Review of: "Solving the naming ambiguity of auditory localization mechanisms: HRTF & HRSL, and ILD, ITD, IPD"

Brian Katz[1]

1 Sorbonne University

Potential competing interests: No potential competing interests to declare.

“Solving the naming ambiguity of auditory localization mechanisms: HRTF & HRSL, and ILD, ITD, IPD”, this manuscript purports to address a nomenclature confusion in the fields of psychoacoustics and spatial audio, fundamentally between physical and perceptual mechanisms. However, the confusion does not appear to be prevalent in the literature, and would rather appear to be a confusion on the part of the author regarding the concrete definitions of the terms in use, and how they are applied in the different fields of study.

While included in the title, ILD, ITD, and IPD are only cursorily mentioned and never even defined.

The list of references employed to support he work is rather weak, with only 1 journal article and 3 conference/workshop proceedings. With the extensive literature, including numerous reference textbooks on the subject, the supporting arguments appear inadequate.

The major points in the review:

Let us consider the following starting points for reference on the subject:

- Blauert, J. Spatial Hearing: *The Psychophysics of Human Sound Localization*, 1983
- Begault, D.R. *3-D Sound for Virtual Reality and Multimedia*, 1994

Subsequent texts can be considered, which provide further examples of common terminology usage:

- Strumillo, P. *Advances in Sound Localization*, 2011
- Xie, B., *Head-Related Transfer Function and Virtual Auditory Display*, 2013
- Katz, B. and Majdak, P. *Advances in Fundamental and Applied Research on Spatial Audio*, 2022

To summarize, the HRTF, or Head-Related Transfer Function, is by definition the acoustic filtering effect that the human head, outer ears, and body have on sounds as they travel from a sound source to the ears. Different acoustic features of the HRTF can be exploited by the auditory system in the aims of localizing a sound, commonly termed the process of “binaural localisation”, or more generally “spatial hearing”. The HRTF can also be measured, or numerically simulated, and employed in an audio processing chain to apply these same features, thereby creating a virtual acoustic reality, or virtual auditory environment, generally termed “binaural synthesis”. The same can be achieved physically, without any audio signal processing, through the use of binaural recordings with in-ear microphones on an individual or using a

dummy-head binaural microphone.

Similarly, the ITD (Inter-aural Time Difference), ILD (Inter-aural Level Difference), and IPD (Inter-aural Phase Difference) are all physical acoustic measures that can be derived from the HRTF. How these are interpreted by the auditory system is of course a matter of study, and there are various auditory models which transform these acoustic cues into more neurological ones. However, there does not appear to be any common confusion in these domains and therefore, at least to this reviewer, no need for the introduction of a resolution to an ambiguity problem which does not currently exist in the domains concerned.

Where some ambiguity exists, we can propose the following precisions:

The HRTF is commonly defined as a function of azimuth and elevation around the listener, for a given measurement distance (a result of the method by which HRTFs are typically acquired). The singular “HRTF” therefore refers to the individual transfer function of one person for a given source radius. The plural “HRTFs” denotes a collection of transfer functions from different measurement conditions (e.g. distances), or more typically from different individuals.

There is much research regarding the individuality of HRTFs and how the auditory system processes auditory cues from a different HRTF, other than from the individual concerned. In these studies, there are various types of HRTFs. We use the term *individual* to identify the HRTF of the listener, *individualized* or *personalized* to indicate an HRTF modified or selected to accommodate the listener best, and *non-individual* or *non-individualized* to indicate an HRTF that has not been tailored to the listener. A so-called *generic* or *dummy-head* HRTF is a specific instance of a non-individual HRTF.

The same modifiers are similarly applied to the ILD, ITD, and IPD.