Clearly Mixed

by Chris & Trish Meyer, Crish Design

When mixing multiple audio tracks, make sure you keep the listener focused on the right one.

Admit it: How many of you mix audio by dragging the dialog, music, and sound effects into your NLE's timeline…and then think you're finished? Okay, you don't , but we've all heard programs that sound as if they were mixed that way. Each component might sound fine individually, but when more than one is playing at the same time, they obscure each other. The solution to this is thoughtful management of audio levels.

If It Sounds Like a Duck…

The most important point to keep in mind is: What should the viewer be listening to at each point in the timeline? That's the audio track that should be the loudest, and the others should be reduced to provide just a supporting role. Obviously, this does not happen automatically – you need to put some effort into mixing their relative levels, and changing those levels over time to help change the viewer's focus from one track to another as appropriate.

The figure below contains just such a mix. To help us visualize our priorities while mixing, we often arrange the audio tracks so that the most important track – the dialog – is on top, and the least important track – the sound effects – is on the bottom. Let's walk through the reasoning behind the details in this mix.

In our sound effects track, an "ambience bed" (a constant background of sound effects related to a scene, such as traffic or waves

Figure 1: In this sample mix, the dialog is the top track, the music is in the middle, and the sound effects – an ambience bed – is on the bottom. We introduce the elements one at a time to build interest, making sure we fade down the less-important tracks as the more important ones take center stage. At the end, the process is reversed.
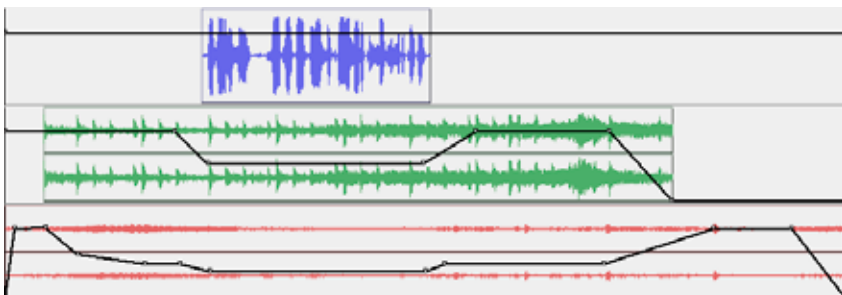
lapping on a shore) sets up the scene, akin to how an establishing shot sets the mood. Rather than coming on at full volume, it is faded up just as the video is dissolved on. It is a good idea to make the duration of the fade match that of the dissolve – this will further lock in the pace as well as the feeling that "we're changing scenes now…"

Next, the music starts on track two. The timing of this fade is optional; an approach we often use is to start the fade down for the ambience bed shortly after the music starts: This way, the fade does not draw attention to itself, but it still gets the background sound effects out of the way. We'll tend to keep the ambience bed playing at a low level in the background, and then fade it up again after the music and dialog tracks stop.

Next comes the dialog track itself. Notice that the music fades down before the dialog starts: This way, it does not obscure any of the words, and subtly cues the viewer that something new is about to happen to pay attention to. This technique is often referred to as "ducking" as the music ducks out of the way of the incoming dialog. Whatever your production may be – television, DVD, the web, or any situation where dialog and music appear together and the dialog is supposed to take the lead – remember to duck the other audio tracks out of its way.

A subtle but important touch is timing the music fades to beats in the music track. You know how people count "ONE Two Three Four" in time with music? That "One" is the downbeat, and it usually marks the most important timing point in the music. (We discussed spotting music in our previous article, Editing to Sound.) When ducking music, it's nice to start the fade either just after the One or the Three, and to be down by the next One – and to start the dialog there. When fading back up, we'll usually start as soon as the dialog stops and get back to the normal volume right at the next One to reinforce that beat.

Timing can affect intelligibility as well. Sometimes words are hard to hear no matter what you do with the levels. Although we will often time important words or starts of sentences to the Ones in the music track for reinforcement, sometimes these beats can step on the words. If you have the flexibility in your project to slide

either the video edits or music track a touch, moving one or the other by as a little as a frame is often enough to get the dialog and music out of each others' way and make the dialog more intelligible.

Figure 2: Here we've slid the timing of the music (the green track) relative to the dialog (the blue track) so that the first word of dialog happens just after – not at the same time as – a major beat in the music. This increases intelligibility.
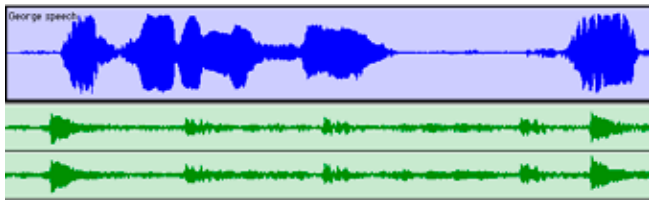
Level-Headed

After getting the overall mix right, we can focus on improving the dialog track. Notice in the adjacent figure that the volume level of the voice track does not stay the same. People rarely talk at a consistent volume – especially if they are not trained voiceover talent. Even if they do, what they are talking over (the music) is rarely consistent in level. And even if everything is consistent, perhaps certain words should be emphasized for expression or because they are key points.

Figure 3: We've refined the mix of the dialog track to emphasize important words that were spoken too softly. We also ducked down the volume of the dialog track to hide some of the background noise that was exposed between sentences.
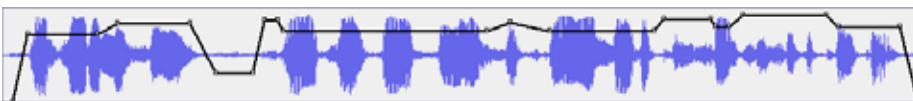
We will constantly tug at the volume of the voice to make sure it appears to be consistent and has the expression we want. Set your work area to overlap the segment you are working on, and don't be afraid to preview the mix again and again and again to get it right.

Editing the volume automation of dialog tracks can quickly become tedious and time consuming. On longer projects, you should bring in some automated help and use an audio "compressor" or "limiter" to help smooth out volume variations in the voice. Compressors look at any sound over a certain threshold loudness, and reduces their level. This may seem counterproductive, but what it means is that you can then raise the overall level of the entire track without worrying as

much about driving it into distortion. A more brute-force device is the limiter, which simply clamps the maximum level without "clipping" (distorting the loudest points in the waveform), which means you can keep cranking up the overall level.

We try to use a compressor when recording our voices to cut down on post production work later. However, you're not always in control of how your footage was shot. If your original audio files were not recorded with compression or limiting applied, you may need to process it after the fact in a dedicated audio program. We personally have a Digidesign ProTools audio editing workstation as part of our studio (the screen shots in this article are from ProTools). There are also several audio programs designed to be used with non-linear editing systems, such as Adobe Soundbooth or Apple Soundtrack Pro.

Zooming Out

It is often a good idea to try to establish the scene before getting into the action. This is where stock footage can come in really handy: Whereas you may need to shoot footage specific to your project to tell your story, stock footage is a good place to turn for this supporting role – especially when you don't have the time or resources to, say, shoot an aerial flyover of the city where your action takes place.

Aurally, we use ambience beds in the same way. You can buy special sound libraries of ambiences from vendors such as Sound Ideas and the BBC. In addition to helping set or reinforce the mood of a scene, keeping a constant background "noise" running also helps cover problems – such as hearing any background location noise in the dialog tracks that may be cutting in and out between edits. A constant low-level noise background also helps cover aural artifacts that may crop up as a result of using heavy data compression when preparing a video program for the web. Once the ears and brain agree upon one explanation for what the background noise is (such as cars on a freeway), they're not going to waste any more time figuring if the noise has changed, meaning they are less likely to notice if something goes awry later. As a result, you can keep your viewer focused on your visuals – not what's wrong with the sound!

###

Chris and Trish Meyer are the founders and owners of Crish Design, an award-winning motion graphics studio that has created a wide variety of work for film, broadcast, corporate events, and special venues. They are one of the original development sites for After Effects, and have two books currently in print: Creating Motion Graphics and After Effects Apprentice.