subjects felt like they could conduct the task well in another

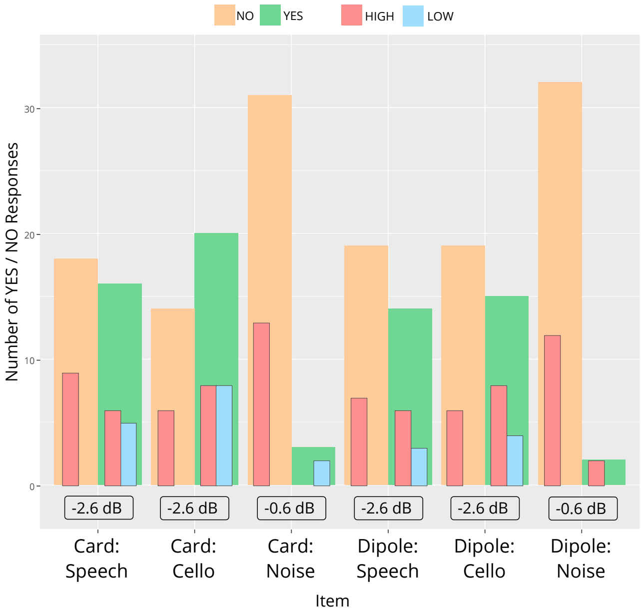


Fig. 8: Number of items listened to with high or low binaural cues overlayed on top of total 'Yes/No' responses. High binaural cues are highlighed in 'Red' and low in 'Blue'.

manner. It is also highly unlikely that in a commercial 6-DoF VR scene there is only one audio source, and that users would walk round it intentionally maintaining low IADs. This raises an interesting question as to whether the same results are obtained if the user remains stationary and the directivity pattern changes due to the user rotating the sound source. This would be similar to Experiment I however, the signal would still be presented binaurally, and subjects are still free to move their heads. If results change such that significantly more subjects could hear a JND, this would provide further evidence to suggest natural body movements *combined* with higher interaural differences makes subjects less sensitive to changes in loudness inside 6-DoF VR. If our perception is effected in such a way, this may help in defining a more perceptually motivated model towards sound source directivity for 6-DoF VR. Furthermore, the scope of this investigation was limited to frequency independent attenuation. Due to interual time and level differences operating at different frequency ranges, loudness variations in specific frequency bands may be more/less noticeable depending on the source relative head orientation of the subject.

7. Conclusion

Two experiments were conducted to investigate if the inclusion of body movements or binaural cues have an influence on our ability to detect broadband changes in loudness, relative to sound source directivity. Using the Method of Adjustment, Experiment I confirmed JND thresholds in literature for stimuli presented equally at both ears. The difference between steady and non-steady state signals was significant, however no difference was observed between the two attenuation functions (based on cardioid and dipole directivity patterns). Experiment II implemented the same directivity patterns inside 6-DoF VR with attenuation limited to the thresholds found in Experiment I for respective stimuli. Subjects were asked to explore the area around the sound source and answer if a difference in absolute loudness could be heard. Results indicate that for the steady-state signal 'Noise', the inclusion of binaural cues and body movements meant that the JND of -0.6 dB presented monaurally in Experiment I was no longer audible and that this threshold is higher inside 6-DoF VR. For both experiments, no significant difference was observed between the two directivity patterns. This may be due to the differences not being large enough over a given time and angular distribution. For future work, the authors aim to investigate various rates of change over angular distances to gain further insight into perceptual thresholds of source directivity inside 6-DoF VR. Finally, cross-correlating subject's head movements with 'Yes/No' responses, showed that spending over 50% of time evaluating the sound source with lower IADs allowed subjects to always hear a just noticeable difference in loudness. This implies that our head orientation with respect to the sound source, and consequently varying IAD, does have effect on how we perceive sound source directivity inside 6-DoF VR. For future perceptual evaluations, this may be an important consideration depending on the task.

8. References

[1] Begault, D. R. *3D-Sound for Virtual Reality and Multimedia*. Academic Press, Boston, 1994.

[2] Blauert, J. *Spatial Hearing: The Psychophysics of Human Sound Localization*, 2nd ed. MIT Press, London, England, 1997.

[3] Churcher, B. G., King, A. J., and Davies, H. The Minimum Perceptible Change of Intensity of a Pure Tone. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 18*, 122 (Nov. 1934), pp. 927–939.

[4] Dalenbäck, B.-I., Kleiner, M., and Svensson, P. Audibility of Changes in Geometric Shape, Source Directivity, and Absorptive Treatment-Experiments in Auralization. *Journal of the Audio Engineering Society 41*, 11 (Nov. 1993), pp. 905–913.

[5] Duda, R. O., and Martens, W. L. Range Dependence of the Response of a Spherical Head Model. *The Journal of the Acoustical Society of America 104*, 5 (1998), pp. 3048–3058.

[6] Google Inc. Resonance Audio Unity SDK API Reference, 2018. https://valvesoftware.github.io/steam-audio/

[7] Hirsh, I. J., Palva, T., and Goodman, A. Difference Limen and Recruitment. *The Journal of the American Medical Association: Otolaryngology 60*, 5 (Nov. 1954), pp. 525–540.

[8] HOARE, S., SOUTHERN, A., AND MURPHY, D. Study of the Effect of Source Directivity on the Perception of Sound in a Virtual Free-field. In *AES 128th Convention* (London, UK, May 2010), pp. 1–8.

[9] KNUDSEN, V. O. The Sensibility of the Ear to Small Differences of Intensity and Frequency. *The American Physical Society: Physical Review 21*, 1 (Jan. 1923), pp. 84–102.

[10] LÜSCHER, E., AND ZWISLOCKI, J. Comparison of the Various Methods Employed in the Determination of the Recruitment Phenomenon. *The Journal of Laryngology & Otology 65*, 03 (Mar. 1951), pp. 187–195.

[11] MEHRA, R., ANTANI, L., KIM, S., AND MANOCHA, D. Source and Listener Directivity for Interactive Wave-Based Sound Propagation. *IEEE Transactions on Visualization and Computer Graphics 20*, 4 (Apr. 2014), pp. 495–503.

[12] MEHRA, R., AND MANOCHA, D. Wave-based Sound Propagation for VR Applications. In *Proc. of IEEE VR Workshop: Sonic Interaction in Virtual Environments (SIVE)* (MN, USA, Mar. 2014), pp. 41–46.

[13] NAMBA, S., KUWANO, S., AND FASTL, H. Loudness of Non-Steady-State Sounds. *Japanese Psychological Research 50*, 4 (2008), pp. 154–166.

[14] PETERS, N., AND SCHMEDER, W., A. Beamforming Using a Spherical Microphone Array Based on Legacy Microphone Characteristics. In *Proc. of International Conference on Spatial Audio* (Detmold, Germany, 2011), pp. 1–7.

[15] RIESZ, R. R. Differential Intensity Sensitivity of the Ear for Pure Tones. *The American Physical Society: Physical Review 31*, 5 (May 1928), pp. 867–875.

[16] ROBOTHAM, T., RUMMUKAINEN, O., HERRE, J., AND HABETS, E. A. P. Evaluation of Binaural Renderers in Virtual Reality Environments: Platform and Examples. In *AES 145th Convention* (NY, USA, Oct. 2018), pp. 1–5.

[17] ROWLAND, R. C., AND TOBIAS, J. V. Interaural Intensity Difference Limen. *Journal of Speech, Language, and Hearing Research 10*, 4 (Dec. 1967), pp. 745–756.

[18] SLOMA, U., AND NEIDHARDT, A. Investigations on the Impact of Listener Movement on the Perception of Source Directivity in Virtual Acoustic Environments. In *Proc. of Annual German Conference on Acoustics* (Munich, Germany, 2018), pp. 1–4.

[19] SOUTHERN, A., AND MURPHY, D. Low Complexity Directional Sound Sources for Finite Difference Time Domain Room Acoustic Models. In *AES 126th Convention* (Munich, Germany, May 2009), pp. 1–10.

[20] STEVENS, S. S. The Measurement of Loudness. *The Journal of the Acoustical Society of America 27*, 5 (Sept. 1955), pp. 815–829.

[21] VALVE. Steam Audio, 2019. https://valvesoftware.github.io/steam-audio/

[22] WANG, L. M., AND VIGEANT, M. C. Objective and Subjective Evaluation of the Use of Directional Sound Sources in Auralizations. Tech. rep., University of Nebraska, Nebraska, Lincoln, (Apr. 2004), pp. 2711-2714

[23] WARD, W. D. Use of Sensation Level in Measurements of Loudness and of Temporary Threshold Shifts. *The Journal of the Acoustical Society of America 39*, 4 (Apr. 1966), pp. 736–740.

[24] WENDT, F., FRANK, M., ZOTTER, F., HÖLDRICH, R. Directivity Patterns Controlling the Auditory Source Distance. In *Proc. of 19th International Conference on Digital Audio Effects* (Brno, Czech Republic, 2016), pp. 295–300.

[25] WENDT, J., WEYERS, B., AND VIERJAHN, T. Does the Directivity of a Virtual Agent's Speech Influence the Perceived Social Presence? In *Proc. of IEEE Virtual Humans in Crowds for Immersive Environments (VHCIE)* (Reutlingen, Germany, Mar. 2018), pp. 1–3.

[26] WIER, C. C., JESTEADT, W., AND GREEN, D. M. A Comparison of Method-of-Adjustment and Forced-Choice Procedures in Frequency Discrimination. *Perception & Psychophysics 19*, 1 (Jan. 1976), pp. 75–79.