Figure 6.10 The people involved with audio setup have to be concerned with not only with sound acquisition, but also with the distribution of the communications gear necessary for the production. Here a reporter's package of microphone and intercom/IFB is placed near a basketball court where it will be used during the show.

Audio Engineering

The Effort and Scope

The A1 and all the A2s on many shoots can have their hands full, depending not only on the size, but also the type of shoot. On an entertainment remote, the placement and number of microphones can be a large effort, and having more than 60 audio sources on a sports shoot is not unheard of. The audio folks are also usually concerned with intercom and IFB requirements, and these communication needs can form a large part of their setup time.

Some high-profile sporting events today can combine high numbers of audio sources with high communications needs. For instance, at Super Bowl XXXVI, over two dozen mics were used just to capture the sounds from the field. Many A1s and their assistants earn their pay by placing mics so the audio coverage equals the video coverage. Coverage today doesn't mean just capturing the sounds, but building a sound image for the viewer. Many viewers now have more sophisticated television listening environments than simply the one or two speakers that are inside their television receivers. Just as virtually nobody watches the video in black and white today, the television audio experience is catching up with the additional "color" that multichannel audio brings. Now the audio effort must result in not just all the sounds, but their placement in space for the viewer. There must be an audio image that goes with the video image.

Figure 6.11 Today microphones are place so as to capture the thumping of the ball, the squeak of shoes on the court, the swoosh of a ball through a basket, the ping of a putter contacting a golf ball, and the plunk of it being sunk.


PLAY BY PLAY
COLOR

Figure 6.12 *Left: The construction and support of the announce booth is tasked to a number of folks, not least the A2s setting up the audio and communications infrastructure throughout the venue. Besides the obvious need for talent mics, the producers and directors must be able to communicate back to the talent, and the talent must be able to not only see what is on air, but also hear it. The box shown center inset allows the announcer to set on-air monitoring, along with IFB interrupt volumes. The white button allows him to take his mic off air and talk to the director over the normal intercom, the red button just takes the announcer off air for coughing or clearing of the throat.*

Even with a monaural feed, poor mic placement could result in poor pickup from desired directions and at some frequencies. With stereo the problems are compounded. Now not only do the video people have to worry about the phase of their signals, so does the audio engineer. An audible single-frequency pressure wave passing by two microphones one-half wavelength apart would produce no audio when combined into a mono signal. In remote television production we talk about giving the viewer the best seat in the house. Often the best physical seat in the house might be a very poor seat from a sound standpoint. Conversely the acoustics of a venue might allow most seats to receive warm, full sound, with the right amount of reverberation at any single location. But collecting and mixing sound from dispersed locations might make the best concert hall sound like a phone booth. Back when it was just a mono audio signal feeding a single three-inch speaker on a television receiver, most viewers filled in mentally what the audio effort left out. Today with our sound expectations catching up with our visual expectations, you probably wouldn't win any Emmys if you created a sound image that made the Cleveland Orchestra sound like its brass section didn't show up, or one that puts the flute section at the back of the hall.

Figure 6.14 *Each multitrack (multichannel) recorder, on the left, records eight tracks of audio, allowing the whole case of recorders to handle 48 tracks.*

Figure 6.13 *A good audio engineer will mic a venue based on its characteristics. Mic placement in a "lively" hall will be different than one with little reverberation. Whereas a hall might give any particular seat a good "listen," badly placed mics might make for a very bad acoustical rendition and a disorienting audio image, especially in a surround sound broadcast.*

A good audio engineer starts by knowing a good microphone for the job at hand. They select mics that have the needed headroom and frequency response. Sound pressures can get as high as 120dB when a basketball

backboard is hit. High crowd noise can reach nearly as high. A good mike should be able to handle up to 130dB before clipping (points of reference = 0dB is the threshold of hearing, 60dB a normal conversation at 1m, 110dB some rock groups, 130dB a jet engine at 50m).

Microphone placement depends greatly on the event being covered. As an example, the mic placement for basketball might be:

• One mounted on each of the backboards to capture the "swish" when the ball goes through the net.

• A couple on the scorers' table to catch the bench, and a third at the center of the table to catch the transition from one court to the other, along with some referee calls.

• A stereo crowd mic

As compared to football:

• Mics on each side of the field at both 20-yard lines and the 50-yard line

• Four portable parabola mic operators

• Umpire mic to catch backfield and quarterback audible calls

• A stereo crowd mic

A football field is about four times the size of a basketball court, so a mic set-and-forget strategy is not possible. Field coverage has to move as the game moves up and down the field. Thus, a number of operators manning parabolic microphones are used. Many football games are also open to the elements; because of this RF interference is often greater as well.

Microphone manufacturers have helped the stereo effort, placing two mikes in the same package. When placed on a camera and on-air, they can produce stereo effects coincident with the observed video. If stereo were from two mics not colocated with the camera, the sound image could be dramatically wrong.

Parabolas and shotgun mics are tools usually only found on sport shoots where you can't place a mic in the middle of the action. Entertainment usually calls for a different scheme of mic placement. Often, as in a concert, be it classical or rock music, the audio becomes a full equal player with regards to its importance. If the show is being taped, a separate sound truck might be called in to record the audio separately, often on as many as 48 tracks. Many, if not all, the individual audio sources will be laid down on a separate audio track. They will be mixed later, although usually a working mix is done during the recording of the show as a reference. This kind of production usually calls for different mic placement, one that puts the mics close to the instrument or person

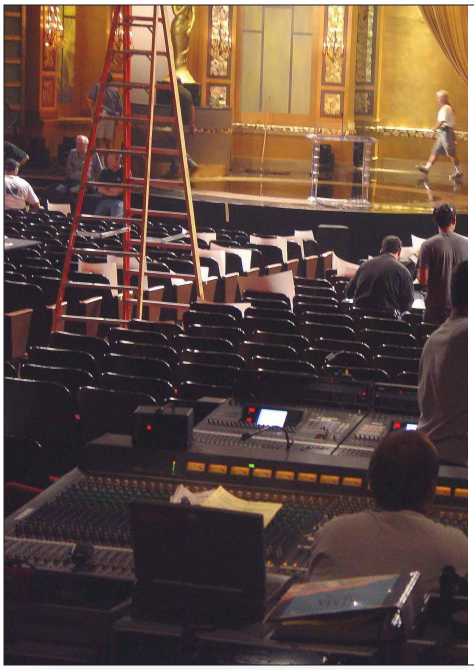

Figure 6.15 *All mixes are different, even ones done at the same venue at the same time. For instance, the audio people creating the mix for the PA feed in the house (shown above) will mix the show differently than the audio folks outside in the television production trucks, even though they are using many of the same sources.*

being recorded. This might mean that in the post-production mix processing, reverberation will have to be added to give the warmth back to the overall mix.

The Mix

Like almost everything else, the audio component of the show has gotten larger. Thirty years ago the output from the truck was a single monologue feed. Twenty years ago that feed generally was stereo; now it is moving to surround sound and even 5.1 Surround Sound.

Generally the A1 will set up a number of sub-mixes for the show. Instead of treating every source individually, they will group types of sources, also known as "stems" in the film world, to be sub-mixed. Stems can be announcers, effects, replays, etc. These sub-mixes then are finally mixed into complete output channels, of which there can be many. Mono, left, right, center, rear, natural sound, and IFB could all be separate audio output channels.

Besides all the separate mixes that could be required today, much processing usually needs to be done. It is not unusual for a limiter/compressor to be installed downstream from the mixer. It is usually set to start to compress audio peaks above "0" VU (+4dB). The higher the levels above the set value, the more it compresses. Only massively large peaks will end up being clipped. This allows the mixes to sound alive and dynamic but not be distorted by clipping. Often a sub-group, such as natural sound or crowd noise, is limited and brought back into the audio board so background noise can't inadvertently drown out the announcers.

Dolby Mixing and Encoding

For NFL done in Dolby Surround Sound, a surround sound image has to be built for the viewer. The field sounds are in front and the crowd noise is in back. Actually there are four audio channels that need to be created: left, center, right and rear surround. These are encoded in a Dolby four-channel matrix encoder, which produces two signals: left-total and right-total. The left-total has in- and out-of-phase audio. The in-phase is left and center audio, and the out-of-phase is the rear. The right in-phase is right and center audio while out-of-phase audio is surround. The rear channel is delayed slightly to mask front channel leakage, based on the fact that if two similar sounds arrive at slightly different times, the listener will not perceive the later sound. To stop rear surround sounds from leaking into the front channels the surround audio is run through a low-pass filter. These two channels can be decoded by any Dolby Pro-Logic decoder back into the four original channels.

Figure 6.16 *The A1 performing the mix has tools other than the ears. The display on the left is a spectral graph of an audio signal. It quantizes exactly how much bass and treble, along with the mid frequencies, a source has. The A1 can adjust how the audio mixer processes that signal. Many things can cause "roll-off" of low or high frequencies besides microphones, cabling or other electronic processing. Mic placement, combined with venue acoustics, can conspire to limit bandwidth. The display on the right is a graphical display of stereo, or multichannel separation. The object of multiple audio channels is to provide a 2-D (in the case of stereo) or a 3-D (in the case of additional channels) audio image. A mono audio feed would produce a single vertical line in the center of the display. Here audio channel separation is robust, providing a stereo image and not simply a single point of audio.*

Anyone listening just in stereo or mono would hear normal stereo or mono audio.

Dolby Surround encodes four-channel analog into two-channel analog. Dolby has introduced Dolby E (for editable, because compression is light enough that the signal can be decoded and encoded a couple of times), which uses AES digital audio streams to pack six channels of audio into a single bit stream. These six channels are actually known as 5.1 because the sixth channel actually provides only low-frequency "rumble" for effects. Also, unlike the Surround system, the rear channels are now stereo and full-bandwidth audio, whereas the Surround rear channel was mono and of limited bandwidth.

A lot must be done before the audio gets to the encoders. Often many sub-mixes must be done first, and it often takes the coordination of multiple people. In the case the multiple parabolic mics used for NFL football, an A2 often sub-mixes these from the announce booth, as he can see where the action is. This mono mix is then sent down to the truck where it is turned back into pseudo-stereo and becomes part of the main mix. Some audio mixing consoles today have joystick or track-ball controls so that the A1 can quickly move left, right, and rear mixes quickly in listener space to capture the sense of listener perception location changes. In essence, left, center, right, rear (left and right in the case of Dolby E) mixes can be made to shift clockwise or counter-clockwise from one channel to the next adjacent to give the illusion of the listener turning their head in different directions.

But often the nature of the event will require that

besides the fully mixed four channels of surround sound, additional instances of background or effects audio must be created. Often a show will be live-to-air or tape (show is recorded but shot as if it is live-to-air), but will need to be edited later for reuse, where the announcer's commentary is not desired. In the case of stereo, or four-channel surround sound on two analog channels, two additional channels would have two additional channels for the "clean audio." But it turns out that it isn't as simple as that. If something has to be "fixed" afterward — for instance, the announcer's track is to be replaced with only natural sound, or if a new voice-over by the announcer was desired — there would be a noticeable change in the sound, as there is some natural sound leakage into the announcers' track. Switching away to just natural sound or to new announcers' tracks without this leakage is usually evident. To mask this, an extra announcers' headset is sometimes hung somewhere in the announce booth to pick up additional natural sound heard in the booth. This is then added to the natural sound tracks.

Digital technology is helping in handling all the extra audio channels required. Serial digital video can hold eight AES programs, with two channels of audio each. Thus 16 audio feeds can be embedded into the video. NBC at the Salt Lake Winter Games sent eight audio channels back from the venues embedded in the video. There are VTRs now that record eight channels of audio. The necessity of producing many audio channels in the field will only increase.

The audio A1 is almost always in a compartment by himself, which is designed to allow good audio rendition, and to allow the audio person some space, as the upper-end boards are so large these days. The folks that provide for the video half of the mosaic, along with the director and producer, are in the production compartment, which we will tour next. ■

Figure 6.17 *Top left is an audio splitter box, where a single source such as a mic can be feed to more than one location; for example, the in-house PA mix and the television truck mix. The cables coming out the bottom of the splitter are audio "mults," single cables with lots of pairs of audio inside, going either to the venue PA feed, the television production truck or, as is often the case, to news media also covering the event. The upper right photo is the production truck's audio ECP panel where mults from inside the venue are connected to it. Bottom left are audio "breakout" boxes, where multiple audio mics or other sources can be combined into an audio mult. Even wireless microphones, as seen bottom right, often end up in mults, once out of the wireless receivers (case sitting on the table next to the mics).*