THE NATURE AND ORIGIN OF OTO-ACOUSTIC EMISSIONS

Graeme K. Yates,
The Auditory Laboratory, Department of Physiology,
The University of Western Australia, Nedlands 6907, WA

ABSTRACT: Transient-evoked otoacoustic emissions have been used for some years now as a screening and diagnostic tool in detecting hearing loss of cochlear origin but still little is known about how these emissions are generated and what information is really carried in them. In short, the basic physiology simply has not been done. Recently, Robert Withnell and I have been investigating emissions from the scientific rather than clinical viewpoint and have shown that, in the guinea pig at least, they are not what has previously been assumed. They are in fact a form of nonlinear distortion and this has some significance for the interpretation of transient otoacoustic emissions.

1. INTRODUCTION

You can probably image the scepticism which greeted the announcement by David Kemp, in 1978, that he had recorded echoes apparently coming from within the inner ear. He had inserted tightly-fitting probes, containing a hearing aid receiver and microphone, into the external ear canals of human volunteers, generated a short click with the receiver and then recorded the sound in the ear canal in the time following. To the great surprise of almost everyone except himself he recorded, in the period after the initial transient had decayed, 'echoes' or re-emissions of sound extending out to as much as fifty milliseconds. A succession of scientific papers in the following two years eventually convinced almost everyone that these echoes were genuinely from the cochlea and that they were evidence for a mechanical amplifier. It is now widely accepted that there is such an amplifier and that it acts within the ear to enhance the vibrations of the basilar membrane, the structure within the cochlea that carries and stimulates the sensory cells of the ear.

In the years since then these and related sounds from within the ear, collectively known as oto-acoustic emissions, have been applied clinically with varying degrees of success. The idea is that if the emissions genuinely reflect the status of the cochlear amplifier then they should also reflect any hearing loss caused by damage to the cochlear amplifier, the most common cause of acquired hearing loss. Today they form an essential part of the audiologist's toolbox, providing a useful adjunct to standard audiology both for screening and diagnostic purposes. Unfortunately, the headlong rush to embrace oto-acoustic emissions, by ambitious political and commercial forces together with well-meaning health-workers, has driven as the 'new technology' of audiology ahead of the basic science. Today it is in widespread use and yet its basic mechanisms are still poorly understood.

Several years back I realised that the need for some basic research here was critical: how could we have full confidence in using oto-acoustic emissions to screen all new-borns (as is now mandatory in some states of the USA), to assess workers for compensation damages, to distinguish between simple hearing loss and acoustic nerve tumours and to support expensive epidemiological studies when we still do not understand even the basics of how they are generated. True, several clinical studies have shown their empirical usefulness, usually in simple pass-fail screening programs such as in pre-term neonatal clinics or population studies, but we can have little confidence in the more subtle interpretations of the various forms of emissions applied clinically. How do we interpret spectral changes in the click-evoked emission, for example? Can we simply look at such an emission and confidently infer the precise location of hearing loss in a patient? And can we accept some of the claims for a 'predictive' ability for oto-acoustic emissions or is there an alternative explanation? Funding from the Australian National Health and Medical Research Council has made it possible for me to make a start on this basic research.

Robert Withnell, a Ph.D. student in this laboratory, and I started with the click-evoked emission first. Almost all research labs world-wide use the commercially-available system widely available and endorsed by the USA National Institutes of Health for use in screening programmes, but I felt that it was too inflexible for basic research. So we put together our own system after searching widely for the best sound generators and microphones we could find for our purpose, and we wrote our own software so that we could vary our experiments as we saw the need. The rest of this paper discusses some of our recent findings and their possible implications.

2. THE CLICK-EVOKED EMISSION

Current wisdom has it that a click stimulus sets the entire length of the basilar membrane vibrating and that the mechanical amplifier is therefore stimulated along the entire length of the cochlea. The emission then results, it is held, from reflection of a small part of the stimulus energy from irregularities along the cochlea; that the vibrations are not perfectly balanced along the basilar membrane and some of the original sound energy, or of the new energy from the amplifier, is sent back towards the middle ear to be recorded in the ear canal as a delayed echo of the original. As such, the spectrum of the emission should contain energy corresponding only to regions of the ear which are working competently and any spectral deficits should reveal problems

with hearing. The problem is that there have been some serious holes in this argument for some time: for example, the work of Paul Avan in France showed that high-frequency hearing loss had an effect on the low-frequency region of the emission spectrum, an entirely unexpected result.

We started work with experimental animals, using guinea pigs to study how the click-evoked emission really was generated. The first work was to use masking tones to inhibit locally small regions of the cochlea. We reasoned that the conventional explanations implied that the tones should function as a local hearing loss (this is certainly what happens in recordings from individual nerve fibres in the same animal) and that emissions should be inhibited in the small range of frequencies either side. In fact we found no such inhibition but instead found a complex pattern of interactions across the emission spectrum, sometimes increasing, sometimes decreasing the emissions (Withnell and Yates, 1998). We were forced to the conclusion that energy at any specific frequency in the click-evoked emission could come from almost any part, and probably from all parts, of the cochlea.

How could this come about? We know that the cochlea is a highly non-linear mechanical system and if we present two tones to the ear simultaneously, a third tone may be heard quite clearly, slightly out-of-key and at a frequency lower than the original two. This new tone may also be detected in the sound field of the external ear canal. It has a simple frequency relationship with the original two and is produced by nonlinear distortion generating the new tone as an intermodulation product of the original two tones. Its frequency is equal to the frequency of the lower tone minus the difference between the lower and the higher tone, or $2f_1 - f_2$. It is another form of oto-acoustic emission and is known as the cubic distortion tone (CDT). It is not the only intermodulation product, however, and a range of other new frequencies are detectable, at frequencies of $mf_1 - nf_2$, where m and n are integers.

Now, since a click is a wide-band stimulus, consisting of all frequencies across the bandwidth of the loudspeaker, it presents many opportunities for intermodulation distortion. Every spectral component of the click could, potentially, interact with every other component, each interaction producing its own range of intermodulation products. If this were in fact what was producing the click-evoked emissions then it would easily explain our perplexing 'suppression' results: simply suppressing one region of the cochlea would not change emissions particularly at that frequency but would only reduce the contribution of the suppressed region to a wide range of emission frequencies. But how to confirm this? In general, if you want to detect intermodulation distortion in a system, you introduce a signal consisting of two or more frequencies and look for new frequencies not present in the stimulus and generated by the system. Since the click has a continuous spectrum there are no 'holes' between frequencies in which we could look for intermodulation distortion, so we had to make a hole in order that any distortion could be seen separately from the stimulus.

In fact, we chose high-pass filtered clicks, not entirely arbitrarily but based on an understanding of cochlear mechanics. We generated a high-pass filtered click by direct software synthesis rather than passing a wide-band click through a filter, so that we could be sure it contained no low-frequency components. When we played this filtered click to the ears of guinea pigs and recorded the total sound, stimulus and potential distortion components, in the ear canal, we found a wide range of additional frequencies present below the 4 kHz cut-off frequency of the click, and at a surprisingly high relative amplitude, well above the 60 dB or greater stop-band of the stimulus waveform. The distortion components of the spectrum were only 30-40 dB below the stimulus components, indicating a very high degree of distortion within the cochlea. Several tests convinced us the distortion was genuinely coming from within the cochlea: first we could find almost no distortion when we tested the transducers in a plastic cavity, second, the phase characteristics told us that the distortion was generated later than the stimulus, by between 300 ms and 2 ms, and third, when we interrupted the middle ear chain, by breaking the ossicles, the distortion all but vanished. Clearly the click-evoked emission consisted of intermodulation distortion at a level much higher than that generated by our equipment.

When we reported these new results at the Midwinter Meeting of the Association for Research in Otolaryngology, in Florida in February 1998, we expected some serious challenges on our claim, but received none, even from David Kemp himself who was in the audience.

So how does this new interpretation influence the raft of existing results on click-evoked oto-acoustic emissions? In fact it doesn't change a lot of the basic confidence in the technique, especially in its role as a simple screening tool. No understanding of basic physiology can ignore the fact that many large studies have confirmed that click-evoked emissions can indeed detect hearing loss. If the cochlear amplifier is not working well in a given subject, then the basilar membrane vibrations will not be great enough to generate distortion components and so little or no emission will be recorded. It is in the more subtle aspects of their use, however, where the results must be more cautiously interpreted. For example, Paul Avan's studies are now easily understood. Remember, Avan found that high-frequency loss in humans resulted in a decrease, on average, in the amplitude of low-frequency emissions. We now see how this comes about. The standard testing equipment generates a click extending up to around 10 kHz, stimulating well into the basal region of the cochlea, and yet it records emissions only up to 6 kHz in frequency. In the case of a normally hearing person, we expect intermodulation products from all regions of the cochlea, including and regions processing the higher frequencies. If the higher frequency regions, say 6-10 kHz, are damaged, however, they will generate little intermodulation and so we expect the emissions to fall, even at lower frequencies around 1-2 kHz. In other words, the changes in the click-evoked emission do not necessarily imply threshold changes in the corresponding regions of the cochlea: they simply imply losses in some region.

3. ORIGIN OF THE $2f_1 - f_2$ DISTORTION PRODUCT

The other cochlear emission which has become of clinical importance is the simple intermodulation distortion component, variously known as the cubic distortion product (CDT, after the polynomial simplification for it's mathematical analysis), the intermodulation distortion product (IDP), $2f_1 - f_2$ (the formula for calculating its frequency from those of the primaries) and, simply, the distortion product (DP). It arises as one of several spectral lines which are generated by the inner ear when presented with two, pure sine waves. The largest, most easily seen and certainly the most easily heard of the lines is the one at frequency $2f_1 - f_2$. It has been found useful in clinical practice but has the perceived disadvantage that it monitors hearing at only a single site along the cochlea. The basic mode of generation, however, is still very poorly understood.

Perhaps one of the biggest mysteries is why this particular spectral line should be most prominent. Theoretically, its symmetrical counterpart, at $2f_2 - f_1$, should be just as prominent but it is only seen at somewhat higher intensities. Des Kirk and I have been studying electrically-evoked emissions and we believe we know the answer. Electrically-evoked oto-acoustic emissions (EEOAEs) are similar to other emissions but are generated by direct electrical stimulation of the cochlea. Of course, we can do this only on experimental animals at the moment, but it tells us a great deal about the mechanisms by which emissions propagate within the cochlea. We have found that energy generated at any particular place along the cochlea will only propagate back to the middle ear, where it emerges into the external ear canal as emissions, will only propagate if its frequency is below that at which the particular site responds best, its characteristic frequency (CF). This is not a clear-cut rule, the separation is not absolute, but there is a very great asymmetry on the magnitude of propagation above and below CF. The explanation lies, however, in the fluid mechanics of the basilar membrane, which analyses the incoming sound signal into its Fourier components. Although its tuning properties are bandpass, its propagation properties are lowpass, i.e., any given place along the cochlea will propagate a wave so long as its frequency is lower than the local CF, but the magnitude will vary. For frequencies above CF, however, the wave motion is evanescent and decays away exponentially and, since the physics is reversible, no energy will propagate as an emission if its frequency is greater than the CF of the site at which it is generated. When we consider the distortion products, it is clear that the frequency $2f_1 - f_2$ is always below the CF of the primary generation site, i.e., somewhere between the f_1 and f_2 sites, whereas $2f_2 - f_1$ is always above the primary site CF.

4. CONCLUSION

Ours is basic research. Our day-to-day efforts are not immediately directed to solving practical problems of audiology. Rather, we are taking the longer-term view, that if we can understand the basic physics and biology behind the hearing process we will then be better equipped to tackle the other, clinically-relevant problems of hearing.

REFERENCES

Kemp, D.T. (1978). Stimulated acoustic emissions from within the human auditory system. *J. Acoust. Soc. Am.* **64**, 1386-1391.

Kirk, D.L. & Yates, G.K. (1994). Evidence for electrically evoked travelling waves in the guinea pig cochlea. *Hear. Res.* **74**, 38-50.

Withnell, R.H and Yates, G.K. (1998). Enhancement of the transient-evoked otoacoustic emission produced by the addition of a pure tone in the guinea pig. *J. Acoust. Soc. Am.* (in press).

Correction

Sound Proofing of a Forge

by Stephen Cooper

Acoustics Australia, vol 26, no 1, page 22

Figures 1 and 2 in original should be replaced by figures below.

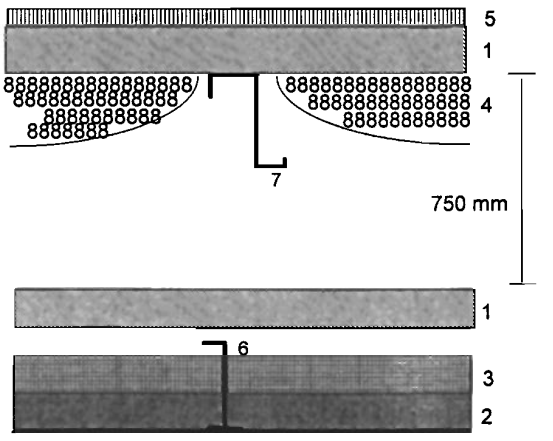

Figure 1. Forge Roof Construction

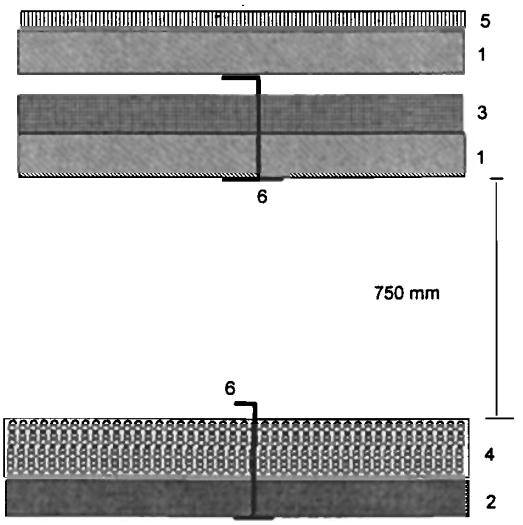

Figure 2. Forge Wall Construction