Sensitivity analysis of pinna morphology on head-related transfer functions simulated via a parametric pinna model

Peter Stitt, Brian F. G. Katz

Sensitivity analysis of pinna morphology on head-related transfer functions simulated via a parametric pinna model

Peter Stitt and Brian F. G. Katz

Citation: The Journal of the Acoustical Society of America **149**, 2559 (2021); doi: 10.1121/10.0004128
View online: https://doi.org/10.1121/10.0004128
View Table of Contents: https://asa.scitation.org/toc/jas/149/4
Published by the Acoustical Society of America

JASA ARTICLE

Sensitivity analysis of pinna morphology on head-related transfer functions simulated via a parametric pinna model

Peter Stitt[a] and Brian F. G. Katz[b]
Sorbonne Université, CNRS, UMR 7190, Institut Jean Le Rond d'Alembert, Lutheries-Acoustique-Musique, Paris, France

ABSTRACT:
The head-related transfer function (HRTF) defines the acoustic path from a source to the two ears of a listener in a manner that is highly dependent on direction. This directional dependence arises from the highly individual morphology of the pinna, which results in complex reflections and resonances. While this notion is generally accepted, there has been little research on the importance of different structural elements of the pinna on the HRTF. A parametric three-dimensional ear model was used to investigate the changes in shape of the pinna in a systematic manner with a view to determining important contributing morphological parameters that can be used for HRTF individualization. HRTFs were simulated using the boundary element method. The analysis comprised objective comparisons between the directional transfer function and diffuse field component. The mean spectral distortion was used for global evaluation of HRTF similarity across all simulated positions. A perceptual localization model was used to determine correspondences between perceptual cues and objective parameters. A reasonable match was found between the modelled perceptual results and the mean spectral distortion. Modifications to the shape of the concha were found to have an important impact on the HRTF, as did those in proximity to the triangular fossa. Furthermore, parameters that control the relief of the pinna were found to be at least as important as more frequently cited side-facing parameters, highlighting limitations in previous morphological/HRTF studies.
 https://doi.org/10.1121/10.0004128
(Received 4 November 2020; revised 8 March 2021; accepted 15 March 2021; published online 12 April 2021)
[Editor: James F. Lynch] Pages: 2559–2572

I. INTRODUCTION

A head-related transfer function (HRTF) is a direction-dependent filter that describes the acoustic path from a sound source to the two ears. In general, the signal at both ears will be different, and these differences (binaural cues) are used by the auditory system to determine the perceived direction of the sound source. The two principal binaural cues are interaural time difference (ITD) and interaural level difference (ILD) (Blauert, 1997). ITD and ILD are used to determine the lateral position of a sound source on a cone-of-confusion (Katz and Nicol, 2019). ITD is the dominant cue at lower frequencies (below approximately 1.4 kHz), while ILD and envelope ITD are dominant at frequencies with a wavelength comparable to, or shorter than, the diameter of the head, leading to ambiguous phase information. Spectral cues are used to evaluate the elevation of the sound source and to resolve so-called *front-back confusions*. Furthermore, dynamic head movements can also be used to resolve where on a cone-of-confusion a source is positioned (Wightman and Kistler, 1992). Perception of distance relies on these and additional cues, often integrating visual (Lewald and Guski, 2004; Rébillat *et al.*, 2012) and dynamic cues (Rébillat *et al.*, 2012) as well as environmental factors. For example, higher ratios of reverberant to direct sound can cause a sound to appear more distant, while spatially coherent reflections can improve localization (Nykänen *et al.*, 2013; Poirier-Quinot and Katz, 2021; Shinn-Cunningham, 2000).

The spectral cues of a HRTF are largely the product of filtering by the outer ears (pinnae). The acoustic waves are reflected from the complex shapes of the pinnae, causing peaks and notches in the acoustic spectrum as the direct sound and reflections interfere (Lopez-Poveda and Meddis, 1996). The peaks and notches arising from the pinna reflections are generally above 5 kHz, due to its size. HRTFs are individual specific, since pinnae are highly unique, like fingerprints (Anwar *et al.*, 2015). Spagnol *et al.* (2013) used geometric analysis to show that notches in the HRTF spectrum correspond to contours on the pinna. However, this analysis concentrated on sources in the median sagittal plane. They then used contour tracing on side-facing two-dimensional (2D) images to generate personalized HRTFs.

A HRTF can be obtained by acoustic measurements, placing microphones in the ear canals of the individual to be measured (Algazi *et al.*, 2001; Jin *et al.*, 2014; Warusfel, 2002). However, this method typically requires an anechoic chamber, specialized equipment and procedures, and time. This makes acoustic measurement of HRTFs highly impractical outside of scientific purposes. Another method for providing personalized HRTFs is through numerical

[a]ORCID: 0000-0002-7273-552X.
[b]Electronic mail: brian.katz@sorbonne-universite.fr, ORCID: 0000-0001-5118-0943.

simulations (Greff and Katz, 2007; Jin *et al.*, 2014; Katz, 2001a; Ziegelwanger *et al.*, 2015). The challenge for numerical simulations is the difficulty in capturing an accurate three-dimensional (3D) scan of the subject (Rugeles Ospina *et al.*, 2015). On top of this, the simulations can be costly in terms of required computational power, although this restriction is being eroded as computing power increases. Other methods for personalization, as opposed to individualization, exist that aim to provide a HRTF (whether processed or not) from a database that is rated or predicted to give good perceptual results (Katz and Parseihian, 2012; Schönstein and Katz, 2010; Zagala *et al.*, 2020).

Listening to a spatial sound scene using a HRTF, one's own or that from another individual, is achieved using binaural synthesis. This technique processes a monophonic sound source by the left and right ear filters corresponding to the desired source direction (e.g., Poirier-Quinot and Katz, 2018). Playback is over a pair of headphones. Using a non-individualized HRTF can cause perceptual problems, such as inaccurate localization (Wenzel *et al.*, 1993) or increased front-back confusions (Zahorik *et al.*, 2006). It is possible to adapt to non-individual HRTFs (Carlile *et al.*, 2014; Hofman *et al.*, 1998; Parseihian and Katz, 2012; Poirier-Quinot and Katz, 2020), but this can be very dependent on the individual, with some unable to adapt even after repeated exposure to training (Stitt *et al.*, 2019; Trapeau *et al.*, 2016).

Several attempts have been made at linking HRTFs to morphology (Algazi *et al.*, 2001; Rugeles Ospina *et al.*, 2015). These have tended to be based on HRTF measurements accompanied by morphological measurements of the subjects. The CIPIC (Algazi *et al.*, 2001) and LISTEN (Warusfel, 2002) databases gather HRTF measurements with morphological parameters, shown in Fig. 1. These parameters define distances, circumferences, and angles of the head, upper torso, and pinnae. These so-called *CIPIC parameters* are not assumed to be comprehensive, and it is possible they include elements that are not strongly linked

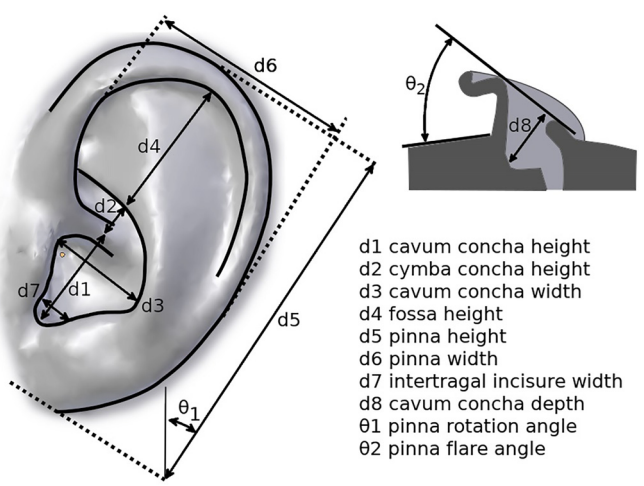

FIG. 1. (Color online) The pinna-related morphological parameters defined for the CIPIC HRTF database. After Algazi *et al.* (2001).

to the HRTF, as well as some that are very strongly linked. There have been attempts to determine the most pertinent parameters for characterizing the pinna (Fels and Vorländer, 2009; Ghorbal *et al.*, 2017; Huang and Li, 2014). Attempts have been made to find associations between morphological and HRTF signal domain parameters. For example, artificial neural networks have been used to generate personalized HRTFs using the listener's morphology data (Grijalva *et al.*, 2016).

For methods based on HRTF measurement databases, it is difficult or impossible to isolate the influence of changing individual or specific parts of the pinna on the HRTF. In contrast, the use of numerical simulations for generating HRTFs was originally envisioned with this goal in mind, over 20 years ago (Katz, 2001a). This study proposes the first evaluation of a parametric ear model, conceived to allow for controlled modifications of pinna morphologies, combined with state of the art computational hardware and software achieving simulations in a reasonable time-frame, for a broad ranging study employing generated HRTFs. The aim of the project is to investigate morphological changes in a systematic manner, by generating a set of HRTFs from the same base ear. This will allow for the most pertinent morphological parameters to be identified to facilitate a more detailed investigation on a reduced data space. A similar approach was taken by Ghorbal *et al.* (2017) and Guezenoc and Séguier (2020), who varied the CIPIC parameters by varying amounts. Their reference ear was generated by their model to have CIPIC parameter values equal to the average of the CIPIC database. While in theory this should give a spatially averaged ear, there is no guarantee that a pinna with a specific set of CIPIC parameter values will be unique. Rather than focusing directly on the impact of changes to the CIPIC parameters, this study investigates the effect of morphological changes by moving specific control points, many of which are linked to the CIPIC parameters.

Since morphological characterisation of pinnae contains a large number of parameters, it would be useful to determine which of these have the largest impact on the HRTF, both objectively and perceptually. This can perhaps allow a reduction in the number of relevant parameters used to characterize the pinna in relation to spatial hearing and the HRTF, facilitating improved HRTF personalization methods. As such, this paper investigates the global impact of the changing control points on the global HRTF to provide such an identification of pertinent parameters.

To provide a framework for the analysis of the results, the following hypotheses are tested:

H1: Relief (or depth) parameters, very often ignored in many morphological studies on HRTFs, strongly influence the spectral characteristics of the HRTF.
H2: Parameters closer to the ear canal entrance have stronger influence than those farther away (e.g., concha-related parameters will have more influence than helix).
H3: Parameters that change the volume of the concha will have the largest effect on changes to the non-directional diffuse field component of the HRTF.

https://doi.org/10.1121/10.0004128

This paper is structured as follows: Sec. II presents the parametric pinna model that was developed and the boundary element method (BEM) simulations used in the investigation. Section III presents the metrics used for the evaluation of the differences between HRTFs. The differences between different simulated HRTFs using these metrics are presented in Sec. IV. An identification of pertinent control points in terms of impact on the HRTF is carried out Sec. V. A discussion of the results is provided in Sec. VI, followed by the conclusions in Sec. VII.

Given the large number of terms and abbreviations used in this paper, a glossary of the most regularly used terms is included. Additionally, anthropometric features are indicated in Fig. 2(a).

- Concha: The cavity containing the ear canal made up of the cavum concha and cymba concha. Its front and rear walls are the tragus and antihelix respectively.
- Helix: The outer edge of the pinna.
- Triangular fossa: The indent in the upper portion of the pinna by the root of the helix at the front of the pinna.
- **CP*X***: a *control point* ***X*** on the parametric pinna model mesh that moves along a vector parallel to the median place.
- **CPd*X***: a *control point* ***X*** that controls the depth/relief of a point on the parametric pinna model mesh and moves along a vector perpendicular to the median place.
- **G*X***: a group of control points with grouping chosen based on the control point location or displacement type to more easily test the above hypotheses.

II. PARAMETRIC PINNA MODEL

A parametric pinna was created to explore the effect changes in morphology have on the acoustic characteristics of the HRTF when varying different parts of the pinna independently. This digital model was based on a high resolution scan of the Neumann KU-100 dummy head with the standard pinnae attached (Greff and Katz, 2007). Variations of the pinna were created by distorting the base mesh (using Blender 2.77). The pinna variation meshes were created from the full head-and-pinna mesh to avoid the need for manual mesh stitching to the head each time a new pinna was generated [see Fig. 2(b)]. The mesh variations for each control point were exported in .stl format and imported (using MATLAB 2019 b), where desired combinations can be generated.

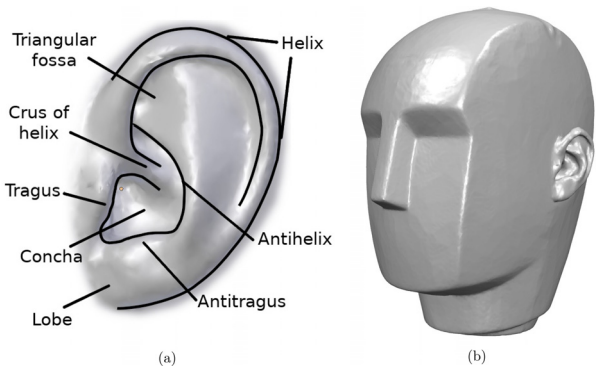

FIG. 2. (Color online) (a) Diagram of the pinna with anthropometric labels for regions referenced in this study. (b) The mesh of the KU-100 dummy head with a parametric model generated pinna.

For each parameter, a control point was selected and displaced such that it and the surrounding vertices in the mesh are moved along the vector defined by the corresponding CIPIC distance (see Fig. 1). Where the control point does not have an associated CIPIC distance, the direction is indicated in Table I. Control points are denoted as **CP*X*-pos** and **CP*X*-neg**, where ***X*** is the index of the control point and pos/neg indicates the magnitude of the displacement from the coordinate system origin. The positions of all defined control points on the base pinna are shown in Fig. 3.

In addition to the parameters moving the control points, there is a parameter that controls the overall scaling of the pinna size and another that controls the angle of rotation (CIPIC θ_1). The origin for the scaling and rotation is defined at the intersection between CIPIC d_1 and d_3. A control point movement is considered positive if it is moved away from the origin of the pinna-based coordinate system, i.e., the norm distance from the origin increases, regardless of the direction. For rotations, the positive direction is in the anticlockwise direction for the left ear viewed from the side.

As well as control points in the plane facing the ear, there are five control points that relate to the relief of the pinna. These control points move along the z axis (left-right) of the pinna-based coordinate system, and their displacement does not alter the shape of the pinna when projected onto the median plane. They are denoted **CPd*X***. Control points are only "fixed" in the axis or plane defined and are not restricted along other dimensions.

Note that control points might be associated with more than one CIPIC parameter, since the CIPIC parameters are not completely independent of one another. For example, distances d_1 and d_2 will change inversely if **CP2** moves. The control points for the parametric pinna model were chosen in such a way that, as much as possible, adjusting them would not change other manipulation points. Figure 4 shows two examples of different pinnae generated by the model. The large pinna has control points that have been moved in the positive direction for approximately half of the total range, and for the small one, they are moved in the negative direction by approximately half the range. These two pinnae are used to illustrate changes possible with the parametric model and its flexibility, particularly if combined with a global scaling parameter, in obtaining a potentially wide range of realistic pinna shapes. It should be noted that the ability to generate any specific arbitrary pinna is outside the scope of this study and will depend on the completeness of the control point parameters used. This will be examined in future work.

To facilitate discussion, control points have been categorized into three morphological groups:

G1: Control points that influence or are in the concha.

TABLE I. Control points of the parametric ear model and the resulting influence due to a change in position on CIPIC parameters. A positive movement of a control point is defined as a movement away from the pinna coordinate origin. The displacement axis of the control points is indicated relative to a coordinate system indicated by the dashed lines on Fig. 3. Control points with a displacement marked *xy* move diagonally. Mean/minimum/maximum values were extracted for 41 subjects of the BiLi database.

Name	Position	Displacement axis	CIPIC parameter	Positive	Mean	Maximum	Minimum	Group
CP1	Base of intertragal notch	y	d_1	d_1+	−14.7 mm	−17.2 mm	−12.4 mm	1
CP2	Crus of helix	y	d_1, d_2	d_1+, d_2-	3.6 mm	5.9 mm	1.5 mm	1
CP3	Crus of antihelix	y	d_2, d_4	d_2+, d_4-	11.1 mm	14.2 mm	7.7 mm	1
CP4	Upper tragus	x	d_3	d_3+	5.9 mm	7.8 mm	3.4 mm	1
CP5	Rear antihelix	x	d_3	d_3+	−10.0 mm	−14.3 mm	−7.4 mm	1
CP6	Inner helix ridge	y	d_4	d_4+	27.6 mm	31.1 mm	24.8 mm	2
CP7	Lobe	y	d_5	d_5+	−28.3 mm	−34.5 mm	−24.2 mm	2
CP8	Upper helix	y	d_4, d_5	d_4+, d_5+	33.4 mm	39.8 mm	26.8 mm	2
CP9	Root of helix	x	d_6	d_6+	13.8 mm	17.6 mm	9.9 mm	2
CP10	Rear helix	x	d_6	d_6+	−17.0 mm	−22.0 mm	−14.1 mm	2
CP11	Lower tragus	x	d_7	d_7+	2.6 mm	4.2 mm	1.5 mm	1
CP12	Antitragus	x	d_7	d_7+	−2.9 mm	−6.1 mm	−0.8 mm	1
CP13	Upper antihelix	xy	—	—	11.2 mm	15.1 mm	7.5 mm	1
CP14	Lower antihelix	xy	—	—	8.4 mm	13.7 mm	5.7 mm	1
CP15	Upper front helix	xy	—	—	33.5 mm	39.5 mm	27.1 mm	2
CP16	Upper rear helix	xy	—	—	29.9 mm	35.8 mm	25.1 mm	2
CP17	Lower-mid rear helix	x	—	—	−21.3 mm	−26.2 mm	−17.5 mm	2
CP18	Lower rear helix	x	—	—	−20.2 mm	−24.2 mm	−16.2 mm	2
CPd1	Root of helix relief	z	—	—	8.5 mm	15.3 mm	3.0 mm	3
CPd2	Upper rear helix relief	z	—	—	15.8 mm	23.7 mm	8.1 mm	3
CPd3	Lobe relief	z	—	—	7.0 mm	12.5 mm	2.4 mm	3
CPd4	Antihelix relief	z	—	—	15.6 mm	21.7 mm	8.9 mm	3
CPd5	Rear helix relief	z	—	—	14.9 mm	22.3 mm	6.8 mm	3
CPθ	Rotation	—	θ_1	θ_1+	38.2°	50.6°	23.3°	—

G2: Control points on the outer pinna regions.
G3: Control points that change the relief of the pinna.

The group to which each control point belongs is listed in Table I.

A. Parameter ranges

To investigate the impact of changing the position of control points on the resulting HRTF, each parameter was varied to a determined *minimum* and *maximum* value. The ranges evaluated in the current study were determined from an analysis of an existing HRTF database with accompanying morphology to be representative of a diverse population sample. These values are not meant to represent the full range of these parameters over the entire population but serve as a reasonable statistical range. Control points for each subject were determined by manual notation of photographs of the left ear of 41 subjects measured for the BiLi database (Carpentier *et al.*, 2014). The points then had their coordinate systems transformed such that the line from **CP1** to **CP6** aligned with the *y* axis. The measured point means

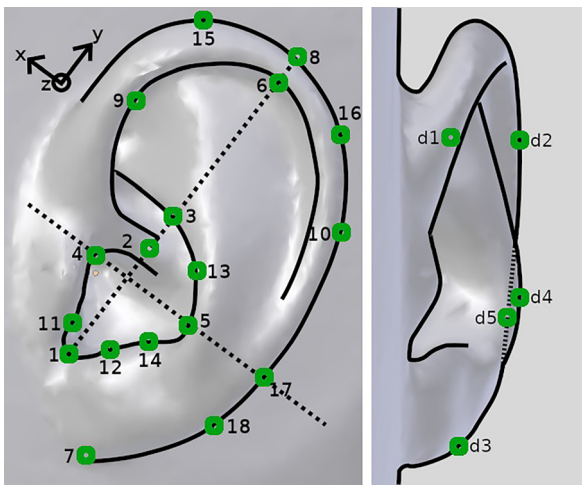

FIG. 3. (Color online) The positions of the 23 control points of the parametric ear model corresponding to Table I. The dashed lines on the left image represent the axis relative to the pinna in which the points are defined.

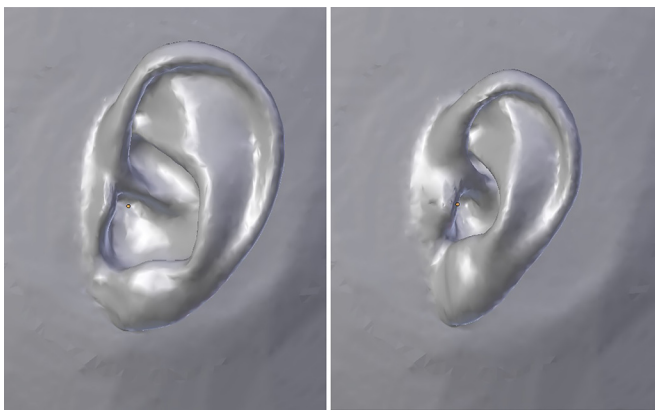

FIG. 4. (Color online) Two pinnae created with the parametric ear model.

https://doi.org/10.1121/10.0004128

and 95% confidence ellipses for **CP1–CP18** of the 41 subjects are shown in Fig. 5, along with those of the KU-100 for comparison. Based on Table I, the *x* or *y* coordinate was used as the parameter value. For parameters indicated as *xy*, the distance from the origin was used. Next, the mean of all these parameters was taken to give the spatial average parameters. Figure 5 also includes the control points from the spatial average pinna mesh generated using the parameters determined from the photographic data.

The pinna generated from the spatial average data was used as the *reference* for each of the pinna variations in the current study. To evaluate the sensitivity of the HRTF to the various morphological control points, each parameter was then varied to the minimum and maximum of the extracted parameter values while keeping all the others as near as possible to the spatial average. There were two exceptions to this. The first was **CP6**, since its maximum value places it above the position of the average value of **CP8**. Instead, the difference between **CP6** and **CP8** was used to determine its range, with **CP6**'s maximum value corresponding to the mean of **CP8** minus the minimum difference and vice versa for the minimum **CP6** value. Similarly, for **CP8**, its minimum value would place it below the average value for **CP6**. As such, **CP6** was moved with **CP8** when it was moved to its minimum/maximum positions.

Due to the perspective of the photographs for the front-facing photographs, which could lead to a systematic underestimation of the depth control points, the spatially averaged pinna used the KU-100 depth parameters. The depth parameters were displaced by the minimum and maximum difference of the measured control point positions from the spatial average value. This leads to the same range of change but with a different reference position.

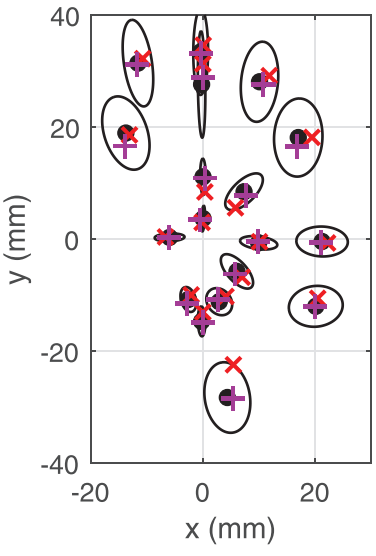

FIG. 5. (Color online) The mean positions of control points **CP1–CP18** (black circles) as measured from a photograph database of 41 subjects' left ears with 95% confidence interval ellipses. Control points for the KU-100 dummy head (red times symbols) and spatially averaged reference HRTF (purple plus symbols) are indicated.

B. Mesh generation

For all parameters (apart from **CP*θ***), the difference between the variation and base KU-100 mesh was computed and stored. This provides a displacement vector for each vertex in the mesh. These displacement vectors can then be weighted with positive or negative values and summed with the base mesh to give a new pinna variation, with a weight of zero resulting in a pinna identical to the KU-100 mesh. A new pinna can thus be generated from a linear combination of the displacement vectors applied to the base mesh. In the case of **CP*θ***, the angle of rotation about the pinna origin was calculated for each point in the mesh. The weighted angle of rotation was then used to calculate a rotation matrix for each point in the mesh with non-zero rotation. These rotation matrices are then applied to the newly generated pinna to rotate it.

Since it was impossible to create pinna variations that were completely independent, as changing one parameter could potentially move multiple control points, a non-linear solver (fmincon in MATLAB) was used to generate pinnae with parameters as close as possible to the desired values for the spatial average reference pinna and all the minimum/maximum variations. As a consistency test, the non-linear solver was run five times to generate the spatial average pinna starting from the KU-100 base ear initial condition. The resulting pinnae were found to be identical. The KU-100 base ear was chosen as the initial condition to ensure that the starting point for the solver was already a feasible pinna shape and because, during initial testing, random initial conditions were found to lead to much longer convergence times or, occasionally, results that converged to a local minimum that had high error.

C. BEM modeling

Once a new pinna had been generated, it was saved in `.stl` format, assuring a closed boundary mesh. It was subsequently remeshed [OpenFlipper 4.1 (Möbius and Kobbelt, 2012)] to reduce the number of elements, since the base scan is of a higher resolution than required for the BEM simulations. The remeshed edge length was chosen as 2.5 mm, such that the limit frequency of the simulation is above 20 kHz, assuming six edge lengths per wavelength. This remeshing resulted in meshes of approximately 40 000 vertices. The mesh was then imported to Blender and exported using the Mesh2HRTF (version 0.1.2) plugins, which prepare the mesh and evaluation grid for the BEM simulation (Ziegelwanger *et al.*, 2015). The evaluation grid used was the same as used for the BiLi HRTF measurement database but extended to include points below the database cutoff elevation angle of $-67°$ as well as the poles at $\pm 90°$.

A multi-level fast multipole method (ML-FMM) BEM simulation was run for each pinna variation, resulting in a total of 49 HRTFs (48 minimum/maximum parameter variations and the spatial average reference). Simulations were run in frequency steps of 100 Hz, up to a maximum of 20 kHz. This provided impulse responses of 400 samples at

a sample rate of 40 kHz. The impedance condition for the mesh was uniformly defined as rigid. While this is in contradiction with measured values (Katz, 2000), which have been shown to have an effect on HRTF simulations (Katz, 2001b), such contributions are not the focus of this purely morphological study.

For analysis purposes, each HRTF was diffuse field equalized by dividing its spectrum by the magnitude of each frequency bin summed over all spatial positions. The resulting diffuse field compensated HRTF is a direction transfer function (DTF) (Middlebrooks, 1999) and contains the direction-dependent characteristics of the spectrum. The diffuse field component (DFC) contains any global resonances or anti-resonances in the spectrum, describing any overall change in coloration in the HRTF. Splitting the HRTF into DTF and DFC allows for analysis of changes to the global coloration in the HRTF separately from direction-dependent changes.

Prior to full modeling and analysis, a preliminary study was carried out to ascertain how critical the microphone element size was to the repeatability of the simulations. This study is detailed in the Appendix. Results show an increasing sensibility with frequency, producing an increasingly present low-pass effect with increasing receptor mesh size. Shifts in notch frequencies occur predominantly above 15 kHz. The mean difference over tested configurations was 3 dB, resulting in a reasonable choice in restriction of the frequency range of analysis to frequencies below 12 kHz for the remainder of this study. Above this, there can be notch frequency mismatches that are strongly related to the position or size of the receiver. In addition, as the typical eardrum is on the order of 8–10 mm in diameter (Faddis, 2008), equating to a half-wavelength of ≈17 kHz, such very high frequencies are unreliable for spatial cues. Unless otherwise noted, the minimum frequency used for analysis was 1 kHz, with a wavelength comparable to head-width, below which the influence of the pinna is negligible.

III. METRICS

A. Objective metrics

Objective analysis allows for a comparison of different HRTFs with respect to the similarity or difference of their spectrum. Common metrics for measuring the difference between HRTFs are the spectral distortion (SD) (Tommasini *et al.*, 2015)

$$SD_p = \sqrt{\frac{1}{N_f}\sum_{k=k_{\text{low}}}^{k_{\text{high}}}\left(20\log\left(\frac{|H_p(k)|}{|\hat{H}_p(k)|}\right)\right)^2} \tag{1}$$

and mean spectral distortion (MSD)

$$MSD = \frac{1}{N_p}\sum_{p=1}^{N_p} SD_p, \tag{2}$$

where k_{low} and k_{high} are the indices of the lowest and highest frequency bins, $N_f = k_{\text{high}} - k_{\text{low}}$, N_p is the number of HRTF measurement positions, and $H_p(k)$ and $\hat{H}_p(k)$ are the generated and reference HRTFs, respectively, in the pth direction. The MSD gives a measure of the overall difference between two HRTFs across the specified frequency range, to give a first approximation of the significance of a control point movement on the HRTF. If the MSD is large, indicating differences between the HRTFs, then further analysis can be performed to determine where in the frequency and/or spatial domain the differences are to be found.

Another commonly used measure is the mean square error (MSE), defined as

$$\text{MSE}_p = 100 \times \frac{||\mathbf{H}_p - \hat{\mathbf{H}}_p||^2}{||\mathbf{H}_p||^2}, \tag{3}$$

where $\mathbf{H}_p$ and $\hat{\mathbf{H}}_p$ are vectors of the pth HRTF magnitudes across a chosen frequency range (Hugeng *et al.*, 2010). The global MSE is defined as the mean MSE across all N_p positions and gives a metric for the difference between HRTFs across all positions and frequencies (1–12 kHz). During analysis, the MSE and SD for each minimum/maximum pinna variation were found to have correlation coefficients of >0.96. Since the purpose of this study is to gain a first sensitivity analysis of parameters that have the largest overall impact on the HRTF, rather than being concerned with the absolute values, only the SD metric is reported here.

The diffuse field of each HRTF variation was calculated to allow comparison to the reference. This gives an indication of any changes in resonant frequencies. This analysis was performed by calculating the SD between the DFC of the reference and modified pinna and is referred to as the diffuse spectral distortion (DSD).

In this analysis, source positions used when calculating the MSD have been limited to ipsilateral sources to avoid, for example, changes in notches on the low-level contralateral side contributing significantly to the metrics when their perceptual weight would be low.

The objective metrics defined here use a linear frequency scale. This choice is made in order to determine the overall acoustic impact of changing different portions of the pinna. As human perception is more adequately represented on a logarithmic scale, the same analysis was repeated for a logarithmic frequency scale using 1/12-octave bands, and the results were found to be highly correlated (>0.99) to the linear scale results. As such, and similar to the results for MSE, the numerical results of this analysis are not reported here due to the high correlation with reported results. In addition, perceptual impact can be better assessed using perceptual models or perceptual testing.

B. Perceptual modeling

Objective measures give a means of classifying the difference between two HRTFs but do not determine whether the changes are perceptually relevant for spatial hearing applications. For example, does a small change in the exact

https://doi.org/10.1121/10.0004128

frequency of a notch have a perceptually significant impact, even though it might result in large SD values? A perceptual localization model in the sagittal median plane (Baumgartner *et al.*, 2014) was used to evaluate this question, as it can perform many test comparisons without experiencing "listener fatigue." The model uses a reference HRTF and evaluates a new HRTF using estimated localization errors. This mimics an internal "learned" HRTF, usually the listener's own, and the degradation of localization quality when listening with a non-individual HRTF. As with the objective metrics, the reference HRTF was the spatial average HRTF.

The metrics provided by the model are polar error (PE) in degrees and quadrant errors (QE) in percent. The PE is a measure of the localization accuracy, while QE is the percentage of predicted responses that are in a different quadrant from the target. This model provides a first approximation of the localization degradation caused by a change in pinna morphology. It does not predict other characteristics that could change, such as the timbre or other perceptual attributes related to HRTF variations (Simon *et al.*, 2016), or adaptation effects of the listener. The results provided by this model should not be considered as completely accurate predictions of possible future perceptual results. Rather, they provide a perceptually inspired alternative to purely objective metrics that may provide additional insight into the HRTF.

The model allows for the sensitivity of the "listener" to be set such that they are a relatively accurate or inaccurate localizer. For this analysis, the sensitivity was set to 0.21, which was the lowest value found by Baumgartner *et al.* (2014) in modeling perceptual experiment results. This ensures that a significant range of parameters can be obtained, since the baseline reference will be as low as possible.

IV. RESULTS

A. Median plane example

To demonstrate the effect of changing one control point on the HRTF, Fig. 6 shows the median plane spectrum for the spatial average reference pinna and variations **CP3-min/max**. For example, the reference HRTF has a deep notch near 7 kHz at polar angle −60° that rises in frequency with polar angle. The **CP3-max** has a similar but more consistently deeper notch that follows the same trend. However, **CP3-min** does not exhibit such a strong notch, although it is still present. At the lowest polar angles, **CP3-min** exhibits a double notch feature, a shallower one near 6 kHz and another, slightly deeper, around 8 kHz. The reference and **CP3-max** exhibit only a single, prominent notch in this frequency range. Considering that this is only one single slice of the full spatial parameter space, it highlights the difficulty of a spatio-frequency analysis over the large number of pinna variations with the aim of finding the most pertinent control points. While this type of investigation is an interesting avenue to explore, it is outside the scope of this study, and analysis will be restricted to the global SD measure defined in Sec. III.

B. Difference in objective metrics

This section presents an objective analysis of the HRTF changes due to changes in pinna parameters. Different metrics have been used to measure the difference in HRTFs. The reference pinna in all cases was the HRTF calculated for the spatial average pinna based on analysis of the BiLi database. In addition to the analysis of the parameter changes, the HRTF of the spatial average pinna was also compared to the KU-100 HRTF to give a measure of the magnitude of differences between the reference and a widely used dummy head. The KU-100 results are excluded from any group analyses.

1. Directional transfer function

Table II shows the DTF MSD and DSD for different parameter variations. The DTF MSD ranges up to 2.64 dB, with a median across all pinna variations of 1.19 dB. The KU-100 compared to the spatial average leads to a MSD of 2.38 dB, which is close to the maximum value of the pinna

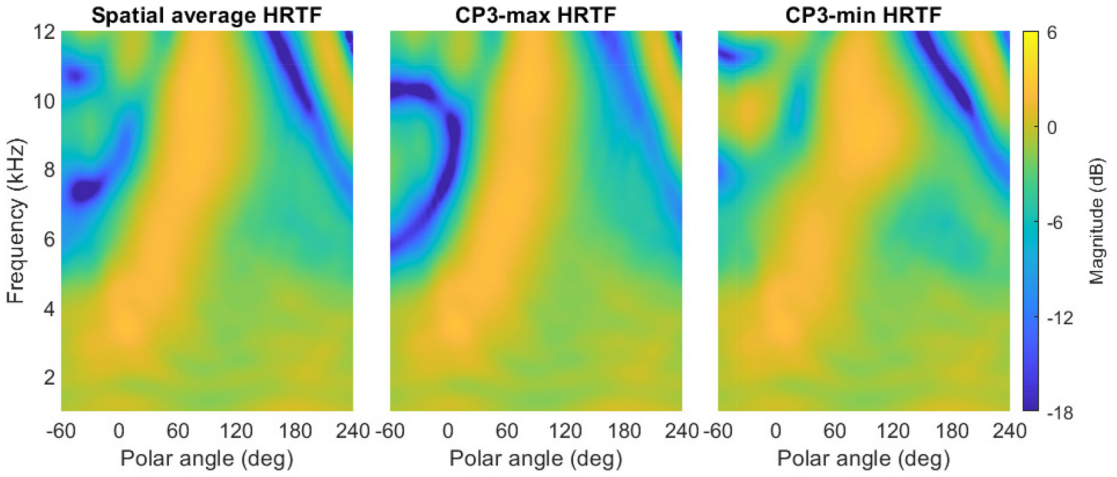

FIG. 6. (Color online) The HRTF spectrum of the reference, **CP3-max**, and **CP3-min** HRTFs.

TABLE II. The values of the objective (MSD and DSD) and perceptual prediction (PE and QE) metrics for each minimum and maximum pinna variation for each control point. PE and QE values for the reference pinna and all metrics for the KU-100 pinna are included.

	MSD (dB)		DSD (dB)		PE (deg)		QE (%)	
Name	Max	Min	Max	Min	Max	Min	Max	Min
CP1	1.26	1.21	1.14	0.87	32.8	31.4	19.2	13.8
CP2	0.82	0.60	0.32	0.42	32.5	31.7	16.2	13.2
CP3	1.74	2.15	0.80	1.14	35.3	34.0	22.6	17.3
CP4	0.42	0.49	0.34	0.47	31.8	31.0	12.9	12.7
CP5	1.93	1.24	0.78	0.61	35.7	31.0	20.8	12.5
CP6	0.34	0.27	0.10	0.07	31.4	31.4	13.0	12.7
CP7	0.66	0.46	0.28	0.25	31.4	31.6	12.8	12.5
CP8	1.17	2.05	0.27	0.32	32.1	33.6	13.4	19.5
CP9	2.43	2.50	0.53	0.75	34.5	32.4	17.0	14.6
CP10	0.65	0.49	0.23	0.28	32.2	31.3	13.4	12.6
CP11	0.67	0.54	0.48	0.28	32.2	31.4	13.5	12.7
CP12	1.10	1.11	0.45	0.45	31.8	32.3	13.8	13.6
CP13	1.97	1.50	0.85	0.83	35.2	33.2	25.1	18.3
CP14	1.78	0.78	1.18	0.47	34.5	31.3	18.9	13.1
CP15	1.26	1.55	0.30	0.58	32.3	33.3	14.7	17.6
CP16	0.65	0.63	0.13	0.26	32.1	31.4	13.2	13.4
CP17	0.57	0.56	0.22	0.14	31.2	31.7	12.5	13.0
CP18	0.61	0.51	0.29	0.17	31.5	31.7	12.8	12.8
CPd1	2.64	2.33	1.39	1.57	35.3	38.7	23.2	21.7
CPd2	2.16	2.22	0.40	0.51	34.1	35.2	19.7	16.9
CPd3	0.92	0.75	0.34	0.19	31.0	32.2	12.4	13.3
CPd4	1.82	2.14	0.83	1.04	38.6	35.8	21.9	22.5
CPd5	1.26	1.31	0.32	0.22	34.0	32.1	15.2	14.6
CPθ	2.20	2.46	0.18	0.19	32.2	35.5	13.5	15.6
Reference	—		—		31.2		12.5	
KU-100	2.38		1.50		36.6		25.5	

variations. This demonstrates that changing a single portion of the pinna can result in a DTF that is as different from the reference as a completely different pinna.

Group **G1** (control points in/around the concha) has a median MSD of 1.16 dB, while **G2** (those on the outer pinna) has marginally less impact, with an average MSD of 0.64 dB. For the non-relief set of control points, this suggests that control points around the concha are more influential than those that are more distant from the ear canal. The minimum/maximum pinna variations for **CP9** are significantly higher than most others in the group and are among the highest across all pinna variations. It is worth noting that **CP15** is in relatively close proximity to **CP9**, and both are near the triangular fossa, and **CP8-min** compresses it. Their relatively large MSD and distance from the ear canal contradict **H2**.

In **G1**, the control points on the front wall of the concha (**CP4** and **CP11**) have some of the smallest impact, with both minimum/maximum variations below the median MSD. **CP2** also has a low impact on the MSD, which could be due to the fact it is in relatively low relief and therefore does not provide a strong reflection surface. Almost all the other pinna variations in **G1** are above the median MSD. The exceptions are **CP12** and **CP14-min**.

The relief parameters **G3** have a median MSD of 1.98 dB, greater than for **G1** and **G2**. Overall, eight of the ten pinna variations in **G3** result in above-median MSD, supporting hypothesis **H1**. Only the relief of the ear lobe, **CPd3**, has MSDs below the median. In fact, the single largest DTF MSD of 2.64 dB belongs to **G3**: **CPd1-max**. **CPd1-min** also has a large (2.33 dB) impact on the MSD. As mentioned above, **CP9** and **CP15** both have relatively large impacts on the MSD and are close neighbours to **CPd1**. The strong influence of **CPd1** on the MSD further suggests that the pinna geometry near the triangular fossa has a large impact on the DTF. **CPd4** also has a very strong impact on the MSD for both minimum and maximum displacements. **CPd2** also has a relatively strong impact on MSD despite being located on the helix. This suggests that the relief of the helix might have more impact than its distance projected in the median plane.

Finally, one of the largest MSDs is for pinna rotation (**CPθ**), which makes intuitive sense, since the main pinna-related notches will be rotated spatially away from their position in the reference ear. Both rotation-modified pinna variations cause similar MSD of greater than 2 dB.

2. Diffuse field component

The DSD, shown in Table II, is lower on average than for the MSD—a median of 0.37 dB compared to 1.19 dB. This lower effect might be due to the DFC having no deep notches, the misalignment of which could cause relatively high SD between HRTFs. Instead, the DFCs of all pinna variations vary smoothly and relatively slowly across frequency. As with the MSD, the KU-100 results in a DSD that is near the maximum of the pinna variations.

The DSDs are generally high for control points around and in the concha in group **G1** with a median of 0.55 dB. Control point **CP4** has low impact on the DSD. Taken with its relatively weak impact on the MSD, this suggests that, overall, it is one of the less important control points. Control points **CP11** and **CP12**, which control the width of the intertragic notch, also have some of the least impact on the DSD. **CP1** has one of the strongest impacts on DSD of all pinna variations for both minimum and maximum displacements. This control point would change the height of the concha, suggesting that this is an important dimension for the non-directional component of the HRTF.

The control points on the outer portion of the pinna, group **G2**, are generally lower than those surrounding the ear canal, with a median of 0.26 dB. This supports hypothesis **H3**, though the difference is relatively small. **CP9** has the most impact on DSD of any in **G2** when averaged over both pinna variations and is the only one to have both pinna variations cause above-median DSDs.

The relief control points **G3** have a median DSD of 0.46 dB, higher than **G2** but lower than **G1**. In fact, **CPd1** has the most impact of any single control point. Figure 7 shows the diffuse field magnitude for **CPd1** with the mean pinna reference. Decreasing the relief of this control point

https://doi.org/10.1121/10.0004128

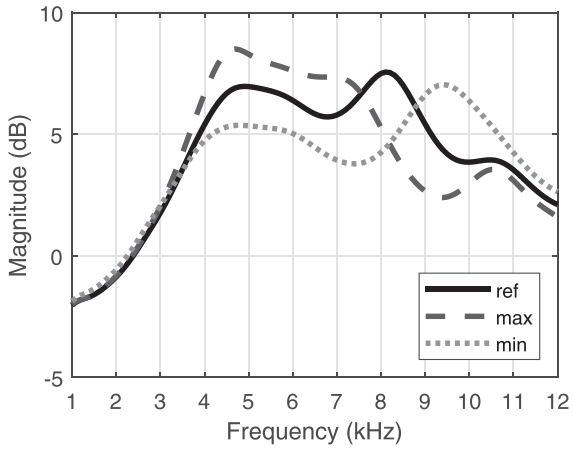

FIG. 7. The diffuse field component of the reference pinna, **CPd1-max**, and **CPd1-min**.

causes the resonant peak at 4.8 kHz to be reduced in magnitude and widened. This peak corresponds to the first mode identified by Shaw (1997). Alongside this, the peak at 8.2 kHz is shifted to higher frequencies when relief is reduced and lower when relief is increased. The magnitude of these peaks remains approximately equal regardless of the direction of the control point displacement. The importance of **CPd1** supports hypothesis **H1** and, along with the impact of **CP9**, further emphasises the importance of the region around the triangular fossa on the HRTF. **CPd4**, which has an impact on the volume of the concha, also has one of the largest impacts on the DSD, which supports hypothesis **H3**. Mokhtari *et al.* (2015) found that increased concha depth leads to increased amplitude of the first peak in the HRTF, which would lead to a change in the DFC as reflected in the result for **CPd4**.

As expected for pinna rotation, the DSD is low (<0.2 dB), since the overall shape of the pinnae has not changed, only their orientation. The differences are possibly due to the head not being completely spherical (see Fig. 2). Were a spherical head model to have been used, then we would expect to see zero difference in the diffuse fields for pure pinna rotations.

C. Modelled perceptual results

Table II shows the perceptual quality measures for each of the modified pinnae using the auditory virtual listener model by Baumgartner *et al.* (2014) for source directions in the median plane. The two perceptual metrics, PE and QE, are strongly correlated, with a correlation coefficient of 0.86. The correlation between MSD and the two perceptual metrics is approximately equally as strong—the correlation between PE and MSD is 0.76, and the correlation between QE and MSD is 0.74. Note that the binaural model uses frequency bands up to a centre frequency of 18 kHz, beyond the frequency limit used in the objective analysis in Sec. IV B.

The maximum PE difference from the reference is 7.4°, with just more than half of pinna variations (25 of 48) exhibiting less than 1° of deviation. The small changes in PE relative to the reference suggest that changing only one element of the pinna does not significantly degrade localization performance, either because the changes are not large enough to be considered perceptually noteworthy or because enough elements of the reference HRTF remain intact. Similar results are obtained for QE, with a large number of the variations exhibiting little change from the reference. The predictions for the KU-100 using the spatial average pinna HRTF as a target are given in Table II as a benchmark for a subject listening with a non-individualized HRTF that differs more significantly. Focus in the remainder of this section is on control points that have predicted performance errors approaching that of using the KU-100 benchmark.

The median absolute changes in PE for **G1**, **G2**, and **G3** are 2.3°, 0.86°, and 3.4°, respectively. The corresponding values for QE are 5.6%, 1.1%, and 6.0%. These group medians indicate a lower perceptual importance for control points located on the outer pinna and helix than those around the concha or that control relief. This might be expected, since control points that affect the concha are likely to change the resonant characteristics of the concha, changing peak (Mokhtari *et al.*, 2015) or notch frequencies and amplitudes. Changes to control points on the outer pinna may not result in such strong direction-dependent resonant changes but can be expected to influence amplification of sound sources to the front and shadowing for sound sources to the rear.

CPd1 and **CPd4** cause degradation in the predicted localization performance, particularly QE. This aligns with the relatively high MSD for both control points. **CP3** also has a relatively strong impact on QE and PE for both minimum and maximum displacement variations, which indicates the perceptual importance of this upper wall of the concha.

Of note is that **CP5-max** leads to a predicted increase in both PE and QE, while **CP5-min** does not. The MSD for **CP5-min** was lower than for **CP5-max** but was still above the median value. This parameter controls the back wall of the concha and might intuitively be expected to have a strong impact regardless of the displacement direction. Nonetheless, its displacement only results in changes that are considered important by the model when the concha breadth increases.

V. IDENTIFICATION OF PERTINENT PARAMETERS

To identify the most pertinent control points with respect to the impact on the HRTF, the four metrics (MSD, DSD, PE, and QE) were all scaled to the range 0–1. This was done for each metric using all minimum/maximum variations considered together for the scaling. Following this, the scaled values for each minimum/maximum pair were averaged to give an average score for each control point. Finally, three scores were obtained by taking the mean of the following metrics: one for the object metrics (MSD and MSE), one for the perceptual metrics (PE and QE), and one

TABLE III. Normalized score [0:1] of each control point for objective, perceptual, and combined metric groups.

Control point	Objective	Perceptual	Combined
CP1	0.52	0.23	0.38
CP2	0.19	0.16	0.18
CP3	0.65	0.54	0.59
CP4	0.15	0.04	0.10
CP5	0.49	0.32	0.40
CP6	0.01	0.04	0.03
CP7	0.13	0.04	0.08
CP8	0.36	0.28	0.32
CP9	0.65	0.29	0.47
CP10	0.12	0.07	0.10
CP11	0.17	0.08	0.13
CP12	0.30	0.12	0.21
CP13	0.57	0.58	0.57
CP14	0.47	0.26	0.37
CP15	0.36	0.27	0.31
CP16	0.12	0.08	0.10
CP17	0.10	0.04	0.07
CP18	0.11	0.06	0.09
CPd1	0.94	0.79	0.86
CPd2	0.54	0.47	0.50
CPd3	0.18	0.06	0.12
CPd4	0.65	0.79	0.72
CPd5	0.28	0.23	0.26
CPθ	0.47	0.27	0.37

combining all four metrics. Table III gives the scores for each of these groups for each control point. The correlation between the objective and perceptual scores is 0.90, indicating a strong correlation. Given the correlation identified in Sec. IV C between objective and perceptual metrics, along with the high correlation between the scaled scores in Table III, the following discussion uses the combined score metrics.

As a subsequent means of analysis, the resulting combined score values were clustered using a basic *k*-means. Following visual inspection of the distribution of values [see histogram in Fig. 8(a)], three clusters were prescribed, considered as having a *strong*, *moderate*, and *weak* influence on the HRTF. The resulting analysis identifies the morphological parameters for which the HRTF is strongly influenced by, or most sensitive to, changes in the defined range as being **CPd1**, **CPd4**, **CP3**, and **CP13**.

Of the four control points identified as having the strongest influence, two are relief parameters. The remaining non-relief control points are located on the upper edge of the antihelix, indicating high pertinence for this region. **CPd4** changes the height of the back wall of the concha, further highlighting the overall importance of the accuracy of the antihelix placement but also demonstrating that the projection in the median plane may not provide adequate information in a HRTF personalization procedure.

The importance of the antihelix is emphasised by the inclusion of **CP5** and **CP14** in the moderate influence group. The region near the triangular fossa is also represented by the inclusion of **CP9** and **CP15**. Similarly scored is **CPd2**, which is the highest scored control point on the rear of the helix and the only one outside of the weak influence group. **CP1**, which is linked to the height of the concha, is also identified as having moderate influence. Combined with the identification of **CP3** as having strong influence, this indicates that concha height has an important impact on the HRTF. See Sec. VI for further discussion on this point.

A total of 11 control points, seen in Fig. 8(a), are identified as belonging to the weak influence group and are therefore the least interesting candidates for future detailed investigation. This set includes all the points on the rear of the helix, with the exception of **CPd2**, as well as both control points on the lobe.

While it is the case that the identified pertinent control points either control depth/relief, control the placement of the antihelix, or are in proximity of the triangular fossa, other control points inside the concha are also identified as having weak influence, despite their proximity. To

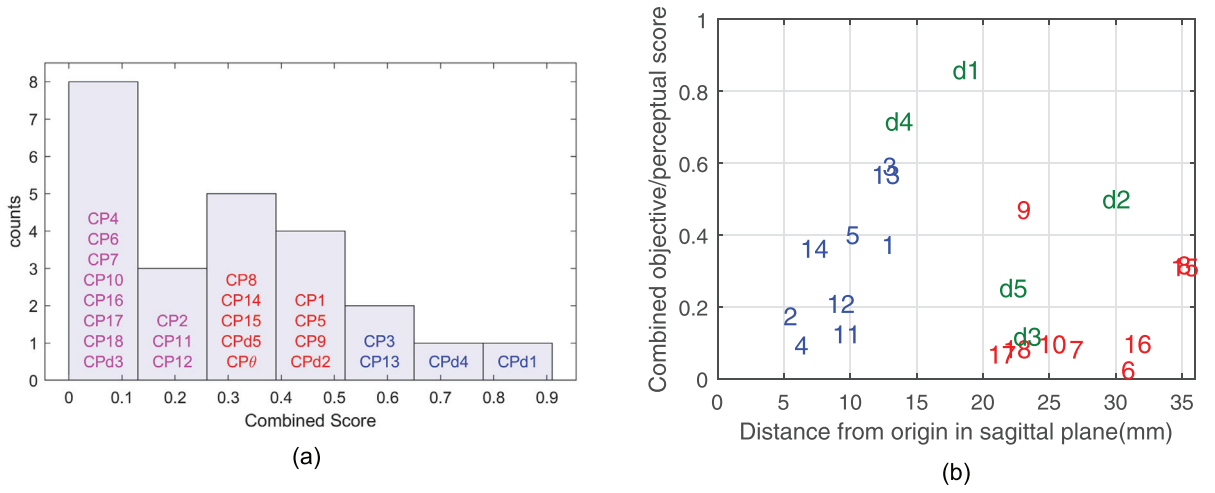

FIG. 8. (Color online) (a) Histogram of combined score values. Each control point is indicated and color coded according to the *k*-means cluster analysis. (b) Scatter plot of control point distance from the pinna coordinate origin against the combined score. Control point groups are color coded: **G1**, blue; **G2**, red; and **G3**, green.

https://doi.org/10.1121/10.0004128

investigate this further, Fig. 8(b) compares the distance of control points from the ear canal (projected in the sagittal plane) against their associated combined scaled score. There is no obvious strong relationship with distance to support **H2**: the global correlation coefficient across all groups is –0.15. Thus, no clear relationship was observed regarding the distance of a control point from coordinate origin of the pinna and its corresponding global pertinence for the HRTF. Groups **G1**, **G2**, and **G3** have correlation coefficients of 0.75, 0.2, and –0.48, respectively, indicating that only the control points in the concha have a relationship between distance and combined score. However, as can be seen in Fig. 8(b), this score increases with distance rather than decreases, refuting hypothesis **H2**.

In should be noted that observation of Fig. 8(b) shows no clear demarcation between the three *k*-mean clusters identified. As such, control points near the border values could potentially be more or less significant with further testing. For this reason, perceptual testing with human subjects is the next logical step in this investigation.

VI. DISCUSSION

The so-called CIPIC parameters define distances between different points on the pinna, but they do not define an absolute reference position. Therefore, it is possible to generate multiple pinnae with the same CIPIC parameter values. For example, consider a pinna where **CP9** and **CP10** are both shifted by the same amount to the front or back. Since they move by the same amount, their distance (CIPIC d_6) remains unchanged. However, Table II indicates that in such a case, it is likely that the MSD would be significantly changed, since **CP9** is an important parameter. While the control points used in the current model are based on the end points used for the CIPIC parameters, they have been supplemented with additional control points. Furthermore, the CIPIC distance values are simply projections onto a 2D plane augmented with a pinna flare angle θ_2. This means that two individuals with the same distance parameters could have very different ears by virtue of having two different pinna flare angles. The relief parameters used here allow more granular definition, and hence control, over the pinna flare, rather than a global parameter that will influence all the side-facing parameters. It should be noted that, in practice, measuring the ear flare angle or the relief parameters of the presented model can be difficult. This is a potentially large source of error in the measurement, which has a knock-on impact on the more frequently cited distance parameters, so it should be taken into account in any HRTF individualization process.

Fels and Vorländer (2009) performed a study of the impact of head, torso, and pinna morphology on the HRTF up to 8 kHz using BEM simulations. Their model used a simplified pinna model, but some comparisons to the current study can be made. Their pinna parameters were concha height, breadth, and depth along with pinna height, breadth, and rotation angle. They found that the breadth and depth of the concha and pinna rotations had the most impact on the HRTF. The closest equivalent parameters in this study are **CP5** (concha breadth), **CPd4** (concha depth), and **CP*θ***. All these parameters were found to have a strong impact on the MSD and DSD in this study. Similarly, they found the pinna height and breadth to have weaker impact, which agrees with the general finding that points on the rear of the helix or the lobe have weaker impact. A point of disagreement is that they found concha height to have less impact than other parameters, while in this study, the equivalent parameters (**CP1** and **CP3**) were found to have a strong impact on the objective metric. The analysis in this study was repeated with a limit frequency of 8 kHz, but the same general trends as presented previously were obtained. The reason for the disagreement between these studies is not clear but could be due to the difference between the detailed parametric pinna model used here and the simplified shape used by Fels and Vorländer.

Huang and Li (2014) used correlation analysis between processed versions of the HRTFs and morphological data from the CIPIC data to obtain a subset of parameters. Their analysis considered d_2 to be unimportant, and it was dropped, and then the remaining parameters were ranked. They rank the remaining parameters as d_3, d_8, d_4, d_7, d_1, d_6, d_5. This corresponds broadly with the results of this study and of Fels and Vorländer (2009). These two studies along with this one indicate the relative importance of concha depth and breadth with the lower importance of the pinna height and breadth (d_6 and d_5), with the caveat that the pinna breadth be varied at the helix, since **CP9** was found to have a strong impact. Huang and Li placed d_4 as their third highest ranked parameter. This is related to control points **CP3**, **CP6**, and **CP8**, so it is difficult to infer an overall weight for this parameter from the current study. However, **CP3** was found to have a strong impact on the HRTF, so it reasonable to consider that this could lead to d_4 having an important weight.

One of the main outcomes of this study is the importance of the relief profile of the pinna, which was shown to have as strong an impact on the HRTF as the more often analyzed side-facing parameters. Some HRTF personalization approaches that utilized the CIPIC (or similar) parameters have not included θ_2, perhaps because of the difficulty in extracting it from photographs (Iida *et al.*, 2014; Torres-Gallegos *et al.*, 2015). Four of the five relief-related parameters in the model presented here have at least moderate impact on the objective DTF MSD, while two were identified as the most pertinent of all 24 parameters in Sec. V. This indicates that, at the very least, a θ_2-like parameter should be included in any HRTF personalization approach. However, the fact that these relief parameters cannot all be described by θ_2 suggests that, like the side-facing parameters, the CIPIC parameters are incomplete with regard to HRTF relevance. It is unclear whether additional parameters would be required to fully characterize the relief

dimensions of the pinna, but the five points in the presented model provide reasonable control over the main regions of the pinna.

The control points nearest to the triangular fossa (**CP9**, **CP15**, and **CPd1**) all have relatively strong impact on the DTF. In pinna parameterisation, this region is generally not considered to have any particular importance. However, Takemoto *et al.* (2012), using numerical simulations, show that strong nodes and anti-nodes can form in the triangular fossa with a dependence on the source direction. Taken with the results of this study, this indicates that the morphology of the triangular fossa is an important region to model when attempting to generate personalized HRTFs.

The number of control points introduced in this model is significantly larger than the number of CIPIC parameters—24 pinna parameters vs 10 for CIPIC (eight distances and two angles). Further tests are required to determine which parameters can be relaxed or discarded while allowing for perceptually well-rated personalized HRTFs to be generated. However, the perceptual scores in Table III give an indication as to which parameters are likely to be found most pertinent.

VII. CONCLUSION

BEM simulations of various systematically adjusted pinna variations were obtained to investigate the sensitivity of the HRTF to morphological changes. The parameters of the model were based on the end points of the CIPIC distances and were augmented with additional points and more granular control over the relief of the pinna.

In agreement with previous studies, parameters that changed the dimensions of the concha were found to have an important impact on the HRTF objective metrics, both the directional and diffuse field parts. The control points on the rear of the helix were found to have among the weakest influence.

An important finding was the importance of the region around the triangular fossa on the HRTF. This region is not frequently considered in HRTF individualization methods but was found to have a relatively strong impact on the HRTF. Furthermore, the relief parameters were found to have some of the strongest impact on the HRTF. Other than the CIPIC pinna flare angle (which is sometimes dropped from analyses), interest is primarily directed to the side-facing parameters. The results of this study strongly suggest that the front-facing profile is also of significant importance, and any HRTF personalization approach cannot rely solely on pinna parameters extracted from side-facing photographs.

It remains to be seen whether the defined parametric model is capable of producing HRTFs that are perceptually close to those of measured HRTFs of individuals. This will require further work, including real subject perceptual testing. Given the number of parameters in the model, future perceptual experiments could also be based on attempting to reduce the parameters to the smallest number of parameters required for perceptual equivalence, starting from the points identified in this study.

ACKNOWLEDGMENTS

This work was funded in part by the RASPUTIN project (Grant No. ANR-18-CE38-0004) and an associated *Innov'up Faisabilité* grant from the Région Île de France. This work is part of SONICOM, a project that has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 101017743.

APPENDIX

To determine how critical the placement of the microphone element is to the repeatability of the simulations, six simulations were run that increased the size of the microphone from a single element to one filling almost the entire concha.

Ziegelwanger *et al.* (2015) showed that the placement of the microphone element was quite free within the concha and that the size of the microphone element did not significantly change the HRTF. This was verified using a binaural model that used the DTF to predict perceptual performance (Baumgartner *et al.*, 2014).

Figure 9 shows that increasing the size of the microphone/receiver leads to, on average, an increasing mean difference between the smallest microphone and the larger ones. The receiver sizes were 1, 2, 6, 24, 40, and 62 elements. The single element receiver size was used as the reference. At 10 kHz, the mean difference is less than 3 dB for all variations, demonstrating a tolerance to receiver size at frequencies whose wavelength is comparable to the diameter of the ear canal.

The spectral difference between the smallest receiver and the others exhibits a general low-pass effect and also large differences caused by notch misalignment. The low-pass effect would be removed from the HRTF when

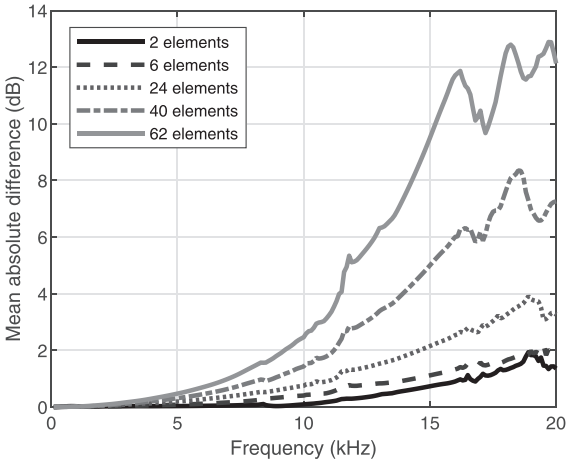

FIG. 9. The mean absolute difference across position between a HRTF simulated with a microphone size of a single mesh element and increasing receiver sizes.

converted to DTF, so it would not have been an important factor in the study of Baumgartner *et al.* (2014). The notch misalignments are largely above 15 kHz for the smallest receiver sizes.

Algazi, V. R., Duda, R. O., Thompson, D. M., and Avendano, C. (**2001**). "The CIPIC HRTF Database," in Proceedings of the 2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (Cat. No. 01TH8575), October 24, New Platz, NY, pp. 99–102.

Anwar, A. S., Ghany, K. K. A., and Elmahdy, H. (**2015**). "Human ear recognition using geometrical features extraction," Procedia Comput. Sci. **65**, 529– 537.

Baumgartner, R., Majdak, P., and Laback, B. (**2014**). "Modeling sound-source localization in sagittal planes for human listeners," J. Acoust. Soc. Am. **136**(2), 791–802.

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT, Cambridge, MA).

Carlile, S., Balachandar, K., and Kelly, H. (**2014**). "Accommodating to new ears: The effects of sensory and sensory-motor feedback," J. Acoust. Soc. Am. **135**(4), 2002–2011.

Carpentier, T., Bahu, H., Noisternig, M., and Warusfel, O. (**2014**). "Measurement of a head-related transfer function database with high spatial resolution," in *Proceedings of 7th Forum Acousticum (EAA)*, September, Kraków, Poland, pp. 1–6.

Faddis, B. T. (**2008**). "Structural and functional anatomy of the outer and Middle ear," in *Anatomy and Physiology of Hearing for Audiologists* (Thomson Delmar Learning, Florence, KY), pp. 93–108.

Fels, J., and Vorländer, M. (**2009**). "Anthropometric parameters influencing head-related transfer functions," Acta Acust. United Acust. **95**(2), 331–342.

Ghorbal, S., Auclair, T., Soladié, C., and Séguier, R. (**2017**). "Pinna morphological parameters influencing HRTF sets," *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx)*, September 5–9, Edinburgh, UK, pp. 353–359.

Greff, R., and Katz, B. F. G. (**2007**). "Round robin comparison of HRTF simulation results: Preliminary results," in *Proceedings of the 123rd Audio Engineering Society Convention*, October 5–8, New York, pp. 1–5.

Grijalva, F., Martini, L., Florencio, D., and Goldenstein, S. (**2016**). "A manifold learning approach for personalizing HRTFS from anthropometric features," IEEE/ACM Trans Audio Speech Language Process. **24**(3), 559–570.

Guezenoc, C., and Séguier, R. (**2020**). "A wide dataset of ear shapes and pinna-related transfer functions generated by random ear drawings," J. Acoust. Soc. Am. **147**(6), 4087–4096.

Hofman, P. M., Van Riswick, J. G., and Van Opstal, A. J. (**1998**). "Relearning sound localization with new ears," Nat. Neurosci. **1**(5), 417–421.

Huang, Q., and Li, L. (**2014**). "Modeling individual HRTF tensor using high-order partial least squares," Eurasip. J. Adv. Signal Process. **2014**(1), 1–14.

Hugeng, Wahab, W., and Gunawan, D. (**2010**). "Effective preprocessing in modeling head-related impulse responses based on principal components analysis," Signal Process. **4**(4), 201–212.

Iida, K., Ishii, Y., and Nishioka, S. (**2014**). "Personalization of head-related transfer functions in the median plane based on the anthropometry of the listener's pinnae," J. Acoust. Soc. Am. **136**(1), 317–333.

Jin, C. T., Guillon, P., Epain, N., Zolfaghari, R., van Schaik, A., Tew, A. I., Hetherington, C., and Thorpe, J. (**2014**). "Creating the Sydney York Morphological and Acoustic Recordings of Ears Database," IEEE Trans. Multimedia **16**(1), 37–46.

Katz, B. F. G. (**2001a**). "Boundary element method calculation of individual head-related transfer function. I. Rigid model calculation," J. Acoust. Soc. Am. **110**(5), 2440–2448.

Katz, B. F. G. (**2001b**). "Boundary element method calculation of individual head-related transfer function. II. Impedance effects and comparisons to real measurements," J. Acoust. Soc. Am. **110**(5), 2449–2455.

Katz, B. F. G., and Nicol, R. (**2019**). "Binaural spatial reproduction," in *Sensory Evaluation of Sound* (CRC, Boca Raton, FL), pp. 349–388.

Katz, B. F. G. (**2000**). "Acoustic absorption measurement of human hair and skin within the audible frequency range," J. Acoust. Soc. Am. **108**(**5**), 2238–2242.

Katz, B. F. G., and Parseihian, G. (**2012**). "Perceptually based head-related transfer function database optimization," J. Acoust. Soc. Am. **131**(2), EL99–EL105.

Lewald, J., and Guski, R. (**2004**). "Auditory-visual temporal integration as a function of distance: No compensation for sound-transmission time in human perception," Neurosci. Lett. **357**(2), 119–122.

Lopez-Poveda, E. A., and Meddis, R. (**1996**). "A physical model of sound diffraction and reflections in the human concha," J. Acoust. Soc. Am. **100**(5), 3248–3259.

Middlebrooks, J. C. (**1999**). "Individual differences in external-ear transfer functions reduced by scaling in frequency," J. Acoust. Soc. Am. **106**(3), 1480–1492.

Möbius, J., and Kobbelt, L. (**2012**). "Openflipper: An open source geometry processing and rendering framework," in *Curves and Surfaces*, edited by J.-D. Boissonnat, P. Chenin, A. Cohen, C. Gout, T. Lyche, M.-L. Mazure, and L. Schumaker (Springer, Berlin), pp. 488–500.

Mokhtari, P., Takemoto, H., Nishimura, R., and Kato, H. (**2015**). "Frequency and amplitude estimation of the first peak of head-related transfer functions from individual pinna anthropometry," J. Acoust. Soc. Am. **137**(2), 690–701.

Nykänen, A., Zedigh, A., and Mohlin, P. (**2013**). "Effects on localization performance from moving the sources in binaural reproductions," in *Proceedings of the 42nd International Congress and Exposition on Noise Control Engineering (INTER-NOISE 2013)*, September 15–18, Innsbruck, Austria, Vol. 4, pp. 3193–3201.

Parseihian, G., and Katz, B. F. G. (**2012**). "Rapid head-related transfer function adaptation using a virtual auditory environment," J. Acoust. Soc. Am. **131**(4), 2948–2957.

Poirier-Quinot, D., and Katz, B. F. G. (**2020**). "Assessing the impact of head-related transfer function individualization on task performance: Case of a virtual reality shooter game," J. Audio Eng. Soc. **68**(4), 248–260.

Poirier-Quinot, D., and Katz, B. F. G. (**2021**). "On the improvement of accommodation to non-individual HRTF via VR active learning and inclusion of a 3D room response," Acta Acustica (unpublished).

Poirier-Quinot, D., and Katz, B. F. G. (**2018**). "The Anaglyph binaural audio engine," in *Proceedings of the 144th Audio Engineering Society Convention*, May 23–26, Milan, Italy, pp. EB431:1–EB431:4.

Rébillat, M., Boutillon, X., Corteel, E. T., and Katz, B. F. G. (**2012**). "Audio, visual, and audio-visual egocentric distance perception by moving subjects in virtual environments," ACM Trans. Appl. Percept. **9**(4), 1–17.

Rugeles Ospina, F., Emerit, M., and Katz, B. F. G. (**2015**). "The three-dimensional morphological database for spatial hearing research of the BiLi project," Proc. M*eet*. Acoust. **23**, 1–17.

Schönstein, D., and Katz, B. F. G. (**2010**). "HRTF selection for binaural synthesis from a database using morphological parameters," in *Proceedings of the 20th International Congress on Acoustics*, August 23–27, Sydney, Australia, pp. 1–6.

Shaw, E. A. (**1997**). "Acoustical features of the human external ear," in *Binaural Spatial Hearing Real Virtual Environments*, edited by R. Gilkey and T. R. Anderson (Taylor & Francis, London).

Shinn-Cunningham, B. G. (**2000**). "Learning reverberation: Considerations for spatial auditory displays," in *Proceedings of the International Conference on Auditory Display*, April 2–5, Atlanta, GA, pp. 126–134.

Simon, L., Zacharov, N., and Katz, B. F. G. (**2016**). "Perceptual attributes for the comparison of head-related transfer functions," J. Acoust. Soc. Am. **140**, 3623–3632.

Spagnol, S., Geronazzo, M., and Avanzini, F. (**2013**). "On the relation between pinna reflection patterns and head-related transfer function features," IEEE Trans. Audio Speech Lang. Process. **21**(3), 508–519.

Stitt, P., Picinali, L., and Katz, B. F. G. (**2019**). "Auditory accommodation to poorly matched non-individual spectral localization cues through active learning," Sci. Rep. **9**(1), 1063.

Takemoto, H., Mokhtari, P., Kato, H., Nishimura, R., and Iida, K. (**2012**). "Mechanism for generating peaks and notches of head-related transfer functions in the median plane," J. Acoust. Soc. Am. **132**(6), 3832–3841.

Tommasini, F. C., Ramos, O. A., Hüg, M. X., and Bermejo, F. (**2015**). "Usage of spectral distortion for objective evaluation of personalized HRTF in the median plane," Int. J. Acoust. Vib. **20**(2), 81–89.

Torres-Gallegos, E. A., Orduña-Bustamante, F., and Arámbula-Cosío, F. (**2015**). "Personalization of head-related transfer functions (HRTF) based on automatic photo-anthropometry and inference from a database," Appl. Acoust. **97**, 84–95.

Trapeau, R., Aubrais, V., and Schönwiesner, M. (**2016**). "Fast and persistent adaptation to new spectral cues for sound localization suggests a many-to-one mapping mechanism," J. Acoust. Soc. Am. **140**(2), 879–890.

Warusfel, O. (**2002**). "LISTEN HRTF database, IRCAM," http://recherche.ircam.fr/equipes/salles/listen/ (Last viewed 13 January 2020).

Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L. (**1993**). "Localization using nonindividualized head-related transfer functions," J. Acoust. Soc. Am. **94**(1), 111–123.

Wightman, F. L., and Kistler, D. J. (**1992**). "The dominant role of low-frequency interaural time differences in sound localization," J. Acoust. Soc. Am. **91**(3), 1648–1661.

Zagala, F., Noisternig, M., and Katz, B. F. G. (**2020**). "Comparison of direct and indirect perceptual head-related transfer function selection methods," J. Acoust. Soc. Am. **147**(5), 3376–3389.

Zahorik, P., Bangayan, P., Sundareswaran, V., Wang, K., and Tam, C. (**2006**). "Perceptual recalibration in human sound localization: Learning to remediate front-back reversals," J. Acoust. Soc. Am. **120**(1), 343–359.

Ziegelwanger, H., Majdak, P., and Kreuzer, W. (**2015**). "Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization," J. Acoust. Soc. Am. **138**(1), 208–222.