

Published August, 2007

There's an art to cutting talking heads. There are also some nifty tricks.

Decapitation

I've worked with a commercial director who has a great technique. We'd do an interview spot, ad-libs from a consumer who liked the product. After we took plenty of good stuff, he'd ask us all to sit quietly for a few minutes. He'd grab a pad and write a script based on the subject's own words. Then he'd do a few scripted takes. These non-actors couldn't read a script convincingly, but it wasn't a waste of tape. Days later, when we were editing and wanted a particular phrase to tie two bites together, he have me grab just the right words from the scripted version. They were short enough that the stiff performance wasn't a problem.

Well, maybe it was a waste of picture, because we'd have to cut away from the subject during the edits. But not the track: those scripted words made the ad-lib commercial work. (The cutaway was also another chance for a client-pleasing beauty shot.) Today, quick dissolves or even jumpcuts are sometimes acceptable for talking heads, leaving fewer cutaways to hide the edits. But digital editing has given us a lot more tricks, whether the words are on-camera or not.

You'll need the right tools. NLE software is often limited to editing on framelines, and 1/30th of a second is a long time for dialog. So move the clip to an audio program. Grab slightly more than you need, or specify handles in the NLE's export dialog (60 frames or so of overlap at each edit point, that the audio program will keep hidden until needed). And you'll need good speakers. This is dialog editing, not mixing, so you don't need a perfect monitoring environment. But it does have to be clear and loud enough for you to hear tiny parts of words.

Virtual Valium

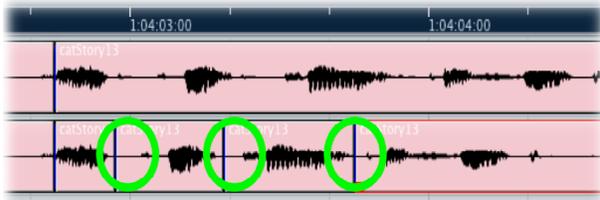
Nervous interviews are less convincing than calm ones. The biggest giveaways that a subject is nervous—or worse, has been prompted to say a line—can be involuntary: glottal shocks, double stops and sucking breaths. We hear them in real life as well as on TV, and take them as cues the speaker isn't relaxed. (If you can't hear these problems in your head or find them on a current project, go to my website, dplay.com/dv, and download a sample.)

- *Glottal shocks* are reserves of breath, stored in the lungs and released suddenly when a



word starts with a vowel (like ‘and’ or ‘I’). These words should start smoothly, with the throat open before breath starts to flow. If breath starts too soon, the sudden burst of air makes a click. Scrub slowly through the start of

the suspect word. The click will be definite, taking maybe a third of a frame where the word starts. Fixing it is easy. If you’re not concerned with lipsync, just cut it out. If the subject is visible and you don’t want to change



the timing, replace it with an equal length of room tone. The subject will sound calmer. Figure 1 shows a before and after, with the original track on the top, and red circles around the shock.

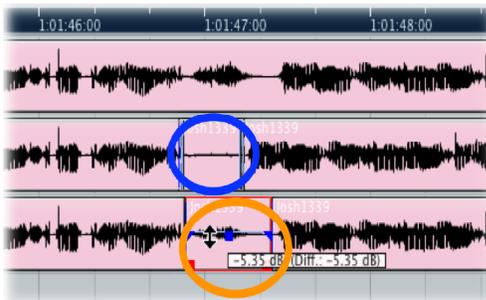


Figure 1: Cutting remarks. Easy edits can make a voice track sound more relaxed. Zapping a glottal shock (red circles); closing double stops (green circles); cutting a

breath (blue oval) or just lowering it (orange oval). The original clips are on top, with the edited versions under them, and these screen grabs don’t share the same time scale.

- *Double stops* can happen when a subject is carefully choosing their words, or an untrained performer wants to sound more formal. Certain consonants, such as /d/ or /p/, normally stop making any sound for the moment that the lips close or the tongue touches the roof of the mouth. In casual speech, when two of these stops are next to each other, we usually combine them. Try saying *that cat tried to sleep*, conversationally. Now say it while paying particular attention to your enunciation. You’ll probably hear extra /t/ and /d/’s, almost like *that-t cat t-tried-d to sleep*. Those are double stops, caused by closing the consonant at the end of one word, then opening and closing it again to start the next. And they’re as easy to fix as glottal shocks: just cut out the first click, and slide the second earlier (the green circles in Figure 1). If lipsync is an issue, you can cover the clicks with roomtone. Be careful with double /b/ or /p/’s are on camera, because you may see the lips close twice... try the edit, but check that it looks natural.
- *Sucking breaths* are fast inhales, when someone knows they’ll soon have a lot of words to say. Even trained announcers take them. They shout “I’m not making this up spontaneously”. For voice-overs or interview cutaways, pull the whole breath out, and replace it with roomtone that runs about two-thirds the breath’s length (blue oval). If the subject is on-camera and you can’t change their timing, fade the breath down. Good audio workstations will let you draw a fade right on the clip; if not, isolate the breath as

its own clip and lower its volume. A quick dissolve may be necessary to get back to normal volume smoothly (orange).

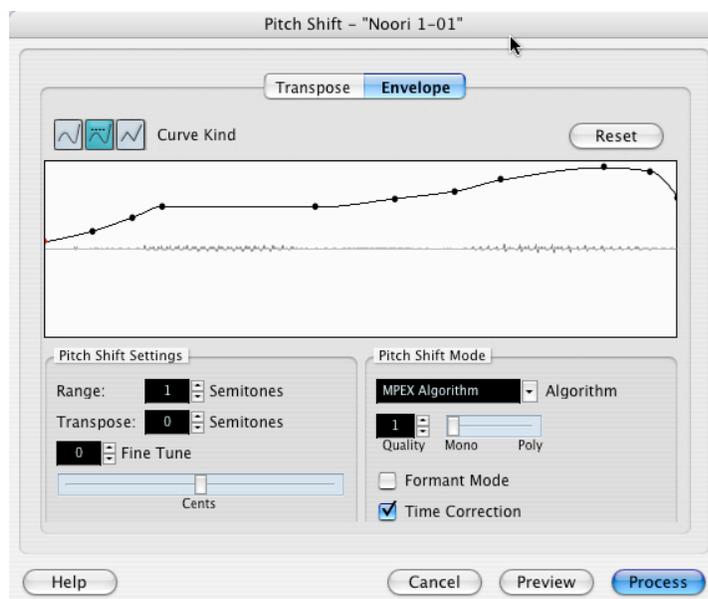
Lick your tails

Sometimes you'll find the perfect phrase in an interview ... but the darned subject keeps on talking, extending the phrase or starting something else without a pause. When you cut the clip to make sense for the story, it sounds abrupt and calls attention to the edit. Again, room tone to the rescue. Scrub exactly to where the next word starts, after the phrase you want—it'll probably be in the middle of a continuous sound—and mark your edit there. But don't cut to silence. Instead, cut to room tone. It doesn't take much—maybe five frames or so—but it has to match perfectly. A separate room tone recording from later in the shoot might not work, but you can often find all you need in a pause somewhere near the desired phrase. Make a short fade at the end of the room tone to silence. Some words that end in /s/ or /z/ can make a click when you cut to room tone, but a short audio crossfade fixes that.

Initial /h/ sounds at the start of an interview bite can get lost, particularly if you're used to cutting by waveform instead of sound, because they're so soft they don't show in the clip viewer. Just move the start of the clip about ten frames earlier (another reason to have handles when you export to audio software) and you'll find the /h/ waiting there for you. If the /h/ wasn't recorded at all—this can happen in a noisy environment or with bad miking—just add some room tone up front. With the right timing, viewers' ears will fill in the missing consonant.

Bend it

Most of these tricks depend on how human speech works, and I used to do them on



1/4" tape and 35mm mag... though they're easier and faster on a computer. One interview-smoothing function wasn't possible until the digital era, unless you had expensive rotating-head audio decks. Pitch-shifters are now built into almost all audio software, and some are absolutely amazing. While they're usually intended to fix a bad singer or create harmonies for good ones, they can also change the melody of an interview clip.

Figure 2: Pitching a curve. A pitch shift like this can make a phrase sound more like a question. Note how the total shift is only about 3%, or one semitone.

Want to turn a statement into a question? Select the last two or three syllables and raise the first a tiny amount, and the second a bit more. Figure 2 shows a typical pattern. If your software doesn't let you draw a variable pitch shift, separate the syllables into clips and process them separately. The biggest shift should be less than about 3%—about one musical half-step—so it doesn't turn the voice into a cartoon. And pitched background noises can give this trick away, so do your filtering and cleanup first.

This also works when taking a bite from a longer question, if you have to cut out before the speaker's pitch naturally raised. And of course you can shift downward instead, to give finality to a statement. If you're working with children's voices, you can go much farther than 3%. That's because kids' upper harmonics vary more than adults'—it has to do with the physical sizes of their heads. I use a variation on this when young children read a script too flatly. (You can hear their typical monotone in many classic Charlie Brown TV shows.) In fact, one time I hooked the pitch shifter to a musical keyboard, and played the pattern of the entire dialog.

Moralists may question whether this kind of trickery is ethical. I don't have a problem with it, so long as you're keeping subject's original meaning. It's no worse than taking time to light them, or using a little makeup so the shot looks better. Besides, it's fun to do magic while the client is watching.

Jay Rose, CAS, has been a dialog cut-up since the days of Movieolas and sync blocks.

© 2013 Jay Rose

*May be reproduced for non-commercial and educational use provided that:
The complete pdf is reproduced intact with no changes or deletions; and
No money is charged for reproduction or copying, or for the editing and assembly of any work containing this material.*

If you need other permission, write to jay@dplay.com.