Audio Post: What You Need to Know

Audio Connections

FireWire from camera to computer, no other equipment (not even speakers)? Skip this section. Otherwise, read on...

Analog wiring

Balanced: usually noise- and trouble-free, used by pros. Often TRS (3-conductor) phone, barrier strip, or anything that supports 2 equal conductors and a shield. The shield isn't strictly ncessary for noise control, and breaking it can help with occasional hum. For even better noise rejection, use shielded 4-conductor "star-quad". Balanced wiring levels are usually specified in dBu (see next column).

Unbalanced: used in most consumer gear, often on RCA or 2-conductor ¼" or mini phone plugs. Prone to ground loops and noise pickup. Requires a single conductor and a shield. Using 2-conductor shielded cable, one conductor for signal and one for ground, with the shield connected to ground at only one end ("telecoping shield"), can help lower noise. Levels usually specified in dBV.

If it hums, the best cure is an isolation transformer in the audio (ProCo, Sescom, or Shure, about $50; Radio Shack 270-054 $15, much lower quality), video or AC supply wiring. A temporary, highly unsafe cure can be disabling the ground pin on the AC connection. *This is not recommended.*

Gotchas:

- Mini TRS connections may be unbalanced stereo (on prosumer and consumer devices) or balanced mono (some wireless receivers). They're not interchangeable.
- ¼" TRS may be unbalanced stereo (headphone jacks), unbalanced bidirectional (mixer inserts), balanced mono (some mixer inputs/outputs). Also not interchangeable.
- **Worst case**: going from a balanced output to a stereo input, either with miniplugs or with a ¼" to dual RCA Y cable. Signal may sound fine in stereo, but will disappear in mono.

Digital wiring

AES/EBU: Used on pro equipment. 5 volts at 110Ω balanced, almost always on XLR connectors.

s/pdif coax: Consumer. 5 volts at 75Ω unbalanced, usually RCA connectors but sometimes BNC.

s/pdif optical: uses light fibers and "Toslink" connectors.

Proprietary: ADAT optical and TDIF multipin electrical, 8 channels, used on some mixers, recorders, and computer interfaces. Some Sony DAT recorders put s/pdif on a 7-pin connector; adapters at www.core-sound.com.

Gotcha:

- AES/EBU and s/pdif are very high frequency and require impedance-matched cables to prevent problems. Treat them like video.

Volts and Decibels

Decibels are ratios between two signals. 0 dB is a ratio of 1/1. -6 dB is a voltage ratio of 1/2. +6 is a voltage ratio of 2/1.

If you want to specify a signal level in decibels, you need a reference (the second number in the ratio). Analog signals (dBu and dBV for wiring, but also dB SPL for sound in air and VU on a meter) use a nominal "zero level" for that reference: the average loudest signal. Analog levels can be louder than zero (+ decibels) or softer (- decibels).

Digital signals (dBFS) use a true zero: the loudest possible signal, with all bits set the same. Nothing can be louder, because there aren't any more bits. Practical digital signals are always below zero, or - decibels.

Nominal analog voltages

Prosumer	Volts	Professional
+6 dBV	2.0v	+8.2 dBu
+1.78 dBV	1.228v	**+4 dBu**
0 dBV	1v	+2.2 dBu
-2.2 dBV	.775v	0 dBu
-10 dBV	.316v	-7.8 dBu
-20 dBV	.1v	-17.8 dBu

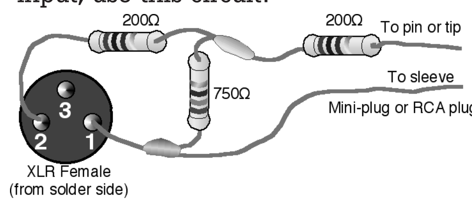
Boldface: nominal Zero levels.

Cross-connecting consumer and pro

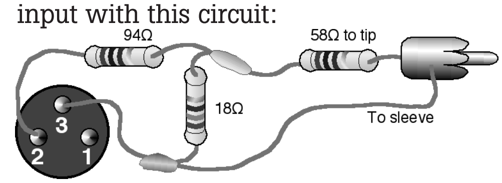
Prebuilt, powered bi-directional converters are available, often with XLR male and female for the balanced equipment and RCA jacks for the unbalanced, starting around $35 per channel (Rolls). They preserve balanced wiring on the XLR side and adapt the levels.

But if you don't need balanced wiring's noise immunity, you can cross-connect directly. Connect balanced "hot" (pin 2 of XLR, tip of TRS, or + of barrier) to pin or RCA or tip of mini. Connect a balanced input's "return" (pin 3 of XLR, ring of TRS, or – of barrier) to sleeve of RCA or mini. For outputs, leave the return disconnected and use the ground (pin 1 XLR, sleeve TRS, or ⏚ . If there's no signal, short return to ground.

This doesn't adapt the voltages. To lower a +4 dBu professional output (1.228v) for a consumer -10 dBV (.316v) input, use this circuit:



AES/EBU to s/pdif: You can connect a pro digital output to a consumer digital input with this circuit:



Gotchas:

- For best results, extend the XLR side of the circuit as needed; keep the attenuator close to the consumer gear.
- Pro output to consumer input is easy. Consumer output to pro input can introduce noise or serious digital problems: use a powered adapter.

Monitoring

You can't make accurate mix or processing decisions unless you're sure of what you're hearing. This always requires good monitors, and your room may also need acoustic treatment. **Monitor speakers are the most important audio post equipment you can buy.**

"Real-world" small speakers aren't good for mixing: they rarely match any other small speakers, and can fool you into making music too hot and missing important problems. But you can use them to check a mix.

You can mix on small speakers if you're positive the audience will be using the exact same small speakers (in a kiosk or similar situation).

Manufacturers of <$500 **powered speakers** often cheat, matching poor components with over-equalized amplifiers. This hypes their response for music listening and game playing but lowers their accuracy and adds distortion. If you're buying mid-price speakers, stick to to unpowered ones with a separate amp; you've got a better chance for honest sound.

Headphones aren't good for mixing. The intensity of the experience often leads to music and sfx being too soft.

Headphones and small, low-cost speakers can be used for editing, if you're not making any mix or processing decisions.

How to buy a loudspeaker

Don't buy from websites or catalogs unless you know how to read specs, and enough to ignore nonsense like "frequency range", distortion ratings without a level in dB SPL, or specs for a powered system that measure only the amplifiers—the speakers are always going to be worse.

Best way: audition at a dealer that lets you switch between two different models at the same acoustic volume. Bring a CD of some DVD or VHS tracks you consider well-mixed, plus some of your own work. Choose selections with dialog and music; it'll tell you more than just music or effects sequences.

Start with the most expensive pair the dealer has, then work your way down until you don't like the sound any more. Go back up one notch and you've found your winners. Before you faint at the expense, remember: you heard the differences yourself... and monitor speakers are an investment that will last your entire career.

If a dealer won't let you audition pairs at the same volume or use your own CD, they're probably trying to cheat you.

Next best way: trust a recommendation from someone experienced, doing the same kind of projects you are (speakers that are good for music mixing aren't necessarily right for soundtracks). Or ask for recommendations on the Community forums at DV.com.

A monitoring must

Keep a constant volume while mixing. If you're mixing for theatrical release, use a sound level meter and set things for 85 dB SPL at the listening postion while playing pink noise at -20 dBFS through a single channel.

Improving a speaker's sound

- Get the speakers at least 3x closer to you than they are to the nearest wall or other reflective surface, with a clear line-of-sight path to your ears.
- If your projects have heavy bass, mount the speakers rigidly.
- Add some absorption to the room. You can use sculpted foam tiles for this, but the cheapest solution is panels of 2" thick Owens-Corning #703 sound absorbing fiberglass, available from acoustic ceiling suppliers.

NLE issues

Choosing a sample rate

If you're capturing from a camera over FireWire, leave the NLE's project and clip settings at 48 kHz, 16 bit. This can prevent sync drift because of sample-rate conversion. (If you shot at 32 kHz, 12 bit—and you'd better have a darned good reason for *that*—the conversion to 48 kHz is automatic and benign.)

If the project is going to be output via FireWire to a digital tape, keep the project at 48 kHz for best quality.

If the captures and output are analog, and a lot of your material is coming from audio CDs in the NLE's CD-ROM drive, keep the project and captures at 44.1 kHz for best quality.

If you're mixing FireWire camera audio with ripped CDs, it can help to convert the CD material to 48 kHz in an audio program before importing it to NLE.

Troubleshooting captures

- Noise and distortion in an analog capture are usually because of improper levels. Listen to the playback deck's or camera's headphone jack: if that sounds good, recalibrate and try again. If it sounds bad, try a different machine... though you may be stuck with badly-recorded material.
- Changes in timbre—too thin, bright, or dull—can be a configuration issues: improperly terminated equipment, bad filters in a sound card, or sample rate incompatibilities.
- Rhythmic ticks are almost always digital audio sync issues. Make sure sample rates agree in the playback deck, any digital capture hardware, and the NLE. Make sure the capture hardware is set to the right reference (usually `INPUT`). If all else fails, transfer through an analog connection.
- Sporadic ticks or splats that always happen at the same place can be level issues. If they're happening in an analog transfer, try FireWire. But they may be on the original tape, because of bad recording or tape damage. If the latter, try a different deck.
- Totally random clicks are usually system issues. Try a different deck or soundcard, look for fragmented memory (reboot) or drives (use a defragment utility), make sure no other processes or network accesses are happening during capture, make sure SCSI chains are properly terminated.

Troubleshooting Lipsync

Longterm drift, where things get progressively worse, are usually system problems.

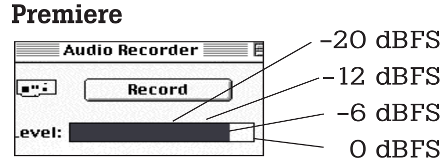
- Check dropframe/non-drop conflicts.
- Is there a break in timecode on the camera tape?
- Are your FireWire drivers the latest version?
- Some cameras don't sample at precisely 48 kHz, and can drift over the long term. Look for a compensation setting in your NLE's capture settings, or transfer via analog.
- If drift happens only when outputting the final project, and you can't find any other cure, break the project into smaller sections. Output each separately, or save each as a QuickTime or AVI and assemble as a separate project.
- Some third-party cards have issues. Check with the manufacturer.

If all else fails, measure the drift over the show's length and apply a speed correction to compensate.

Erratic sync problems can be systemic also. First ask yourself if this combination of camera, capture or FireWire card, software, and tower has ever had perfect sync; if not, prepare to spend a lot of time on the phone with customer support (and if there are multiple manufacturers involved, prepare for finger pointing.) Otherwise:

- Try the suggestions under "long term drift".
- Look for dropped *video* frames.
- Make sure audio and video are going to the same output device.
- Don't stress your system with too deep a color setting. 32-bit color is overkill for digital video.

How to read the record level meter in Premiere



Standard digital levels:

-20 dBFS: Believe it or not, this is the standard "zero tone" for network programs, and the nominal level for professional cameras. Most NLE clip windows are virtually unreadable at this level.

-12 dBFS: Nominal level / zero for miniDV. It's necessary because of the cameras' poorer noise performance, but sacrifices safety margin.

-10 dBFS: Network standard peak level. If you're sending a digital master to a dub house for VHS or other analog dubs, don't let your mix get any louder than this.

-6 dBFS: Halfway point of most clip windows.

-2 dBFS: Practical absolute maximum for miniDV audio

Why are some NLEs calibrated so that -20 dBFS looks so wimpy?

Probably because NLE programmers don't understand that audio is logarithmic. But it works in your favor: While the professional standard assumes a -96 dBFS noise floor, most sound cards, prosumer cameras, and factory-supplied NLE filters are much worse. So it makes sense to work louder to overcome internal noise.

Work louder, and knock the mix down to -20/-10 standards if you're feeding a pro digital recorder. If you're feeding an analog deck, don't worry about it.

Editing tips

Voice

1) Listen for phonemes, not words or phrases. That's what you'll be editing.
2) Don't try to cut from one continuous sound to another.
3) You can usually cut from a continuous soft sound into a hard consonant.
4) You can almost always cut from the start of a phoneme in one syllable to an identical or cognate phoneme elsewhere.
5) You can frequently replace an unvoiced consonant with the same consonant from a different person — even someone with a very different voice!
6) Try not to edit where it's expected.
7) A "soft cut" doesn't work in dialog.
8) Intonation can be tweaked with varispeed — but only over about a 3% range (slightly more for children).
9) Most breaths can be replaced with a shorter pause or roomtone, about 60% as long as the original breath.
10) Zap double stops and glottal shocks! They only make your talent sound amateurish.

English Phonemes

There's an official way to write these—the International Phonetic Alphabet—but writing them isn't as important as learning to hear them in dialog, particularly when scrubbing.

Stop Plosives as cognate pairs; you can always edit in the tiny silence that accompanies them.

UNVOICED	*VOICED*
p (paint)	*b (barbell)*
k (cat)	*g (gut)*
t (tamarind)	*d (delight)*

Fricatives as cognate pairs; listen for the "sss" sound while jogging.

f (fellow)	*v (very)*
s (silly)	*z (zebra)*
sh (shoe)	*zh (leisure)*
th (thin)	*TH (then)*
h (horse)	*(no cognate)*

Nasals; must be matched

m (mighty) *n (nap)* *ng (lung)*

Vowels; they're all different.

ay *(never by itself)*	ee *(eat)*	ih *(sit)*	
i *(never by itself)*	ae *(hat)*	eh *(lend)*	
aw *(all)*	o *(note)*	u *(bull)*	oo *(tool)*
uh *(up)*	er *(worker)*	ah *(father)*	

Glides

w *(willing)* y *(yes)* l *(locate)* r *(rub)*

Diphhongs; double-dipping phonemes

t-sh *(church)* d-zh *(judge)* ay-ih *(play)*
i-ih *(high)* aw-ih *(toy)*

Some Music-cutting Tips:

1) Learn to trust your ears, not your screen!
2) Feel the beat and respect the pattern.
3) Edits with an even number of bars usually sound best. If it doesn't sound good, try an odd number.
4) Realize that melodies often start off the downbeat. You still have to mark bars, but before you edit, roll the points to catch the melody.
5) If a barline is almost in sync with a picture cue, try shifting the entire clip a few frames to match. You may find that the previous edits still work against pix.
6) Don't be afraid to varispeed slightly — no more than 3% for acoustic instruments, but as much as 10% for synths. Of course, you'll have to do this to the entire song.
7) If an edit *almost* works, try spreading it to two tracks and cross-fading. You can also crossfade on sustained notes to make them shorter or longer.

When you think you've found the right music, try muting it while playing the other tracks and pictures. If you don't immediately feel, in your gut, that something's missing… then it's the wrong piece of music. Keep looking.

Freq out

How to tune a parametric equalizer to a specific sound:

1. Set the bandwidth as narrow as possible (or Q as high as possible).
2. Set the boost to maximum… even if you're tuning to get rid of a sound.
3. With audio playing, sweep around the frequency range where you think the sound is. When you hit the exact frequency, you'll hear the sound suddenly jump up or even go into distortion.
4. Now that you've found the frequency, adjust the boost/dip control. If you're trying to get rid of a noise, leave the bandwidth narrow and set the dip to maximum. If you want to gently emphasize a sound, set the bandwidth fairly wide and boost about 3 dB.
5. Boosting an otherwise clean track more than 6 dB is almost always a mistake.

"My NLE doesn't let me sweep frequencies while sound is playing!"

Sorry. That step is a must: even experienced engineers do it. Invest in some plug-ins, or move the clip to an audio program. Or try this workaround:

Do the first two steps above. Then set keyframes about ten seconds apart, and about an octave (2:1 frequency ratio) different around the frequency you suspect. Preview the effect, noting where the frequency adjustment is when you hear the peak.

Important frequencies for dialog:

< 80 Hz: Usually just mud. Most dialog tracks are improved by cutting here.

~150 - 300 Hz: A gentle boost here can add warmth. However, this is also where most practical rooms get boomy, so a sharp dip—tuned with the technique above—can help.

~1.75 kHz: Fundamental of most consonants. A 2 dB boost can improve intelligibility.

~3.5 kHz: First harmonic of consonants. Boosting here can add presence, if you're not boosting the fundamental.

~4 - 5 kHz: Sibilance.

> 10 kHz: Very little useful dialog above here.

Things to remember about equalizers:

1. They go down as well as up. Boosting ~1.75 kHz to make dialog stronger over music isn't as good as dipping the same frequencies in the music track.

2. Despite the pretty pictures in your NLE, 10 kHz *isn't* the middle of the audio band. 1 kHz is the middle.

3. Most factory-supplied NLE equalizers aren't very precise and can affect sounds an octave away. Third-party plugins are usually much better.

Noise reduction

Good noise reduction software breaks the audio band into tiny slices. Each slice is turned up when signal exceeds a certain level, and down the rest of the time. Calibrate the bands by selecting noise with no dialog, then process the entire clip. Where voices are louder than noise at specific frequencies, they'll mask the noise. Where they're softer, noise (and voice) is turned off.

- The less noise in a track, the better NR software will sound.
- The more echo in a track, the harder it is to remove noise. That's why close miking and acoustic control on the set are important.

Too much NR sounds artificial. Let some noise back in. It's better to do this by changing the NR's settings, rather than by adding a separate background track.

As adjacent bands open and close, the differences can cause squeaks or chirps. Good NR progams let you link bands together, as well as control attack and decay time, to reduce this.

My favorite NR tool this week is Arboretum's Ionizer (www.arboretum.com). But there are plenty of others, from the free cross-platform Audacity (audacity.sourceforge.net) all the way up to Cedar's $3,200 Retouch (www.sadie.com), which can do things that are impossible in any other program. Check the feature on NR, coming in this September's DV Magazine.

Before using NR software, fine-tune a parametric equalizer to notch out any whines, whistles, or hums. Use a comb filter to get rid of power-line or dimmer buzz.

If you don't have NR software, try your NLE's *downward expander* audio filter.

Edit by the numbers

1 Edit the dialog first. It tells the story.

2 Edit the narration second.

3 Put in just the large, important sync effects. Ignore the others for now.

4 Add music.

5 Add background ambiences where needed. Often, music makes them unnecessary.

6 Add smaller sync effects only if needed. Often, random sounds in the ambience track can take their place.

Be a Mix Master

Don't try to mix while you're editing. It may be tempting to adjust rubber band levels while you're in each sequence, but mixing in bits and pieces seldom yields a smooth overall track. In fact, if you've just finished editing, don't mix at all. Take a long break (preferrably at least a few hours), so you can approach the track with a fresh perspective.

- Take a few minutes, before starting, to think about the elements and which should predominate. It's usually dialog, but some sequences might be driven by music or sound effects.
- Think about the medium as well. A TV mix can't have more than a few things going on at once (dialog, music, and an effect or two), and must be mono-compatible. A theatrical mix can be richer. A trade-show video can't be complex at all: it has to compete for the viewer's attention in a noisy environment.
- Mix from the top down. Get the principal dialog sounding perfect first, including any equalization or compression. Then adjust any ADR to match. Leave the voices up while you adjust any hard sfx, then the music, then background sound effects.
- Mix on good speakers, and check surround in stereo, stereo in mono. If you don't have good speakers, it may be more economical to take just the mix to a video sound professional... who'll also be able to spot and fix problems that could haunt you later.

Resources

Digital Video Magazine www.dv.com Information resources, hard-hitting reviews, active video-community forums on a variety of subjects, and back-issues of the magazine (including my monthly audio tutorials).

Digital Playroom www.dplay.com My website, including an index to my DV Magazine articles and other tutorials (many online), sound files and extras for DV readers, streaming audio and video of projects I've done, some humorous takes on our industry, plus other stuff of interest to filmmakers and broadcasters. **Need stock music?** The tutorials section of my site has actual reviews (not PR re-hashes) of dozens of libraries, with contact information.

Location Sound Corp. www.locationsound.com Where Hollywood audio pros shop. Lots of good info.

videouniversity.com and **videoemporium.com** Both good, but some articles are dated.

Equipment Emporium www.equipmentemporium.com Rental/dealer's website, with lots of useful articles on audio.

MC Squared System Design www.mcsquared.com Audio design company; their website features Javascript calculators for decibel loss, attenuator values, and other useful stuff.

Rane Corporation www.rane.com Equipment manufacturer's site features the Pro Audio Reference, a dictionary of everything from AAC to Z-weighting—with links to more source material for most entries.

Tomi Engdahl www.hut.fi/Misc/Electronics/audio.html Tons of indexed links to articles on every technical aspect of audio.

Sounddogs www.sounddogs.com Immense library of downloadable, fully licensed, high-quality sound effects. You have to pay for the sounds, but only a few dollars for all but the longest ones.

Hollywood Edge, 7080 Hollywood Blvd, Hollywood, CA 90029 USA, 322/292-3755. Their demo CD/CD-ROM also includes about a hundred clean sound effects, fully licensed for you to use in productions. It's free for potential customers; others can download the effects from www.hollywoodedge.com.

 <jrose@DV.com>. I'm a Clio- and Emmy-winning sound designer, equipment design consultant, and writer for *Digital Video* Magazine. My new books are *Audio Postproduction for Digital Video* and *Producing Great Sound for Digital Video (2nd Edition),* both of which are category best-sellers at Amazon. Details, downloadable excerpts, critical and reader comments, tables of contents, and links to the web's best discount sales are at www.dplay.com/book.