

Inter-Laboratory Round Robin HRTF Measurement Comparison

Arete Andreopoulou, Durand R. Begault, and Brian F. G. Katz

Abstract—Head-Related Transfer Function (HRTF) measurements underlie the signal processing used in binaural auditory displays, but measurement techniques, equipment, and post-processing vary substantially between laboratories. This variation can result in significant differences in measured spectral and timing data taken from the same subject for the same sound source locations. An ongoing project for comparing databases from laboratories across the world (colloquially titled “Club Fritz”) employs a single dummy head microphone for measurements (Neumann KU-100) at various sites. The current study examines magnitude and timing differences between left and right ear data from 12 different HRTF sets taken from 10 different laboratories. Results revealed spectral magnitude variations up to 12.5 dB for frequency bands below 6 kHz and up to 23 dB above that, as well as large spectral left/right asymmetries ($d_{corr} \leq 0.4$) for high-frequency content. Further subjective studies are necessary to determine the perceptual relevance of these findings. Nevertheless, the observed ITD variations of up to 235 μ sec are alarming as they often exceeded reported JND values. Such findings highlight the potential impact of physical spaces, measurement routines, and equipment types on the collected HRTF data.

Index Terms—HRTF/HRIR, ITD variations, measurement, repeatability, spatial symmetry, spectral variations.

I. INTRODUCTION

ANALYSES of the characteristic variation in Head-Related Transfer Functions (HRTFs) as a function of sound source location allow an understanding of the perceptual cues to spatial hearing. HRTFs are frequency-domain transforms of time-domain Head Related Impulse Responses (HRIRs), which can be captured using specialized configurations of loudspeakers and microphones. Digital filters based on these measurements constitute the principal signal processing used for implementing binaural auditory displays. The HRTF characterizes the transformation of sound from a specific source location to the listener's ears in the free-field. Common measurement locations include near the eardrum or at the entrance to the ear canal (blocked meatus).

Manuscript received July 12, 2014; revised November 07, 2014; accepted January 26, 2015. Date of publication February 05, 2015; date of current version July 14, 2015. This work was supported in part by the French project BiLi (“Binaural Listening,” www.bili-project.org, FUI-AAP14). Additional support was provided by the Space Human Factors Engineering Project of NASA's Human Research Program. The guest editor coordinating the review of this manuscript and approving it for publication was Prof. Akio Ando.

A. Andreopoulou and B. F. G. Katz are with the Audio Acoustics Group, LIMSI-CNRS, 91400 Orsay, France (e-mail: andreopoulou@limsi.fr).

D. R. Begault is with the NASA Ames Research Center, Mountain View, CA 94035 USA.

Digital Object Identifier 10.1109/JSTSP.2015.2400417

Although binaural auditory displays and means for measuring HRTFs have existed for several decades, there is no single methodology used by laboratories worldwide. It is to be expected that differences in hardware, measurement resolution, microphone placement, test signal generation, equalization, and other factors could yield inconsistencies for the same purported sound source to receiver location transfer function. The current investigation highlights inter-laboratory differences while controlling several significant sources of variability. Specifically, the approach compares measurements using a single unique dummy head (*kunstkopf* in German). Head geometry and microphone type and location remain consistent, thereby highlighting differences due to factors unique to the measurement systems.

The “Club Fritz” is an open project, initiated in 2004 by authors Katz and Begault, with a current collection of over 60 submissions from 14 research labs from around the world. According to the project's call for contributions “Club Fritz” is a collection of researchers interested in HRTF measurement and measurement system comparisons. The goals are two-fold; respond to the fundamental question of how comparable are HRTF measurement systems, and address the need to establish a reference or standard for data quality. The database consists of physical HRTF measurements and numerical simulations performed on the same dummy head microphone system (“Fritz II” model KU-100, serial number 40, manufactured by Georg Neumann GmbH) [1], thereby controlling for morphology variations (fabrication differences between mannequins of the same make and model) and microphone position errors due to different microphone size and placement protocols.

Requirements for participation include these directives:

- 1) Participating institutions should perform HRTF measurements following their standard procedure.
- 2) The required deliverables are (a) raw Head-Related Impulse Responses (HRIRs) using the dummy head's internal microphones and (b) impulse responses of the reproduction equipment to be used for free-field equalization.
- 3) Additional deliverables may include measurements using other binaural microphones, multiple HRIR measurements for testing repeatability in the data, and free-field compensated post-processed data.

A parallel effort also exists concerning the acquisition of 3D geometrical mesh data and numerical simulation of HRTFs. Repeated submissions of data by a single institution, especially when changes are made in the measurement setup of a certain lab, are possible and encouraged. Preliminary comparison results regarding measured HRTFs have been presented in [2] while results for numerical simulations were presented in [3].

In this paper, an analysis of some of the observed differences between 10 different laboratories is highlighted by comparing 12 HRTF datasets from the “Club Fritz” database. One laboratory's submission included 3 different data sets. First, an overview of common HRTF measurement techniques and processing is presented. A review of prior inter- and intra-HRTF comparisons is provided. A consistent means of post-processing is described with corresponding results.

II. OVERVIEW OF HRTF MEASUREMENT TECHNIQUES

Acquiring direct measurements from a specific subject is the most accurate and explicitly documented means for HRTF data acquisition [4]. HRTFs are typically sampled on a spatial grid of varying resolution, taking advantage of both the finite localization resolution of humans and the ability to interpolate between measurements in an auditory display [4], [5]. Grids used are typically set to a constant distance equivalent to the longest wavelength of interest. Under the assumption of linearity and time-invariance, HRTF “pairs” (ipsilateral and contralateral measurements from both ears) can be measured and transformed into digital filters that capture the magnitude and timing differences at the two ears and then subsequently used in a binaural auditory display. Despite logistical complexity and lengthy duration, direct measurement is the most popular means for collecting binaural data, resulting in growing numbers of publicly available HRTF databases [6]–[11].

Various signal processing methods exist for measuring transfer functions, such as swept-sine, maximum length sequences, Golay codes, etc. Over the past decade several attempts have been made to improve the efficiency of the direct measurement method by minimizing recording time, without compromising on spatial resolution. Examples of such approaches are the Measurement via Reciprocity Method [12], the Multiple Exponential Sweep Method (MESM) [13], or the Continuous Measurement Method [14].

Post-processing of impulse response measurements allows removal of the transfer function of the measurement system. Such factors can include the responses of the microphone and loudspeaker as well as reflections from the measurement apparatus or room. The most direct method of processing consists in deconvolving the measured impulse response of the measurement system without the head present from the measured impulse response with the head present. The resulting data adheres to the definition of the HRTF, representing the transfer function due to the head only. Subsequent time windowing is applied to remove the presence of reflections. Alternative post-processing methods can be found in the literature, with the most common being division of the measured impulse responses by the mean (spatially weighted) response with the head present, resulting in what is termed the “directional transfer function” (DTF) [15] rather than the HRTF as it represents spatial variations of the head response relative to the global mean.

III. PRIOR RESEARCH COMPARING HRTF MEASUREMENTS

Currently, most spatial audio-related research focusing on the analysis and objective or subjective evaluation limit their data pool to HRTFs originating from a single database. The reason is twofold: first, not a great deal of data have been

made publicly available, and second, the lack of recognized standards for HRTF acquisition has resulted in databases of varying spatial resolutions, filter length, and sampling rates. Yet, even though several post-processing steps are necessary when combining databases to assure uniformity in the data format, their potential impact on the perceptual quality of the data is questionable, making inter-database comparisons a very challenging task [16].

A. Intra-Database Single Subject Repeatability

Several studies have been carried out concerning the quality and repeatability of measured HRTFs. Riederer [17] examined the repeatability of HRTF measurements for a given system, monitoring the degree of variance due to a number of factors such as signal noise floor, reflections from equipment, accidental movements or misalignment of subjects, and placement of the measurement microphones in a subject's ears. Similarly, Wersényi and Illényi [18] carried out repeated measurements on a computer motorized dummy-head system, exploring the effects of near-the-head objects like hair, caps, and glasses.

Intra-subject variations were explored in a comparison study of repeated HRTF measurements on four human subjects [19]. Spectral variations were observed as a function of frequency, direction, and distance of the ear to the sound source. This study was accompanied by a perceptual evaluation study which attempted to reveal the extent to which such variations were perceptually relevant.

B. Intra-Database Inter-Subject Variations

Previous research on inter-subject variations in HRTF databases has broadened our understanding of the degree of variability in HRTF data. Several metrics have been proposed [20]–[22] for a variety of temporal and spectral attributes. Algazi proposed two measures: the maximum Interaural Time Difference (ITD) and the pinna notch frequency, and investigated possible correlations between those and anthropometric data [6]. Zhong investigated the spatial symmetry of 52 HRTF datasets using cross-correlation analysis. Results revealed that symmetry varied as a function of frequency as well as source position (ipsilateral vs contralateral) [23].

C. Inter-Database Variations

Zhong *et al.* recently explored inter-database variations of binaural data by comparing five sets of KEMAR HRTFs available from three different research institutes' databases; the Massachusetts Institute of Technology Media Lab, the University of California Davis CIPIC Lab, and the South China University of Technology Acoustic Lab [24]. The datasets selected for this analysis had the advantage of sharing similar spatial and temporal resolutions, minimizing the need for additional processing for comparison. A common model mannequin allowed for minimal variation in the “subject's” morphology. Differences of greater than 10 dB were observed in the magnitude of frequencies above 6 kHz.

Preliminary results of the “Club Fritz” project were reported by Katz & Begault on data from five different labs, focusing on spectral variations and comparisons of ITD differences between sets [2]. Results compared both measurements done with the

TABLE I
OVERVIEW OF DATABASE CONTRIBUTORS AND ASSOCIATED MEASUREMENT SYSTEMS. # INDICATES THE NUMBER OF SPEAKERS

Institution	Country	Year	Positions	Azimuth	Elevation	f_s (Hz)	#	Dist. (m)
Institut de Recherche et Coordination Acoustique/Musique (IRCAM)	France	2004	2016	$\Delta 5^\circ$	$-45^\circ : 5^\circ : 90^\circ$	44100	1	1.95
University of Maryland (UMD)	USA	2007	823	$\Delta 5^\circ - \Delta 10^\circ$	$-35^\circ : 5^\circ : 90^\circ$	44100	6	0.90
Advanced Controls and Displays Group, NASA Ames Research Center	USA	2007	432	$\Delta 10^\circ$	$-40^\circ : 10^\circ : 70^\circ$	96000	12	0.90
Institut de Recherche et Coordination Acoustique/Musique (IRCAM)	France	2007	1944	$\Delta 5^\circ$	$-40^\circ : 5^\circ : 90^\circ$	44100	3	2.00
Institute of Technical Acoustics, RWTH Aachen	Germany	2009	2016	$\Delta 5^\circ - \Delta 10^\circ$	$-80^\circ : 5^\circ : 90^\circ$	44100	1	2.00
Department of Signal Processing and Acoustics, Helsinki University of Technology	Finland	2009	180	$\Delta 10^\circ$	$-40^\circ : 20^\circ : 40^\circ$	48000	1	1.50
Science & Technology Research Laboratories, NHK	Japan	2009	936	$\Delta 5^\circ$	$-30^\circ : 10^\circ : 90^\circ$	96000	10	1.30
Multisensory Cognition and Computation Laboratory, NICT	Japan	2009	1008	$\Delta 5^\circ$	$-40^\circ : 10^\circ : 90^\circ$	48000	1	1.00
Takeda Laboratory, Nagoya University	Japan	2009	2017	$\Delta 5^\circ$	$-50^\circ : 5^\circ : 90^\circ$	48000	1	1.00
RIEC, Tohoku University	Japan	2010	648	$\Delta 10^\circ$	$-80^\circ : 10^\circ : 90^\circ$	48000	35	1.50
Institut de Recherche et Coordination Acoustique/Musique (IRCAM)	France	2014	1560	$\Delta 6^\circ$	$-39^\circ : 3^\circ - 6^\circ : 86^\circ$	96000	1	1.95
Acoustics Research Institute, Austrian Academy of Sciences	Austria	2014	1550	$\Delta 2.5^\circ$	$-30^\circ : 5^\circ - 10^\circ : 86^\circ$	48000	22	1.20

KU-100 internal microphones as well as those obtained using binaural microphones unique to each institute. No interpolation was applied to the data, with the analysis focusing solely on a limited number of common measurement points across sets. Results showed ITD variations of up to 200 μsec for positions on the horizontal plane and revealed the presence of asymmetries for certain datasets. In addition, spectral variations such as shifted peak and notch frequencies for the same spatial position were found to be greater between datasets recorded with different laboratory mics than between those recorded with the built-in internal microphones.

A comparative study on numerical HRTF simulations using a 3-dimensional mesh generated from a Computerised Tomography (CT) scan of the Fritz II has also been carried out [3]. The resulting datasets were compared to select measurements from the “Club Fritz” collection. Spectral variations between numerically computed HRTFs were minimal, but exhibited a shift above 5 kHz when compared to physically measured sets, examined for the 0° azimuth position on the horizontal plane. Such differences resembled the variations previously observed between measurements made with the built-in and other house mics, and were therefore attributed to errors in the definition of the microphone position in the 3D mesh. ITD comparisons were more consistent, leading to differences $\leq 100 \mu\text{sec}$.

The current study offers a global comparison, rather than a point-by-point comparison, of a total of 12 “Club Fritz” HRTF datasets. Only the internal built-in microphone responses have been considered in an attempt to focus on inter-database variations independent of the type of microphone related variations mentioned above. The criteria examined concerns spectral variations, ITD comparisons, and left/right symmetry assessments between datasets.

IV. DESCRIPTION OF DATA ANALYZED

A. Contributors

The “Club Fritz” database HRTFs have been converted to the open standard SOFA file format for HRTFs [25] to enable ease of data navigation.¹ 12 HRTF sets were selected from the full database according to the following criteria:

- existence of measurement positions within the range of $\pm 30^\circ$ in elevation,
- minimum azimuthal resolution within the above mentioned range of $\leq 15^\circ$, and
- existence of measurements on the horizontal plane.

Table I offers an overview of the selected contributions. To preserve anonymity, all presented data and analysis are not identified. It is evident that there are significant variations with regards to spatial extent and grid density, with HRTF sets containing between 180 and 2017 points. Such differences in spatial resolution render direct comparisons between datasets difficult. In addition, due to varying grid definitions there exist very few points in common across the entire collection. Some representative examples of the measurement grids utilized by different labs can be seen in Fig. 1.

Other differences in measurement systems and protocols include the acoustic conditions of the measurement room, loudspeaker setup, test signal, processing method to obtain the impulse responses, and data acquisition sampling rate, f_s . Differences in loudspeaker setups included single speaker or different multi-speaker array configurations, with loudspeaker characteristics ranging from small 3 in. single drivers to 6.5 in. coaxial drivers, and physical spaces ranging from anechoic chambers to sound treated rooms.

¹These datasets are available for research use at the SOFA HRTF repository <http://sofacoustics.org/data/database/ClubFritz/>.

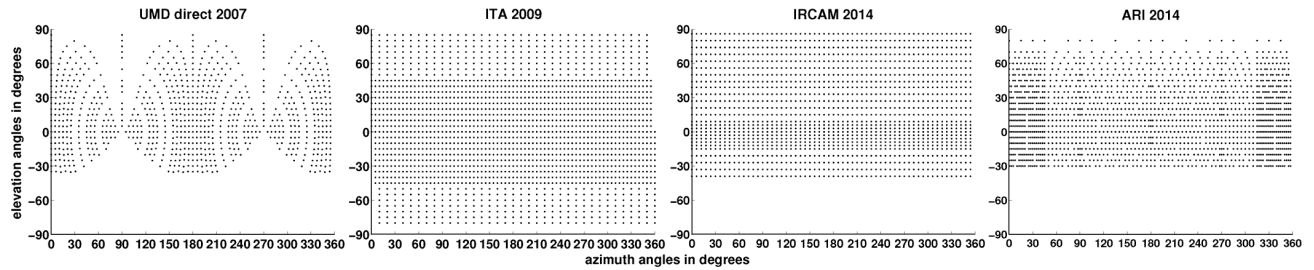


Fig. 1. Examples of measurement grids from four different labs. Points indicate angles measured locations, the x and y axes azimuth and elevation angles respectively.

B. Post-Processing

In order to achieve uniformity in the data used in this study, a series of post-processing steps were applied to the datasets. First, all measured responses were free-field equalized as described previously to remove the system responses in order to obtain actual HRTFs. It was preferable that this process be carried out in the course of this study, and not prior to submission by the separate institutions, due to possible variations in signal processing steps (Fourier transform length, smoothing, etc.) and methods used. However, one must have the following points in mind when reviewing the presented results. During the earlier stages of the “Club Fritz” project (prior to 2014), researchers were forbidden from removing the KU-100 built-in microphones to record the equalization responses. Rather, they used external reference microphones. As a result, the free-field equalized responses of these datasets have the equipment and room responses removed, but still contain the characteristics of the KU-100 mics. This is true for all datasets, except for the latest IRCAM and ARI submissions, which have the internal mic responses equalized.

Upon free-field equalization, all HRTFs were low-pass filtered at 18 kHz and sample-rate converted to 44.1 kHz. Limiting the range to 18 kHz was necessary for uniformity purposes. Any potential direct current (DC) offset was then removed and impulse responses were truncated using a 2.7 msec rectangular window in order to eliminate the presence of any room reflections. The window start point was set at 40 samples (0.9 ms) prior to the first detected “onset” as evaluated across the entire dataset. The onset was defined as the first sample greater than -10 dB relative to the peak value.

Left, right, and overall root mean square (RMS) values were computed across all datasets. Results revealed hemisphere imbalances of less than 1 dB, which, even though probably insignificant, were still corrected. In addition, the median global RMS value was selected as the reference level, and all HRTF datasets were gain adjusted to that common value. Finally, in order to situate the results of this study relative to human auditory perception, and not simply pure signal comparisons, data relevant to symmetry and spectral variations were processed according to the Equivalent Rectangular Bandwidth (ERB) scale, as defined by Moore [26], [27].

V. INTER-AURAL TIME DIFFERENCES

The first method of inter-dataset comparison is based on Inter-aural Time Difference variations. Several methods for the extraction of such information have been proposed over the years in numerous studies. Recent studies carried out on a single

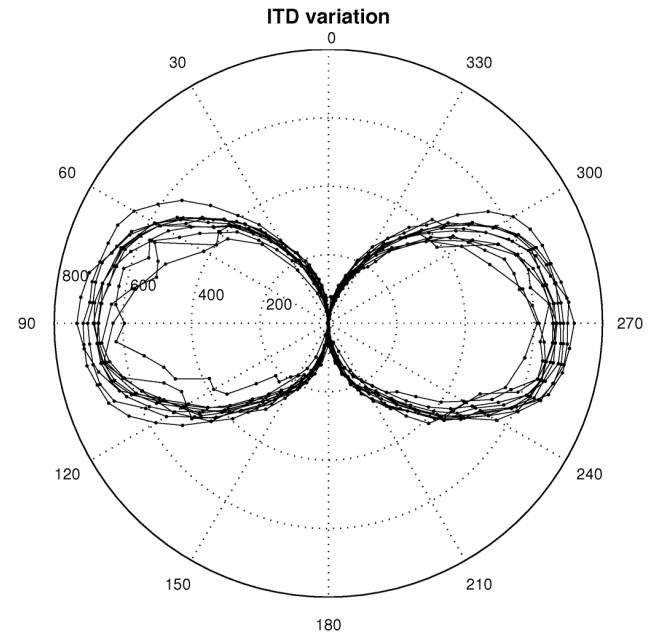


Fig. 2. Variability over horizontal plane estimated ITDs across all submissions, presented in μ secs.

HRTF dataset have shown that there is a significant level of variance between the various proposed methods [28].

For the purposes of this study, the threshold method described in [28] was selected to estimate ITD values. According to this method the ITD is calculated as the difference between the time of arrival of the impulse response (IR) at each ear for a given source direction. The time of arrival of each ear's impulse response was determined according to a relative threshold. A threshold value of -10 dB relative to the peak magnitude value of each response was employed to determine the time of first onset. The use of a relative threshold value is employed to eliminate the effect of the Inter-aural Level Difference (ILD) on the onset detection.

Fig. 2 presents extracted ITD values for horizontal plane source positions across all datasets. Results show variations increase as a function of distance from the median plane. More specifically, for the frontal region between 40° and 320° as well as for the rear region between 160° and 210° variations range between $\sim 30 \mu\text{sec}$ and $\sim 100 \mu\text{secs}$, while for points around the inter-aural axes variations reach up to $\sim 235 \mu\text{secs}$. Such variations can either be the result of positional misalignment of the dummy head or are related to the measurement setup (equipment type and arrangement) or the room itself. In an

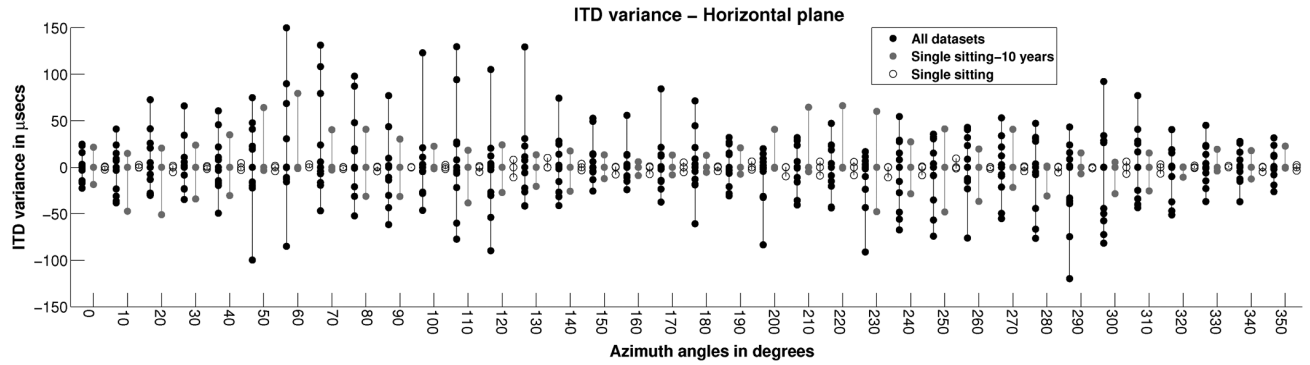


Fig. 3. Horizontal plane ITD variance across three HRTF subsets, relative to median values for each angle & dataset: all 12 datasets, 3 datasets from the same sitting (institution) carried out over a span of 10 years, 3 datasets from the same sitting carried out over a span of a week.

attempt to quantify the impact of each of those parameters on the HRTF measurement procedure, additional ITD comparisons were carried out on two specific subsets of the “Club Fritz” database each containing repeated measurements from a single institution. A description of the data and a comparative analysis follows.

A. Repeatability Comparisons

The first subset consists of a series of three HRTF datasets collected from the same institution, using the exact same system setup and room, over the span of a week. This type of comparison is typical in studies evaluating the repeatability of HRTF measurement systems. The second subset consists of three datasets recorded in the same physical space (same institution, anechoic room, turn-table, and loudspeaker positioning system) over the course of ten years, with associated evolutions in hardware and processing. Over this long period, even though the system setup had changed, the KU-100 dummy head and the means of positioning the head remained the same, leaving head misalignment and the equipment as the two principal variables that could potentially affect the result.

A comparison of the ITD variances on the horizontal plane was run for the three aforementioned HRTF subsets: (a) all 12 datasets, (b) 3 datasets from the same institution over 10 years, and (c) 3 datasets from the same institution over a week. Due to differences in spatial resolution between sets (see Table I) results were spline-interpolated in 10° azimuthal increments. This resolution was selected because it maintains the largest number of common points across datasets, keeping the offset between measured and interpolated points to $\leq 5^\circ$. Fig. 3 shows the comparison graph.

Various observations can be made. As already observed, there is less variance for azimuth locations closer to the median plane than for locations near the interaural axis. ITD variances for the 3 measurements repeated over the course of a week, when compared to the other two cases, are minimal, never exceeding $25 \mu\text{secs}$. Such observed variations are possibly due to dummy-head misalignment (position or rotation), changes in reflection positions from movable elements (cables), or other unpredictable experimental errors.

Over the span of 10 years, the maximum variance was $\sim 110 \mu\text{sec}$. This may indicate variations due to changes in processing and measurement equipment, as well as changes

in the dummy head over time, such as small variations in microphone and pinnae position after successive demounting and remounting. The 10 year repetition subset exhibits generally smaller variances than those observed over all HRTF datasets, clearly indicating the effect of variations in measurement apparatus and protocol, despite the common method used across all data sets in the present analysis for estimating ITD.

VI. SPECTRAL VARIATIONS

This section presents inter-dataset comparisons based on HRTF spectral variations. The purpose of this comparison is to reveal the potential impact of physical space, equipment, and measurement routines on the spectral content of HRTF data. Each of the 10 institutions participating with their data in this analysis utilized different routines for capturing the binaural IRs. This resulted in datasets of various spatial resolutions, rendering point by point comparisons of HRTFs challenging without the employment of some interpolation method. Nevertheless, interpolated data can introduce additional sources of spectral distortion, especially when the measurement grids are too sparse [29].

In order to limit the number of variables responsible for spectral content variations, no interpolation method was employed. The spectral comparison was limited to the 0° and 180° azimuth locations on the horizontal plane, the only common measurement locations across all of the analyzed data sets. Fig. 4 plots the left ear spectral content of these locations across all datasets. In both cases the magnitude responses are similar between most of the measurements. The general contours are very similar across sets, even though some features might not be apparent in every one of them. The pinna-notches centered around 8 kHz for the front and 9.5 kHz for the rear locations are generally stable. The observed minor shifts in frequency could have been caused by misalignment of the dummy head during the measurement procedures. There is also a question concerning misalignment of the internal microphones during reassembly of the head, though this is unlikely due to the design of the microphone support. In addition, for the frontal location there appears to be an entry whose magnitude response for frequencies up to 8 kHz are on the order of 5 dB lower than the remaining sets, and another that exhibits a strong frequency roll-off above 14 kHz. For the rear location, there appears to

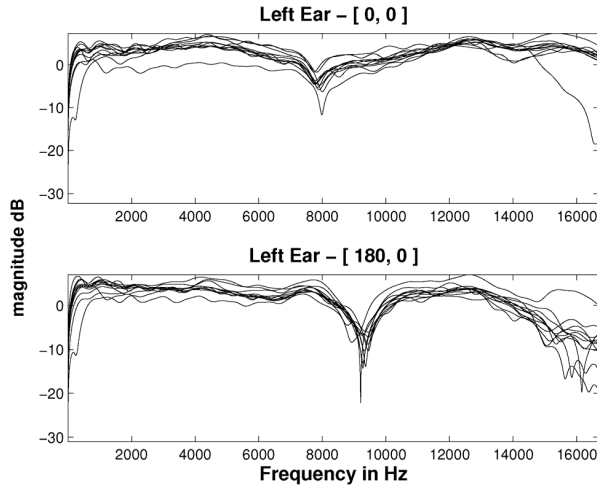


Fig. 4. Left ear spectral content of the 2 common locations across the 12 HRTF datasets.

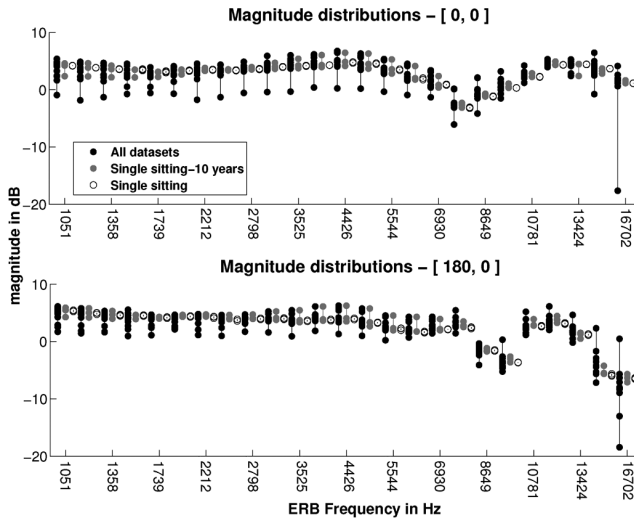


Fig. 5. Comparative plot of the magnitude distributions between all 12 “Club Fritz” datasets, 3 repeated measurements over 10 years, and 3 repeated measurements over a week, for 0° az/0° el and 180° az/0° el.

be a single set which above 12 kHz follows a different spectral pattern than the remaining sets.

The reasons for such inconsistencies are possibly related to variations in the measurement setups of different institutions. In an attempt to reveal the impact of such factors on the measured data, the range of spectral content variations across the 12 “Club Fritz” datasets was compared to the subset of the 3 measurements conducted by a single institution in the same anechoic room over the course of 10 years, as well as to 3 repeated measurements conducted in the same sitting over the span of a week. Fig. 5 shows the spectral difference for the three subsets at the two selected locations. Results are presented per ERB band for a relevant frequency range of 1 kHz to 16 kHz.

In general, the subgroup of repeated measurements conducted over a week demonstrated minimal spectral variations, less than 0.35 dB (mean 0.17 dB) for both locations. Such differences can be attributed to misalignment of the dummy-head during measurement setup, or to unpredictable non-linear behaviors of the measuring system. The range of variations was ≤ 3 dB for mea-

surements conducted in the same physical space over 10 years (mean 1.37 dB in the front and 1.42 dB in the rear). Larger variations were observed for frequency bands between 1 kHz and 1.5 kHz and between 3.5 kHz and 7 kHz in both locations. The increase demonstrates the impact of the equipment setup and measuring routine on the collected data, as the anechoic space, turn-table, and loudspeaker positioning system had remained the same across all three measurements. The resulting variance could not be completely eliminated even after free-field equalization and further processing the collected IRs.

When all 12 datasets were considered, spectral variations were much larger. For the frontal location and frequency bands below 6 kHz they ranged between 4.5 dB and 6.7 dB (mean 5.8 dB), while the rear were between 2.5 dB and 5 dB (mean 4.8 dB). Up to 10 kHz frontal variations ranged between 3 dB and 6 dB (mean 5 dB) and rear between 2.8 dB and 5.5 dB (mean 4 dB), while variations for very high frequency content reached up to 22 dB in the front and 19 dB in the rear.

This overview concentrated on two commonly measured points across 12 sets from the “Club Fritz” database, revealing the cumulative effect of factors such as the physical space, equipment, subject alignment, and measuring routine on the data. The following points should be taken into consideration with respect to the results:

- All HRTFs were collected on a dummy head, and are therefore presumably free of errors related to accidental movement during the measurement procedure.
- The perceptual relevance of these variations needs to be examined before any conclusions are drawn regarding the range of acceptable spectral variations between HRTFs.
- The current analysis is limited to a very small number of locations (2) and can only reveal possible trends in the spectral variations caused by different measurement protocols. For a more comprehensive comparison these variations need to be observed over a wider range of locations, which can be achieved by either imposing a common spatial resolution, or by using an elaborate high-precision interpolation method.

VII. LEFT/RIGHT SPECTRAL SYMMETRY

This section examines symmetric properties as a function of frequency and source position. The purpose of such analysis is twofold; first, to determine the overall symmetric properties of the Neumann KU-100 dummy head evaluated across datasets, and second, to investigate how the measured symmetry may vary between datasets from different institutions. The selected metric was correlation distance (proposed by [23]), defined as:

$$d_{corr} = 1 - \max \left(\frac{\sum_{k=1}^N H_l(\theta, \phi, f_k) H_r(\theta, \phi, f_k)}{\sqrt{\sum_{k=1}^N H_l(\theta, \phi, f_k)^2 \sum_{k=1}^N H_r(\theta, \phi, f_k)^2}} \right) \quad (1)$$

where $H_l(\theta, \phi, f_k)$ and $H_r(\theta, \phi, f_k)$ denote the left and right ear HRTF power spectra respectively. Computations were made within the ranges of $\pm 45^\circ$ in elevation.

Authorized licensed use limited to: Northwestern University. Downloaded on June 12, 2024 at 04:58:13 UTC from IEEE Xplore. Restrictions apply.

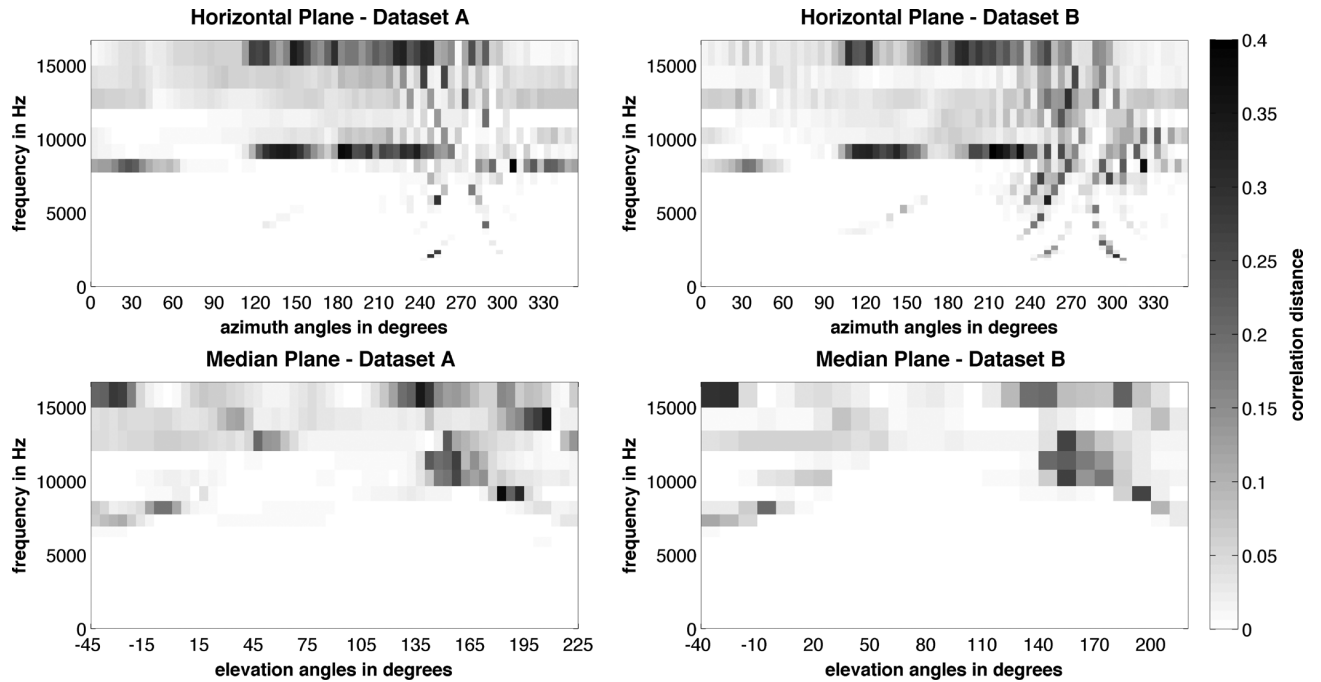


Fig. 6. Left/Right correlation distance of two HRTF datasets for locations on the (upper) horizontal and (lower) median planes. Azimuth angles $\leq 180^\circ$ correspond to ipsilateral and $> 180^\circ$ to contralateral content data respectively. Elevation angles $< 90^\circ$ correspond to positions in the front, while angles $> 90^\circ$ to positions in the rear. Correlation distance between left and right hemisphere data is coded according to the color-bar.

A. Left/Right Symmetry on the Horizontal and Median Planes

Fig. 6 compares the symmetric properties in the HRTF data as measured by two different labs across locations on the horizontal and median planes. Results for ipsilateral positions on the horizontal plane (upper two graphs) appear to be symmetric, in general, across the whole frequency range subject to three exceptions. There exist a peak in the computed correlation distances centered around 8 kHz for angles between 0° and 90° with d_{corr} values ranging between 0.2 and 0.25 and two peaks for locations between 90° and 180° for frequency bands centered around 9 kHz and above 15 kHz with d_{corr} between 0.25 and 0.4. The same three peaks propagate in the complementary contralateral positions (azimuth angles $> 180^\circ$). Nevertheless, the corresponding patterns for contralateral positions appear to be a bit more complicated. For azimuths between 210° and 330° there exist various peak d_{corr} values ranging between 0.1 and 0.4, across all frequencies above 2 kHz. The lower values of symmetry for contralateral positions could simply be a result of the more complex nature of this data, due to its reduced signal-to-noise ratio (SNR) caused by head shadowing.

For positions on the median plane, results show HRTFs are left/right symmetric across all frequencies below 7 kHz. For locations measured in the range of 20° to 160° in elevation, there appears to be a direct relationship between angle increase and presence of d_{corr} peaks at high frequencies. More specifically, as elevation increases HRTFs retain high symmetry for a much wider frequency range. As a result, HRTFs measured above the head (elevation angles around 90°) appear to be symmetric across the whole frequency range. In addition, locations measured towards the rear of the head (elevation angles $\geq 120^\circ$) exhibit increased d_{corr} peaks for frequency bands above 7 kHz.

Authorized licensed use limited to: Northwestern University. Downloaded on June 12, 2024 at 04:58:13 UTC from IEEE Xplore. Restrictions apply.

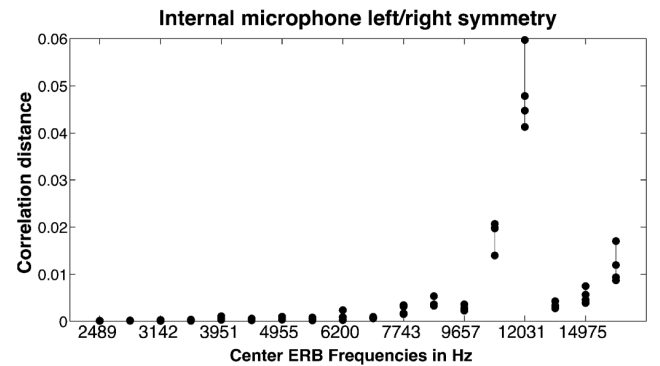


Fig. 7. Correlation distances for on-axis free-field responses, with binaural head removed, of the left/right internal microphones, per ERB bands. Results presented for 4 repetitions using 4 different speakers of a single measurement array.

Imbalances in left/right symmetry between locations in the frontal and rear hemisphere will be discussed in Section VII-B.

While variations were observed in symmetry levels of HRTF data from two selected submissions (Fig. 6), similar patterns are apparent across all submitted datasets. Therefore, one can infer that there exist some left/right asymmetric properties inherent in the dummy head used. To investigate this assumption, variability in the left/right internal microphone responses were examined separately using the free-field calibration data from one lab. Fig. 7 shows the correlation distance for the calibration responses of each of 4 speakers of a movable array recorded with the microphones removed from the head and oriented on-axis to each speaker. Results are displayed only for frequency content above 2 kHz, as low-frequency correlation data was relatively flat.

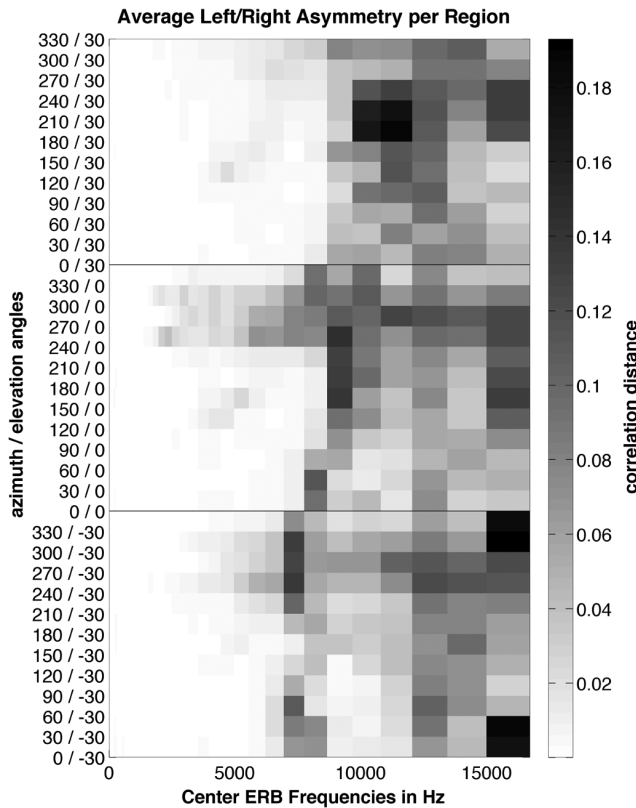


Fig. 8. Average left/right correlation distances per region cross all datasets. Regions are labeled according to the azimuth elevation coordinates of their center points (y-axis). The three different elevations are separated with solid lines. Asymmetry labels are presented across ERB frequency bands.

Results show that the measured data is symmetric across the whole relevant frequency range, with only minor fluctuations above 10 kHz, with $d_{corr} < 0.06$. Such correlation distance levels are minimal when compared to observed values of 0.4 in Fig. 6, and can neither be considered perceptually relevant, nor a possible cause of the observed patterns, which must be the result of asymmetries in the geometry between the left and right side of the Fritz II dummy head or mounting configurations of the pinnae and microphone assemblies, and not due to the microphones themselves.

B. Left/Right Symmetry Across Regions

As discussed earlier, variations in the spatial resolution of HRTF data originating from different labs prevent direct point-by-point inter-dataset comparisons. As a result, in order to monitor the left/right symmetry variations across the whole HRTF collection, data were divided into regions and symmetry was examined collectively for each subdivision. More specifically, the area within $\pm 45^\circ$ in elevation was divided according to a 12×3 grid of 36 non-overlapping regions each spanning 30° in azimuth and elevation, beginning at the 0° azimuth & 0° elevation location. This specific partitioning was selected so that all datasets had at least one measured point represented in each region. Data was divided according to the ERB scale and correlations were performed on the filters' power spectrum. The overall

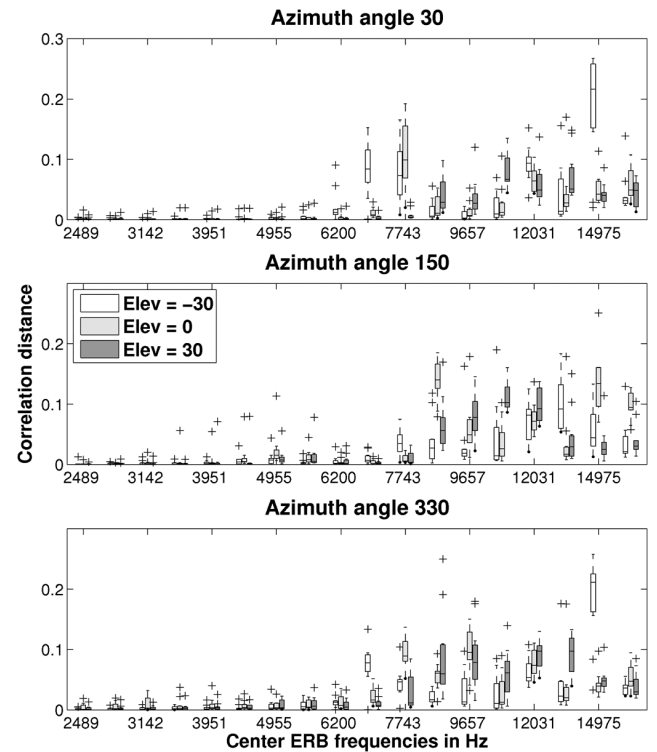


Fig. 9. Variations in the left/right symmetry for 9 regions. (Top) center locations: 30° azimuth at elevations -30° , 0° , and 30° . (Middle) center locations: 150° azimuth across the same elevations. (Bottom) center locations: 330° azimuth across the same elevations. Boxes mark data within the 25th and 75th percentiles, whiskers extend to $\pm 2.6\sigma$, (+) indicate outliers.

left/right correlation distance for each region was computed by averaging the left/right correlation distances of all binaural filter pairs present in that region. This necessary operation resulted in smoother representations of the data but the most prevailing patterns are still visible.

An overview of the computed values can be found in Fig. 8. The graph contains the mean left/right correlation distances for all 36 regions, across the relevant frequency range of 2 kHz to 16 kHz. The main observations noted in the previews section regarding symmetries on the horizontal and median planes still hold. For example, contralateral content data (210° to 330°) exhibit higher asymmetry in frequencies below 7 kHz for regions centered around -30° and 0° in elevation. These levels become less significant, however, for higher elevations. In addition, one can see that the center frequency of the first significant correlation distance peak ($d_{corr} > 0.1$), increases as a function of elevation angle. For regions centered around -30° in elevation these peaks appear at ~ 7 kHz, for regions at 0° in elevation at ~ 8 kHz to ~ 9 kHz, while for those at 30° in elevation above ~ 10 kHz.

To facilitate a more comprehensive investigation, Fig. 9 presents variations in the left/right symmetry for 9 selected regions across all datasets, showing ipsilateral regions centered around 30° in azimuth, ipsilateral regions centered around 150° in azimuth, and contralateral regions centered around 330° in azimuth respectively, each across the 3 elevations.

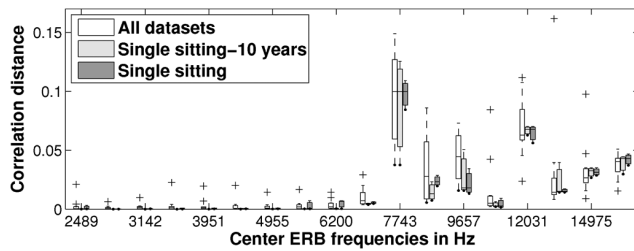


Fig. 10. Left/right asymmetry distributions of the 0° azimuth & 0° elevation region across: all 12 labs, 3 measurements from the same lab over the course of 10 years, and 3 measurements from a single sitting. Boxes mark data within the 25th and 75th percentiles, whiskers extend to $\pm 2.6\sigma$, (+) indicate outliers.

The following observations were made:

- Left/right asymmetries arise for frequency content above 7 kHz regardless of region position and condition (ipsilateral or contralateral).
- For ipsilateral results in the frontal plane asymmetries in the data increase as elevation decreases. Asymmetries in the region of 7 kHz are more pronounced at -30° and 0° elevations.
- For ipsilateral results in the rear plane asymmetries follow a similar behavior to those in the frontal plane, but the variation distribution is wider, indicating stronger patterns, generally increasing with elevation.
- For contralateral results asymmetry patterns are similar to ipsilateral rear results, though initiating at lower frequencies and evolving less evidently with elevation.
- For regions centered on the horizontal plane, the correlation distance peak at 8.6 kHz corresponds to the Neumann KU-100 pinna notch range. The reasons for this left/right asymmetry are either related to the geometrical differences of the pinnae or due to repositioning variations of the microphone and pinnae assembly.

It has been demonstrated that the origin of the observed left/right asymmetry patterns is not related to the internal microphone responses (Fig. 7). It was rather suggested that the physical geometry or assembly of the dummy head was a more probable cause. In an attempt to further explore this matter, Fig. 10 shows the asymmetry distributions for a single region centered around the 0° azimuth & 0° elevation point across: all 12 labs, 3 measurements from the same lab over the course of ten years, and 3 measurements from a single sitting. Observations made for this subset are generalizable to the remaining HRTF regions.

While for frequencies below 7 kHz the data is strongly left/right symmetric, for higher frequencies variations increase as a function of the number of factors affecting the quality of the measurement procedure. Examples of such factors are the physical space that the measurements had taken place, equipment setup, spatial resolution, dummy head alignment, precise microphone and pinnae assembly alignment and position in the head, excitation signal, method employed for capturing the impulse responses, etc.

As seen in Fig. 10, the distribution of correlation distances across all sets is greater than that observed for the 10 year repetition measurements conducted in the same anechoic chamber. This change demonstrates the possible effects of the physical

measurement space on the left/right symmetry. Similarly, the possible effects of measurement setup and capture procedure can be estimated by comparing the subset of 3 repeated measurements conducted in the same anechoic space over a week, which included remounting but not disassembly of the dummy head (i.e. microphone pinnae assembly), to the group of measurements conducted over 10 years.

Nevertheless, even though the distribution of the variations changes as a function of the aforementioned factors and conditions affecting the measurements, another interesting point arises from the fact that the median correlation distance values for frequencies above 7 kHz remain at the same levels across all three data subsets (except for the 9.7 kHz band). It is our understanding that this reflects the inherent left/right asymmetry properties, particular to the binaural dummy head microphone used in this project. The true levels of that asymmetry are hard to determine without some interpolation that would allow direct point-by-point comparisons across all datasets. In addition, it remains unclear whether the frequency range and levels of these inherent asymmetries have a relevant perceptual effect on binaural rendering.

VIII. DISCUSSION

Even though HRTF measurement has been a widely used solution to binaural audio for several decades, there has been no single, widely utilized method for data capture. An overview of different publicly available HRTF databases [6]–[11] and their methods for data collection reveals differences in hardware, measurement resolution, microphone placement, test signal generation, equalization, and other factors which could result in variations in the collected location dependent transfer functions. The repeatability of the measurement procedure, as well as the effect of such factors on measurements, has been extensively discussed in a series of studies analyzing data collected from dummy-head binaural microphones [17], [18], [24]. However, the main limitation of such works is that their data were collected from a single or very small number of sittings, hence not fully exploring the potential impact such factors have on HRTFs.

“Club Fritz” is an on-going project for HRTF measurement system comparisons containing over 60 datasets from 14 research labs around the world. The uniqueness of this collection lies in that by controlling for several variables like head geometry and binaural microphone type and placement, emphasis in the analysis can be given to differences due to factors particular to every participating institution. The current study offers a global comparison of a subset from the “Club Fritz” database consisting of 12 KU-100 HRTF datasets from 10 laboratories. Only the built-in microphone responses were considered in an attempt to focus on inter-database variations independent of microphone-related variations. The analysis focused on spectral and interaural time difference comparisons as well as left/right symmetry assessments between datasets.

The spectral variation overview between the 12 HRTF datasets was based on the two common points measured across all institutions: 0° and 180° azimuth on the horizontal plane. The results were compared to two HRTF subsets measured in the same physical space over the span of a week and 10 years.

The variability in HRTF data has been explored previously. Riederer [17] examined the repeatability of HRTF measurements for a given system. Data was collected in the same physical space using the same equipment setup and measurement routines over the span of a few months. Overall, the reported intra-subject variations were up to 2 dB for frequencies below 6 kHz and between 3–5 dB for frequencies below 10 kHz. This collection is comparable to the two Club-Fritz subsets of repeated measurements. For the first, collected in the same space over the span of a week, the computed variations were up to 0.3 dB for frequencies up to 6 kHz (mean 0.23 dB), between 0.1 dB and 0.33 dB (mean 0.15 dB) for frequencies up to 10 kHz, and up to 0.34 dB for very high frequency content. For the second, collected in the same space over 10 years, the computed variations were up to 3 dB for frequencies up to 6 kHz (mean 1.58 dB), and 0.5–2 dB (mean 1.13 dB) for frequencies up to 10 kHz. However one should keep in mind that observations on both subsets were affected by the small number of locations and repetitions studied. It is possible that a more global comparison including more repetitions and locations would have yielded numbers closer to those reported by [17].

Inter-database variations of dummy-head HRTFs have also been explored on a smaller-scale study by Zhong *et al.* looking at differences between five KEMAR HRTF sets collected by three institutions [24]. Spectral variations were calculated for a frequency range of 0.5–20 kHz and averaged to provide a single overall metric of inter-database spectral variance, ranging from 0.7–3.2 dB. In the current study, comparisons across all 12 datasets resulted in average spectral variations of 6 dB (min 1.4 dB, max 22 dB) and 5 dB (min 2.5 dB, max 19 dB) for the front and rear locations respectively across a frequency range of 1–16 kHz.

The differences between the two studies can be attributed to the larger number of HRTFs in the Club-Fritz database, which cover a much wider range of measurement systems, as well as to the difference in the number of positions analyzed. While both studies focus on commonly measured locations, hence eliminating the need for interpolation, in [24] comparisons were run across 72 locations on the horizontal plane, while in the present study across only 2. However, both studies discuss ideal situations as they analyze HRTF data of dummy-heads, which are free of accidental subject movements, offer better control of subject alignment, and are not affected by the binaural microphone placements, though [24] did not use the same head and microphones, only data from the same model dummy head, so microphone placement and response are not assured to be the same.

In the HRTF variability study of Andreopoulou *et al.*, presenting repeated measurements on human subjects, the range of reported spectral variations differed. For ipsilateral content spectral variations reach up to 12.5 dB for frequency bands below 6 kHz, and up to 23 dB above that, while for contralateral content observed variations increase by up to 4 dB. This drastic increase reveals the impact of the human subject factor on the measured data.

Comparisons of the extracted ITD information from the 12 HRTF datasets revealed variations up to $\sim 235 \mu\text{sec}$ for locations on the horizontal plane and an average variation

of $\sim 120 \mu\text{sec}$. In an earlier preliminary study on 5 HRTF datasets from the “Club Fritz” project, Katz & Begault utilized 3 methods for ITD information extraction (Threshold, Maximum of the Inter-Aural Cross-Correlation, and Centroid of the Inter-Aural Cross-Correlation) and reported variations within a slightly smaller range ($\sim 200 \mu\text{sec}$), even though the responses were not free-field equalized [2]. Variations were attributed to factors related to potential misalignment of the head, the measurement setup, or the physical measurement space itself.

To further investigate the cause of these variations 2 additional HRTF subsets were analyzed in this work, each containing repeated measurements from a single institution. In the first case, where the repeated measurements were conducted using the exact same system and room over the span of a week, ITD variations were found to be minimal, never exceeding $\sim 25 \mu\text{sec}$ (avg. offset $\sim 7 \mu\text{sec}$) across locations on the horizontal plane. In the second case, however, where the repeated measurements were conducted in the exact physical space over the course of 10 years, ITD variations were found to be considerably larger (max. offset $\sim 110 \mu\text{sec}$, avg. offset $\sim 50 \mu\text{sec}$), but still much smaller than when compared to the variations across all 12 datasets.

This comparison highlights the impact of the physical space and measurement setup on the collected data, even after the necessary processing steps have been applied. However, it still presents the ideal case as the data was collected from a dummy head, and hence is free of accidental subject movement artifacts, as was well as variations due to binaural microphone placement. In a similar HRTF variability study Andreopoulou *et al.* compared ITD variations on repeated HRTF measurements on 4 human subjects. To maximize variance, two different rooms, speaker setups, and tracking systems were utilized, but the capture and transfer function extraction methods remained the same [30]. The ITD analysis revealed variations up to $\sim 160 \mu\text{sec}$ across subjects.

This work also demonstrated the tendency of ITD variations to change as a function of distance from the median plane, with angles in the frontal and rear regions exhibiting smaller variations ($\sim 30 \mu\text{sec}$ to $\sim 100 \mu\text{sec}$) than those around the interaural axis ($\sim 235 \mu\text{sec}$). Similar behaviors have been reported in the literature [19], and were also indicated in the ITD extraction plots of [2] even though not explicitly discussed.

The left/right symmetry of the 12 KU-100 “Club Fritz” HRTF datasets was also a subject of investigation. Correlation distance calculations were performed for positions on the horizontal and median planes, as well as for 36 regions between $\pm 45^\circ$ in elevation. Similar work was conducted by Zhong *et al.* on a database of 52 Chinese subjects [23]. Before a direct comparison is made between the two studies, the following two fundamental differences must be acknowledged; 1) the current study is based on HRTF data from a single dummy head collected by various laboratories, while [23] used measurements on human subjects in a single laboratory. Hence, while in this study variances should be attributed to external factors, in [23] they are primarily the result of differences in the head geometry of subjects; 2) the spatial homogeneity between HRTF sets in [23] enables point-by-point comparisons, while in this work data comparisons are made collectively by region.

Analysis of the results revealed symmetry variations as a function of position (ipsilateral vs contralateral) and frequency. More specifically, left/right asymmetries were significant ($d_{corr} \geq 0.1$) for frequencies above 7 kHz for ipsilateral content and 2 kHz for contralateral. In comparison, results reported in [23] indicated lower corresponding symmetry frequency limits (5 kHz and 1.1 kHz respectively).

However, it should be noted that this study presents the range of asymmetries per frequency-band and region, while [23] reported only averaged asymmetry values cross all 52 HRTF subjects, one for frequency bands below 5 kHz ($d_{corr} = 0.02$) and a second frequencies above ($d_{corr} = 0.22$). Applying an equivalent analysis with the “Club Fritz” data collected here, the corresponding distance correlation values are significantly lower: $d_{corr} = 0.005$ for frequency bands less than 5 kHz and $d_{corr} = 0.06$ for frequency above. This difference is expected as [23] investigated the symmetric properties of different heads recorded in the same room, using the same method and equipment, whereas the current study explores the effect of the latter attributes on the symmetry of HRTF data collected on a single binaural dummy head.

Analysis also revealed an inverse relationship between left/right symmetry and elevation, according to which correlation distances decrease as elevation angle increases. This finding is in agreement with past literature indicating that at higher elevations (above 30°) spectral variations become flatter and repeated measurement accuracy improves significantly [18]. In the “Club Fritz” database the left/right symmetry on the median plane for elevations between 60° and 120° increased ($0 \leq d_{corr} \leq 0.1$), verifying the above observation.

Finally, analysis was carried out according to the previously described subsets of repeated measurements conducted by a single institution to determine the effects of the room and measurement setup as well as the stability of the given dummy head, over time. As expected, left/right symmetry exhibited fewer variations as the number of factors affecting the measurement procedure decreased. More precisely, for the subset of repeated measurements conducted using the exact same system and room over the span of a week the pattern variations were almost minimal, while for the subset of repeated measurements conducted in the same physical space over the span of 10 years, the variation levels were less or comparable to the ones of the full database. Nevertheless, this comparison revealed that the median correlation distance values at high frequencies remained at the same levels across all three data subsets. This finding implies the presence of inherent left/right asymmetry properties particular to the binaural dummy head used in this project, which is not related to the responses of the built-in binaural microphones, but rather is the result of the physical geometry. The actual levels and angles of this asymmetry cannot be investigated without the use of elaborate interpolation methods which would allow for direct point-by-point comparisons between all datasets in the “Club Fritz” database. In addition, subjective studies are necessary to determine whether such values have a perceptually relevant impact on the rendered binaural data.

IX. CONCLUSIONS

This work highlighted variations observed in measured binaural impulse responses in the absence of a widely accepted and utilized method for HRTF measurements. Emphasis was put on variability in extracted Interaural Time Difference information, spectral variations across commonly measured locations, and left/right spectral symmetry of HRTFs in the “Club Fritz” collection. Future work will involve a spherical harmonic decomposition of the HRTF data, which will allow for elaborate point-by-point comparisons between different measurement grids, as well as subjective evaluations of the perceptual relevance of the reported variations.

Results of this study have shown that despite the use of a single subject a certain degree of variability is observed between “identical” HRTF measurements. While the perceptual importance of the spectral and left/right symmetry variations remains to be examined, these differences must be taken into account as an quantifiable degree of “measurement error” when comparing datasets. Variations in ITD can be considered more troubling, as there are more established studies concerning ITD perception and the observed variations are in excess of what one would expect for repeated measurements on a single subject. Additional studies are required to determine the origin of these variations as being due to positional errors, measurement protocol, signal processing architecture, etc.

ACKNOWLEDGMENT

The authors thank all those who have contributed to the “Club Fritz” database over the years on its numerous voyages around the world.

REFERENCES

- [1] The Dummy Head—Theory and Practice. Berlin [Online]. Available: <https://www.neumann.com/download.php?download=docu0009.PDF>
- [2] B. Katz and D. Begault, “Round robin comparison of HRTF measurement systems: Preliminary results,” in *Proc. 19th Int. Congr. Acoust. (ICA '07)*, Madrid, Spain, 2007, pp. 1–6 [Online]. Available: https://www.researchgate.net/publication/236213216_Round_robin_comparison_of_HRTF_measurement_systems_preliminary_results
- [3] R. Greff and B. Katz, “Round robin comparison of HRTF simulation results: Preliminary results,” in *Proc. 123rd AES Conv.*, New York, NY, USA, Oct. 2007, no. 7188, pp. 1–5 [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=14246>
- [4] D. R. Begault, *3-D Sound for Virtual Reality and Multimedia*, 2nd ed. New York, NY, USA: Academic, 1994.
- [5] J. C. Makous and J. C. Middlebrooks, “Two-dimensional sound localization by human listeners,” *J. Acoust. Soc. Amer.*, vol. 87, no. 5, pp. 2188–2200, 1990.
- [6] V. Algazi, R. Duda, D. Thompson, and C. Avendano, “The CIPIC HRTF database,” in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, New Paltz, NY, USA, Oct. 2001, pp. 99–102.
- [7] T. Carpentier, H. Bahu, M. Noisternig, and O. Warusfel, “Measurement of a head-related transfer function database with high spatial resolution,” in *Forum Acousticum*, Sep. 2014, pp. 1–6 [Online]. Available: http://www.fa2014.agh.edu.pl/fa2014/_cd/article/RS/R19_3.pdf, Krakow: European Acoustics Association
- [8] N. Gupta, A. Barreto, M. Joshi, and J. Agudelo, “HRTF database at FIU DSP lab,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Dallas, TX, USA, Mar. 2010, pp. 169–172.
- [9] T. Nakado, T. Nishino, and K. Takeda, “Measurements of head-related transfer function in sagittal and frontal coordinates,” *J. Acoust. Soc. Amer.*, vol. 123, no. 5, pp. 3652–3652, 2008.

- [10] O. Warusfel, Listen HRTF Database. 2003. [Online]. Available: <http://recherche.ircam.fr/equipes/salles/listen/>
- [11] B.-S. Xie, X.-L. Zhong, D. Rao, and Z.-Q. Liang, "Head-related transfer function database and its analyses," *Sci. China Ser. G: Phys., Mechan. Astron.*, vol. 50, no. 3, pp. 267–280, Jun. 2007.
- [12] D. Zotkin, R. Duraiswami, E. Grassi, and N. Gumerov, "Fast head-related transfer function measurement via reciprocity," *J. Acoust. Soc. Amer.*, vol. 120, no. 4, pp. 2202–2215, 2006.
- [13] P. Majdak, P. Balazs, and B. Laback, "Multiple exponential sweep method for fast measurement of head related transfer functions," *J. Audio Eng. Soc.*, vol. 55, no. 7/8, pp. 623–630, 2007.
- [14] K. Fukudome, T. Suetsugu, T. Ueshin, R. Idegami, and K. Takeya, "The fast measurement of head related impulse responses for all azimuthal directions using the continuous measurement method with a servoswiveled chair," *Appl. Acoust.*, vol. 68, no. 8, pp. 864–884, Aug. 2007.
- [15] D. Hammershoi and H. Moller, "Methods for binaural recording and reproduction," *Acta Acust. United With Acust.*, vol. 88, no. 3, pp. 303–311, 2002.
- [16] A. Andreopoulou and A. Roginska, "Towards the creation of a standardized HRTF repository," in *Proc. 131st Audio Eng. Soc. Conv.*, New York, NY, USA, Oct. 2011, p. 8571 [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=16096>
- [17] K. A. J. Riederer, "Repeatability analysis of head-related transfer function measurements," in *Proc. 105th Audio Eng. Soc. Conv.*, San Francisco, CA, USA, Sep. 1998 [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=8334>
- [18] G. Wersényi and A. Illényi, "Differences in dummy-head HRTFs caused by the acoustical environment near the head," *Electron. J. Tech. Acoust.*, vol. 1, pp. 1–15, 2005.
- [19] A. Andreopoulou, A. Roginska, and H. Mohanraj, "Analysis of the spectral variations in repeated head-related transfer function measurements," in *Proc. Int. Conf. Auditory Display (ICAD)*, Lodz, Poland, 2013, pp. 213–218.
- [20] C. Avendano, R. Duda, and V. Algazi, "Modeling the contralateral HRTF," in *Proc. AES 16th Int. Conf. Spatial Sound Reproduction*, Rovaniemi, Finland, Apr. 1999, pp. 313–318.
- [21] Q. Huang and K. Liu, "A reduced order model of head-related impulse responses based on independent spatial feature extraction," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 281–284.
- [22] A. Bondu, S. Busson, V. Lemaire, and R. Nicol, "Looking for a relevant similarity criterion for HRTF clustering: A comparative study," in *Proc. 120th Audio Eng. Soc. Conv.*, Paris, France, May 2006, no. 6653 [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=13457>
- [23] X.-L. Zhong, F.-C. Zhang, and B.-S. Xie, "On the spatial symmetry of head-related transfer functions," *Appl. Acoust.*, vol. 74, no. 6, pp. 856–864, Jun. 2013.
- [24] X.-L. Zhong and B.-S. Xie, "Consistency among the head-related transfer functions from different measurements," in *Proc. Meetings Acoust.*, Montreal, QC, Canada, 2013, vol. 19, no. 050014.
- [25] P. Majdak, Y. Iwaya, and T. Carpentier, "Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions," in *Proc. 134th Audio Eng. Soc. Conv., Rome: Audio Eng. Soc.*, May 2013, no. 8880. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=16781>
- [26] B. Moore and B. Glasberg, "A revision of Zwicker's loudness model," *Acta Acust. United With Acust.*, vol. 82, pp. 335–345, 1996.
- [27] B. C. Moore, *An Introduction to the Psychology of Hearing*. Leiden, The Netherlands: Brill, 2012.
- [28] B. F. G. Katz and M. Noisternig, "A comparative study of interaural time delay estimation methods," *J. Acoust. Soc. Amer.*, vol. 135, no. 6, pp. 3530–3540, 2014.
- [29] X.-L. Zhong and B.-S. Xie, "Maximal azimuthal resolution needed in measurements of head-related transfer functions," *J. Acoust. Soc. Amer.*, vol. 125, no. 4, pp. 2209–2220, Apr. 2009.
- [30] A. Andreopoulou, A. Roginska, and H. Mohanraj, "A database of repeated head-related transfer function measurements," in *Proc. Int. Conf. Auditory Display (ICAD)*, Lodz, Poland, Jul. 2013, pp. 287–291 [Online]. Available: http://www.icad2013.com/paper/38_S9-04_Andreopoulou.pdf



Areti Andreopoulou is a Post-Doctoral Researcher at the LIMSI-CNRS laboratory in the Audio & Acoustics group. She has a background in music technology and obtained a Ph.D. in music technology from New York University. Her fields of interest include spatial audio, immersive environments, auditory displays, and data sonification.



Durand R. Begault received his Ph.D. from the University of California San Diego in 1987, where he developed 3D audio technologies at the Computer Audio Research Laboratory. He has been associated with the Human Systems Integration Division of NASA Ames Research Center (Moffett Field, California) since 1988. He works in the area of research and development of new audio and multi-modal technologies for aeronautic and space applications, including psychoacoustic research, human factors evaluation, sound quality, acoustical modeling and communications engineering. He is a Fellow of the Audio Engineering Society and an Adjunct Professor at McGill University.



Brian F. G. Katz is a Researcher at the LIMSI-CNRS laboratory, coordinator of the Sound & Space research team, and head of the Audio & Acoustics group. His fields of interest include spatial 3-D audio rendering and perception and room acoustics. With a background in physics and philosophy, he obtained his Ph.D. in acoustics from Penn State in 1998 and his HDR in engineering sciences from UPMC in 2011. Before joining CNRS, he worked for various acoustic consulting firms, including Artec Consultants Inc., ARUP & Partners, and Kahle Acoustics. He has also worked at the Laboratoire d'Acoustique Musical (UPMC) and IRCAM in Paris.