Why Do We Hear What We Hear?

An appreciation of the psychoacoustical character of hearing can increase the understanding a sound engineer brings to his job. Following is a partial transcript of a panel discussion held before the New York Group of the AES this past winter.

The panelists were: Dr. H. Newby, City University of New York. Prof. James Lang, professor of speech and hearing science at Brooklyn College; Jurgen Tonndorf, M.D. Professor of Otolaryngology, College of Physicians and Surgeons, Columbia University; The moderator was Dr. Leo L. Beranek, president, Bolt, Beranek and Newman.

Because of space limitations we have not reproduced the talks of Dr. Newby and Dr. Tonndorf, who spoke on the medical-physiological aspects of hearing. Thus our transcript begins with Dr. Lang's talk and concludes with a condensation of the question and answer period.

DR. LANG: As I looked through my reasonably good library on the topic of hearing, it was only in a fragmentary way that I was able to find some material on a topic about hearing which has interested me for many years and which I have never before attempted to pull together in one very brief discussion. These are aspects of hearing about which we are all very familiar, but for which a great deal of research has not been done. These are the aspects of hearing which are somehow not suggested to us when we consider the ear from the anatomical viewpoint or from the physicist's or engineer's view-

Figure 1. Left to right—Dr. Newby, Dr. Tonndorf, Arthur Gruber, New York Section, AES chairman (standing), Dr. Leo Beranek, Professor James Lang.

point. These are the aspects of hearing which are, in part, a function of learning and experience—and so on. But even that is too restrictive a suggestion. So, may I bring your attention to some aspects of hearing with which you are all familiar, but for which I can not find any name.

For example, let us consider the fact that the ear really, unlike a microphone, is not a constant parameter device. It is a device which changes as a function of experience, which can be *deliberately* changed as a function of training.

May I give you some examples. I think all of you who are in the recording industry are aware, sometimes to your distress and chagrin, that the hearing mechanism of John Doe is one which has become accustomed to distorted sounds: I am referring to the chap who definitely prefers the kind of booming jukebox bass sound. If you expect to sell records to him you have to put bass boost into the recording; otherwise he won't buy it. On the other hand, if you put bass boost into your recordings, then all the hi-fi enthusiasts swear at you. This individual's hearing mechanism, the chap who likes the jukebox sound, does not have something the matter with his ear, it is simply that this is the kind of thing he has become accustomed to listening to. Experience has given his particular hearing mechanism a set of parameters different from those of the individual who is a hi-fi enthusiast.

Let us consider another example, which I think all of us somehow or another know but have not perhaps thought about in a structured way. This is the fact that you have been born and raised with a particular language, the English language, which has conditioned your hearing a particular way.

But consider the individual whose native language is Pakistani. It seems difficult for us to realize that someone whose native language is Pakistani would not be able to differentiate the two words *park* and *bark*. To him these two words sound the same. He would not be able to differentiate the words *gross* and *grows*, nor would he be able to differentiate between *sweet* and *Swede* nor *fine* and *vine*, nor *pluck* and *plug*. The Japanese would not be able to differentiate the words *rake* and *lake*. To them these words are identical. To us they are quite different words. Now, are we then to conclude that these folks have defective hearing? Not at all, it is simply a question of their experiential background.

And really, we could think of the thing the other way around, because when we take the two words *peel* and *pool*, the initial *P* sound of those two words, to us, are identical. But to the individual whose native language is Pakistani those two sounds are to him very different in his language. If we were to attempt to speak Pakistani we would have to learn to hear the difference between the *P* in *peel* and the *P* in *pool*.

By spectrographic analysis we can show that indeed those two sounds are acoustically different but we hear them as the same. We could expand further examples of this, but I think perhaps it is not necessary.

Let us ask ourselves for a moment what kinds of changes can be brought about to an individual's hearing as a function of training and practice. There are two things that can be done. The individual's *sensitivity* can be improved. (There is a little bit of controversy in the literature about this, but I think the preponderance of experimental evidence is that you *can* improve an individual's sensitivity at least a bit, as a function of training.) I use *sensitivity* in the same sense that you talk about the sensitivity of a microphone. There is one thing that you can clearly do as a function of training and that is to improve the individual's discrimination ability.

A second thing that can be done (especially if you are interested in research as opposed to practical application) is to reduce the variability of responses which an individual gives to the stimuli you are presenting to him as a function of training.

One of the problems that we, in the field of speech and hearing encounter, is dealing with the individual who is assigned the duty of helping another person to correct certain speech defects or a foreign accent. We get someone who came to this country whose native language is not English, and who speaks with a very heavy foreign accent. So we assign this individual to a speech therepist to train him to speak English somewhat more intelligibly. After several weeks of work, the therapist swears up one side and down the other that the individual is improving. Actually, however, if a tape recording was made of his voice at the beginning and at the presumed end of therapy, it would show that the individual had not improved very much at all. What actually has happened is that the therapist has learned to discriminate the speech sounds that his student is producing. So the therapist is *convinced* that the client is improving when actually it is the therapist that is improving.

I remember also that some years ago I became interested in infinitely peak-clipped speech. Now back in the late 1940's and early 1950's, it was demonstrated quite conclusively that infinitely peak-clipped speech is remarkably intelligible. It doesn't sound very much like speech at all, it's more like static in an electric transmission. We were playing around in a laboratory with some infinite speech clippers and getting strangers to come in and listen to our infinitely peak-clipped speech. To us the infinitely peak-clipped speech sounded really remarkably intelligible, but strangers didn't even recognize it as speech, let alone understand what was being said. When we pointed out to them that it was in fact speech, though highly distorted, and that if they would just kind of listen and pay attention a bit that they would be able to understand it, lo and behold, with very little practice they were able to understand the speech quite well. Now in part, this was a function of practice and in part it was a function of expectation.

This brings me to my next topic. I've been discussing the effects of experience and training on hearing: now I would like to move to a little different topic, the topic of expectation.

Rather largely we really do hear what we expect to hear. To say it another way, the *a priori* probability of the acoustic events to which you are going to be listening has a rather remarkable effect on an individual's hearing. An associate of Dr. Beranek's has recently been working with a very interesting mathematcal theory, a kind of mathematical model of the human hearing mechanism, referred to as the *theory of signal detectability*. One of the things that has been shown is this. If you present a stimulus to the subject in 50 per cent of the trials and fail to present stimulus in the remaining 50 per cent of the trials, you will get one measure of the individual's sensitivity. And if you change nothing, absolutely nothing about the experimental procedure except the percentage of time that the stimulus is presented, and now present the stimulus only 10 per cent of the time, or 90 per cent of the time, the individual's sensitivity will shift as the function of the probability of the stimulus. Now, somehow or other, this seems very strange to us; we wouldn't predict this. Certainly a microphone wouldn't behave this way. So I hope you're beginning to see the kind of topic I'm attempting to talk about—the kinds of aspects of hearing that usually are

not suggested to us when we think about hearing in a fairly mechanistic way.

Let us take the question of discrimination and its relationship to hearing what you expect to hear. I've only lived in this area a relatively short time and I ride the Flatbush car out to Brooklyn College. They announce as the train is moving—with a good bit of signal to noise—the stations over the p.a. system on the train and when I first started I could not understand one thing that was being said. But you know, the funny thing was that after I had been riding the Flatbush line for a while, I got so I could understand every announcement that was made. I got so I could predict what stop was coming and I knew ahead of time what the stop was, so that what was coming over the p.a. system sounded perfectly intelligible to me which is just another way of saying that the individual who least needs to know what the next stop is, is the only one who can understand what is coming over the p. a. system.

There is another problem, a very personal one. Just recently I visited my physician to get a shot in the arm (I take a series of these shots each year and this was the first one this year) and the nurse said to me (I discovered later) "Would you like us to bill you once a month." I didn't think that was what she had said and I responded. "Oh no, I have to come twice a week for these." I was *expecting* her to say something about the frequency with which I would make my visits to the doctor and instead she was saying something about the frequency of the billing. And so my response was related to my expectancy of what she said, not to the acoustic signal itself.

Well, so much for the topic of expectancy.

Let us consider the matter of auditory memory. We utilize an auditory memory, especially a long-term auditory memory, in a way which has the effect of reducing the information content of whatever acoustic ensemble is presented to us. Let us consider the vowel sound—aaaah—Obviously that is a complex sound. Now the ear really is not quite able to analyse, even if the ear were highly trained,that particular complex signal into its basic components.

However, it has been shown that an individual with training can learn to identify the center frequency of the three or four formants, the resonant peaks, of the various vowels. This can be done only after considerable training because when we listen to the sound—aaaah—we have learned to identify it as an acoustic pattern, as an ensemble. It is only through considerable un-learning that we could begin to utilize the analytic abilities that our auditory system has. Only with un-learning could we analyse that vowel sound into the several formants of which it is composed.

Our memory, however, has not found that to be a useful way of dealing with the particular acoustic phenomena which we call—aaaah—so by memorizing the vowel as an ensemble, we reduce the information content of the message by grouping it. On the matter of short-term acoustic memory as opposed to long term, if you attempt to learn a new language, and you have to produce a new speech sound which is not part of your native language, you must repeat this word immediately after you have heard it. Apparently it takes quite a while for an acoustic memory to become reasonably well established. If you don't repeat this sound immediately, somehow or other that memory trace is not sufficiently with you as yet.

Well, let me summarize what I have attempted to inform you about. The hearing mechanism, unlike a microphone, is not a constant parameter device, its variables are very much a function of experience and it can be modified in ways important to you, by training. Then there is the matter of expectancy. We hear what we expect to hear. Finally I have mentioned ever so briefly some of the peculiarities of auditory memory. This is the fact that we seem to remember things by reducing the information content of the signal so that it can be perhaps more economically stored.

Dr. Beranek: Thank you Dr. Lang.

I would like to comment on one other kind of hearing experience. I attended some hearings recently conducted by Senator Ted Kennedy. This hearing was held in the Post Office Building in Boston and was attended by all the brass of aviation on one side of the room and on the other side all the irate citizens from around Logan Airport that could get into the room. Several things of interest come out in this kind of situation; where people are listening to signals that do not bear useful information, but rather are annoying.

These were people with presumably relatively normal hearing. If you take a thousand of them and question them about their reaction to noise, you will find about 25 per cent of the people in a city will say that noise dosen't bother them—they can sleep near the runways of an airport, or next to an elevator train that is going along, or next to a highway.

There is another roughly 15 per cent of the people (these are of course very rough numbers) who cannot stand noise that is generated by someone else, and it dosen't really matter very much how loud that noise is just as long as it doesn't convey any useful information to them.

This of course complicates the job of someone like Senator Kennedy who is trying to come up with useful legislation, because if 25 per cent of the people don't care what noise environment they are in and 15 per cent cannot be satisfied anyhow—then who are you quieting for.

This is one of the auditory problems which is facing us today because the word's noise pollution problems are becoming more and more common, and the legislators are trying to figure out what kind of laws might be useful and sensible and will protect people's comfort. Now if we were to look to Dr. Tonndorf's charts of what happens in the brain we would see that the signals are carried up to a rather high level so this whole sense of annoyance is somehow integrated in the higher levels in the brain. It's not something that happens peripherally in the ear. The facts are that sounds or noises that are rich in frequency content at the same points that hearing is most sensitive, are the areas in which you get the most annoyance. As a result, on the standard sound-level meter if you use what is called the A scale—which discriminates against the low frequencies and also discriminates against the high frequencies—readings correlated highly with peoples' annoyance with sounds that do not carry any useful information. This is another aspect of the whole hearing problem that we have before us. Now as moderator I would like to ask the panel members if they would like to comment.

Dr. Tonndorf: I would like to emphasize something both of you have said. Take the situation in which a man tells you that he can sleep with trains running by; you very frequently find that when there is a train failure and the trains stop running, he awakens. So this is a negative reinforcement. And the second point I would like to make is in the line of a little joke. I am of German origin and came to this country in 1947. When I first came, of course, I couldn't speak much English. I had a fairly strong accent. I was the only one in my department, so the people there were lucky, they had only one to face. In the eye department there were three

Germans and the Colonel who ran this department was asked by a friend how his German was coming along with these three Germans in his department. Said he, "I am not learning how to speak German, I am learning to speak English with a German accent."

Questions from the audience:

Q: It is reputed that persons of middle age can only hear to 10,000 cycles, yet two years ago I was working with an extended range tweeter up to 26,000 cycles. We all heard it, painfully, and we distinguished between that and 20,000 and 15,000.

Dr. Newby: We are learning more all the time about presbycusis. This is the term given to the effect of aging on the hearing. Curves on the average hearing levels and populations will show an increasing hearing loss to the higher frequencies as you increase each decade of life. This effect statistically can be shown to exist as early as the third decade of life, although I am sure a thirty-year old doesn't consider that he is aging any.

But recently, with some information that has been acquired from civilizations far removed from modern-day noise, certain tribes over in Africa, it has been shown that there is very little effect of aging *per se* on the hearing since men of sixty and seventy years of age have as acute hearing as people in their twenties in this country; so that now we are beginning to wonder if that which we have called the curves of presbycusis are not actually what might more properly be called socio-cusis. These are the effects on the hearing of all of the insults that we incur, as a result of illness, infection, climate, noise, etc.

Q: Could I have the panel address itself to the *young* sound or the *hard rock* sound which is with us today? I happen to work for a recording studio. We had a group of five musicians in the studio the other day that were so loud that I thought I'd bring out a sound survey meter. I cranked it up to the 110 scale and found that we were reading transient peaks of 118 dB or better. This is not the same loudness that's incurred in some of the discotheques and certainly not the levels incurred by the players themselves who are located some inches away from the instruments.

Dr. Lang: I must confess that I don't know too much about this topic except to say that I read a brief statement somewhere just the other day that sound levels in a discotheque had been recorded with some peaks somewhere around 124 dB. I think Dr. Beranek is the expert here on damaged-risk criteria and unless I am very much mistaken there is the distinct possibility of damage to the hearing mechanism at these levels.

Dr. Beranek: And that is what is happening. I know some young neighborhood boys whose hearing I tested and they are not doing so well. I think we will have a high incidence of deafness among these young 'boomphiles'.

Dr. Newby: We had an example in our clinic of a student who makes a living outside of class by playing in one of these bands who was afraid he was losing some hearing. So he came in for some tests and in truth he was down, he was down even in the speech range. We encouraged him to think of some other way of earning his way through college.

Q: In a recording studio I have measured sound-pressure levels of 110 dB. The explanation that is given is that sound engineers can hear the nuances and innuendos better when it is so loud. What is the response of the ear and its linearity to such greatly increased levels.

Dr. Tonndorf: The ear distorts. The funny thing is you are not aware of it.

Q: Is there anything gained by it?

Dr. Tonndorf: Actually no. Any of you who have flown as navigators or people who have used ground-to-air intercom in aircraft will probably bear me out on this, when I say a novice will set his headset much too high and after about two or three weeks he will come down in level because he has learned the language. I remember, many years ago when I was working at Randolph Field, we were trying to do a salvage job for the Air Force. They began to realize that as the pilots got older they should not be thrown away, that these people had valuable experience. True they couldn't see so well anymore, some of them started wearing glasses and some of them couldn't hear so well. I remember one fighter pilot quite well. He was deaf as a doorpost. We took him up in the air and the only test we could do quickly was to go up and start ground-to-air communication and see how he made out. Well, I can tell you, he did a lot better than I did, because he knew what he was *expecting*.

Speech, and especially a limited vocabulary, is so highly redundant, as Dr. Lang pointed out, that the man who knows what he is expecting, (there are a limited number of choices in a given situation, say an approach to a landing-strip) gets them. *I* didn't get them, I didn't get a single word. I didn't know what the choices were.

Dr. Lang: May I comment on this also. When the audiologist draws a configuration of the sensitivity of human hearing, the frequency-response curve appears to be a straight line. But this is a distortion of the way the ear really works. It is a distortion which is deliberately done for clinical purposes by the instrument called an audiometer. What I would like to point out to you, however, is that the sensitivity curve, and the frequency-response curve of the human hearing is not a straight line but a kind of a U shaped curve. I think probably most everyone here is familiar with the Fletcher-Munson curve. The lowest line, the zero phon line, on this graph is the frequency-response curve and sensitivity of the ear. Now let us consider playing a message—be it music or speech at a relatively low sound-pressure level. Now some of the low-frequency and some of the high-frequency sounds as you come down in sound pressure level are going to fall below the threshold of the individual simply because his sensitivity curve is non-linear. As you increase the sound-pressure level of the signal he will begin hearing more and more and more of the low- and high-frequency sound since they are now rising above his threshold. So it seems reasonable to me that the man who is listening to a monitor in a studio, and is listening for what he calls the nuances, would want to turn it up fairly high. This would bring the low-frequency and high-frequency sounds sufficiently above his threshold so that he would be able to make meaningful discriminations with them. Now, those of you who work with high-fidelity equipment correct for this non-linearity of sensitivity by making loudness controls in hi-fi sets. Then, when you want to play concert music at a dinner-listening level, you add bass and treble boost to the system so that the frequencies you would normally hear in an auditorium where sound levels are up around 80 or 90 dB come back up into the region into which the individual can hear them.

Dr. Tonndorf: I doubt that you need 110 dB for this, 70 or 80 is enough. This level is used in testing; we call it the comfortable-listening level and most people are pretty consistent in selecting this, especially after they have had a little bit of training.