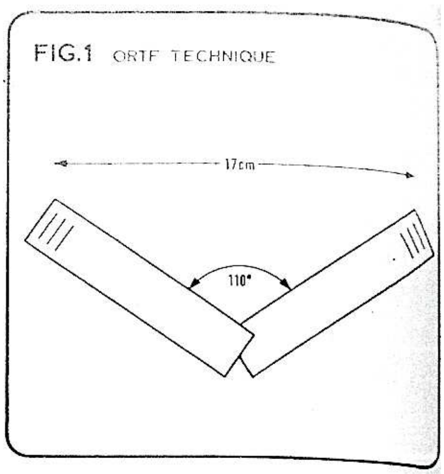
Dummy Head Recording

Michael Gerzon

Recently, there has been a lot of interest in dummy-head or binaural recording for reproduction via headphones. This has been presented by some as 'the answer to quadraphony', and some ill-informed comment has thoroughly confused people as to the advantages and disadvantages of binaural techniques. Issues and available information are summarised, together with an indication of areas of doubt.



FIRST, IT SHOULD BE emphasised that binaural recordings, i.e. recordings made for reproduction via headphones, contain three main types of sound-localisation cue that is absent from conventional stereo, and that there has been some conflict of opinion as to which cue is most important. The first cue is that of time delays between the ears. It is clear that a sound from, say, the left will arrive at the left ear before the right ear. For a sound on the extreme left, this time delay at the right ear is about 0.62 ms, and for a sound arriving from (say) 30° from the left of front (or back, or up, or down), the time delay is about 0.24 ms. Clearly, the time delay cue cannot distinguish front from behind from above from below. All it indicates is the angle of arrival of the sound from the axis of symmetry of the ears. One technique of binaural recording makes use mainly or only of this cue, and that is the ORTF technique of using a pair of microphones spaced apart by about 17 cm. In order to improve compatibility with stereo loudspeaker reproduction, the microphones are directional ones pointing to the left and right respectively, angled about 110° one from the other (see **fig. 1**). For this application, cardioids are to be preferred, as the anti-phase lobes of hypercardioids tend to give exaggerated width requiring a smaller spacing.

It is important that the spacing not generally exceed 17 cm, as otherwise the time delays are too large and everything is concentrated at the extreme left or right. Certainly, ear-spaced binaural recording gives much sharper defined images via headphones than ordinary stereo, even when recorded with a pair of omni-directional microphones. An indication of the importance of time delays between the ears is given if one makes one channel of a binaural recording 10 dB louder than the other. Surprisingly, the image shift thus caused amounts at most to a few degrees, since the time delays are unaltered. Time delays around 0.5 ms are much less important in speaker reproduction, and this leads to the possibility of an amusing paradoxical recording in which a sound appearing on the left via headphones appears on the right via speakers. Simply record a sound on the left of a pair of omni mics spaced 17 cm apart, and turn up the gain of the right channel about 10 dB.

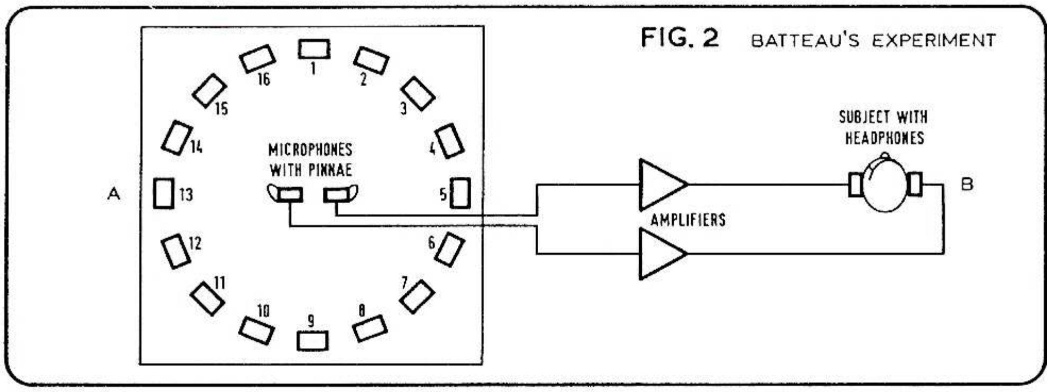
The second cue for headphone reproduction is the fact that the head casts an acoustical shadow across each ear for sounds from the opposite side. This effect is significant only in the treble (above about 500 Hz). Many early workers in binaural sound, such as de Boer and Blumlein, considered that this head obstruction effect was important in sound localisation, and some experimenters still do. However, as will be explained, the evidence seems to suggest that this is by far the least important cue. Indeed, experiments in which a mono sound is fed to both ears but with differing gains show that a relative gain of as much as 15-23 dB is required to create an illusion of a sound coming from 45° from the front, which is much more than the difference in the level at the two ears caused by a live sound from this direction. Moreover, if the ears have just previously been exposed to sounds with natural time delays between the ears, such a panned mono sound can seem to come from only about 15° off front. In other words, not only do the ears make poor use of differences in level, but when they are provided with other cues, they almost entirely disregard these level differences.

The third cue is the effect of the pinnae, which is the name for the flaps on the ears. The various ridges on the pinnae reflect and refract the sound waveform before it enters the ears, and the coloration thus produced varies according to the direction of arrival of the sound. This coloration, which mainly affects frequencies above 5 kHz, is now known to be of vital importance to the ears in localising and positioning sounds, although the way in which this coloration is used by the ear to provide information is still not understood.

An intriguing experiment of Batteau (described in *ref. 1*) demonstrates this in no uncertain fashion. In one room, he set up 16 loudspeakers (see **fig. 2a**) in a circle around a pair of microphones spaced apart by ear distance. The outputs of these were fed to a subject sitting in another room via headphones. The various speakers were then fed with sound and the subject was asked which of the 16 directions the sound appeared to be coming from. This test was performed both using ordinary omni microphones, and with microphones fitted with accurate replicas of human pinnae, but with no dummy head used in either case. When no pinnae were used, the subjects found it difficult to localise the sounds, assigning them to more-or-less random positions. However, with the pinnae fitted to the microphones, localisation was correct with no confusion between front and rear.

Other experiments have also demonstrated that pinnae are of vital importance for correct localisation. Roffler and Butler (*ref. 2*) describe experiments in which a subject's head was fixed, so that he could derive no clues from head movements, and in which a sound source was moved in the plane of symmetry of the subject's head, so that it could be above, below, in front or behind. Since the sounds reaching the two ears is then identical, conventional theories of stereo hearing would suggest that height effect cannot be heard under these conditions. However, Roffler and Butler found that a change of the sound source elevation as little as 5° could be clearly heard. On the other hand, if the subject wore a 'pinna mask' which covered up the pinnae but allowed sound to enter the ears, then no change of elevation could be heard.

So we see that the pinnae play an essential role in locating sounds, and that they should therefore be accurately modelled (preferably by taking moulds from human pinnae) if used with a dummy head. It seems that most of the recent commercial dummy head recordings have used inadequately-accurate pinnae for optimal effect. We also see that the complication of an actual dummy head between the microphones-with-pinnae may be omitted, thereby improving the visual appearance of the microphones as well as reducing some of the



coloration if the sound is played via speakers. Alternatively, a very idealised 'dummy head', such as a simple baffle to separate the microphones (as suggested by Blumlein, *ref 3*), may be used.

Dr Edmund Rolls, of the Department of Psychology, University of Oxford, has recently been conducting experiments in dummy head recording, using small microphones placed in the ears of actual people (although one conjectures that they may resent being termed "dummies"). This microphone technique, so purist that advocates of Blumlein technique must blush with shame, is capable of giving very superior binaural results, as would be expected with such accurate dummy heads. Recognising the importance of the pinnae described above, Dr Rolls has suggested a simple, ingenious and effective method of reproducing dummy head recordings via loudspeakers.

The trick is to reproduce the dummy head recording via stereo loudspeakers placed either to each side of the listener (A in **fig. 3**), or at least angled widely apart (B in **fig. 3**). I have found that angles **θ** (see **fig. 3**) of more than 110° work well. The listener listens wearing a pinna mask. (For listening tests, it is adequate to use the hands to cover the back part of the pinnae.) Since the sound has been past pinnae once during the recording, and since it is prevented from going over them again by the pinna masks, the ears hear just the pinna coloration inherent in the recording, and hence hear a correct directional effect, including sounds form behind or above. I have used this to demonstrate dummy head recordings to an audience of about a dozen via loudspeakers.

Correct localisation is not the only benefit produced by pinna coloration. It is well-known that headphone reproduction always gives the effect of in the head localisation (ihl). This has been

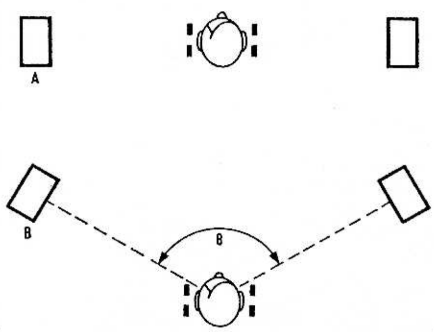


FIG. 3 LOUDSPEAKER LAYOUTS FOR LISTENING TO BINAURAL RECORDINGS WHILE WEARING A PINNA MASK.

explained as being due to the fact that a dummy head cannot move in the original sound field in the same way as the listener's head is moving, and it has been supposed that it is the information produced by such movements that prevents ihl and allows front and back to be distinguished. While head-movement information is undoubtedly of some importance in these regards (see *ref.* 4), the pinnae also are capable both of localising sounds (as we have seen) and of externalising them outside the head, without any help from head movements. Thus the conventional explanation of ihl is wrong (see also *ref.* 5, if you can read German).

A dramatic illustration of the ability to place sounds outside the head is obtained if one takes one channel only of a good binaural recording, and feeds it to *both* earpieces. Despite the fact that both ears are now hearing the same thing, the pinna coloration still allows front and back to be distinguished to some extent. Even more intriguing is what happens if such one-eared recording is made to pick up a sound to the side of the dummy head. Since the information reaching the two ears of the listener is identical, it is impossible for him to place the sound at either side, and it is difficult to say precisely where it is. Yet despite this, the sound is heard as being definitely external and not in the head at all! The experiments of Batteau mentioned above showed that this externalisation occurred if no dummy head was used so long as pinnae were affixed to the microphones.

However, dummy head recording is not without its serious problems, both in its imperfections and in the technical and commercial problems. The worst problem, assuming that an accurate dummy head is used (or at least accurate pinnae), is that the least accurately defined positions tend to be at the front, just where accuracy of location is most required. In the absence both of the visual cue present live or the cues given by the effects of small head movements, frontal sounds given half a chance tend to appear to be either in the head or even slightly behind the listener. This pulling in of the frontal sound stage is disliked by most listeners, who find difficulty in being sure that frontal sounds are front or back, although back sounds are quite unambiguously at the back. If the recording contains strong clues as to when a sound is at the front (eg marked differences in room acoustics, or a commentator telling you where he is), then the ambiguity disappears. This is why on the Sennheiser Dummy Head recording No. 1 (see *ref. 6*) it is better that you listen first to the German side (assuming *now* that you *don't* understand German!) before listening to the English.

This tendency of front sounds to be localised behind is not unique to dummy head; recording anyone with practical experience of quadraphony will have experienced similar difficulties. However, with a well-made quadraphonic or ambisonic recording, suitably reproduced, the ability to move one's head often provides sufficient extra information to lock sounds at the front without ambiguity.

We can obtain some understanding of why front sounds are so unstable, and of why headphone reproduction tends to pull sounds behind the listener, if we study the effects of the pinnae in more detail. If we examine the pinna (**fig. 4**), we see that there are two main ridges from which incident sound is reflected or refracted before entering the ear, marked **1** and **2** in **fig. 4**. The effect of these ridges is for a sound impulse to arrive at the ear followed a few tens of microseconds later by delayed impulses reflected off the ridges. The delays, of course, depend on which direction the sound arrived from in the first place. These delays have been measured by Batteau and others (see *ref. 1*) for various sound directions and the results are shown in **fig. 5.** This shows the delay of the reflected impulse after the arrival of the original impulse both for sounds (**fig. 5a**) in the horizontal plane, and (**fig. 5b)** in the side to side vertical plane. It will be noted that the vertical displacement of sounds causes much larger delays (of the order of 200 μs) than horizontal displacements, which cause delays only of the order of 50 μs.

Because 50 μs is the duration of only half a cycle at a frequency of 10 kHz, the ear gets rather little information about horizontal position from the pinna effect, and so we would expect ambiguities to be worst in this plane. Moreover, sounds from the back involve no delayed impulse reflected from ridge *1* in **fig. 4**. Thus, if a sound is not perceived as having a delayed impulse delayed by around

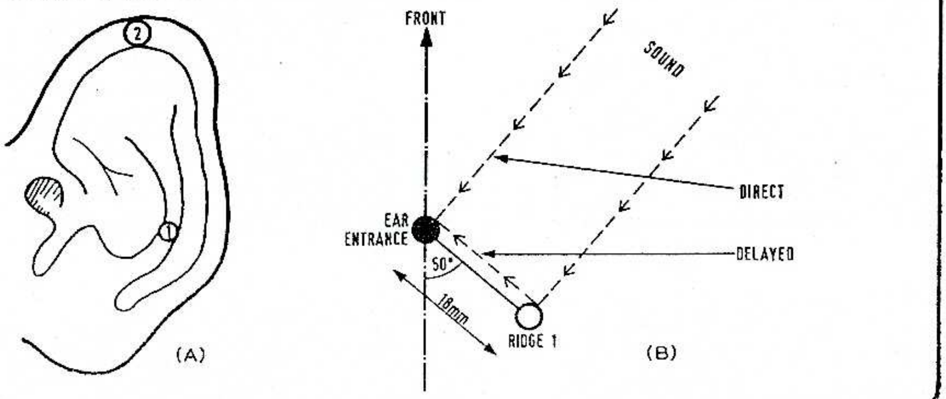


FIG. 4 (A) THE PINNA, SHOWING REFLECTING RIDGES 1 AND 2 THAT RESPECTIVELY DETERMINE HORIZONTAL AND VERTICAL SOUND DIRECTION THROUGH THE DELAY BETWEEN THE DIRECT SOUND AND THAT REFLECTED FROM THE RIDGES SHOWN IN (B)

15-100 μs, then it will be heard as coming from behind.

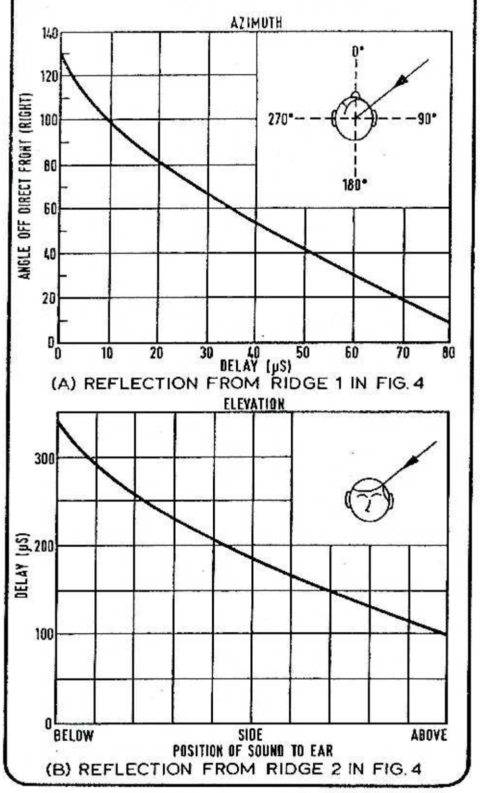
This explains why normal stereo reproduced via headphones tends to seem to be slightly behind the listener in many cases, because such sound lacks any delayed impulses. However, ordinary stereo via headphones is not very convincingly right behind the listener, but rather in his head, which presumably is a result of such a sound not having the second vertical information reflection from ridge 2 (**fig. 5b**) either. One presumes that if suitable delayed impulses according to **fig. 5** were supplied in such cases, then the headphone reproduction would tend to be externalised.

However, since the short delays of the first delayed impulse for horizontal frontal sounds are difficult to disentangle from the complexities of the sound waveform, one expects the ear to miss the presence of the first delayed impulse altogether in many cases. When the delayed impulse is not detected by the ear, one would expect the ear to assume that the sound is behind the listener, since back sounds lack such a first delayed impulse together (see **fig. 5a**). This explains why back sounds are always heard at the back, but front sounds tend to be heard at the back with some degree of uncertainty. The much larger delays involved in vertical discrimination (**fig. 5b**) are much easier to detect and thus give more reliable localisation.

We thus see that the poor localisation of front sounds is inherent in headphone reproduction. For live sounds, the extra clues derived by moving one's head seem to be vital in confirming that a sound is in front.

Reproduced from *Studio Sound*, Vol. 17, pp 42-44 (May 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk

FIG. 5 THE DELAY OF REFLECTED IMPULSES FROM THE PINNA (A) IN THE HORIZONTAL PLANE AND (B) IN THE SIDEWAYS VERTICAL PLANE.



We do not have a complete understanding of how the delays caused by reflection from the pinna are actually pulled out of the sound waveform information by the ear. The processing involved is still something of a mystery, so that the above explanation must still be regarded as incomplete. In effect, we are saying that *if* the ear can use the delay information, then we can explain the behaviour of binaural recordings, but we don't know *how* the ear can use this information.

It is a matter of experience that if one adds the two channels of a good binaural recording together to get mono, then the overall quality of the mono obtained is very poor, certainly poorer than the mono consisting of one ear channel only. As explained earlier, the one–eared mono fed to both ears during reproduction still retains all the pinna reflection information required to externalise sounds correctly. The sum-signal mono, however, combines two separate signals, one from each ear, each with its own time delays. The extra time delays thus introduced not only cause unpleasant signal colorations, but also so confuse the listener that no sense of externalisation is obtained. Thus we see that binaural recordings inherently have very poor mono compatibility, which virtually rules out the use of binaural recording techniques for most public broadcasting applications, unless the majority of mono listeners are to be sacrificed. (Indeed, some say that the ORTF are doing just that with their preferred classical microphone technique.)

When reproduced via loudspeakers, binaural recordings also tend to give a poor stereo effect, which is unstable in the bass and rather unsharp and colored in the treble. This is partly caused by the very frequency-dependent polar diagram of dummy heads in the treble. Since we have seen that the precise form of the dummy

Reproduced from *Studio Sound*, Vol. 17, pp 42-44 (May 1975), by permission of IPC Media Ltd, publishers of *Hi-Fi News* (www.hifinews.co.uk

head is unimportant providing that the pinnae and the intermicrophone spacing is correct, one could presumably choose the form of the microphone baffling very carefully so as to optimise the sharpness and quality of stereo speaker reproduction in the treble. Clearly, the design of a suitable intermicrophone baffle is very complex, and is probably as much an art as conventional loudspeaker design. For this reason, we have to leave to the interested reader the problem of designing a dummy head baffle with good stereo compatibility and retaining good binaural reproduction.

One might consider getting a good stereo image by fitting an ear-spaced pair of *directional* microphones (such as those of **fig. 1**) with replica pinnae, but there is a serious problem with this proposal. For correct effect, all the sound should enter the microphones after first having passed over the surface of the pinna. But directional microphones obtain their directionality by having more than one entrance through which the sound gains access, and they lose this directivity if some of the sound entry points are covered up. Since the pinna only has one point at which sound is allowed to gain access, we can only use replica pinnae effectively in conjunction with omni-directional microphones.

Despite their overall effectiveness, binaural recordings are seen to pose severe problems as regards mono and stereo compatibility. Added to these problems is the poor localisation of frontal sounds binaurally (unless additional clues are given to the listener), and the difficulty of achieving a binaural mixdown of multimic recordings.

References

1. William Mills, A. *Auditory Localization*, which forms chapter 8 (pages 303-348) of *Foundations of Modern Auditory Theory* (ed. J. V. Tobies) Volume II, Academic Press, New York, 1972. Mills' article is the best survey that I know of concerning how stereo hearing works.
2. Roffler, S.K. & Butler, R.A. *Factors that Influence the Localization of Sound in the Vertical Plane*, Journal of the Acoustical Society of America, Volume 43, pp. 255-9 (1968).
3. Blumlein, A.D. UK Patent 394, 325 (1931).
4. Bernfeld, B. Psychoacoustics of sound Localization – A New Approach, Audio Engineering Society Preprint, 47th Convention, Copenhagen, March 1974.

5. Plenge, G. Uber das Problem der Im-Kopf-Lokalisation (in German), Acustica, Volume 26, pp. 241-252 (May 1972).
6. Sennheiser, Kunstkopf-Stereofonie, 45 rpm record.

Further references

The above Sennheiser record and the Sennheiser Dummy Head Stereo Record 2 are both available from Hayden Laboratories Ltd, Hayden House, 17 Chesham Road, Amersham, Bucks HP6 5AG. An article on dummy head recording has appeared in the September 1974 'Wireless World'.

Two more recent articles of interest on dummy head recording have been published in French and German:

Wilkens, H. *Kopfbezugliche Stereophonie–Ein Hilgsmittel fur Verglach und Beurteilung verschiedener Raumeindrucke,* 'Acustica' Volume 26, pp.213-221 (April 1972).

Céoen, Carl. *La Perception Periphonique–Casque ou Haut-parleurs?,* Conférences des Journées D'Etudes, Paris Festival International du Son Haute Fidelité Stéréophonie Editions Radio, Paris, 1974.