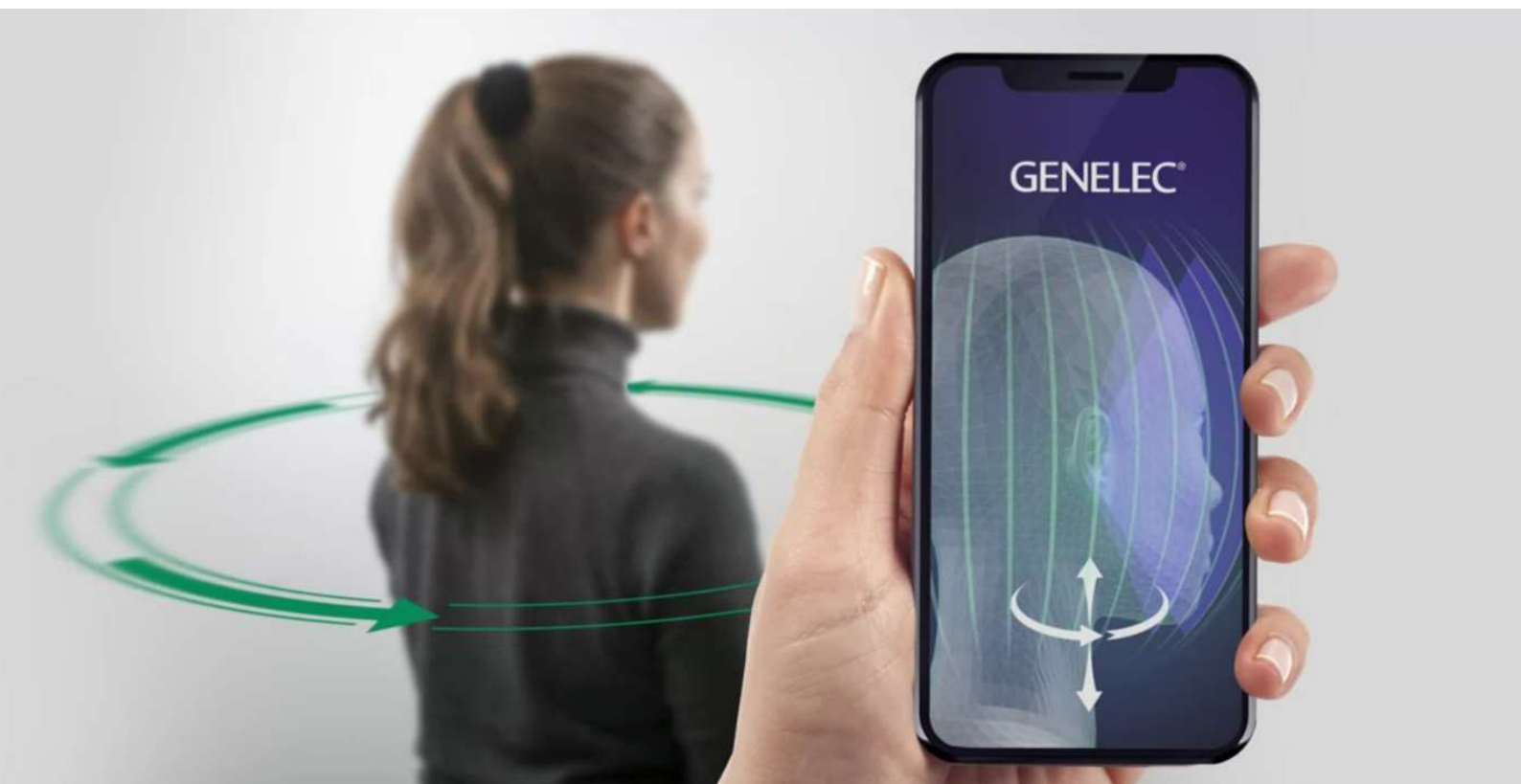# GENELEC®

YOUR AURAL ID. YOUR SOUND.

How to obtain and use the Genelec Aural ID, what benefits it gives, how it is calculated and why it is needed.

Quick Intro

The purpose of audio monitoring is to evaluate audio presentations in a neutral way to ensure good translation to other reproduction systems. The head and outer ear shapes with head movements, the main localization mechanisms of our auditory system, provide our wonderful ability to localize sound sources and enable loudspeaker monitoring to work. Headphones break the link to these natural mechanisms we have acquired over our lifetime. Because of these reasons on-ear and in-ear headphones have not been the best choice for monitoring.

Normal headphones make it difficult to set levels, pan sound locations and equalise important sources, like the human voice or tonal instruments, because headphones do not have well-controlled frequency responses in the midrange and headphone-to-headphone sound character variation is large. This matter is complicated by individual differences between persons. What you hear on headphones can be quite different from what the other persons hear, even with the same set of headphones.

These characteristics are entirely different in good in-room monitoring loudspeaker systems. Work done using good loudspeaker monitoring systems translates precisely to other loudspeakers and sounds the same for all listeners, and also works well on headphone reproduction.

Aural ID is a reliable path to enable stereo, surround and immersive audio monitoring using headphones. Aural ID computes how your head, external ear and upper body affect and colour audio arriving from any given direction. This effect is called the Head-Related Transfer Function (HRTF).

Aural ID delivers your unique personal HRTF in the SOFA file format. The SOFA (Spatially Oriented Format for Acoustics) file format has been standardized by the Audio Engineering Society (AES) and is widely accepted and supported by audio software. Aural ID requires a processor or audio workstation for audio rendering.

This document covers
- what is Aural ID?
- how to obtain Aural ID
- how to use Aural ID
- tutorial on Aural ID and why it is needed for accurate monitoring on headphones.

Table of Contents

What is Aural ID?

Aural ID is your personal HRTF

The Head-Related Transfer Function (HRTF) is a filter for each ear separately, with a different value for each direction of arrival on the horizontal plane (azimuth) and vertical plane (elevation), incorporating all spectral and time-domain effects. HRTF describes all changes to the sound produced by the sound approaching an ear from a given orientation. To create complete understanding of the directional effects to sound we need measurements of a large number HRTFs, for all the directions of arrival over a sphere surrounding the listener's head.

Genelec Aural ID is your unique and personal HRTF information. It contains the HRTF information expressed in the time domain. Technically this data is called the Head-Related Impulse Response (HRIR).

What is in the Aural ID file?

The Genelec Aural ID .sofa file contains HRIR data for both ears given in 836 different directions of arrival for the audio signal. The HRIR impulse responses are 1024 samples long and use the sampling rate 48, 44.1 or 96 kHz, depending on your preference. The azimuth angular resolution in the file depends on the elevation and is given in Table 1. The highest azimuth resolution is at small elevations where the human hearing system also has the highest angular resolution.

Aural ID has the SOFA file format

Genelec Aural ID file has the standard SOFA file format. Spatially Oriented Format for Acoustics (SOFA) is an internationally standardized file format for storing multiple HRTFs. The files of this format have the extension '.sofa'.

Table 1. The angular resolutions at different azimuths in the Genelec Aural ID file.

Elevation	Azimuth resolution	Explanation
±90°	360°	direct up and direct down
±80°	30°	
±70°	15°	
±60°	10°	
±50°	8°	
±40°	6°	
±30°	6°	
±20°	5°	
±10°	5°	
0°	5°	on the ear level, in front of you

Why Aural ID is useful

The key to enabling the process of rendering audio over headphones is access to accurate HRTF information. Scientists have found that although we all share in the same broad principles of how the HRTF looks like, we are all individual with slight differences in details. These details, they have found, do not successfully translate from one person to another. Borrowing

someone else's HRTFs for listening will not be successful. You need your own, personal HRTF information.

Aural ID is calculated using numerical modelling of the acoustic field at the head. This method of acquiring your personal unique HRTFs brings significant advantages:

- Instrumentation is simple and low-cost – a standard mobile phone is used.
- Acquiring the necessary data is simple and the acquisition can be taken in any room and it does not require the listener to sit still for a long time.
- Silence is not necessary. This method does not require the use of an anechoic room although it offers anechoic HRTFs.
- Aural ID eliminates the uncertainty related to placing the microphones in the listeners ears.
- The effects of listener movements do not play a role in the quality of the final result.

What Aural ID is not

Aural ID is not a simplification of reality.
There are methods of offering HRTFs that use data not uniquely measured from the person in question. Such data may come from a mannequin or a dummy head. Typically, this would leave a lot to be desired in terms of quality.

There is also data available in databases and originally measured from totally different persons. Methods have been devised for selecting the best match in such databases, but this typically does not result in reliable rendering as there are still significant individual differences and perfect selection methods do not exist yet. Such selection is usually based on measuring a set of dimensions in the person, such as the size of the head and dimensions of the ear, commonly called anthropometrics. Aural ID is 100% based on the actual video and photo footage you provide and 100% of this data is used in creating the 3D model of your head and torso - and finally in calculating the acoustic effects. No anthropometric short-cuts are used. Aural ID is 100% personal data from you.

Efficiency of calculation is frequently increased by applying simplified physics, such as ray-tracing. Aural ID uses a full acoustic solution of the acoustic field around the head and ear, without simplifications, and therefore works accurately at all frequencies, including very low and very high frequencies.

How Aural ID is calculated

Aural ID uses photogrammetry to map the 3D shape of the head, external ears and upper torso, and has only become really feasible through modern computer technology that is now able to take a video, detect and match the features in the video and then generate the 3D shape of the object shown in the video. As you know, this technology is increasingly used in mapping terrain, charting archaeological sites, and creating renderings of apartments, statues and other objects. Using this technology for mapping the personal detailed shapes of a listener opens vastly new possibilities in using a computer to calculate the way in which an audio signal is interacting with a person's head and ears. We can calculate the exact characteristics of the audio

signal reaching the ear and hence bring all the personal features to determine what changes sound experiences when it arrives to your ears.

This method requires a simple video to be taken of the listener. This video includes the head and upper torso (shoulders) of the listener while the footage circles the listener, showing these details in all angles. After this, the Genelec calculation process analyses the video and builds a three-dimensional model of the listener using photogrammetric methods, by looking at the differences that occur between different images in the video. This model is then used in calculating the impact sound experiences when it arrives at the two ears from different orientations. Using this information, finally, the HRTF filters can be obtained in more than 800 different orientations of audio arrival.

Who uses Aural ID?

Genelec's HRTF measurement method is aimed at audio professionals, opening new possibilities to the whole audio industry. It does this in a fully personal level, maintaining complete accuracy and full personal detail, with the robustness and convenience needed to bring the HRTF to all professional applications.

Aural ID is needed by all those who have information about the direction of arrival for sound. Examples of these potential users are Virtual Reality and Augmented Reality systems, game engines calculating the audio presentations dynamically as part of the gaming process, and researchers working with 3D audio.

Ultimately, all stereo, multichannel and immersive audio formats can benefit. The methods of presenting stereo, surround and immersive audio are developing to include the processing needed to plug in the information about the direction of sound arrival. Once in place, the Aural ID personal HRTF information can be used to obtain correct presentation of these audio formats over headphones. This will elevate headphone listening to become more reliable and accurate.

Video Capture for Aural ID

Aural ID is generated using a video of your head and upper body. First, we need you to shoot a video of yourself. After that we need photos of your ears with a ruler or a tape measure next to the ear to ensure we get the dimensions just right.

Please follow the steps and advice below. Before starting video capturing, watch the tutorial video to see the proper filming conditions and correct scanning procedure.

Preparing the room for video capture
- Study the example video before starting, paying attention to the room, lighting, arrangements and the procedure.
- The open area in the room for the scan should be at least 3x3 m.
- Ensure good lighting in the room. Lighting should be bright, monochromatic and diffuse. Fluorescent lights in the ceiling usually work well. Ensure that there are no shadow regions. Ensure you do not have bright windows appearing in the image as they tend to make everything else dark. Avoid large mirrors appearing in the video.
- Mark your intended path of movement on the floor to ensure keeping the shooting distance constant.
- Clean the background in the room as much as possible. Single colour walls work the best.

Video settings
- Use a camera that is capable of producing high quality Full HD video. Use a camera with 1080p video resolution (1920 x 1080 pixels).
- Use portrait mode in video capture. Hold the phone upright.
- Turn on the camera light.

Preparing yourself for video capture
- Study the example video before starting. The link to the demo video is on the Aural ID ordering page.

Before starting view the video here:
https://s3-eu-west-1.amazonaws.com/genelec-webstore-tutorial-assets/VideoCapture_Tutorial.mp4

- Remove eyeglasses and big earrings.
- Wear a T-shirt or thin long-sleeved shirt without large collars. Do not wear a hoody or other thick clothing that significantly changes the appearances of the shapes you have.
- Avoid single-color shirts, especially black. Checker or stripe pattern with good contrast works the best.
- There should be no hair covering the pinna. Push away and collect your hair if possible. Use a tight beanie hat to flatten your hair if possible.
- When the video is being shot, the person being captured must not move. Aim your eyes on the same point and keep your head still during filming.

Shooting the video
- Start at one metre (3 to 4 feet) and in front of the person you are capturing.
- Keeping your distant constant, move around the person while keeping the person's head in the middle of the image.
- At the ear, stop and approach the ear slowly.
- Come close to both ears with the camera. Carefully point the camera to all directions close to the ear to ensure that they get recorded on the video. We want to see all the 'nook and crannies' in the ear as we want to calculate their acoustic effects.
- After this, keep the ear in the middle of the image and move away, back to the same distance (1 metre) where you started your approach. After this, keeping the head in the middle, start again moving around the person.
- Repeat the approach for the other ear.
- Complete the circling until you are back in front of the person. You should finish at the same spot, at exactly the same distance, where you started from.
- The length of the video should be from 3 to 4 minutes. About 60% (about 2 minutes) should be used to shoot the close-up details of the ears.
- Rehearse the video capturing. Walk slowly and keep the camera steady. Overall movement during video scan - and especially close-up scanning part of the ears - should be fluent and pretty slow.
- Watch the Aural ID video shooting tutorials.

Video capturing process
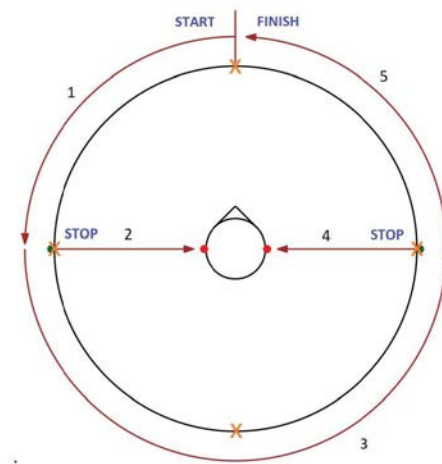
Starting the video shoot in front of a person

Approaching the ear and shooting the details of the ear – note the use of the phone video light

Taking the dimension photos

After capturing the video take photos of each ear with a ruler placed in front of the ear. We show an example of the photos we need. The ruler is transparent and enables us to see the ear dimensions. Point the camera straight at the ear. Do not shoot high up or down. This preserves the dimensions well in the photos.

Photos needed for size scaling

Ordering Aural ID

Aural ID ordering site

Go to the Genelec Community web site at https://www. community.genelec.com/ and go to the Aural ID ordering section. You may have to create a username and password for the Genelec Community. You will need your username later when you access the Aural ID.

Instructions

First, click on the "Create New Profile" button.
You will see an ordering process indicator like this
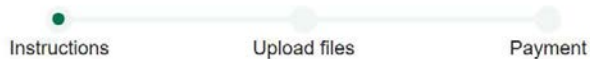
Take time to read the instructions. In particular, pay attention to watching the example videos to understand about the correct way of creating the video and taking dimension photos.

Once on the Aural ID ordering page, you will perform the following steps
  • upload your video and photos
  • pay for the Aural ID (most credit cards are accepted)

Video Upload

On the upload page
  • give your Aural ID profile a name you like
  • click on the three upload boxes to upload your video and dimension photos
  • read the end-user licence agreement and click the check box to indicate you agree with the licence terms

* The video should be at least Full HD (1920×1080) and should last at least 10 seconds.

☐ I understand that Aural ID is for my personal use only. I have read and I accept the Aural ID End-User Licence Agreement.

After you have completed the upload, you will see your video and photos as shown in the example below. Once everything looks good, click the "Next" button to move to the payment stage.

Example of completed upload of the video (left) and dimension photos (middle and right)

Payment

The payment can be made using common credit cards. Please allow the system to process your payment and return to the payment page reporting that the payment has been completed.

Your Aural ID files now move to production. The uploaded video and photos will be checked for quality. If we think there are unsurmountable issues with them, we will send you a message. You can also find these messages on your Aural ID ordering page. In this case, you can chat with Aural ID support to clarify the issues and enable successful Aural ID design.

Aural ID delivery takes typically 2 to 3 working days as we also quality check your video before delivery.

You will receive an email that informs you that the Aural ID can be downloaded on your personal page. Log in with your credentials and download the file.

Your download page will be available to you always, enabling you to return to download your Aural ID at any time later. The video and photos you have uploaded can be removed by Genelec after we are convinced your Aural ID works as intended and there is no need to return to this information any longer.

How to use Aural ID

Compatible plugins and software

In case you want your software in this list below please contact Genelec at auralid-support@genelec.com.

To use Aural ID in a Digital Audio Workstation (DAW) you will need a plugin that supports .sofa files. How the plugin is used in your DAW depends on the plugin you are using, and you should check the instructions that come with your plugin.

In the list below we give some examples of software that supports SOFA files. Note that this is not a comprehensive list as more and more systems start supporting the SOFA file format. Also note that Genelec does not guarantee that all these will work perfectly as these are not Genelec products. At the end of this document we have an example of using the Aural ID.

- SPARTA/COMPASS binaural plugin collection (Aalto University, http://research.spa.aalto.fi/projects/sparta_vsts/)
- Noise Makers AMBI HEAD renderer for DAWs (https://www.noisemakers.fr/ambi-head/)
- SSA plugins (https://www.ssa-plugins.com/demo-plugins/)
- Harpex-X plugin (https://harpex.net/)
- SOFA for MAX (https://github.com/APL-Huddersfield/SOFA-for-Max)

These are also available but have not been thoroughly tested by us:

- VLC Media Player (with SOFAlizer module, by Wolfgang Hraud)
- Steam Audio game engine, works e.g. in Unity, Ffmod Studio, and Unreal Engine 4
- S3A Binaural Synthesis Toolkit (Franck, A., Costantini, G., Pike, C., Fazi, F. M.)
- Panoramix by IRCAM, a post-production workstation for 3D-audio content

Example of using Aural ID for headphone monitoring

This example uses Reaper as an example of a DAW installation. Aural ID can be installed in your DAW using these sample principles. The steps you must take will vary depending on the DAW you are using.

The first step is to create an audio project. In our example, we just create a stereo track that is sent directly to the master output.
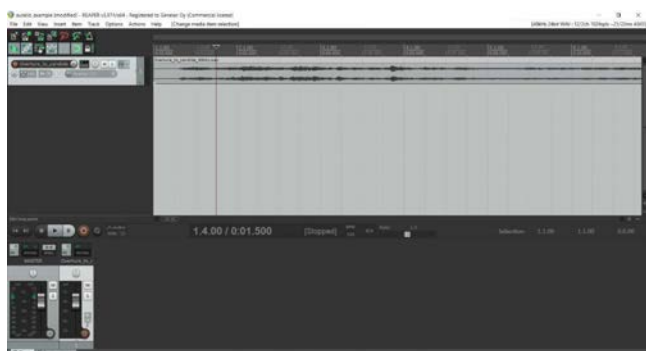
To virtualize this output, an effect is added to the master track. This effect will perform the audio processing needed to apply SOFA information.

In our case, we are using the free SPARTA Binauraliser plugin. The effect selection is shown below.
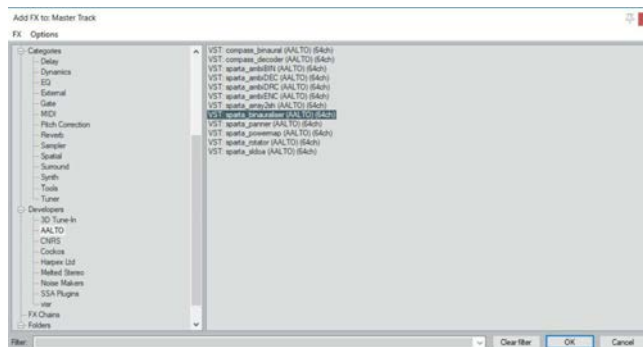
Once the effect has loaded in the DAW, the personal Aural ID SOFA file is loaded in the plugin software (the red arrow points to the load button). The following step shows the selection of the 'example_ID.sofa' Aural ID file.
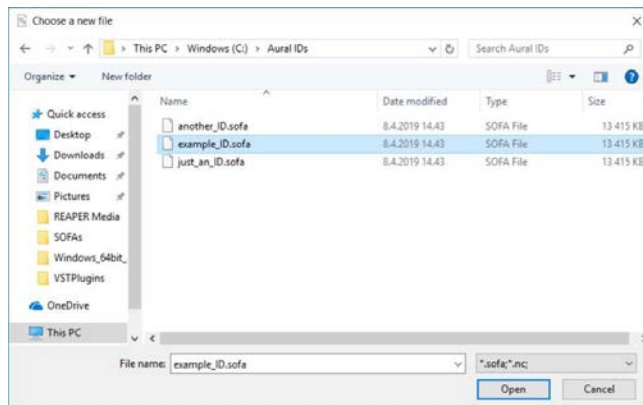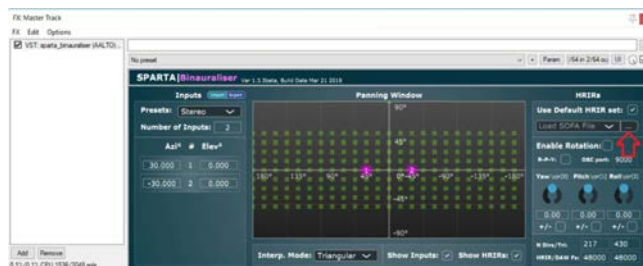
Once the Aural ID file has loaded, the SPARTA plugin is set to create two virtual channels at the standard stereo angles. One channel is set at -30 degrees and the other channel at +30 degrees azimuth and at 0 elevation. These represent the concept of having two loudspeakers at these orientations from the listener. In the same way, you could place any number of such virtual loudspeakers at any orientations form the listener.
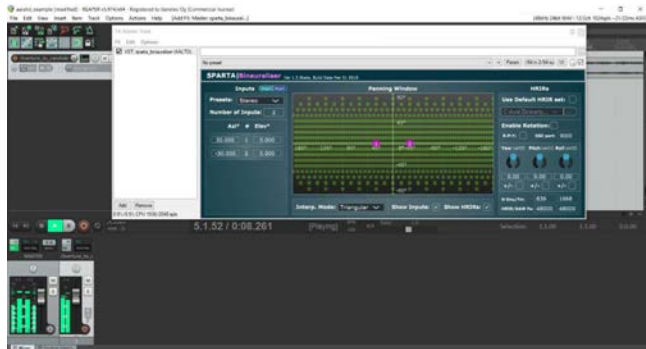
When we finally play back the track, the output we hear will now be the virtualized output. The sound appears outside of the head and in front of the listener. The sensation is more natural and will cause less listening fatigue than normal headphone monitoring.

Learn More about Aural ID

Monitoring of Audio

When audio is recorded, edited and mixed, the result is monitored typically using loudspeaker and increasingly also using headphones. Recorded sound is usually monitored using loudspeakers because audio presentations are normally designed for loudspeaker reproduction. Loudspeakers work together, interacting with the room, to create the experience of acoustic space and auditory images appearing in the presentation.

Frequently mixes use virtual sound images. Virtual sound images are created on the sound stage between two or more loudspeakers. Level panning is the typical method used for placing virtual sources. This happens when the sound presentation is designed in a studio. Stereo or multichannel microphone techniques for on-site recording pick the acoustics of space. This acoustic space presentation can be delivered using loudspeakers.

Experiences of sound stage, virtual sound images, and acoustic space are created by exploiting the mechanisms of human directional hearing. When we are listening to loudspeaker, sound radiated by loudspeakers is reflected by the walls and surfaces in the room and arrives to the listener with slight delay and from different directions. This reflected sound is called 'early reflections'. Early reflections contribute to experiencing a sensation of space and they support understanding the distance to the loudspeakers.

Two ears hear the same sounds but of course slightly differently. The brain automatically combines these inputs, and this enhances the sense of direction for sound. We also spontaneously slightly turn our head while we try and locate sound. This enables us to understand how sound changes when the head turns, and this partly unconscious process greatly enhances our ability to localize sound.

Mechanisms for Localization of Sound

We localize sound using

- Interaural Time Difference (ITD)
- Interaural Level Difference (ILD)
- HRTF – sound colour changes depending on the direction of arrival
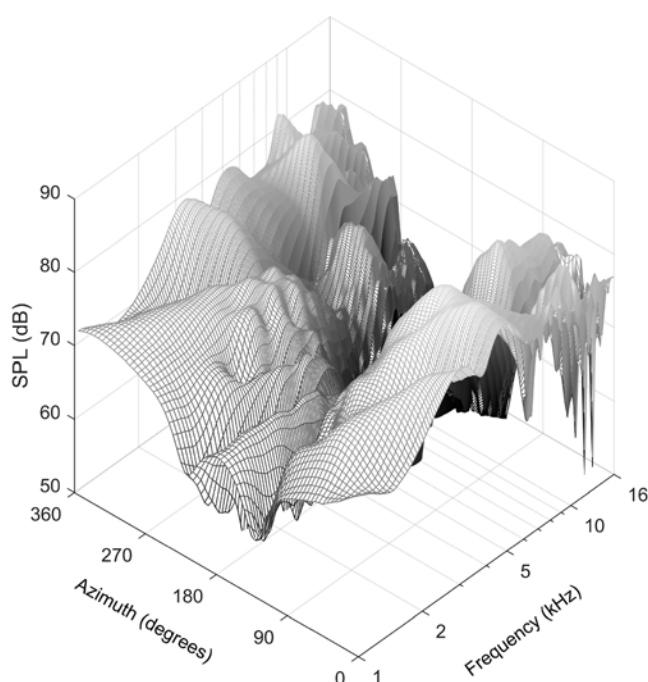- head movements

The localization of sound for humans is traditionally explained using the concepts of the Interaural Time Difference (ITD) and Interaural Level Difference (ILD). ILD is suggested to be the main localization mechanism above the frequency where the human head becomes a significant shadow for sound (above 1500 Hz) and the ITD below those frequencies. However, both of these are part of a more elaborate description of the directional hearing mechanism, that of the Head-Related Transfer Function (HRTF).
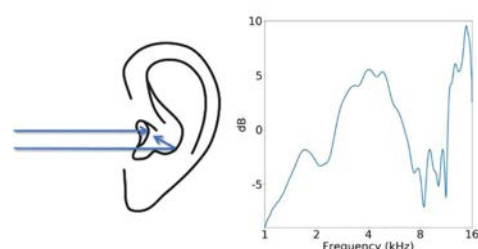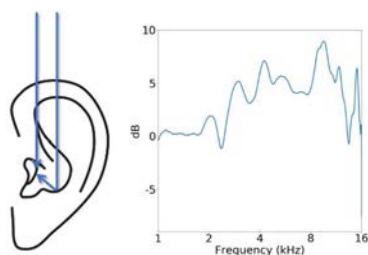
Example of the direction-dependent filtering of head-related transfer function in the external ear, sound arriving from left and sound arriving from above – our brain automatically decodes the colour given by the external ear as the direction of arrival for audio.

Headphone Monitoring

Headphones are typically used to monitor smaller details in a mix because headphones can offer excellent signal-to-noise ratio and isolation for in-room background noises, so low-level audio becomes more clearly audible. Headphones are also used for checking that a loudspeaker mix also sounds good through headphones, as a lot of music these days is listened through headphones.

When headphones are placed over the ears (or inside the ears), the sound is directly delivered into your ears. Then, head-related sound colour changes cannot occur. Also, the earphones that are on your head remain at the same location always - and it is no longer possible to use head movements to help in sound source location. These are the reasons why conventional headphone sound seems to happen inside an earphone listener's head.

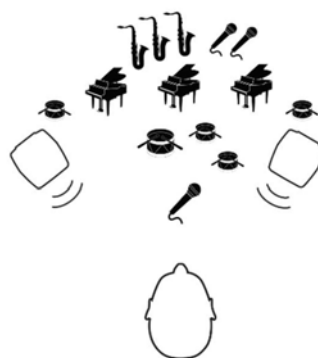Loudspeakers offer a sound stage with virtual sound sources

Examples of the direction-dependent level variations for the Left ear located at about 80 degrees, seen on the horizontal plane going around the head; straight forward direction is at zero degrees, straight back is at 180 degrees, and the rotation with increasing angle value occurs rotating to the left.

Virtual sound sources appear inside the listener's head with standard headphone listening

To give you some examples of what the head-related transfer function is doing to sound, let's look at the direction-dependent level variations. In the figure below, you will see the frequency-dependent level at the ear when the audio is arriving to you from different directions on the horizontal plane and going straight up and down. Specific filtering exists for all orientations on the sphere surrounding you. We only show a small subset here, for two planes, for the sake of giving an example. In real life, you will not hear changes in sound colour but instead you understand that there is a certain direction of sound.

Aural ID can recreate the correct sense of imaging for headphone listening

Processing audio to recreate direction and space

Once the information to localize sound, described by the HRTF, is reintroduced by signal processing, sound recovers the correct direction of arrival and no longer seems to be happening inside your head. In order to strengthen the sensation of sound source distance, some room simulation is usually also needed. All signal processing typically happens inside an audio workstation and there are software applications already available enabling you to do this. These typically come in form of either stand-alone software or as plugin software. These work by adding functionality to your audio workstation.

Using HRTF based signal processing it is possible to place one or more sound sources in your desired virtual locations. This means that HRTF processing can be applied to all audio presentation systems, from monophonic to high channel count immersive.

Processing a complete virtual immersive audio presentation means doing this to all the components of audio forming the full presentation. These components could be sound objects or sound channels, depending on the method of immersive sound presentation.

You enable the processing by importing/opening the HRTF contained in your Aural ID file in your audio processing system or plugin. Before using Aural ID you may need to install plug-ins that will enable you to do this processing. Setting the virtual loudspeaker orientations and creating early reflections to model a room requires some setting in your audio processing plug-in.
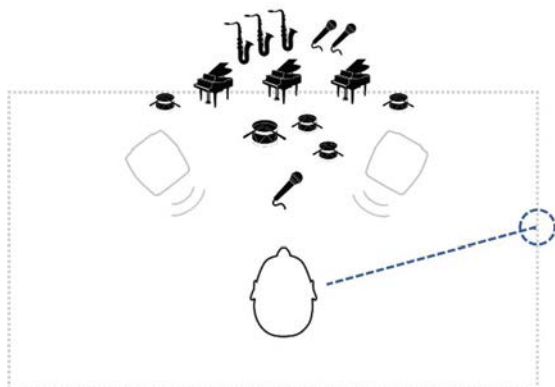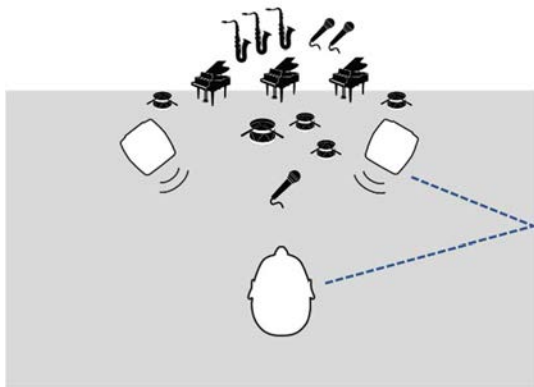
Aural ID can also create an impression of a room in which the audio is being presented by delivering early reflections associated with room acoustics in a real room (left) with virtual early reflection sources having the same directions, level and timing in headphone listening (right)

Use Case Q and A

Can the Aural ID file be moved between systems, if so how?
Aural ID is a standard format SOFA file. It can be moved to and used in any system, but it is important to understand that the content of an Aural ID file is specific to the person whose data it is.

Does Aural ID work with any type of headphone?
In principle, yes. Certain headphone frequency responses can skew localisation. It is beneficial to use headphones with neutral sound characteristics. The reason for this is that the headphone frequency response acts as another filtering effect on top of the information contained in Aural ID.

What difference is the user likely to hear when switching Aural ID in/out?
The directional cues become accurate and the audio presentation of space and direction starts making sense, as they become more natural and accurate. You get the sensation that the audio is no longer happening just inside your head and between your ears.

I'm a researcher and want to know what my HRTFs look like and analyse them. How can I read the data from the file?
There are many APIs available at https://www.sofaconventions.org/ for accessing information in standard-compliant SOFA files. Depending on your preference on the programming language pick one and use it.

I'm a developer and want my software to support .sofa files. How can I do that?
See the previous answer. In case you also want to develop your own .sofa file content access method please refer to the SOFA-standard available at the AES web site, http://www.aes.org/publications/standards/search.cfm?docID=99

I want to enjoy music over headphones using my personal Aural ID. How can I do this?
At the moment the simplest way to do this is to route your music through software that can support .sofa compatible plugins. You could also use your DAW to do this following the principles given in the example earlier.