DECOMPOSITION OF THE HRTF FROM A SPHERE WITH NECK AND HAIR

Bradley E. Treeby, Roshun M. Paurobally, Jie Pan

Centre for Acoustics, Dynamics and Vibration
School of Mechanical Engineering
University of Western Australia
35 Stirling Highway, Crawley, WA 6009, Australia
treebs@mech.uwa.edu.au

ABSTRACT

Sphere scattering models are commonly used with binaural synthesis as they provide a convenient approximation of the acoustic characteristics of the human head. However these models suffer from being an over simplification of human geometry, and if used in isolation provide ambiguous source location cues. Such models also over exemplify the lobe of increased pressure (bright spot) that occurs at the rear of the sphere due to symmetrically diffracted waves arriving in phase. This paper uses decomposition to examine how the addition of a cylindrical neck and hemispherical hair covering alters the azimuthal head-related transfer function (HRTF) from a rigid sphere (up to 5 kHz). Neither anthropometric feature provides a major perturbation of the sphere HRTF. The neck produces a reduction in the bright spot magnitude in the order of 2-4 dB. The hair produces asymmetrical changes to the HRTF for ipsilateral angles in the order of 1-2 dB. Additional asymmetric reductions in the order of 2-4 dB are seen for contralateral angles when the source is near the interaural axis.

[Keywords: HRTF, decomposition, human hair, neck]

1. INTRODUCTION

In a natural listening environment external sounds are coupled to the auditory system via the pinnae and ear canals. Source location information is then extracted from the binaural level and timing differences between the sound signals at the two eardrums. Additional information is also decoded from the detailed spectral information contained within the monaural signals. In the azimuth plane these localisation cues are dominated by the presence of the head in the wave path between the two pinnae. For sources offset from the median axis this produces a path length and hence timing difference between the ipsilateral and contralateral pinnae (see Fig. 1 for a visual representation of spatially descriptive terminology). For an average sized human head, the maximum interaural time difference (ITD) occurs for sources located on the interaural axis and is in the order of 700μs. At higher frequencies (when the wavelength of the incoming sound wave is comparable to the size of the head) offset sources will also produce a reduction in sound level at the contralateral pinna. This frequency dependent head shadowing causes an interaural level difference (ILD) that becomes more significant as the wavelength decreases.

Implementations of binaural synthesis seek to exploit these fundamental mechanisms of human hearing to create artificial listening environments in which three dimensional (3D) audio can be experienced. In its simplest description, a virtual audio signal is created by the convolution of sound stimuli with impulse responses that encompass the desired directional and environmental information. This is then presented to the listener such that the resulting waveform at their eardrums matches the corresponding natural situation. Aside from the contribution of visual and *a priori* cues, the auditory percept is extracted entirely from these waveform. Correct replication of the various acoustic processes that occur in the natural environment will thus result in a virtual sound projection that is indistinguishable from its archetype.

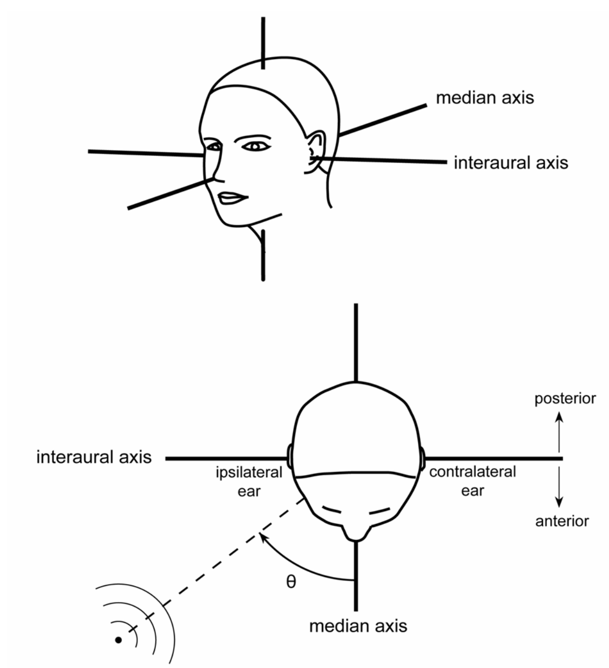

Figure 1. *Visual representation of spatially descriptive terminology. The lower panel shows an overhead view of the horizontal (azimuth) plane in which the HRTF are measured. The HRTF data shown within this paper correspond to the heads' right ear (denoted ipsilateral for the source location shown).*

The use of analytical or approximate models of sphere scattering with binaural synthesis is commonplace, particularly to recreate the appropriate time delays for sources displaced from the median plane. Woodworth and Schlosberg presented a simple ray tracing approximation for the path length and hence time delay introduced due to the presence of the head (modelled as a sphere) in between the pinnae [1]. Many implementations of binaural synthesis will simply use minimum phase head-related transfer function (HRTF) reconstructions and reintroduce ITDs using such a model [2, 3]. Similarly the solution for diffraction around a single sphere has been used to describe the broad properties of azimuthal hearing [4], and to augment experimental HRTFs with low frequency information [5, 6]. In these implementations the HRTF below 3-4 kHz is gradually replaced with information from a scattering model. Such augmentation provides stable azimuth cues and overcomes problems with measuring coherent HRTF information below 500 Hz. The sphere model also commonly forms the basis for structural modelling of the HRTF [7].

Although spherical diffraction solutions can account for the basic ILD and ITD variations in line with duplex theory, they suffer from being an over simplification of human geometry. If used in isolation a particular set of interaural differences do not specify a unique source direction in virtual space, but rather a curve of possible directions coined the 'cone of confusion' (e.g., [8]). This commonly results in localisation confusions about whether stimuli are positioned in the frontal or rear sector. ILDs extracted from sphere models also over exemplify the diffraction 'bright spot' which occurs due to symmetrically diffracted waves arriving in phase at the rear of the sphere. The bright spot for HRTFs measured from human subjects is reduced due to the presence of the neck and the non-symmetrical nature of the head. The over typified bright spot arising from sphere models can produce unnatural artefacts in virtual sound [5].

In addition to these debilities, spherical head models also remove any secondary low frequency elevation cues. These arise due to the non-spherical nature of the head (including pinnae offset) in addition to reflections from the upper torso. The torso reflections produce elevation dependent comb filter interference patterns in the HRTF in the order of 5 dB. Geometric models which match the broad features shown in experimental HRTF have been proposed [9]. Likewise elliptical ray tracing models for ITD have also been previously discussed [10]. Whilst geometric scattering studies have been extended to account for torso reflections and offset pinna, previous studies into the effects of other peripheral scatterers are limited.

This paper presents an experimental investigation into the contribution of the neck and hair to the azimuthal HRTF derived from a rigid sphere. Both the neck and hair are sufficiently significant anthropometric features that changes in the HRTF would intuitively be expected. Experiments are performed using a rigid spherical head to which a semi-infinite cylindrical neck and a hemispherical synthetic hair covering are added. Subsequent decomposition (examination of the changes in relation to the rigid reference condition) of the experimental HRTF gives insight into the individual contribution of both the neck and the hair covering. Although the experiments are performed using a spherical head, such decomposition also aids in understanding the effect of such peripheral scatterers on the human HRTF.

The effect of the neck on the spherical HRTF has not previously been formally investigated, however the expected changes may be considered intuitive. As mentioned human HRTF do not exhibit a strong posterior bright spot like those from spherical models, and a reduction of this feature would thus be expected. In this regard filter approximations of scattering models, whilst removing some salient features, do not over exemplify the bright spot [5]. The effect of hair has been investigated in a previous experimental study using a mannequin fitted with a variety of wigs [11]. The hair coverings produced an increased shadowing effect of the incident wave, with changes in the HRTF above 2 kHz in the order of 5 dB. These results are in agreement with a previous systematic investigation into the effect of a uniform impedance covering on the azimuth cues derived from an analytical sphere scattering model [12]. Additional studies utilising boundary element models with non-rigid head surfaces have shown similar results. Katz [13] observed up to 6 dB changes in the HRTF due to the inclusion of hair, the effects again being most pronounced in the rear (shadowed) region of the head.

2. EXPERIMENTAL METHOD

To classify the effect of the neck and hair, two series of HRTF measurements using an enlarged wooden sphere (diameter 0.248 m) without pinnae were completed in an anechoic chamber. The enlarged diameter of the wooden sphere (constructed for a previous experiment) was to allow enough space including cable relief for two approximately diametrically opposed internal microphones. The HRTF shown are not frequency scaled and any future data comparison should take this into consideration. The wood material was verified using an impedance tube to be approximately rigid over the test frequency range. For the current study the sphere was equipped with one internal microphone (BSWA Tech - MP205 tip, MA211 inline preamplifier) which was positioned flush with the outside of the sphere surface for all experiments.

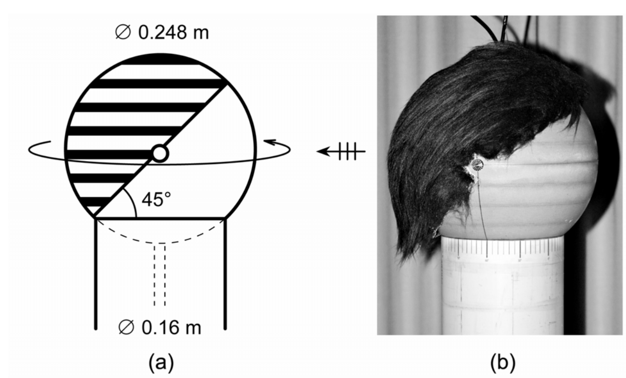

Figure 2. (a) *Diagram of experimental setup shown at a rotation angle of 0°. The hair covers one hemisphere and is elevated 45° from the median axis. For tests including the neck the sphere is supported by the piping, otherwise the sphere is supported by a thin steel rod.* (b) *Experimental sphere setup including the neck and hair.*

The first series of measurements investigated the rigid sphere without the cylindrical neck, both with and without hair. The sphere was supported by a thin steel rod which allowed rotation. Reference measurements were taken of the rigid sphere. These were then repeated with the sphere hemispherically covered by a short synthetic hair material. The hair was elevated 45° from the median axis, with the internal microphone located along the hair boundary as shown in Fig. 2(a). The hair covering was trimmed so that it did not cover the microphone as can be seen in Fig. 2(b). The symmetrical hair and pinna alignment (corresponding to a 90° pinna offset from the frontal median axis) was chosen so that any asymmetries arising from the hemispherical hair covering would be easier to distinguish. The specific impedance characteristics of the synthetic hair material used are shown in Fig. 3 where ρc is the characteristic impedance of the propagation medium. The results correspond to a surface reference plane coincident with the inner rigid boundary. Due to the specifications of the impedance tube used for this measurement, results are only shown over the lower part of the frequency range investigated.

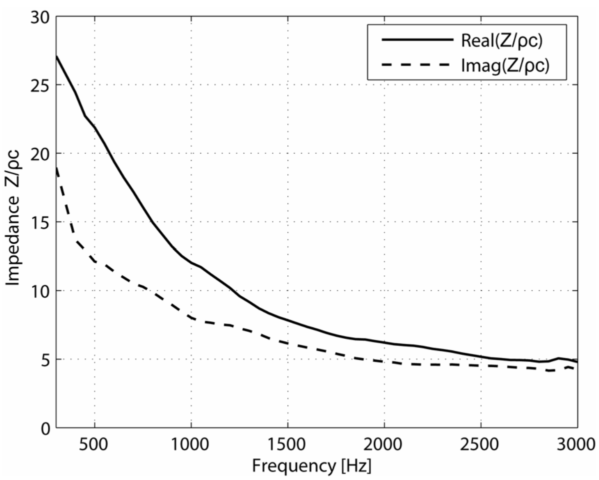

Figure 3. *Specific acoustic impedance characteristics of the synthetic hair covering used to investigate changes in the HRTF.*

The second series of measurements utilised a 1.4 m long PVC pipe 'neck', 0.16 m in diameter with 9 mm thick walls. Whilst such a long neck is unrealistic, it was chosen in favour of a shorter length to eliminate the effects of end scattering on the response which also do not occur in human HRTF. The internal radius of the neck was mitred so that the sphere sat flush within it. The diameter of the neck was chosen to be within the anthropometric range of possible head breadth or length to neck diameter ratios [14]. If the sphere diameter is taken to be the interaural distance, the neck radius used categorically corresponds to the lower echelon of realistic values. Material availability ultimately decided the final diameter value. HRTF measurements were repeated for the sphere and neck arrangement (shown in Fig. 2), both with the hair covering and without.

For each test impulse response measurements were obtained using maximum length sequences produced by the Brüel & Kjær DIRAC software and a Brüel & Kjær HP1001 unidirectional sound source. A sequence length of 2^{14}-1 (the shortest available sequence length) with 10 averages and a sampling frequency of 48 kHz was used. To remove the effects of the imperfectly anechoic measurement environment the impulse response peak onsets were located and the tails truncated to 128 samples. The complete impulse responses were then shortened to 256 samples (with the timing information preserved) and converted to the frequency domain using a 512 point FFT. All measurements were taken at 5° increments of sphere rotation, starting with the internal microphone facing the frontal incident wave direction. Rotations of the sphere were done in an anticlockwise direction corresponding to source movement as shown in Fig. 1. The measurements shown throughout this paper correspond to the HRTF from the sphere's right ear location. Whilst the sphere used is pinna-less, the term 'ear' will be used to denote the microphone location in lieu of a more satisfactory description. Given the same rotational angle definition, the left ear HRTF is simply a reflection of these measurements about 180°. For the experiments without the neck, the rotation angle was aligned using a laser level positioned at the base of the sound source (located approximately 3 metres from the sphere) in conjunction with degree markings on the rotating sphere stand. For the experiments including the neck, angles were aligned using degree markings on the upper neck. Symmetric tests without hair were taken up for rotation angles up to 180° and the data mirrored accordingly, asymmetric tests including hair were completed up to 360°.

3. EXPERIMENTAL HRTF DECOMPOSITION

The individual contributions of the hair and neck to the HRTF were decomposed by subtracting the HRTF for the rigid sphere from the corresponding HRTF containing either the neck or hair. The same was done for the HRTF containing the rigid sphere with both the neck and hair. These decomposed results are shown in the right panels of Fig. 4. The corresponding measured HRTFs are depicted in the left panels. From the top these are the sphere, sphere with neck, sphere with hair, and the sphere with hair and neck. The changes due to the peripheral scatterers do not immediately appear obvious, illustrating that the neck and hair produce subtle rather than major perturbations of the sphere HRTF. This is perhaps surprising as both the hair and neck are reasonably significant anthropometric features.

The experimental HRTF for the single rigid sphere is shown in Fig. 4(a) and exhibits a prominent posterior bright spot as expected. The bright spot results from symmetrically diffracted waves arriving in phase at the rear of the sphere. This produces a major lobe of increased pressure in the posterior region exactly opposite the source direction. This is visible in Fig. 4(a) at 270°, the angle at which the ear and source are located in this symmetric position. Other lobes of increased pressure can be seen on either side of the most prominent bright spot resulting from the same phenomena. The angular location of these ancillary lobes is dependent on frequency. The characteristics of the experimental results match closely with previously published analytical HRTF derived from a rigid sphere (e.g., [9]). This provides an informal verification of the experimental and data processing techniques utilised in the current study.

The change in the HRTF due to the presence of the neck is shown in Fig. 4(c). There is a clear reduction of the contralateral bright spot. The addition of the neck disturbs a segment of the symmetrical wave path which proportionally reduces the effect of the posterior in-phase interference. For the sphere and neck radius used in this study, the neck covers approximately 12% of the sphere surface area. This produces a bright spot decrease in the order of 2-4 dB which becomes more prominent as the frequency increases. The addition of the neck also provides small magnitude comb filtering effects. These are most significant for contralateral angles and arise from the additional symmetries provided by the arrangement of the cylindrical neck and spherical head. Additional spectral changes are also seen for ipsilateral angles. These are extremely similar in nature (although much smaller in magnitude) to the arch shaped notches provided by torso reflections using either ellipsoid or spherical models for the torso (ref [9], Fig. 13). An increase in the extent of these HRTF changes would be expected using a larger neck diameter in relation to the spherical head.

The changes due to the addition of the synthetic hair covering are shown in Fig. 4(e) and are of similar magnitude to those produced by the neck. On first inspection this result may well be surprising given the general neglect of the effect of hair in hearing and auditory investigations. For the present study however the neck, whilst providing an additional scattering body, only perturbs 12% of the sphere surface area. Conversely whilst the supplementary scattering effects produced by the hair are much smaller, the hair covers a significant 50% of the sphere. Given that the pressure on the rear surface results from waves diffracted around all sides of the sphere, a large perturbation of these wave paths would obviously produce changes to the surface pressure. This comparison at least anecdotally provides an explanation for the similar magnitude HRTF changes.

Unlike the changes due to the neck which are symmetric about the contralateral bright spot, the azimuthal HRTF changes due to the hair covering are asymmetrical. Two separate characteristics of these changes necessitate discussion. Firstly for ipsilateral angles a general decrease in the HRTF is seen. This becomes more prevalent as the ear is rotated away from facing the source (90°). The decrease is asymmetrical about this rotation angle, and is greater for rotations that move the hair covering to the frontal surface (angles greater than 90°). For these angles the increased frontal absorption produces a reduction in the magnitude of the scattered wave. This combines with the incident wave that is diffracted to other regions of the sphere. For a sphere that has a hemispherically divided boundary, there is a significantly larger reduction of the surface pressure in the regions away from the posterior bright spot for an absorbent front-hemisphere and rigid rear-hemisphere compared to the other way around. The asymmetries may also be visualised by considering whether the shortest wave path to the ear is over the rigid or the hair surface as shown in Fig. 5. For the hair covering used in this study the changes in the ipsilateral region of the HRTF are in the order of 1-2 dB. Informal investigations using denser hair materials with stronger absorption characteristics show a systematic increase in the HRTF reduction for ipsilateral angles that move the hair covering to the frontal surface.

Secondly, whilst there is not a significant perturbation of the sphere scattering bright spot, other asymmetrical HRTF changes are clearly evident for contralateral angles. These are most apparent between 285° and 315° where the decrease is in the order of 2-4 dB. For posterior contralateral angles there is a small increase. Overall this creates a noticeable asymmetry about the

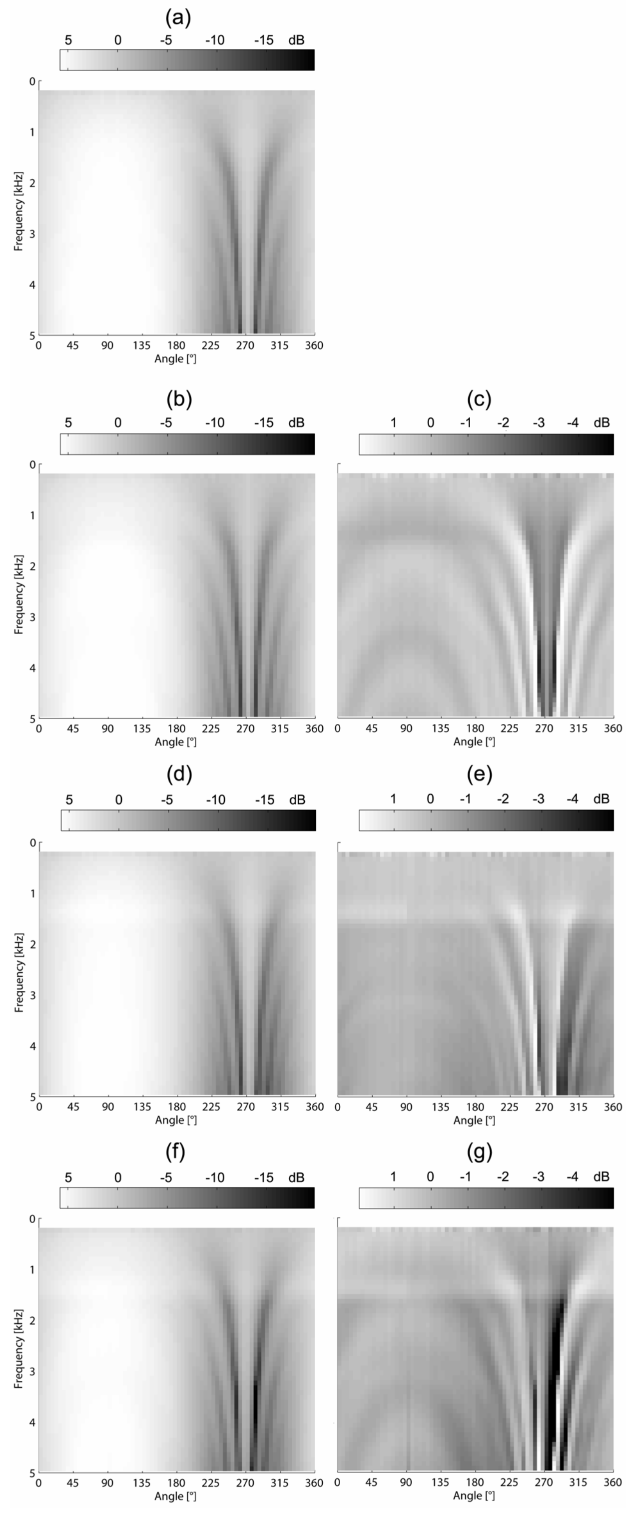

Figure 4. *Experimental HRTF measurements* (a) *sphere* (b) *sphere with neck* (c) *change due to neck* (d) *sphere with hair* (e) *change due to hair* (f) *sphere with neck and hair* (g) *change due to neck and hair.*

contralateral bright spot. A detailed explanation of these phenomena from a description of the interaction of the incident wave with the spherical surface is somewhat complex. Both the real and imaginary parts of the impedance of the hair covering dictate distinct changes to the surface pressure [12]. The non-uniformity of the surface boundary also introduces additional coupling between incoming and scattered wave modes that is not present for uniformly covered or rigid spheres. A further explanation of these properties is difficult without mathematical accompaniment and this is beyond the scope of the current investigation.

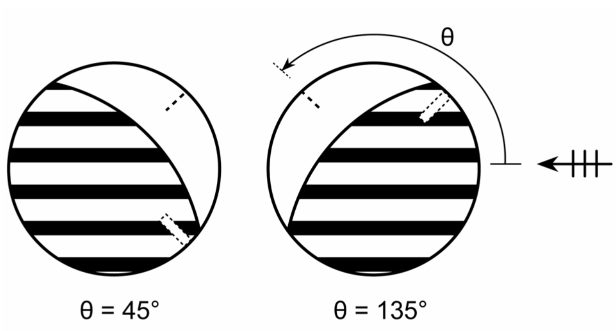

Figure 5: *Overhead view of the head surface showing the rigid facial (unshaded) and hair (shaded) regions. The dashed rectangle illustrates the heads right ear (corresponding to the HRTF shown in Fig. 4). Two angles of sphere rotation are shown that result in the principal wave path to the ear to be over either a primarily rigid or primarily absorbent sphere surface.*

The changes due to the addition of both the neck and hair are shown in Fig. 4(g). The primary changes to the HRTF are well explained by the individual contributions of the neck and hair as previously discussed. A reduction in the contralateral bright spot along with ancillary arch shaped notches can be seen due to the addition of the neck. The other asymmetrical HRTF changes are consistent with the discussion on the effects of the hair covering, both for ipsilateral and contralateral angles. Any differences may be attributed to experimental errors in maintaining exactly consistent rotation angles, and any coupling between the scattering effects of the neck and hair that does not appear when they are investigated separately.

4. DISCUSSION AND CONCLUSION

Overall both the neck and the hair produce small but distinctive changes to the azimuthal HRTF. The addition of the cylindrical neck disturbs a proportion of the symmetrical wave path to the posterior region of the sphere. This reduces the magnitude of the bright spot in the order of 2-4 dB. The neck also provides ancillary arch shaped notches for ipsilateral angles that are similar to those provided by a geometric torso. The addition of the hair material produces reasonably complex asymmetrical changes to the HRTF. For ipsilateral angles a general reduction is seen which is slightly greater for angles where the hair surface is moved to the frontal hemisphere. Asymmetrical changes in the order of 2-4 dB are also seen for contralateral angles, particularly near the bright spot. These result in the anterior ridge of decreased pressure (adjacent to bright spot) being noticeably more apparent than the posterior ridge.

Although only azimuthal HRTF are investigated, given the 45° alignment of the hair covering changes in the frontal plane HRTF due to hair may be considered analogous to those discussed here (with the appropriate angular remapping). Changes of similar order would be expected for median plane HRTF, particularly for rear sources. The discussion given for the asymmetrical ipsilateral HRTF changes is particularly relevant with regards to the effect of hair in the median plane.

The broad spectral changes exhibited due to the addition of the hair surface used in the current work are consistent with those presented previously. The addition of a wig to a mannequin head investigated by Riederer [11] produces similar asymmetric changes to those discussed here. For the right ear this corresponds to a decrease in the HRTF for anterior contralateral angles, and a slight increase for posterior. Whilst modifying the hair does not produce highly idiosyncratic HRTF changes, untrained listeners are able to perceive a difference when HRTFs are used that have substantial hair covering modifications [11].

In the absence of methodical and robust psychoacoustic studies it remains difficult to comment on the perceptual effects of the hair and neck to sound localisation. Obviously there are a multitude of environmental and intrapersonal anthropometric variations that cause subtle changes to the waveform present at the eardrum for a particular source direction. Variations in clothing, headwear, and complex room interactions all produce changes in the spectral HRTF features, but this does not necessarily mean that the ability to localise sounds is significantly perturbed. The addition of the neck instinctively reduces the bright spot that is over exemplified by spherical approximations of human hearing. Additionally the asymmetrical changes provided by the hair may contribute as a secondary cue to the discrimination of front-back ambiguities, particularly in the absence of other high frequency spectral changes provided by the pinnae. If these asymmetries could be modelled and easily replicated, their inclusion with binaural synthesis may resolve some of the directional ambiguities that currently afflict sphere scattering models assuming a rigid head surface.

5. ACKNOWLEDGEMENTS

The first author would like to acknowledge the assistance of the Robert and Maude Gledden, and F S Shaw Memorial Postgraduate Scholarships.

6. REFERENCES

[1] R.S. Woodworth and H. Schlosberg, *Experimental Psychology*, 2nd edn, Henry Holt and Company, Inc., New York, 1954.

[2] D.J. Kistler and F.L. Wightman, "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction," *J. Acoust. Soc. Am.*, vol. 91, no. 3, pp. 1637-1647, 1992.

[3] A. Kulkarni, S.K. Isabelle and H.S. Colburn, "On the minimum-phase approximation of head-related transfer functions," in *IEEE Workshop on Applications of Signal*

Processing to Audio and Acoustics, IEEE, New Platz, New York, pp. 84-87, 1995.

[4] G.F. Kuhn, "Model for the interaural time differences in the azimuthal plane," *J. Acoust. Soc. Am.*, vol. 62, no. 1, pp. 157-167, 1977.

[5] V.R. Algazi, R.O. Duda and D.M. Thompson, "The use of head-and-torso models for improved spatial sound synthesis," in *Proceedings of the 113th AES Convention*, Audio Engineering Society, Los Angeles, CA, preprint 5712, 2002.

[6] D.N. Zotkin, J. Hwang, R. Duraiswami and L.S. Davis, "HRTF personalization using anthropometric measurements," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, IEEE, New Paltz, NY, pp. 157-160, 2003.

[7] C.P. Brown and R.O. Duda, "A structural model for binaural sound synthesis," *IEEE Trans. Speech Audio Process.*, vol. 6, no. 5, pp. 476-488, 1998.

[8] B.G. Shinn-Cunningham, S. Santarelli and N. Kopco, "Tori of confusion: Binaural localization cues for sources within reach of a listener," *J. Acoust. Soc. Am.*, vol. 107, no. 3, pp. 1627-1636, 2000.

[9] V.R. Algazi, R.O. Duda, R. Duraiswami, N.A. Gumerov and Z. Tang, "Approximating the head-related transfer function using simple geometric models of the head and torso," *J. Acoust. Soc. Am.*, vol. 112, no. 5, pp. 2053-2064, 2002.

[10] R.O. Duda, C. Avendano and V.R. Algazi, "An adaptable ellipsoidal head model for the interaural time difference," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, IEEE, Phoenix, Arizona, pp. 965-968, 1999.

[11] K.A.J. Riederer, *HRTF analysis: Objective and subjective evaluation of measured head-related transfer functions* in Electrical and Communications Engineering, Helsinki University of Technology, Espoo, 2005.

[12] B.E. Treeby, R.M. Paurobally and J. Pan, "The effect of impedance on interaural azimuth cues derived from a spherical head model," *J. Acoust. Soc. Am.*, vol. 121, no. 4, pp. 2217-2226, 2007.

[13] B.F.G. Katz, "Boundary element method calculation of individual head-related transfer function. II. Impedance effects and comparison to real measurements," *J. Acoust. Soc. Am.*, vol. 110, no. 5, pp. 2449-2455, 2001.

[14] M.D. Burkhard and R.M. Sachs, "Anthropometric manikin for acoustic research," *J. Acoust. Soc. Am.*, vol. 58, no. 1, pp. 214-222, 1975.