Excerpts from
Producing Great Sound for Film and Video

It's Not Just the Writer's Responsibility

If you're creating the script, you've got a great opportunity to write for sound. But if you're shooting someone else's script, you must still go through the process of hearing it in your head, marking sound cues, and possibly breaking things up to make room for audio. It's the best way to assure that the track will go together smoothly and predictably.

Mark the Script

If the script is already written, you can still apply careful analysis. But the time to do it is before the shoot. Walk through the script, hearing every scene in your head and making notes about possible sound treatments. It's usually easiest to make these notes on the script itself, since there probably will a lot of them referring to specific actions. If the script is on a computer and you don't want to print it out, see whether your software supports sticky notes. Figure 4.1 shows these notes on a script in Adobe Acrobat.

Some of those notes are reminders of things to do in post, both real-world sounds and textures, and special effects that might be worth trying. Others are things that should be decided before the shoot, so actors can perform or react to them. These notes will also often suggest extra sounds or wild lines you can grab quickly at the shoot (and a lot more cheaply than trying to recreate later). They may also uncover some audio conflicts that can be resolved before you roll camera.

Once editing has started, sound notes are still important, particularly on film-length projects where audio post can last weeks or more. But once editing starts, the script might become irrelevant: While shorter projects are usually pretty good at sticking to the script, long form documentaries and narrative films often get rewritten during the edit. I've found the best way to keep track of notes on long post jobs is with a spreadsheet, organized by timecode (Figure 4.2). Start a couple of columns for minutes and seconds, and one for hour or reel if it's a long show. Then watch the edit. When you hear something that could be improved, stop and note its location, its cue or dialog, and what you want to do. Go through the film more than once, but don't worry if the notes aren't in order—it's easy enough to sort them by the time columns when you need to. Leave the minutes and seconds blank for ideas that don't have a particular address; when you sort, they'll fly up to the top of the list.

Figure 4.1 Notes on a feature film script.

Figure 4.2 Notes for the same film, after it's shot and editing has begun.

You can also add columns to identify a note, dialog, effects, what kind of work is needed, which characters are involved, and even the director's responses. It'll all save time when you're actually doing the work. In the figure, the columns represent:

Time (columns a/b/c)—in hours, minutes, and seconds. (Features are often broken down by reels of fifteen to twenty minutes, with the timecode hour designating the reel number.) It isn't important to worry about frame numbers at this stage, and obviously, you don't need to worry about hours on short projects.

Cue (d)—a quick note about what's happening, so you can refer to it easily.

Scene (e)—the scene number on the original script and as shot. Even though scenes might not be in this actual order after the film is edited, the slates and captured scenes in NLE refer to these numbers.

By (f)—whose idea this particular note was, or its priority. When deadlines approach, it helps to know which note were specific requests by the director and which were random thoughts of the sound designer.

Done (g)—a simple checkmark when a task is completed, very helpful if you're using conditional formatting (below).

Dialog/Effects/Foley/Special (h/i/j/k)—these columns can be used to sort notes by task or order of importance.

Character (l)—which actors are involved, in case ADR will be needed.

Note and Decision (m/n)—what we want to do, and what we finally decided to do about it.

(I wasn't responsible for scoring on this film; if you are, add appropriate columns for musical ideas.)

Note the page tabs on the bottom. One of the nice things you can do with spreadsheet macros is use them to automatically generate Foley or ADR lists that appear on other pages, or keep a "to do" list for chores that still need to be done. A spreadsheet's conditional formatting function can make the notes change their font format based on the kind of sound or operation, or gray out their times once an operation is checked off as done.

Make Room for Sound

Real people stop talking every now and then. When they do, we hear the world around them. As a filmmaker, you can help this along by avoiding dialog over noisy actions: Sue shouldn't be talking to another character while she hangs up the phone. It complicates mic placement and makes a good mix more difficult.

Even if a sound effect will be added in post, leave room for it in the dialog. Characters shouldn't start screaming at the same moment there's supposed to be a car crash or explosion—that just detracts from both sounds. Let them scream in anticipation as the car careens toward the wall, or react after the explosion settles.

Since different sounds have different frequency ranges, consider how their brightness or deepness might conflict. The metallic crunch of car against wall has a similar timbre to an adult male shouting, so you won't hear both if they're happening at the same time. But substitute either breaking glass or a female scream and they'll coexist together.

Jumping about 40 pages...

Chapter 6
Microphones and Room Acoustics

Rose's Rules

✓ What you hear isn't what you get. Mics and acoustics conspire to make recordings sound different than reality.
✓ But with a little knowledge, you can choose the right mic and make any room sound better.

We listen to the world through rose-colored ears (mine are sort of pink). Within a few minutes of entering a room, most people learn to ignore its acoustics. The hum of machinery and computer fans that normally surrounds us seems to disappear as soon as we start paying attention to something else. It's no problem to pick out your child's voice in a crowded schoolyard or your friend's conversation at a noisy cocktail party. We use subtle timing and level variations between how the sound strikes each side of our head, and how it's affected by the convoluted shape of the outer ears, to figure out where a sound is coming from. We use our eyes, and our experiences of being in different spaces, to focus our hearing even further.

But the real world doesn't sound like our perception of it. When you squeeze dialog through a microphone and speaker, most of those subtle cues disappear. All the noise and reverb that your ears automatically ignored at the shoot can come crashing back, polluting what you thought you were recording.

The only way to avoid this is to take special care in how you set up the room and the mic.

ABOUT MICROPHONES

A camera's lens defines a precise rectangle. Point it in the right direction, and you can frame a shot that includes only what you want the viewer to see. Everything else is out of camera range.

With a lens, we could take a basic two-shot (Figure 6.1) and zoom in on just the horn-playing trickster. His bear friend would be completely gone (Figure 6.2).

But a mic doesn't work that way. When we talk about a mic's pickup pattern, we're merely indicating in which direction the mic is slightly more sensitive. It actually continues to hear things from all around the room. There is no such thing as "just out of microphone range."

Figure 6.3 is a visual analogy of what a good shotgun mic would hear, focused on just one subject.

Three visual aspects of Figure 6.3 translate to what you'll hear with real-world directional microphones:

Figure 6.1 A basic 2-shot.

- A directional mic will make sounds coming from the front seem slightly closer. That's why our horn player is slightly larger, compared to the bear, in Figure 6.3.
- It does this by lowering the volume of sounds from other directions. That's why the rest of the scene is darker.
- A directional mic will change the timbre of sound coming from other directions, adding a coloration that emphasizes some frequencies and lowers others. Note how the bear's colors are subtly changed.

Figure 6.2 A longer lens lets us focus on just one character.

The more directional a mic is, the more you hear these

effects. Better-quality mics will have less of the off-axis coloration than cheaper ones with the same directional pattern, but it will always be there.

Actually, Figure 6.3 is a simplification. What that directional mic *really* hears is more like Figure 6.4. The mic's pickup continues beyond the photo—picking up all of my desk, me, my camera, and the room behind me. Sounds from those directions might be softer than our star horn player, but they'll be there. Depending on a lot of factors, they might even be louder.

Figure 6.3 Unfortunately, the mic isn't as discriminating. Even the most directional shotgun won't exclude the bear.

Try This

Take your best directional microphone, camera, and good headphones. Turn on a radio, at conversational volume, in an average room of your house or office. Stand about three feet away from the radio, and listen to it through the mic.

Now turn the mic around, so its back is pointing to the radio. Listen again and you'll still hear the radio. It'll just be a little bit softer and not sound as good.

Try the same thing at 6 feet. The difference between front and back will be less, and they'll both sound bad.

Types of Microphones

A microphone is just a pickup element to convert sound pressure into electrical voltages, mounted in a box. The construction of the pickup element determines how it translates pressure into electricity, but only a couple of methods—dynamic and condenser—are appropriate for video production. We'll cover those in a few pages.

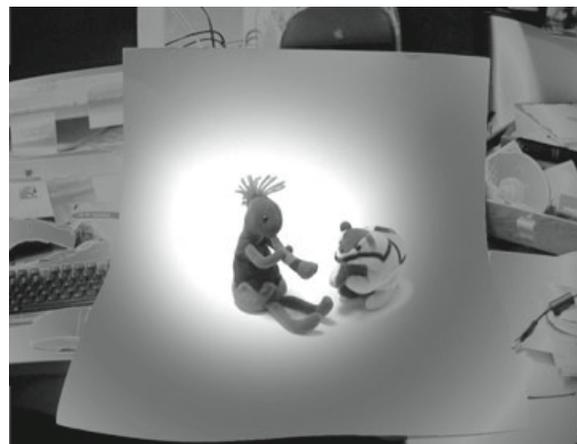What's much more important is how the microphone picks up pressure from different directions. This is determined

Figure 6.4 Real-world mics continue to pick up, far beyond what we think they're pointing at.

by the shape of the box, its vent holes, and internal baffles. But it's also influenced by frequency: sounds of different frequencies will work their way through those holes and baffles differently.

Microphone Directionality

Polar Patterns

It's fairly easy for mic manufacturers to measure this phenomenon. They put the mic in a room without echoes, walk around it with a tone generator, and read the mic's output at different frequencies and directions. They plot this on a circular graph like Figure 6.5. Directly in front of the mic is considered 0°, which is at the bottom of this figure (some manufacturers put it at the top). The voltage for a sound coming from directly in front is considered 0 dB. As the voltage drops for sounds from different directions, it's plotted on the inner circles. Multiple plots are used to show how the mic reacts to different frequencies. The graph is split in half because mics are symmetrical: you can assume there's a matching half circle, facing in the other direction, for each side.

All of the polar patterns in this section are based on actual measurements of some highly-regarded, expensive mics, and re-drawn for clarity. Figure 6.5 shows the polar pattern of an *omnidirectional* mic. On the left side, you can see that the mic is equally sensitive from any direction, for any frequency between 125 Hz and 1 kHz. On the right, it's also equally sensitive from any direction between 2 kHz and 8 kHz. But the gray line on the right shows how this mic becomes slightly directional at 16 kHz: sounds at that frequency, coming from the rear, will be –5 dB softer than those from the front. This is insignificant for most purposes. It's caused by the mic's body casting an acoustic shadow. (Sound lab often use very thin omnidirectional mics to measure noise sources, so the mic's body doesn't cast much of a shadow.)
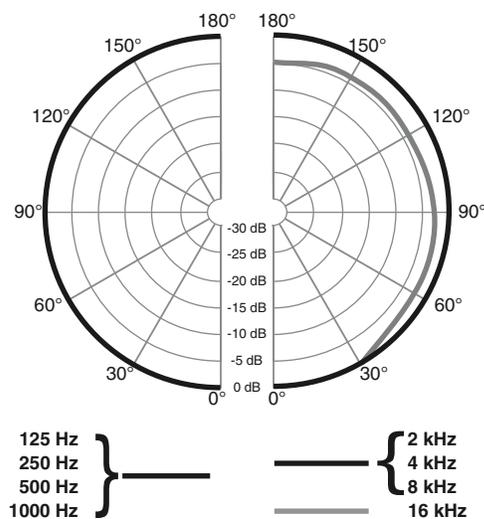
Omnidirectional

Omnidirectional mics (or *omnis*) like the one measured in Figure 6.5 have a solid box around the element's side and back. There's only a tiny hole in the back, to equalize barometric pressure (Figure 6.6). Air-pressure waves strike the front of the element and generate a signal. The air inside the box acts as a spring to push the diaphragm out when the sound wave is in rarefaction. Any sound waves coming from the back refract around the mic and hit the element. Omnis are sometimes called *non-directional* for this reason.

Since an omnidirectional mic is mostly a sealed box, it's less sensitive to wind noise than other designs. This kind of mic is also sometimes called a *pressure*

Figure 6.5 A graph of the polar response of a

Jumping about 200 pages...

pay to this factor. If you're cutting between takes of the same copy, or changing words in a performance by a trained actor or narrator, the pitch will be remarkably consistent. Any slight variations you cause by editing will probably sound like natural intonation, particularly if the edit is smooth and the viewer is following the content. Wide variations will be immediately apparent, so you can undo the edit and try something else. One of the things you should try is a few percent of pitch manipulation or varispeed, if your editing system has this feature. A pitch shift—no more than 3% higher or lower—may be all you need to join two unfriendly phonemes.

Projection Levels

A performer's projection level affects more than how loud they are. As you raise your voice, the throat tightens, and the vocal buzz that forms voiced sounds loses some of its lower frequencies. At the same time, most trained speakers will also push their voices "into the mask," directing more energy to resonating cavities in the front of the face and adding high harmonics. You can't compensate for these timbral changes by simply adjusting volume. In most cases, you can't compensate for them at all, and editing between two takes of widely varying projection levels can be difficult.

The envelopes of individual words also frequently change as projection levels rise. The beginnings of individual words get stressed more, to separate them, than they'd be in normal speech.

- If you must cut between two takes with different projection, keep the takes on two separate tracks as you move their phonemes around. Then experiment with volume and equalization to make them match. A little reverb on the softer track can also help. Once the sound matches, mix the two tracks to a single composite voice.
- This idea of keeping "almost but not quite" matching clips on separate tracks, so they can get different equalization or volume treatments, is one of the keys that separates well done theatrical film dialog from the voices in corporate videos or many TV shows. In fact, in Hollywood, every scene of dialog is often split into three or more tracks.

EDITING IV: THEATRICAL FILM DIALOG

Most of what's in this chapter applies to spoken words in any project: training video, TV commercial, wedding album, feature film. They all usually require some dialog editing, and the rules for that editing are always the same.

But theatrical film deserves an extra level of care. That's partially because when a track is played in a large auditorium through a good system, we can concentrate on it and hear minor flaws that would be acceptable on the small screen. It's also because the budgets and traditions of feature film often let you take more time with the dialog editing, and our expectations are higher. I'm not complaining. I wish we could put this kind of work into every project, no matter what the medium.

Track Splitting

The main technique that separates theatrical dialog from other forms is *track splitting*. It serves two primary purposes:

- Hiding subtle jumps in timbre and background sound. There can be thousands of these in a film. They're inevitable when you cut from one camera angle to another, because the positions of microphone, walls, and noise sources change with each setup. They're also the result of condensing a morning's worth of out-of-sequence shooting into a single scene: the differing levels of traffic, airplanes, and other noises change abruptly, even when the dialog is supposed to be continuous.
- Simplifying the mix. This surprises some people, since you're taking a single track—needing just one fader at the mix—and splitting it to a half dozen or more tracks and faders. But it really does help: It lets you (or a re-recording mixer) preset equalizers and other processes for elements that need them, rather than having to switch the processes on and off between words.

Figure 13.7 shows some of the track splitting in a scant fifteen seconds of a fairly simple scene.

- On the bottom (letter A in this drawing) is how the picture editor cut the scene in an NLE, and passed it to audio post as an OMF. This scene actually consists of a master shot and a couple of close-ups.
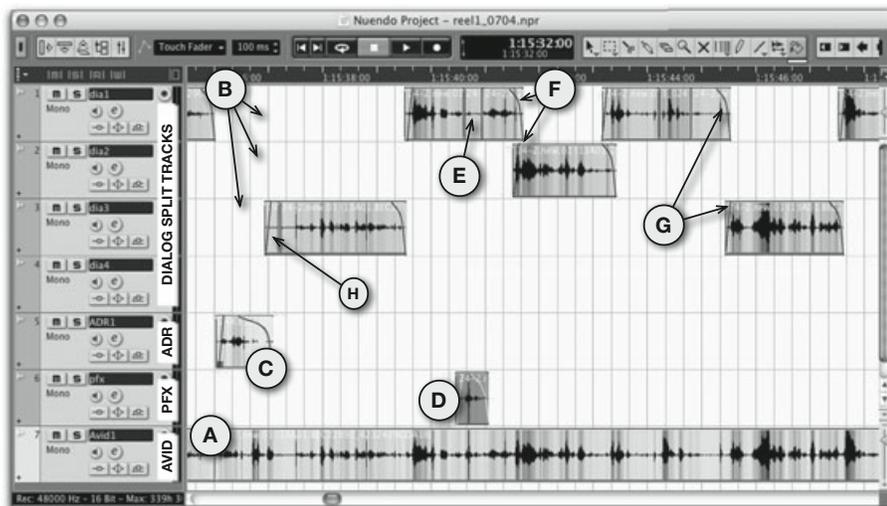

Figure 13.7 Dialog editing in fifteen seconds of a simple scene.

- The scene has three different characters. Letter B shows how their individual lines have been copied from the NLE's version, and are split among the top three tracks. This way, the re-recording mixer can apply different volume and equalization settings to each character, to compensate for different distances from the boom or differences in background tone when the close-ups were shot. This also makes it easier if the director wants to add subtle emphasis to just one character during the mix.
- One line had to be replaced with ADR and will need completely different processing to match, so it's separated at letter C.
- One sound effect, a door slam, was picked up by the boom during the action. It's not dialog but might be useful in the mix, so it's moved to a production effects track (at D). That slam might be replaced with a different one, recorded separately. So a little bit of room tone is copied from elsewhere in the take, and pasted where the slam used to be in the track (E, on the top track). This way, the background can continue if a different slam is used.
- Note how most of the clips on the top three tracks overlap by a half-second or so (such as at F), even though there are hard cuts between them on the original. These clips were extended by grabbing their *handles*, extra pieces of each clip that aren't in the original edit but were carried on the OMF.
- The slanted or curved lines drawn over the ends of most of the clips (G) are fades. They work in pairs as cross-fades during the overlaps, to disguise differences in room tone. Most good audio programs have the ability to add and edit these fades, which work separately from any volume adjustments you make during the mix.
- We were lucky editing this scene, in that most of the overlaps could be done with handles: the actors paused between their lines, so there was room ambience around them. But you can see (at H) how a small clip of extra room tone had to be stuck to the front of a dialog clip. There was no pause before the line in the original performance, so if we'd opened up the handle there, we'd hear extra words during the fade. Instead we found some clean ambience, where no one was speaking, elsewhere in the take and used it for the cross-fade.

✺ *Why Not Just Process the Clips?*

Most audio software lets you apply volume, equalization, and other processing to individual clips on a track separately, or apply them to the entire track as a whole. The former is handy for repairs and effects where just one line, or sometimes one word, needs to be treated specially. But the latter is faster, easier to change, and usually leads to a smoother mix.

Continued

> Also, pre-processing locks you into a sound, and restoring the original can be difficult or impossible. It's seldom a good idea to make this kind of decision until all the elements are in place, and you're monitoring in a proper environment.
>
> If you are tempted to pre-process a clip to solve a problem, keep an unprocessed version of it on an unused track. You can mute this backup or *mirror* track so it doesn't confuse things. There's a good chance you'll reach back to it at the mix. In fact, most Hollywood re-recording mixers won't touch a pre-equalized project unless there are mirror tracks also available.

There's another feature film technique that results in better tracks, even though it also adds time to the process: the *dialog premix.* We'll cover it in Chapter 17.

EDITING V: THE TRICKS

Breathless

Go back to Track 16 on the book's CD, the one you loaded at the start of this chapter, or look at Figure 13.3. There's a big puff of air in the middle. Most announcers will take a gigantic breath between paragraphs, and many loud, fast ones throughout the paragraph. These noises don't communicate anything other than "I'm reading from a script," and—unless you're trying for a comedy effect—should come out of voice-over tracks. Breaths in front of a take are easy to get rid of; just move the in-point a little later. But those catch-breaths during a read are more troublesome. While you can spot breaths on a waveform and hear them when you scrub, you can't just cut them out; that would pull the surrounding words too close together. You also can't just erase them or turn their volume down because that'll leave the words too far apart.

In general, a breath in a voice-over can almost always be replaced by a pause *two-thirds* its length. If it takes one second for the announcer to gasp, use 20 frames of silence instead. The result sounds cleaner, more energetic, and completely natural. This ratio has worked for me in thousands of projects with hundreds of different announcers. Even though I edit by ear—erasing the entire pause, playing the line in real time, and tapping the Mark key where I think the next phrase should start—it almost always turns out to be two-thirds. I have no idea why that number is magic.

If you're starting with a clean voice-over recording and it will be used on a limited medium such as broadcast TV, or be mixed with music, you can replace breaths with digital silence. But if it'll be played at theatrical levels, use room tone; otherwise, the finished edit will sound choppy. On-camera breaths are trickier, since the two-thirds trick would destroy sync. Even replacing the whole breath with room tone may be a mistake; if we see the talent's mouth open, we should hear something.