An evaluation of 3D printing for the manufacture of a binaural recording device

Article *in* Applied Acoustics · October 2020

DOI: 10.1016/j.apacoust.2020.107610

CITATIONS
13

READS
494

2 authors:

John Kennedy
Trinity College Dublin
121 PUBLICATIONS **1,685** CITATIONS
SEE PROFILE

Daragh Oconnor
Delft University of Technology
2 PUBLICATIONS **45** CITATIONS
SEE PROFILE

Applied Acoustics 171 (2021) 107610

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

Technical note

An evaluation of 3D printing for the manufacture of a binaural recording device

Daragh O'Connor, John Kennedy *

School of Engineering, Trinity College Dublin, Ireland

ARTICLE INFO

Article history:
Received 28 April 2020
Received in revised form 25 July 2020
Accepted 17 August 2020
Available online 2 October 2020

Keywords:
3D printing
Spatial or Binaural Recording
Open-Source

ABSTRACT

With increasing availability of low-cost, high fidelity 3D printers, complex designs which previously proved costly to manufacture can be made readily available for research and public engagement. This work reports on the development of a binaural recording device which is tested for compatibility on 4 commercial 3D printers, each offering different print capabilities. The primary objective is to demonstrate the potential for open access designs that can be reliably 3D printed but maintain a high fidelity on the recorded HRTF. The influence of both the overall manufacturing strategy and pinna print material were investigated. A dummy head was printed and used to generate a database of HRTF's in the horizontal plane. The HRTF's are compared per material and relative effects are observed. A numerical simulation for the rigid boundary case provides validation data and serves as verification of the manufacturing strategy and acts as a baseline to inspect influence of pinna print material.

1. Introduction

Using both ears the human brain processes both monaural and interaural cues resulting from the scattering of sound by the head and torso in order to localise the source of a sound. A common method to replicate these cues for playback has been to build a dummy head and torso simulator or HATS, and place a microphone at each ear canal to replicate the effect the human anatomy has on the sound field. The sound recordings taken are then suitable for presentation to human listeners with increased realism due to the 3D spatialized experience. This has led to binaural audio being used for environmental noise assessments [1,2] and various virtual reality applications[3].

With increased the popularity and availability of 3D printing, an open source framework for individualised binaural presentation may prove beneficial.

The Head Related Transfer Function (HRTF) describes these spectral cues received by the brain. These monaural cues are then combined into the inter-aural time (ITD) and level differences (ILD) that allow for the localisation of a source.

Given that these vary with anthropometry a single HATS will only satisfy a portion of the population who match the HATS suitably.

The first step in the generation of a binaural recording system is the definition of the head and torso geometry. The objective being to best match the target audience recognising that this will change as a function of age, sex and ethnicity.

After initial attempts were made to refine the primary parameters influencing the HRTF, Hu et al. [4] proposed a method were correlation analysis was performed upon the CIPIC database where 27 measurements were reduced to 8 primary parameters. Two stages of linear regression analysis were performed on a HRTF and then personalization was performed by correlation analysis. Bomhardt et al. [5] later attempted to compare reconstructed HRTF's with individualised versions to observe the spectral differences observed per subject. They go on to discuss the number of necessary components where it was shown that "anthropometrically-estimated HRTF's with six dimensions differ only slightly compared to the ones with thirteen dimensions." [5] Fels [6] has published many other studies including work into using the BEM to simulate the HRTF's similarly to the method used in this work.

Zhang et al. [7] took a slightly different approach, they investigated the relative effect of each parameter of the CIPIC database.

In 2010, Hugeng et al. [8] proposed a strong argument for their improved method to further reduce these parameters. They developed a process which improved upon their previous results and results in the literature at the time.

Seeing as the chosen parameters were of lesser importance to the primary goal of developing a framework for 3D printed binaural technology, the parameters defined by Hugeng et al. were

* Corresponding author.
E-mail address: jkenned5@tcd.ie (J. Kennedy).

https://doi.org/10.1016/j.apacoust.2020.107610
0003-682X/

deemed sufficient for this study. Although it should be noted that Ghorbal et al. [9] published a comprehensive review of the different methods which have been implemented in the search for these critical parameters and have highlighted issues in the work done by Hugeng et al.

There is no clear agreement within the literature at present to define the most influential parameters and many approaches, based on various data sets, have failed to result in a conclusive answer.

Within this work Section 2 will focus on the process of adapting the head and ear morphology along with a description of steps to allow for print-ability on a range of commercial 3D printers.

Section 3 will discuss the steps undertaken to experimentally measure the HRTF of the printed device. The method described by Algazi et al. [10] to create the CIPIC database of impulse responses was adapted for use in this testing. The CIPIC database contains left and right ear impulse responses for 1250 different spherical locations centred about a subject. Section 3 will also define the steps taken to simulate the experimental setup in COMSOL multiphysics. The Boundary Element method was used as per guidelines outlined by Katz [11]. Katz [12] indicates that both pinna and hair impedance is required when modelling a real human HRTF but with skin having an average reflection coefficient of 0.97 [13] it could likely be modelled as rigid. Given this, a hard boundary wall model was determined to be suitable as a benchmark for the performance of the manufacturing strategy and the investigation into the effect of pinna print material.

2. Design

2.1. Morphological adaptation

The objective was to create human head model according to anthropometric parameters defined by Hugeng et al. [8]. Table 1 shows the mean values between male and female dimensions from various different sample populations and will act as a guideline to future work where alternate anthropometric data is available.

CAD models of a human head and ear from the online community CG Trader formed a basis for the design. The mesh files were transformed according to the dimensions from Table 1 in meshing software Blender, special attention was paid towards maintaining realistic curvature. The design was limited to above the shoulders due to the size constraints of standard 3D printers. Note that the absence of the torso will likely effect the performance of the head for locating sound sources of varying elevation and should be considered if this is a required for the end application. The microphone diaphragms are inclined to match the morphology of the ear models chosen.

2.2. Materials

The objective here was to choose a range of printable materials to investigate the influence they have on the HRTF. The design features a rigid head component combined with flexible ears using materials available for low cost commercial 3D printers. Polylactic acid (PLA) is the standard choice for non-stretch parts that do not require high temperatures or stresses and was chosen as the primary material used for the head. The printer used was the Creality Ender 3, an FDM printer with a build volume of 230 mm × 230 mm × 250 mm which was sufficient to print the head in two halves. There are numerous options for the flexible ears and the following were investigated (in order of increasing cost): Thermoplastic polyurethane (TPU) printed using the Creality Ender 3, PhotoCentric 3D Flexible resin for the Anycubic Photon, MonoFlex 100 resin for the Prusa SL1 printer and the Formlabs Photopolymer Flexible Resin for the Formlabs Form 2 printer.

Table 1
Eight primary anthropometric parameters and corresponding mean male–female measurements.

Parameters [8]	Mean Male–Female Dimensions (mm)
Head Width	149.08 [14]
Head Length	193.71 [14]
Neck Width	111.5 [15]
Shoulder Width	N/A
Ear Canal Height	17.05 [16]
Concha Width	15.88 [17]
Ear Length	59.45 [17]
Ear Width	30.88 [17]

2.3. Design for manufacture

The aim was to create a design following an open-source philosophy that could be printed reliably on a range of commercial 3D printers. To address common print failures such as shrinkage and expansion of large parts a custom brim was added onto the head component to widen the footprint of the first deposited layer. This can been seen in Fig. 1 where chequered material was removed post print. Threaded inserts and groves were added to align and mount components in position. The two ear components are designed to hold standard 1/4 inch microphones with a small form factor to be contained within the head.

3. Methods

The ear is known to begin affecting the HRTF at frequencies of 4 kHz and above [18], hence to fully capture the influence of pinna material the HRTF was inspected from 220–16000 Hz.

3.1. Experimental

The tests performed in this procedure were limited to the horizontal plane at 50 locations which could be directly compared to azimuth angles chosen by Algazi et al. An impulse response test was performed by playing white noise from a loudspeaker over a 15 s interval at 44.1 kHz, while synchronously recording the output of each microphone. The microphone signal is deconvolved with the source signal sent to the loudspeaker to obtain the impulse response. The microphone data were acquired through a National Instruments DAQ and the signals were recorded in Matlab. In this case the head and ears were instrumented using GRAS 40PL microphones as they have a frequency response of ±1 dB within 50–5,000 Hz, ±2 dB within 5–20 kHz and upper limit of the dynamic range of 150 dB re 20 μ*Pa*. A Genelec 8020B loudspeaker with frequency response of ±2.5 dB within 66–20,000 Hz was placed 1 meter from the inter-aural axis of the dummy head, in a large, non-anechoic room shown in Fig. 2. In order to insure the validity of the test results the room radius, the location at which the direct and reverberant fields are equal, was measured as a function of octave band and reported in Table 2. Tests were performed at a 1 m distance with the loudspeaker playing at approximately 115 dB, as measured 10 cm in front of the speaker, to ensure that levels were significantly above the background. There is a compromise to be made between near field directivity effects of the loudspeaker and entering the reverberant field at a greater distance from the speaker. The 1 m distance is at or within the room radius for all frequencies above 125 Hz. This insures that the direct field is at a greater amplitude than the reverberant for all test frequencies while remaining sufficiently far to avoid the near field effects. The fundamental room mode is at 23 Hz meaning that only high order room modes exist at the test frequencies. Porous absorbing panels

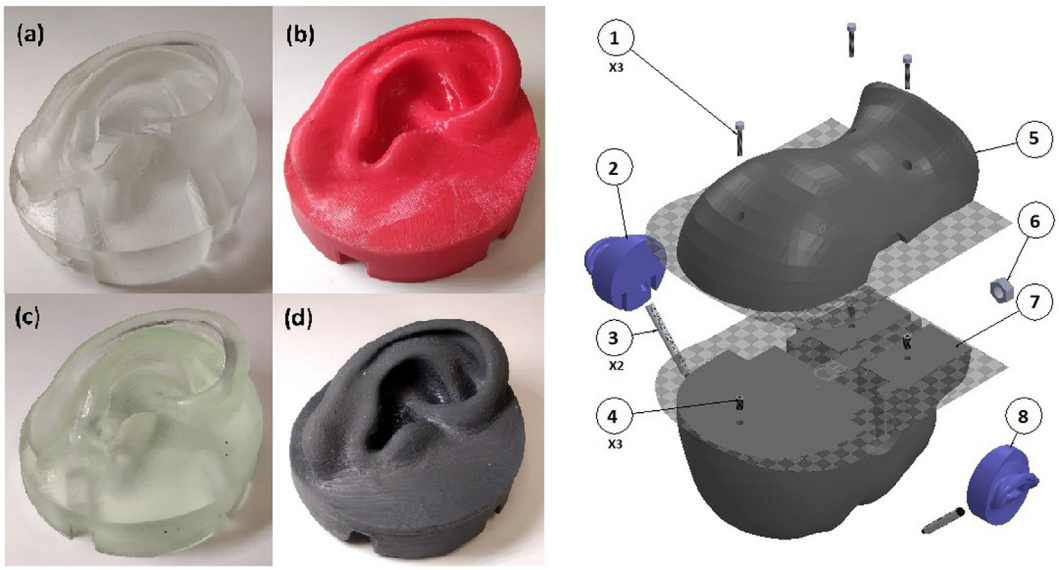

Fig. 1. Left: 3D Printed ear models. (a) Monoflex 100 (b) TPU (c) PhotoCentric Flexible (d) Formlabs Flexible. Right: exploded view of head: (1) M4 Bolts, (2) Ear left, (3) Microphones, (4) M4 Threaded Inserts, (5) Head Back, (6) M9 Nut, (7) Head Front, (8) Ear Right.

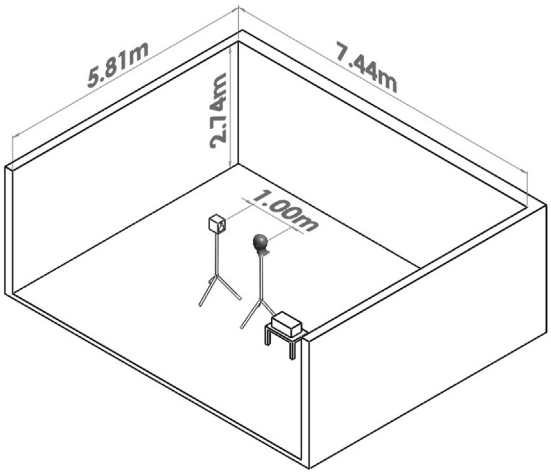

Fig. 2. Experimental setup.

placed at irregular locations were added to the test room during the experimental campaign in order to further reduce the influence of any room modes.

The test rig rotates to the chosen azimuth increments using an automated stepper motor and measurements were taken at each location. Fig. 3(a) below shows the estimated arrival time of the direct sound and the first reflection based on a calculation of the distance and speed of sound. For the calculations the onset of the impulse is determined by a threshold value above the background noise. As can be seen the estimated arrival time and the detected impulse align well. There is no position in which the first reflection arrives within the 200 points from the onset threshold location. Smoothing was required in order to estimate the SNR of the impulse response. This was achieved using a Hilbert envelope computation and shown in the Fig. 3(b). The background noise level in the room is over 60 dB below the peak value of the impulse. At the point of the truncation of the impulse the level has decayed 40 dB from the peak value. Again the arrival of the first reflection can be clearly seen beyond the truncated region Fig. 3(b).

Table 2
Measured values of the room radius as a function of octave band.

Frequency	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
Room Radius	107 cm	104 cm	100 cm	125 cm	159 cm	177 cm	180 cm

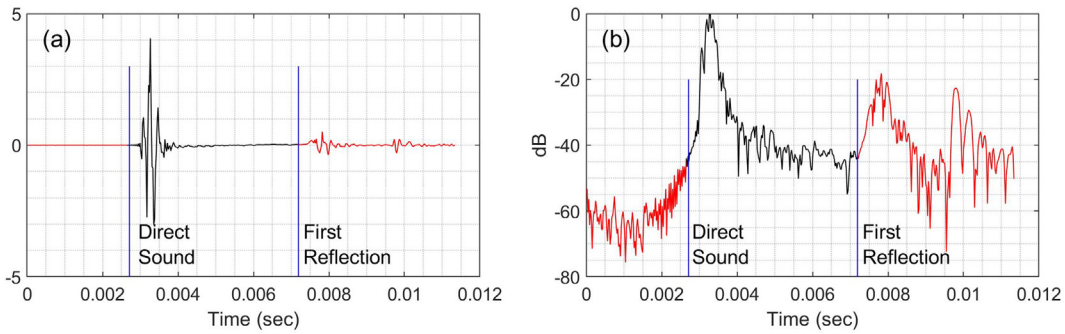

Fig. 3. Truncated region of 200 points shown in black (a) impulse response (b) signal to noise ratio.

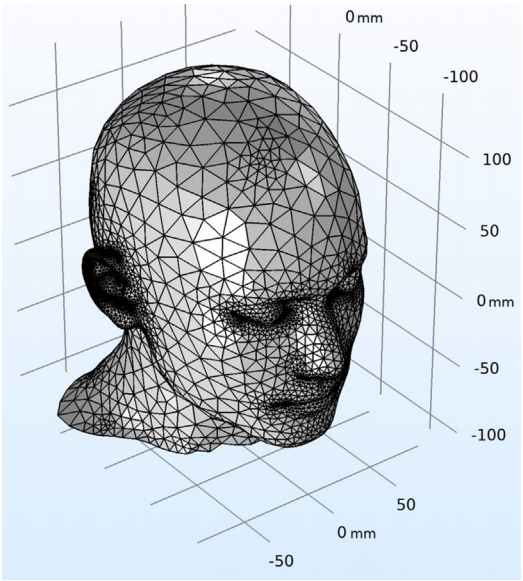

Fig. 4. Surface mesh for boundary element method at 1102.5 Hz.

Each of the 4 materials chosen for the ears were tested 3 times for repeatability over 3 separate days where the experimental setup was dismantled and reconstructed each time. The HRTF was then calculated using the impulse to frequency response function provided in the CIPIC database[10], an NFFT equal to the sample length was applied. The ILD was computed as the difference between the HRTF's of the left and right channels and the ITD as the difference in arrival time between the HRIR's of both channels.

3.2. Numerical simulation

Numerical simulation in was setup in the commercial software Comsol v5.4 as the package allows for easy importing and meshing of complex.stl files. The geometry and domain is displayed in Fig. 4.

The pressure acoustic boundary element interface that is based on the Boundary Element method (BEM) solved with the generalized minimal residual iterative solver was implemented [19]. Both the method and solver were utilised for their ability to handle large problems. The method uses the Helmholtz equation in the form of:

$$-\frac{1}{\rho_c}\nabla^2 p_t - \frac{k_{cg}^2}{\rho_c} p_t = 0 \tag{1}$$

$$p_t = p + p_b \tag{2}$$

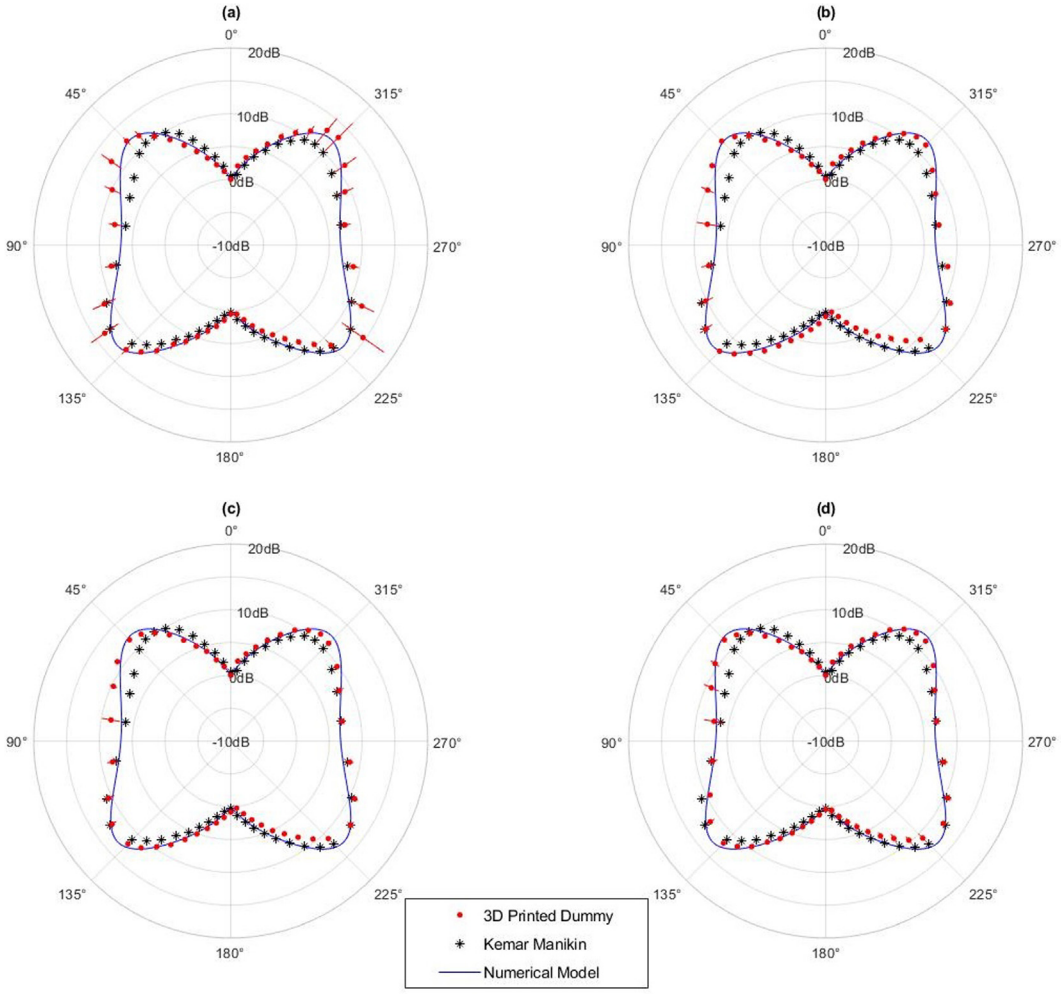

Fig. 5. ILD's in the horizontal plane at 1102.5 Hz. Mean Experimental measurements for four pinna materials: (a) MonoFlex 100, (b) PhotoCentric Flexible, (c) TPU and (d) Formlabs Flexible versus numerical model and data from the CIPIC database [10] for the Kemar manikin.

$$k_{cq}^2 = \left(\frac{\omega}{c_c}\right)^2 \tag{3}$$

where p_t is the total acoustic pressure, k_{eq} is the wave number, ρ_c is the density and c_c is the speed of sound. The subscript c denotes that the values can have complex components for a model that includes dampening. A far field approximation is used for matrix assembling as interactions defined in the near field are expressed explicitly while those in the far field can be approximated as an outgoing wave. To save computation time the reciprocal method is employed, a harmonically oscillating inward velocity boundary condition is applied at the entrance to the ear canal defined as:

$$-\mathbf{n}\cdot\left(-\frac{1}{\rho_c}\nabla p_t\right) = i\omega v_n \tag{4}$$

where ω, the angular frequency, is the dependant variable of a parametric sweep from 220.5 Hz to 16097 Hz in increments matching those generated by the inverse fourier transform performed on the truncated impulse responses described earlier. The surface of the model is treated as a hard sound boundary wall. Air was used as the medium with a density of 1.204 kg/m^3. An unstructured triangular adaptive mesh was used where the maximum node size is defined as $\lambda/6$ on the basis that six cycles of a waveform is adequate to describe the period of an acoustic wave [11]. The minimum size for the mesh was 0.35 mm. The initial step solves for the potentials for every element such that all source and boundary conditions are met. Using these results the pressure field at 1 m radius in the horizontal plane at 360 locations is calculated through the superposition of all the contributions [11]. The HRTF is then simply the ratio of the sound pressure level at the ear canal entrance relative to the quiescent sound pressure level, when the head is not present. Finally for the ILD, the HRTF was assumed symmetric and the same process was applied as for the experimental results.

4. Results and discussion

By their nature the experimental results have a finer frequency resolution for each of 50 source locations tested while the numerical results are based on 73 iterations in the frequency domain but have a much finer spatial resolution. All experimental data are presented as the mean and standard deviation over 3 days of testing. The ILD and HRTF data is normalised such that source location at 0°, the position directly ahead of the manikin, has a magnitude of 0 dB.

Figs. 5 and 6 shows a near exact match between the experimental and numerical values of the printed head's ILD in the horizontal plane. Good symmetry is observed for experimental ILD data. The material chosen for the pinna component appears to have little discernible effect at these frequencies as predicted by[18]. The

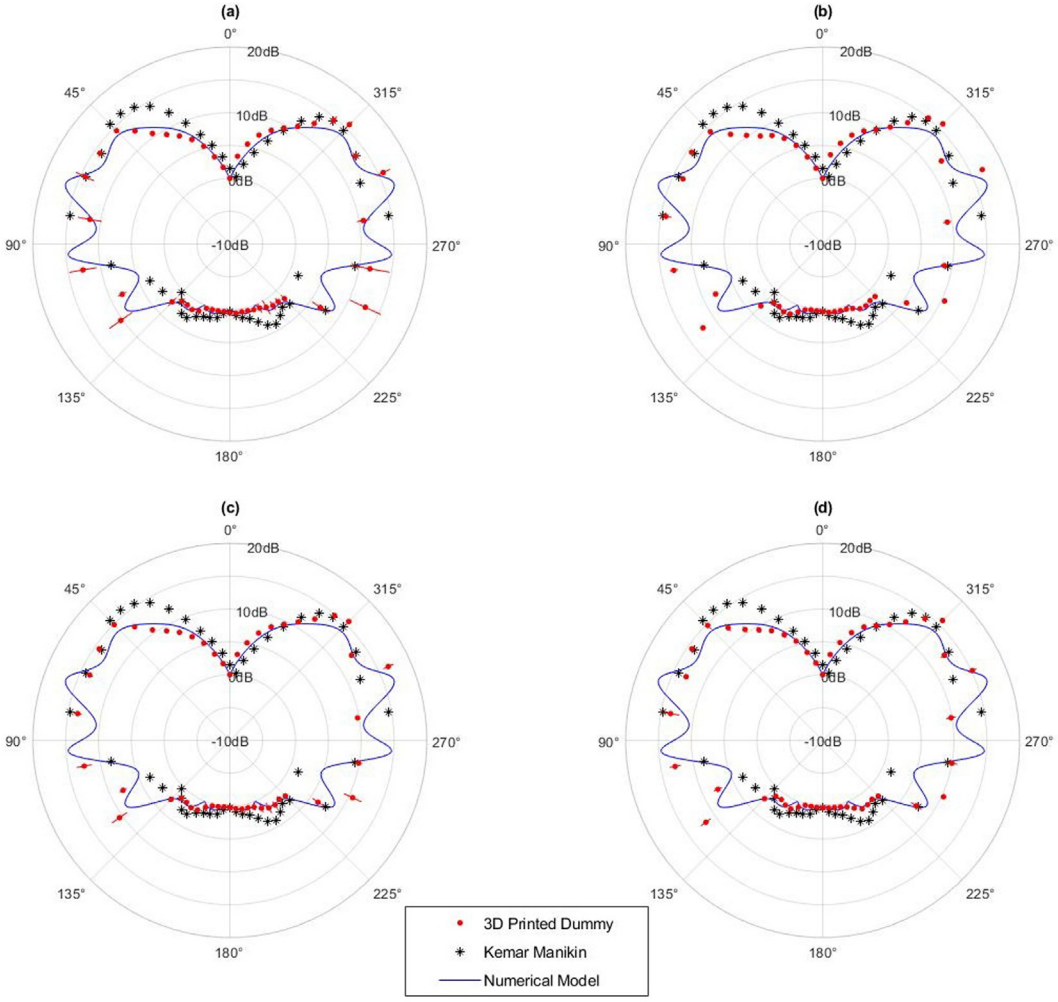

Fig. 6. ILD's in the horizontal plane at 3969 Hz. Mean Experimental measurements for four chosen pinna materials: (a) MonoFlex 100, (b) PhotoCentric Flexible, (c) TPU and (d) Formlabs Flexible versus numerical model and data from the CIPIC database [10] for the Kemar manikin.

dummy head shows very similar behaviour to the Kemar manikin's data extracted from the CIPIC database.

Fig. 7 shows excellent agreement between experimental and numerical HRTF values across the 6 values chosen near octave band centre frequencies displayed for the MonoFlex ears. For source positions behind the ear attenuation starts to be observed in the numerical data nearing 4 kHz and above. These deviations from the experimental results show the hard sound boundary may cause issues when simulating the HRTF at frequencies above 4 kHz, especially for positions behind and to the far side of the head.

For each of these frequencies the median and maximum deviation between the numerical and experimental results has been calculated. The angle at which the maximum deviation occurred is also noted. This information is presented in Table 3, which includes the results for each material. Generally there is excellent agreement between the numerical and experimental results with a median deviation of less than 1 dB for all frequencies up to 4 kHz and all materials. At high frequencies greater deviations occur with the Monoflex material having the lowest median deviations across the full frequency range. The maximum deviations occur for angles between 220° and 330°, when the sound source is on the opposite side of the head. Significant deviations of greater than 5 dB begin to occur at frequencies above 4 kHz with the Monoflex material again providing the closest match to the numerical results. From these results it can be seen that the material choice for the ear is not a dominant factor at frequencies below 4 kHz. At higher frequencies the deviations may also be due to small geometric inaccuracies in addition to the material choice. The geometric accuracy of the 3D printed part is dependent on both the process and material. As each ear model was manufactured using a different process it is not possible to determine if the material choice or geometric accuracy is the main source of these deviations at higher frequencies.

Experimental ITD measurements for the printed head, reported in Fig. 8, show good agreement with those of the Kemar manikin in the horizontal plane. Pinna material does not have a strong effect on the ITDs as values agree within experimental uncertainty.

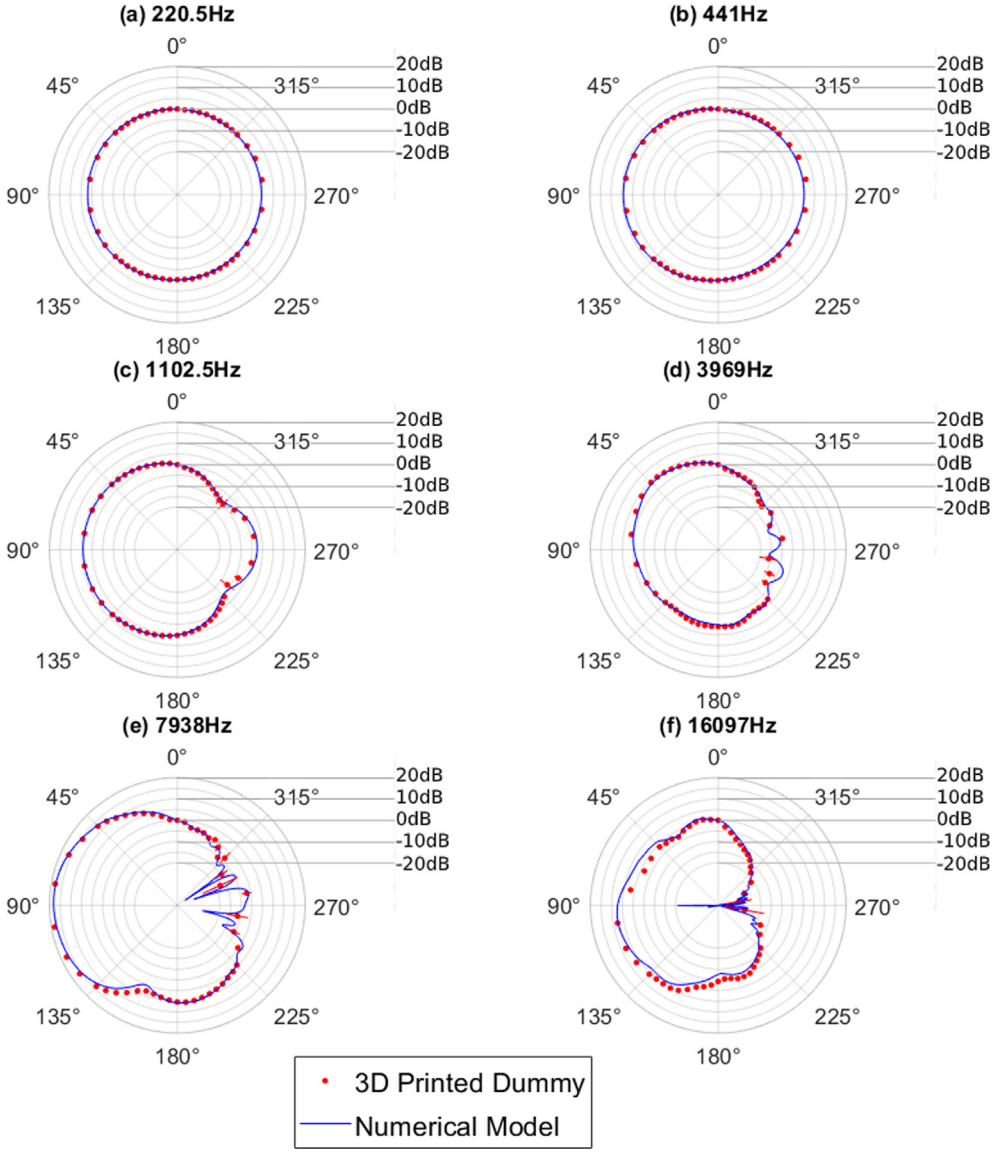

Fig. 7. HRTF for left ear in the horizontal plane for MonoFlex ear components. Mean Experimental measurements for selected 6 frequencies: (a) 220.5 Hz, (b) 441 Hz, (c) 1102.5 Hz, (d) 3969 Hz, (e) 7938 Hz and (f) 16097 Hz versus numerical model.

Table 3
The median deviation, maximum deviation and angular location of the maximum deviation between the numerical and experimental results for each pinna material.

	FormLabs			Monoflex		
Frequency (Hz)	Median (ΔdB)	Max (ΔdB)	Angle (°)	Median (ΔdB)	Max (ΔdB)	Angle (°)
220.5	0.2	1.1	330.0	0.2	1.3	245.0
441.0	0.2	1.3	280.0	0.4	1.9	280.0
1102.5	0.2	1.2	225.0	0.3	2.9	245.0
3969.0	0.6	5.0	245.0	0.7	6.7	245.0
7938.0	3.4	9.5	295.0	0.5	9.2	260.0
16096.5	0.6	19.8	280.0	1.7	15.7	280.0
	Photocentrix			TPU		
Frequency (Hz)	Median (ΔdB)	Max (ΔdB)	Angle (°)	Median (ΔdB)	Max (ΔdB)	Angle (°)
220.5	0.2	1.6	245.0	0.3	1.5	200.0
441.0	0.2	1.7	280.0	0.2	1.4	260.0
1102.5	0.3	1.6	220.0	0.3	1.4	220.0
3969.0	0.7	5.7	245.0	0.5	4.7	245.0
7938.0	1.6	13.2	295.0	1.2	10.1	295.0
16096.5	1.8	20.6	280.0	1.7	18.8	280.0

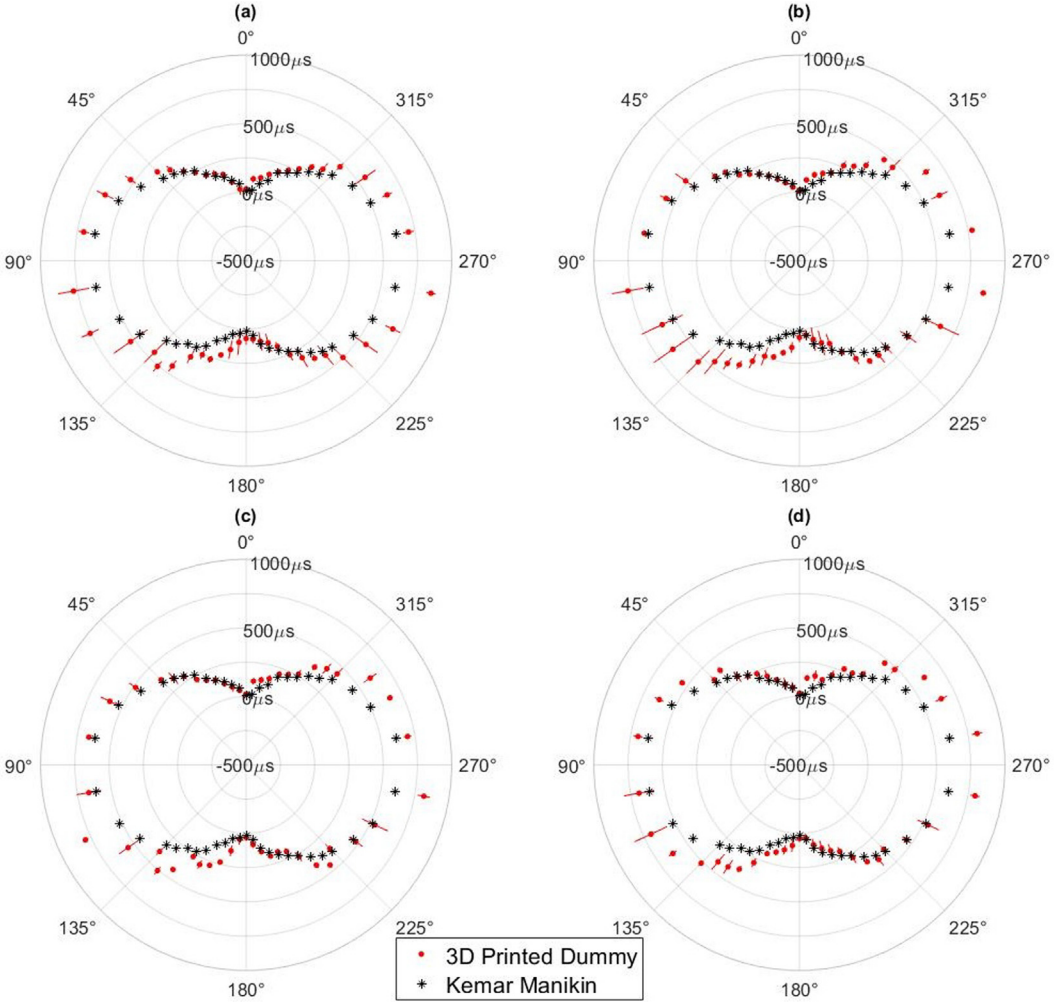

Fig. 8. ITD's in the horizontal plane for the four print materials and the Kemar manikin. Mean experimental measurements for four chosen pinna materials: (a) MonoFlex 100, (b) PhotoCentric Flexible, (c) TPU and (d) Formlabs Flexible versus data from the CIPIC database [10] for the Kemar manikin.

5. Conclusion

In conclusion, a binaural device was successfully designed for manufacture on a range of 3D printers. The device performed well in the range of 220 Hz to 4 kHz across the four printers tested. Excellent agreement was obtained between numerical results and experimental values determined in a non-anechoic space. Deviations of less than 1 dB were observed up to 4 kHz which implies high fidelity binaural recordings can be achieved with low cost 3D printed parts. At higher frequencies significant deviations between the numerical and experimental results were observed for angles between 220 and 330°, with the source on

the far side of the head. The numerical HRTFs in this region are complex in character and minor geometric inaccuracies in the test setup could contribute to these deviations. The 4 print processes used have different geometric accuracies and further work would be required to determine if the observed deviations are due to differing material properties or geometric differences. The results indicate that an authentic replication of the ear's acoustic properties is essential for accurate spatial audio recording at high frequencies above 4 kHz. Future work should aim to develop a more skin-like 3D printable material which can be printed on commercially available printers with a high geometric accuracy. With more realistic materials higher fidelity reproduction of spatial audio will be made available for research and public engagement. The excellent agreement between the developed system and the BEM simulations up to 4 kHz demonstrates the validity of the system for many acoustic problems. At present the developed system is being used by the authors for subjective listening tests in the context of environmental noise annoyance. Mesh files for the printable head are available in the Supplementary material along a database for all tests performed in this report.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, athttps://doi.org/10.1016/j.apacoust.2020.107610.

References

[1] Brambilla G, Maffei L. Responses to noise in urban parks and in rural quiet areas. Acta Acust United Acust 2006;92(6):881–6.
[2] Kang J, Schulte-Fortkamp B. Soundscape and the built environment. Boca Raton: CRC Press; 2018.
[3] Lentz T. Binaural technology for virtual reality. J Acoust Soc Am 2008;124(6):3359. 3359.
[4] Hu H, Zhou L, Zhang J, Ma H, Wu Z. Head related transfer function personalization based on multiple regression analysis. In: 2006 International conference on computational intelligence and security. Vol. 2; 2006. p. 1829–1832.
[5] Bomhardt R, Braren H, Fels J. Individualization of head-related transfer functions using principal component analysis and anthropometric dimensions. Proc Meet Acoust 2016;29(1). 050007.
[6] Fels J, Vorlaender M. Anthropometric parameters influencing head-related transfer functions. Acta Acust United Acust 2009;95:331–42.
[7] Zhang M, Kennedy R, Abhayapala T, Zhang W. Statistical method to identify key anthropometric parameters in hrtf individualization, 2011.
[8] Hugeng WW, Gunawan D. Improved method for individualization of head-related transfer functions on horizontal plane using reduced number of anthropometric measurements. CoRR abs/1005.5137.
[9] Ghorbal S, Seguier R, Soladié C, Auclair T. Pinna morphological parameters influencing hrtf sets. In: 20th International conference on digital audio effects - + DAFx 2017; 2017.
[10] Algazi VR, Duda RO, Thompson DM, Avendano C. The cipic hrtf database. In: Proceedings of the 2001 IEEE workshop on the applications of signal processing to audio and acoustics (Cat. No. 01TH8575); 2001. pp. 99–102.
[11] Katz B. Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation. J Acoust Soc Am 2001;110:2440–8.
[12] Katz B. Boundary element method calculation of individual head-related transfer function. II. Impedance effects and comparisons to real measurements. J Acoust Soc Am 2001;110:2449–55.
[13] Treeby BE, Pan J, Paurobally RM. An experimental study of the acoustic impedance characteristics of human hair. J Acoust Soc Am 2007;122(4):2107–17.
[14] Young JW. Head and face anthropometry of adult U.S. Civilians; 1993.
[15] Vasavada A, Danaraj J, Siegmund G. Head and neck anthropometry, vertebral geometry and neck strength in height-matched men and women. J Biomech 2008;41:114–21.
[16] Ji X, Zhu Z, Gao Z, Bai X, Hu G. Anthropometry and classification of auricular concha for the ergonomic design of earphones. Human Factors Ergon Manuf Service Ind 28 (2);90–99.
[17] Engineer P, Japatti S, Tiwari MRBA, Siddegowda C, Hammannavar R. Anthropometric assessment of the normal adult human ear. Ann Maxillofac Surg 8 (1);2018:42–50. doi:10.4103/ams.ams_183_17.
[18] Spagnol S, Geronazzo M, Avanzini F. Fitting pinna-related transfer functions to anthropometry for binaural sound rendering. In: 2010 IEEE international workshop on multimedia signal processing. 2010. p. 194–199.
[19] Inc C. Comsol multiphysics reference manual version 5.4; 2018. URL www.comsol.com.