

Naturalness and artificiality in recording

Prof P B Fellgett

An academic, writing about industry, begins at a disadvantage. How, it may be asked, can anyone have valid opinions unless he has earned his living in the industry for many years? In the present instance there are fortunately two circumstances which immediately help meet this objection. First, a leading practitioner of the audio industry expounding on just what is, or is not, 'commercial' in records, may sound convincing when these commercial records are justifying their name by selling well, but becomes less credible when there is a recession in sales. Second, I was brought up in the rag trade, and to this day the whiff of hot ironing and a rail of garments being wheeled along the Commercial Road will bring back nostalgic memories. The relevance of this is that, like recording, it is a fashion industry. So I know from early experience that fashion can afford only so much froth of artistic temperament at the interface with the public, and depends on hard detailed technical routine underneath the surface appearances. Moreover a fashion industry may 'decree' what the public shall like, but the public has a habit of thumbing its nose and following the dictates of its own preferences.

There is also a deeper reason why someone from outside the mainstream can sometimes make a

Getting back to basics need not be a retrograde step. A fresh approach to problems helps put them in perspective and can often provide a solution.

suggestion that the industry can take up with profit. Everyday operations compel a practitioner to think mainly about what is necessary for tomorrow or next week, and amid these pressures it is almost impossible to give enough attention to what is best for next year or the next decade. A researcher, however, has the duty and privilege of going back to fundamentals, and so has a proper basis for thinking about long-term optimisation. Nor should this be thought of as 'mere theory'; research requires theory and experiment to be harnessed in a single yoke, and my department now has almost a decade of experience of ambisonic recording using a soundfield microphone, beginning at a very experimental level and progressing to state-of-the-art innovation.

The writer A A Milne, as I recall, was once asked to contribute to an anthology under the title 'My naughtiest story', and offered a tale about a princess and a frog, the whole point of which was that all the characters in it knew the original story of the princess and the frog, and acted accordingly. Probably most of us have known at one time or another a colleague who would have

listened solemnly to A A Milne's version of the story, and then tried (as he thought, having missed the central point) to cap it by telling the original and well-known version! A rather similar misunderstanding seems to be triggered by any article which attempts to question the current fashion in audio. A 'reply' is published which merely reiterates the well-known arguments to which the article in question has already addressed itself. There have been several examples of this happening already, and it seems unnecessary to increase their number.

The past fifteen years or so have seen a rapid trend towards artificiality in recording. Use of a multiplicity of microphones is now common, indeed it may be thought of as the norm, in all kinds of music, while popular music is almost always built up from many separately recorded tracks. This has been in fashion for sufficiently long for a generation of studio personnel to grow up never knowing any other method of working. It is salutary therefore to recall that this is indeed a recent phenomenon which is not typical of the history of the industry. During the first half of this century

the normal method of recording was direct-cut using a single microphone (or what served as a microphone at the time); magnetic tape had yet to be invented.

Let us remember also that history has many examples of enthusiasm for change going into reverse, with a concomitant alteration in the emotional implications of the words used. The Victorians and Edwardians referred to as 'improvements' what we now call 'property development' and regard this as, if not quite a dirty word, at least ambivalent. So when alteration of the raw audio signals is called 'enhancement' let us remember that it would be at least as accurate, but just with the emotional implication reversed, to call it 'distortion'.

It is beyond the scope of the present article to argue the relative merits of naturalness and artificiality in various circumstances. It aims only at raising some questions, and recalling some facts, which may help some people to begin thinking critically about assumptions that they have previously taken for granted.

A sound starting point (yes, both senses) is to enumerate the characteristics we hear when we listen to the natural sound of a musical performance. These include:

- (a) the direction from which each sound comes;

- (b) a sense of the distance of the instrument or other source;
- (c) knowing whether the instrument is being played loudly or softly;
- (d) a sense of the size of the instrument;
- (e) a feeling that the performance is taking place in some recognisable place;
- (f) continuation of the sound after its source finishes due to reverberation;
- (g) a sense of cleanness, and fine discrimination;
- (h) loudness which when measured on a meter is less than we might expect subjectively;
- (i) frequency weighting (we can hardly call it response) which except in extreme cases, our ears accept as correct;
- (j) a sense of excitement from fine performance whether the sound is loud or quiet.

Comparing this list with a typical multitrack recording, shows that the industry has concentrated on items (c) and (f), realised by panpots and echo devices. Items (c) and (d) are falsified by the use of faders; frequency response (i) is often deliberately modified by filtering; items (e) and (j) are hardly represented at all in the recording; and if everything is made as loud as possible, it is an admission of failure with respect to (h). There is enough subjective judgement involved in (g) to admit differences of opinion, but there does seem to be a general consensus that direct sound does in this respect surpass anything heard on a monitor. Surprisingly, reports of this observation (some in *Studio Sound*) have tended to ascribe it to low background noise although this is not of course an inevitable property of natural sound, and indeed is not crucial to the observation (yet the industry devotes much effort to noise-reduction). It is generally accepted that (h) is because our ears expect a distorted sound as louder than it really is physically. Actually (i) is quite surprising in view of the interference effects that can result from multipath transmission in any but an anechoic room, and this is discussed further, later on.

To approach the general question from another point of view, in the early days of commercial stereo when records were released in separate mono and stereo versions, it was sometimes assumed that it would be pointless to have other than the mono version of a recording of a solo instrument. Of course that naive idea did not last long; or did it? For what are panpot input signals but mono?

The early mistake of thinking that mono was all that is required for a solo instrument was made because it was treated as a point-source of sound, without either size or the ability to evoke reflected sounds having directions of their own; and

of course a spot or solo microphone treats it in precisely this mono manner.

Natural sound in fact has two properties which imply all the characteristics which we have listed. First, it is naturally free from any artificial non-linear distortion. Second, it is rich in information about the relationship between the direct sound and the sounds reflected from floor, walls, ceiling and other objects. This information can in principle be made available in reproduction if a stereo-pair 'main balance' is used, but it is absent from panpotted signals and falsified in artificial reverberation.

In the long course of evolution, our ears have acquired exquisitely refined powers of discrimination and interpretation. We use these all the time in everyday life, usually without being consciously aware of what we are doing. In particular, the ear is very good at relating the direct sound of an instrument or voice to the whole pattern of delays, directions and intensities in the indirect reflected sound. One illustration of this is that it is almost hopeless to attempt, using normal quasi-steady-state methods, to measure the frequency response of loudspeakers or mics in an ordinary room, but our ears are not put off by the interfering reflections and can react to an error of as little as 1dB in equalisation. As mentioned earlier, it is only when the reverberation time (more precisely, the room constant) varies badly with frequency that we feel disturbed by room-colouration.

This analytical ability of the ear enables us to derive a lot of information about the origin of the sound. Early reflections, particularly from the floor and any surface immediately behind the performers, seem to be particularly associated with the sense of depth and distance. Later reflections seem to tell us mostly about the general size, shape and character of the room, studio or hall. Together, these impressions add up to a labelling of the sound characteristic of the place and where within that place the sound originated. This ambience labelling is one of the very important clues we use to unravel the complexities of a musical structure, to hear each instrument, voice or section separately, and to distinguish inner lines in the music.

Ambience labelling is found to have a powerful stabilising effect on stereo images. Particularly when the classical Blumlein technique is used (as discussed later on), satisfactory images may be obtained over a wide area in front of the loudspeakers, whereas panpotted images tend to disappear into one loudspeaker or the other with small movements of the listener from the central position. Even when binaural clues are suppressed, either in reproduced sound or live by suitably masking the ears,

ambience clues can enable us to gain an impression of the true direction of the sound.

Multitrack recordings need to make up, in various ways, for their lack of this kind of information present in live sound. With nothing more than sonority and panpotted directionality to rely on, it is much more difficult to hear separate instruments or other components, so that in practice it is necessary to make all of them (or at least the important ones) of nearly equal loudness. This requirement has resulted in the current preoccupation with balance, even to the extent of recordists being called 'balance engineers' which is a poor reflection of the wide knowledge and diverse skills they actually need to deploy. The result often falls down through being over-balanced and not reflecting the musical intentions of the composer.

Other stratagems include 'presence' filtering, and the use of close-miking to pick up a 'dry' sound. It is then necessary to add a controlled amount of artificial reverberation to improve blend, and this practice has led to a misunderstanding. Since ambience clues are absent from artificial reverberation it can contribute only to blend, and it is easily forgotten that natural ambience can contribute both to blend and to discrimination by the mechanism of ambience labelling. This misunderstanding can become self-reinforcing if it leads to natural ambience being treated only as a contribution to blend, as for example when signals are mixed-in from one or more microphones placed towards the back of the hall.

A more truly balanced approach would be to direct due attention to other properties than loudness, for example the correct representation of depth, scale and perspective, in general allowing the eventual listener to hear the same acoustic as that to which the musicians instinctively adapted themselves in the very act of performance.

When the aim is to preserve natural-sound information, the choice of microphone technique is clearly crucial. Here there is something of a national dichotomy; British recordists usually follow the pioneering work of A D Blumlein by using a pseudo-coincident directional pair, while American and some continental European colleagues tend to follow a tradition of spaced omnidirectional mics which seems to go back to the famous Philadelphia-Washington DC relay of 1933. For reasons that are now beginning to be understood theoretically, the British tradition is on the whole the better subjectively, but two qualifications are necessary. First, the classical Blumlein arrangement of figure-of-eights crossed at 90° is notably better-sounding than any of its modifications; in particular,

cardioids at 120° (despite some practical conveniences of this configuration) sound sufficiently inferior to make the phrase 'cardioid colouration' spring to mind. Second, and again there is now some theoretical understanding of the reasons for this, omnidirectional mics seem to be specially well able to give a good sense of depth; and indeed American recordists do seem to prize this quality more than we tend to in Britain. (Interestingly, if the attempt is made to find a stereo format that combines the virtues of Blumlein and omni techniques, the ambisonic UHJ 2-channel specification seems to be quite a good choice, even apart from its surround-decoding capability.)

Of course, even the best mic can only respond to the information that reaches it, and careful attention needs to be given to the acoustic environment, particularly the optimisation of the important early reflections. This is best done by direct listening on site, as has been described by M A Gerzon.

High mic positions may make it easier to obtain approximately equal loudness from all the performers, but they produce a distorted perspective which is very disturbing. This is often noticeable in BBC concert broadcasts where the mic has to be high to avoid audience sight-lines; the ear cannot make sense of the perspective until the listener forces himself to imagine he is looking steeply down on the stage, when everything clicks into place. But in any case, if one climbs up and listens directly from above the performers, the sound is distinctly less pleasant than from a normal listening position.

Meticulous attention to detail is indeed necessary at every stage. No part of the chain is ideal; even the best modern mics, electronics and loudspeakers all have imperfections and distortions which can adversely affect what we hear in recordings. The recording medium itself presents special problems which it is hoped to discuss in the future.

It is conventional business wisdom that in a depression the survivors are those who adopt well-chosen innovations. Progress is not always along a straight line, but may involve spiralling back to reconsider old ideas in new forms appropriate to changed circumstances and new technological resources, and so remembering things that have become overlaid and forgotten in the enthusiasm for current fashions. The present may be an especially appropriate time for the audio industry to take a new critical look at its fundamental technical assumptions.

Further references can be found in *Studio Sound* as follows: 'The Echo of Fashion', P B Felgett, January 1977; 'Why Coincident Microphones?', M A Gerzon, March 1971; 'Stabilising Stereo Images', M A Gerzon, December 1974.