A Preview of the Broadcast Audio Processing Course

About the Author
Stephen M. Poole, CBRE, AMD, CBNT, is Market Chief Engineer for Crawford Broadcasting's Birmingham cluster of radio stations. A musician and broadcast veteran with three and a half decades of experience, he has also done professional sound and lighting, as well as contract programming and custom electronic design. He and his wife Sandy live near Warrior, Alabama.

Course Description
Those of us who care about broadcasting are dismayed at the smashed, squashed, and distorted audio that is often heard on both radio and television -- and especially on radio. But the answer is to learn how to do it right so that you can take the lead. Show them how it's done!

This course takes a practical, real-world approach. An entire chapter is devoted to human hearing and how it actually responds in the real world. All of the familiar terms are here: attack and release times, compression, peak limiting ... but emphasis is given to how these sound to the listener.

The audio processing that is available now is an order of magnitude better than the standards of just 10 years ago. It continues to improve, too. Many of the assumptions and techniques that were common with older processors are no longer valid. A state-of-the-art digital multi-band processor can execute billions of floating point mathematical operations per second, meaning that very complex algorithms (for example, exhaustive Fourier analysis) can be applied to the audio in real time. These units can be driven (and can drive) much harder before audible distortion sets in and offer a phenomenal range of control over the final sound. You can literally adjust these for anything from no processing to obnoxious on-air loudness.

Who Will Benefit
This course is for the intermediate-to-experienced broadcast engineer who is already familiar with the basics of good broadcast engineering. It will be assumed that you know how to mount equipment with proper grounding and shielding, and then connect the inputs and outputs as needed -- whether AES, S/PDIF, analog, composite, or otherwise. You are proud of your studios and transmitter site(s), and keep everything on air and operating properly, but would like to know more about that "dark art" known as audio processing.

This course is especially targeted to radio engineers, for whom sound is the primary marketable product. The principles discussed here can be applied across the board, but given that most of the innovations in audio processing have come from the radio community, that will be our focus.

Course Content
1. Introduction
2. Loudness and the Human Ear
3. The Basics of Broadcast Audio Processing
4. The Basics of Gain Control
5. Limiting and Clipping
6. Adjusting a Wideband Processor for Voice
7. Multi-Band Processing
8. Other Considerations
9. Summary

SBE Recertification Credit
The completion of a course through SBE University qualifies for 1 credit, identified under Category I of the Recertification Schedule for SBE Certifications.

Enrollment Information

SBE Member Price: $80
Non-Member Price: $115

Loudness and the Human Ear

In this first chapter, let's take a look at how the human ear perceives audio in general and loudness in particular. This is invaluable when making the compromises inherent to audio processing.

The Decibel Revisited

Broadcast engineers are familiar with the decibel and with the formulas used to calculate the values:
Figure 2: Calculating decibels from power, voltage or current ratios

$$\text{Decibels (Power)} = 10\log\left(\frac{P1}{P2}\right) \qquad \text{Decibels (Voltage or Current)} = 20\log\left(\frac{E1}{E2}\right)$$

The difference between a power of 1 watt and 10 watts is 10 dB, as is a ratio of 3.16 volts or amperes. An increase of voltage by a factor of 10 across a constant impedance represents a 20 dB change and 100 times the power.

The original measure was the Bel, but this unit was found to be too large, whence the "deci-bel" (or 10th of a Bel). The decibel, in and of itself, is a dimensionless quantity that simply describes a logarithmic ratio. But references are sometimes specified with a suffix to indicate the standard. For example:

- 0 dBm = 1 milliwatt into 600 ohms (our standard for analog audio levels)
- 0 dBμ= 1 microvolt (Greek letter "μ")
- 0 dBK = 1 kilowatt
- 0 dB SPL = 20 μPa (micropascals) RMS sound pressure

That last one is the standard measure of sound pressure level, or acoustic energy. The 0 dB SPL reference is considered the absolute threshold, at mid-frequencies, of a person with excellent hearing. Normal conversation in a quiet room averages around 60-70 dB SPL; someone operating a jackhammer should wear hearing protection, because he or she can be exposed to 100 dB SPL or greater.

Perceived Differences in Loudness

Not surprisingly, our ears are optimized for conversation: for human speech. But they must also adapt to a phenomenal range of sound pressure levels in real life, from a whisper to an explosion. To accomplish this, our sense of hearing compresses that huge dynamic range logarithmically.

The original Bel wasn't pulled from thin air. it represents a factor of two in perceived loudness.

A typical listener will judge that a sound at 60 dB SPL is roughly twice as loud as one at 50 dB. This relationship holds electrically as well. Assume a quality loudspeaker that produces 70 dB SPL at the listener's ears with 1 watt RMS of applied power. If you gradually increase the RMS drive into that speaker, here is how most people would respond:

- 1-2 dB (1.26-1.6 watts): only a very perceptive listener will notice.
- 3 dB (2 watts): the smallest increase that most people will notice.
- 6 dB (4 watts): most people will call this is a "noticeable" change.
- 10 dB (10 watts): most people will consider this "twice as loud."

This is not accurate over the entire audio range, especially at deep bass and extreme treble. Nor is it accurate at very high SPLs. But these deviations tend to cause the perceived loudness for a given power increase to be less, not more.

For example, at very high sound pressures (a rock concert, a jet plane taking off, or an explosion), the human ear actually requires an even larger increase before it will notice the difference. What this means is that, if your heavy metal band is already at 118 dB SPL and you wish to be twice as loud, unless you're Metallica and have deep pockets, you probably won't be able to afford the needed equipment.

The non-technical person will not believe this at first because it is non-intuitive. However, it can be easily demonstrated. Because the original Bel was designed to measure speech, start with that. (The air monitor for a talk radio station will work.) Feed the audio into a studio mixer and speakers and set the average level to about -10 dB on a standard VU meter. Give the listener time to get acclimated and ask them to close their eyes. Slowly push the fader up until they indicate that it sounds "twice as loud." Do this several times to average out errors in perception.

On average, the meter at "twice as loud" will be 10 dB higher, around 0 dB.

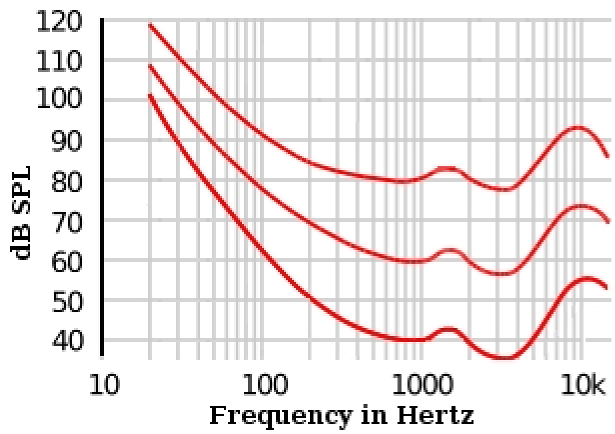
Perceived Loudness vs. Frequency

Audio Frequencies are usually defined as those falling between 20 Hz to 20 KHz. This represents the absolute limits of hearing for a young, healthy male or female. As we age (or if we suffer injury), this range diminishes, especially at higher frequencies. But even a healthy ear doesn't have a linear frequency response. We hear frequencies between 1 and 5 Khz the best.

Suppose you feed a reference tone of 1KHz into a set of calibrated headphones and adjust for 80 dB SPL. A 50 Hz tone would have to be at 100 dB SPL for the average listener to consider it of equal loudness. At the opposite end of the spectrum, a 10 KHz tone would need to be at 92 dB SPL to be called "equal."

Figure 3 is an edited version of the well-known "equal loudness contours;" it only shows three sound pressure curves. But it is obvious at a glance that low and high frequencies require much more acoustic power to be considered equal in loudness.

Figure 3: Equal loudness countours



A Fine Point ...

Someone might object: "if this is the case, why doesn't everyone demand tons of bass and treble from all audio sources, including live music?" The argument could be made that they do (especially the bass). Just listen to any teenager driving by with an aftermarket stereo system in his/her car.

We have been talking about hearing and perception in this chapter. When you listen to a live piano, you don't necessarily expect the bass to be as loud as the midrange. You expect it to sound like a piano. But when playing back recorded music at lower than live levels, you might indeed want to increase the bass and treble to compensate. That was the original purpose of the "loudness" control on consumer audio equipment, though many listeners do tend to use it all the time. (Hey, people like bass.)

... And The "500,000 Rule"

The operative phrase above is, "what your ears expect." Listeners want a balanced sound, and this is where the "500,000 rule" comes in. Audio engineers have known for decades that, for a natural balance with an otherwise-flat frequency response, the sum of the low and high rolloff frequencies in an audio system should roughly equal 500,000. For example, if the bass extends to 30 Hz, the high limit should be at least 16 Khz (30 x 16,000 = 480,000).

This is anything but a hard and fast rule, and a quick Web search will show that even that "500,000" figure isn't universally agreed upon (you'll probably see values from 400,000 to 1,000,000). Also, this is only true with a flat frequency response. For example, you can sometimes "cheat" by cutting the deep bass and then boosting the mid bass a little to compensate.

Finally, common sense applies. You could limit response from 200 Hz to 2.5 KHz and satisfy the 500,000 rule, but only someone who lives inside a telephone would call that a "natural" sound.

Still, this is a useful rule of thumb, if only because it helps demonstrate what you can't or shouldn't do.

Let's contrive an example. Suppose your old audio processor just can't protect your FM transmitter against high treble frequencies. The 75uSec preemphasis curve keeps the peak light flashing on your modulation monitor. You might decide to sharply reduce the treble into that processor. This would indeed allow you to drive the transmitter harder, but your listeners won't judge that you are "louder," because the bass will be unbalanced relative to the treble. You'll just sound "bassy" to them. They expect that treble and will miss it if it isn't there.

Psychoacoustics

We use terms like "perceived" and "apparent" for a reason: our sense of hearing is the result of nerve impulses being sent to the brain, where they are interpreted. The study of how we perceive sound is called psychoacoustics. It has been essential to the development of so-called "lossy" audio codecs such as MP3, where hard choices must be made between absolute fidelity and reduced data size/bandwidth.

It's completely beyond the purview of this course to cover this in any detail, but we can mention several things that are useful to know when adjusting an audio processor.

To start with, our ears integrate (or average) loudness over time, essentially following a Root Mean Square (RMS) law. They will tend to ignore brief, non-repetitive events, even if they're extremely discordant or distorted.

Our hearing will tend to ignore the softer of two sounds. The closer they are in frequency and the greater the difference in level, the more likely that the softer one will be "masked." Most interestingly, because of the averaging effect of our hearing, masking can take place even if the softer sound occurs immediately before or after the louder one. We will assert later that high-end digital audio clipping is vastly superior to methods used in the past. One reason is because the better designs split the audio into narrow bands and use psychoacoustic masking to 'hide' the artifacts of clipping.

Finally, our brains will also "fill in" missing information. The best example of this is the phenomenon of phantom fundamentals. In a system with limited bass, a very low tone won't be reproduced. But as long as the expected harmonics are present, you will "hear" the missing fundamental.

Of course, this doesn't mean that a wide frequency response is a waste of time. For one thing, deep bass is "felt" as much as heard and those listeners with better receiver systems will notice if it's removed, phantom fundamentals or not. Nor do we want to take advantage of masking and other psychoacoustic effects to reduce the quality of our air sound to something no better than a highly-compressed MP3.

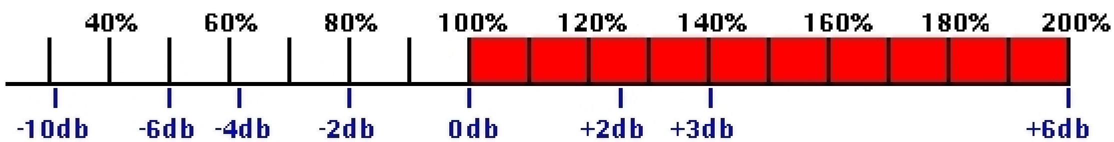
We mention these principles because a judicious and strictly-limited use of them can help achieve an excellent compromise that sounds great to the listener.

A Practical Application: Overmodulation

You will run across people in broadcasting who think that illegal overmodulation is necessary to be competitive. The equal-but-opposite side of this coin is that they might insist that a competitor is overmodulating, when in fact, the competition may simply have a better audio processor and someone who knows how to adjust it.

You can use the facts covered in this chapter to prove them wrong. Unless they are willing to overmodulate to great excess, they would be interfering with neighboring stations and courting a fine for no real benefit.

Figure 4: An expanded VU scale showing dB values at high modulation levels



The expanded VU scale above (figure 4) clearly shows that you would have to modulate 200% to obtain a marked increase in loudness (6 dB). The more "modest" 125% modulation level is almost pointless: you'd still risk being fined ... and *most listeners will barely notice that 2 dB difference.*

When listening to typical program audio on an entertainment system, if you wish to increase the sound pressure level such that it is perceived as being "twice as loud" as before, you will need

- Approximately 10 times the power
- Approximately 2 times the power
- To ensure that the loudspeaker has adequate cooling
- To shift all harmonics beyond the audible range of human hearing

A modulation level of 125% represents an increase over 100% of about

- 6 dB
- 10 dB
- 14 dB
- 2 dB