Monitoring loudness levels

You know the problem: you're watching the television with the sound turned up to a comfortable level when along comes a commercial break and suddenly you're being deafened. RENÉ MOERCH, technical director of DK-Audio explains why it happens and asks why is it still allowed to happen.

OVER RECENT YEARS THERE has been a lot of talk about the differences in the audible level of broadcast sound. This isn't just between various broadcasters; often it extends to programmes on the same channel, indicating that loudness levels are being set by postproduction houses before the programme or commercial even gets to the broadcaster.

The fact that this happens is something we can all agree on because you don't have to be a skilled audio engineer with golden ears to be irritated by the huge differences in audio levels. Bearing this in mind, surely a measure of success in broadcasting is to come up with a loudness level with which most people are comfortable, and to have some way of measuring it so that those producing and broadcasting programmes know what they are aiming for.

This may sound obvious, but it's with these criteria that we have the biggest challenges because the preferred listening level actually depends on a great many, very hard to measure parameters. The age and sex of the listener, the time of day and the type of job they do are just a few of them. Others include the impact that the audio is supposed to be making in relation to the picture – in some cases it needs to be loud to create effect. A good example of this is the opening scene of Saving Private Ryan. The noise of the battle is vital to its impact. If it wasn't loud it just wouldn't work.

When audio is linked to picture — in a film or on television — it becomes the less obvious part of the overall equation because we humans are usually more interested in pictures than we are in sound. Most of us are visually orientated and prefer to assimilate our surroundings through visual impressions. Because we pay less attention to sound, we will often only notice the audio if something is wrong. This explains why we notice when audio levels shift dramatically or when the dialogue is fuzzy, out of sync, or just hard to understand.

While most sound engineers focus only on the audio, for broadcasters it's a different story. They must think about the overall impression they are conveying and what kind of impact they want to have on their customers. This is of prime importance, particularly at a time when viewing figures often dictate a station's revenue. Thanks to cable and satellite, there are now so many channels to choose from that broadcasters can't afford to do anything that disappoints viewers in case they vote with their remotes and switch channels.

Commercials are one area where we have seen major misunderstandings on how best to approach the customer. Advertisers want their commercials to be noticed, therefore they want them to be as loud as

Figure 1

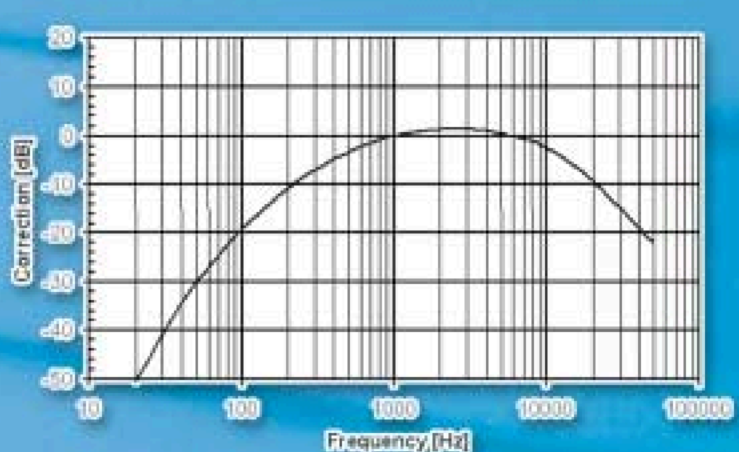

Figure 2

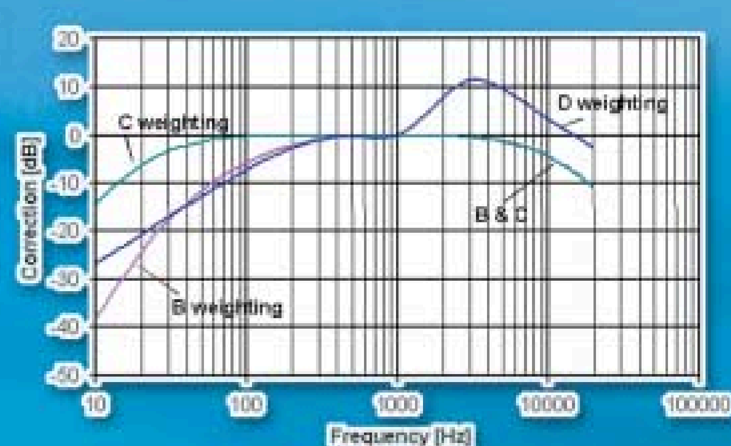

Figure 3

possible to underline their brand messages. The problem with this approach is that everybody at home simply grabs the remote and turns the sound down whenever there is a commercial break. In terms of pure volume, the commercial may not be louder but it seems louder to the viewer because of the high compression levels applied in the mix. If every commercial is compressed to a dynamic range of only a few dB (or even less), the viewer will notice how unnatural the audio sounds and will get annoyed. He's then likely to change channel or, at best, just turn the sound down. Surely this isn't the effect advertisers want to have? Instead of reinforcing their message in a very provocative way, they are actually making their commercials so irritating that they turn customers off.

It isn't just advertisers who are facing confusion over loudness levels. Broadcasters are also culpable because they insist on believing that the higher the output level, the better. This is why so many are locked in battle trying to offer the highest output levels, despite the fact that the audio becomes so noticeable and it is no longer a natural part of the viewing experience. The same trends exist in radio where once again everything is being mixed to within a few dBs and compressed to the point where it sounds too loud.

When you toggle through the channels of your TV set, you'll soon notice that the levels of the various channels are all over the place. Some attempts have been made to solve this. One approach is via automated level control, where a machine tries to balance the level via prepared algorithms. This method uses gain control as well as compression/limiting to balance the signal. However, the output is only as good as the algorithm on which these automated systems are based. There is no human interpretation at all.

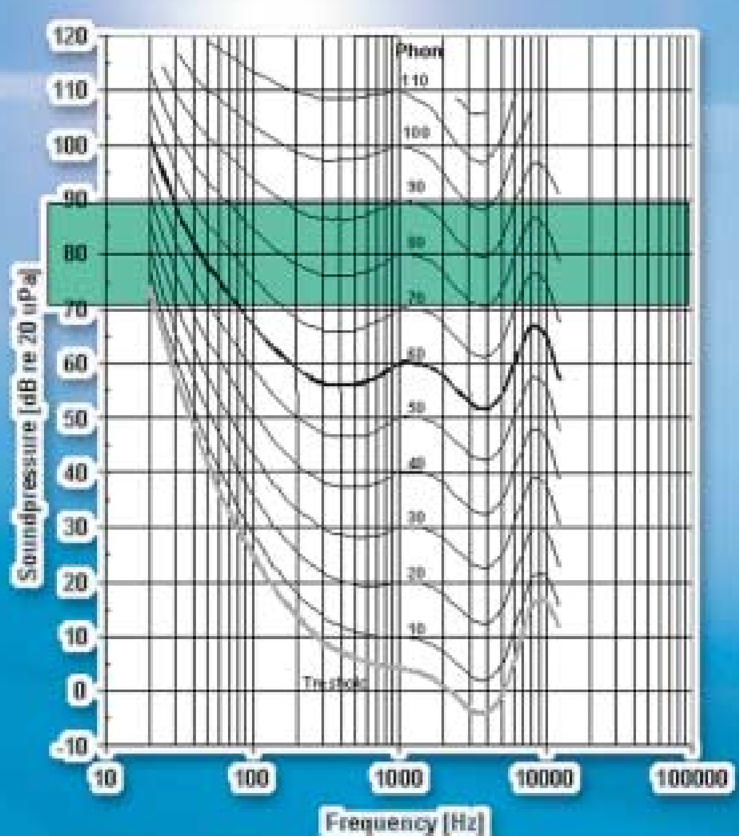

Figure 4

Another approach involves using level meters that are designed to give good correlation between the reading on a scale and what is actually being heard. Attempts are being made to create such meters, but at present most people are relying on standard PPM technology for information about level, which sadly doesn't provide all the information they need. To understand why this is the case, you need to look at the various IEC scales and how they fit in.

Traditional PPM level meters were designed when everyone worked in analogue. They were merely meant as a tool to drive the AM and FM transmitters as far as possible without reaching the point of full modulation. PPM meters don't offer any measure of how loud a listener will perceive a specific piece of programme. All they offer are numbers on a scale, which must be interpreted by an operator to establish a measurement of loudness.

When an engineer uses his ears and combines this information with the reading he's getting from the PPM meters, it is entirely possible to create good levelled programmes that sound great when they are broadcast. An example of this occurs in the BBC where various programme types are aligned to a BBC scale (IEC type IIa) to give an equal perceived loudness level at the end of the broadcast chain. According to the BBC, pure speech should have a higher meter reading than a piece of recorded music. This is due to the fact that the PPM meter will only give an idea of how the media you are driving is saturated and will not give an impression of how loud it sounds. The PPM must therefore be interpreted by the audio engineer and evaluated together with the output of the speakers to judge the loudness. PPM level meters are meant to help engineers navigate, not to lead them by the nose. They can't accurately determine loudness without the help of human ears because of the way in which loudness is measured. There is only one standard way of measuring perceived loudness and this is ISO standard 532-1975.

ISO 532-1975 is built around the work of Zwicker, a physicist who was conducting his research into sound during the late 1960s and 1970s. His basic idea was to use the corresponding value between 1/3-octave bands and printed curves to give a value to perceived loudness. The trouble is that only equipment developed for acoustical measurement has this true loudness scale implemented. Most meter manufacturers, including some who claim to offer a 'loudness' meter, use instead some kind of frequency weighting curve (Figure 1). This measurement is based on theory and praxis from straightforward noise measurement. In the best cases the weighting curves follow the standardised A, B, C or CCIR468 parameters (Figures 2 and 3), but in many cases the curve is actually invented by the manufacture. This gives rise to serious concern as there is no standardisation between manufacturers, so one meter might be very different from another even though both claim to do the same thing. Apart from that, the whole idea of using a frequency weighting filter in front of the meter could well prove to be wrong, according to the latest research.

There is another problem with PPM metering. Depending on where you are based geographically, you have to mix in accordance with certain specific broadcast standards. For instance, in the UK IEC type IIa is the standard while in Germany it's the DIN scale. As the world becomes smaller, there is a need for a standard scale or, at the very least, the ability to mix in any of the various standards. This makes commercial sense for everyone because programming is no longer created for just one territory. Music, films, commercials, and documentaries are now produced for the whole world and therefore have to work regardless of where they are played, screened or broadcast.

At the moment, the best we can do is try to understand the relationship between the various scales so we can avoid big variations in level when the programme is broadcast. However, a better solution would be to implement a standardised loudness scale, one that is agreed in all territories and accepted worldwide.

Right now, work is being carried out on this by the ITU, whose SRG-3 group is trying to establish a common standard. At the last SRG-3 meeting, many highly qualified suggestions based on loudness tests carried out on approximately 100 people were evaluated. Many of the suggestions were based on the frequency weighting theory, but further investigation seemed to indicate that, when looking at the energy content of the signals rather than pure filtering, there was a higher correlation between the meter readout and perceived loudness levels.

When the RMS value of a complex signal was measured, it seemed that the best results were achieved in praxis. The reason for this is probably down to phon curves. When looking at the phon curves (Figure 4) of our preferred listening level at home (here stated to be 70-90dB SPL) we see that the curves for the three lines 70, 80 and 90 are almost identical. This means that within this area we can judge the ears' frequency dependency to be linear. This, of course, is based on the assumption that during the recording session the engineer used his ears to compensate for the natural non-linearity of the human ear! If we look at higher dynamic ranges and compare the 40 phon curve with the 70 phon curve, the linear principle doesn't work. But as the programme is normally mixed within the 15-20dB dynamic range, it is easy to conclude that we will have no need for a frequency weighting curve in front of our measuring device.

It is easy to place a filter in front of a meter and just as easy to take an RMS measurement, but the outcome of the two approaches is very different. This is why it is important for engineers to question loudness meter manufacturers about their test methods and praxis, because the concept on which the meter is built might not prove to be the best in application.

The subject of loudness has been under discussion for a long time and many attempts have been made to design a meter that gives a reading that correlates to our hearing. Many broadcast stations have had their own laboratories trying to make such a meter, but as yet no one has succeeded. It is indeed no easy task.

But at least the work of the ITU SRG-3 group is going a long way towards setting a practical standard – and that's something we all should welcome. ■

DK-AUDIO, DENMARK:
Website: www.dk-audio.com