


The Dummies Guide to Binaural Sound

👤 Ivor Taylor 📅 November 25, 2016 📁 Audio, Technology 💬 No Comments Yet

 Like 14 people like this. Be the first of your friends.

Binaural is a recording technique or processing method that can be used to capture or create sound in a human perception realistic way. It does this by mimicking the way the shape of our head affects the sounds that we are listening to. To explain, you have two ears and a unique head and body shape. Your 'shape' uniquely and very importantly differently for each ear alters the sound each ear receives. These alterations depend on where a sound is located in relation to your head – left – right – behind – above – below.



The three affects called localization clues which your brain uses to localize a sound source.

Affect one: Someone is speaking on the left side of your head. Your right ear will hear a lower level of sound than your left ear. Your head is located between the speaker on your left and your right ear, it 'shades' the sound source from the right ear reducing the level.

Affect two: the speaker's sound wave has to travel further to reach your right ear because your head is placed in between and your ears are approximately 12 centimeters apart. This delay means the right ear receives the sound fractionally later the left ear; a miniscule time delay which allows the brain to derive a bearing for the sound source.

Affect three: the shape of your head and ears causes very subtle changes to the frequency spectrum of the sound. These changes are different for each ear and are dependent on where the speaker is located in relation to each ear both in angle and in height.

In summary each of your two ears hears sound differently unless the sound is dead in front or dead behind and then the instant that that sound or your head moves the ears hear subtly different sounds.

A binaural or post processed recording attempts to capture/create all of the three affects discussed above. You can make a binaural recording in two ways. Either by using tiny microphones in each ear or using a head mannequin with exact human like ears instead.

Because binaural recordings are recorded from inside each ear they capture exactly what you hear complete with all three head related affects. Because of this when listened to you get an amazing sense of presence. This sense of immersion is possible because the binaural recording has captured the localisation cues that you hear continuously everyday.

So why haven't all recordings been binaural?

To this point in the article I have been talking about live recordings, but almost all of the sound you hear on your iPod, TV or computer has been 'produced'. This means it is assembled and edited together from individual mono sounds and then positioned in stereo to be left-centre-right. The result in localisation terms is very crude. The old technology had no way of adding correct localisation cues. Computers weren't fast or powerful enough and the underlying maths required for adding localisation cues hadn't progressed far enough. So the technologist of the time did the best they could, which is a very simple left/right panning control. Those technical obstacles are falling away at an ever increasing pace. At GCRS we have started producing some amazing synthesized 'binaural' style recordings. We have been using both old and new technologies blended together. This has allowed to implement a practical workflow which is a first step on the road to incorporating enhanced binaural style localisation cues into the stereo sounds we hear every day.

How accurate is your hearing's localisation?

Awful would be a quick answer. When judged by the standard set by your vision or your smartphone camera your hearing has atrocious resolution of direction. Its resolution is roughly 10 degrees in the horizontal direction and 30 degrees in the vertical and it often just gets plain confused as to whether the source of a sound is located in front or behind.

How does this compare to your Smartphone which takes 2D pictures of what is in front of you? Your hearing's resolution equates to locating 18 sound sources horizontally and 6 sound points vertically. For sound points read pixels which gives a stunning pixel resolution of 108 pixels. You have two ears so we have a two channel 108 pixel per channel 3D audio camera. Who would buy one of those?

But... Your head is always moving as often are the sound sources around you. Just hearing a sound will often induce a slight change of your head's angle - vertical and horizontal. This changes the left/right localization difference cues, which in return allows your brain to fit in missing bits and resolve the 'is it in front or behind'dilemma. So your 2 x 108 pixel hearing, when analyzed by the best known analytical machine in the world - your brain - and then fitted to your visual perception of the world - is pretty awesome.

Where does binaural hearing fit with vision?

Current neurological studies suggest that vision can utilize as much as 60% of the brain's resources which is a good clue that it is the dominant sense. A thought model that works well is that the hearing process disassembles the sound environment into individual sound sources complete with localization information. When a sound's localization clues correlate to an in-vision sound source, the brain's vision and hearing cognition align. But if nothing can be seen that matches then maybe the source is located behind. The person turns to look but as they turn they improve their hearing's localization clues. Once a visual source is identified that fits the sound's localization clues the listener relaxes unless, of course, it was a tiger hiding in the long grass in front of them.

Proof of the power of vision is that no matter how technically perfect and matched to the listener a binaural recording is the listener will rarely if ever get a clear frontal sound image without a clear coherent visual cue. Your brain's survival evolved logic challenges that if you can't see the sound source it may be behind you which makes you look behind - just in case the source of the sound might kill you.

Why is now the time everyone's going on about binaural?

Virtual reality (VR). The fundamental concept of VR is suspending reality for the subject and replacing that with an alternate reality. Your grasp of reality keeps you alive and your brain is very hard to fool. Vision is the overriding cognitive sense and sound interplays and supports vision. A sound out of place in a VR environment tasks the brain to ask us what is going on and the illusion will never be fully immersive when that happens.

So for sound to be utilized to its fullest extent in the VR world, it has to play by the same rules as sound in your everyday day life. And those rules are binaural, and the fall out of that is that the tools to add localization cues to sound have been recognized as essential to creating the ultimate VR experience.

How far will binaural change audio?

In time it will change everything. Unfortunately on our way there, there will be many bad applications of the technology but when executed correctly it can be stunning and not just when using earphones but also when listening with two or more speakers.

At GCRS we experiment a lot with sound and pride ourselves on pushing the boundaries in our field. Raja – our Director of Sound, combined a whole bunch of emerging plugins, Ambisonic – Binaural and then the kitchen sink. In our iconic Studio 5, a great but plain vanilla stereo recording of the King singing 'My Way' was transformed into the orchestra actually in the room not behind the walls, all from a stereo mix whilst Raj was able to move the King around the room at his will. Awesome – a word I rarely use.

Finally a word about the dirty acronym – HRTF

This stands for [Head Related Transfer Function](#) and crudely put it's the scientific equivalent of the three affects that I started with at the beginning of this article. Everyone's head has a different HRTF, so unless a recording is processed or made with your unique HRTF you will not get the exact localization's of the original recording or production. This problem is sited as why binaural sound will never take off in the wide public domain. I fundamentally disagree. For me the irrationality of this argument is demonstrated when you put a hat/helmet/glasses/goggles on or (for men) grow a beard. All the aforementioned change your HRTF but have you ever heard of firemen, soldiers, pilots, snowboarders or skiers complaining they can't localize sounds when they put their headgear and helmets on? Your brain's hearing adapts very quickly to changes in your HRTF and, within reasonable limits, will rapidly update its localization abilities as your HRTF changes.

By Ivor Taylor, Grand Central Recording Studios ([GCRS](#))