

DENSE CLARITY - CLEAR DENSITY

by *Walter Murch*

Simple and Complex

One of the deepest impressions on someone who happens to wander into a film mixing studio is that there is no necessary connection between ends and means. Sometimes, to create the natural simplicity of an ordinary scene between two people, dozens and dozens of soundtracks have to be created and seamlessly blended into one. At other times an apparently complex ‘action’ soundtrack can be conveyed with just a few carefully selected elements. In other words, it is not always obvious what it took to get the final result: it can be simple to be complex, and complicated to be simple.



Murch editing *Cold Mountain* at Cinelabs in Bucharest, Romania, 2002.

The general level of complexity, though, has been steadily increasing over the eight decades since film sound was invented. And starting with Dolby Stereo in the 1970's, continuing with computerized mixing in the 1980's and various digital formats in the 1990's, that increase has accelerated even further. Seventy years ago, for instance, it would not be unusual *for an entire film* to need only fifteen to twenty sound effects. Today that number could be hundreds to thousands of times greater.

Well, the film business is not unique: compare the single-take, single-track 78rpm discs of the 1930's to the multiple-take, multi-track surround-sound CDs of today. Or look at what has happened with visual effects: compare King Kong of the 1930's to the Jurassic dinosaurs of the 1990's. The general level of detail, fidelity, and what might be called the “hormonal level” of sound and image has been vastly increased, but at the price of much greater complexity in preparation.

The consequence of this, for sound, is that during the final recording of almost every film there are moments when the balance of dialogue, music, and sound effects will suddenly (and sometimes unpredictably) turn into a logjam so extreme that even the most experienced of directors, editors, and mixers can be overwhelmed by the choices they have to make

So what I'd like to focus on are these ‘logjam’ moments: how they come about, and how to deal with them when they do. How to choose which sounds should predominate when they can't all be included? Which sounds should play second fiddle? And which

sounds - if any - should be eliminated? As difficult as these questions are, and as vulnerable as such choices are to the politics of the filmmaking process, I'd like to suggest some conceptual and practical guidelines for threading your way through, and perhaps even disentangling these logjams.

Or - better yet - not permitting them to occur in the first place.

Code and Body

To begin to get a handle on this, I'd like you to think about sound in terms of light.

White light, for instance, which looks so simple, is in fact a tangled superimposure of every wavelength (that is to say, every color) of light simultaneously. You can observe this in reverse when you shine a flashlight through a prism and see the white beam fan out into the familiar rainbow of colors from violet (the shortest wavelength of visible light) - through indigo, blue, green, yellow, and orange - to red (the longest wavelength).

Keeping this in mind, I'd now like you to imagine *white sound* - every imaginable sound heard together at the same time: the sound of New York City, for instance - cries and whispers, sirens and shrieks, motors, subways, jackhammers, street music, Grand Opera and Shea Stadium. Now imagine that you could 'shine' this white sound through some kind of magic prism that would reveal to us its hidden spectrum.

Just as the spectrum of colors is bracketed by violet and red, this sound-spectrum will have its own brackets, or limits. Usually, in this kind of discussion, we would now start talking about the lowest audible (20 cycles) and highest audible (20,000 cycles) frequencies of sound. But for the purposes of our discussion I am going to ask you to imagine limits of a completely different conceptual order - something I'll call *Encoded* sound, which I'll put over here on the left (where we had violet); and something else I'll call *Embodied* sound, which I'll put over on the right (red).

The clearest example of Encoded sound is speech.

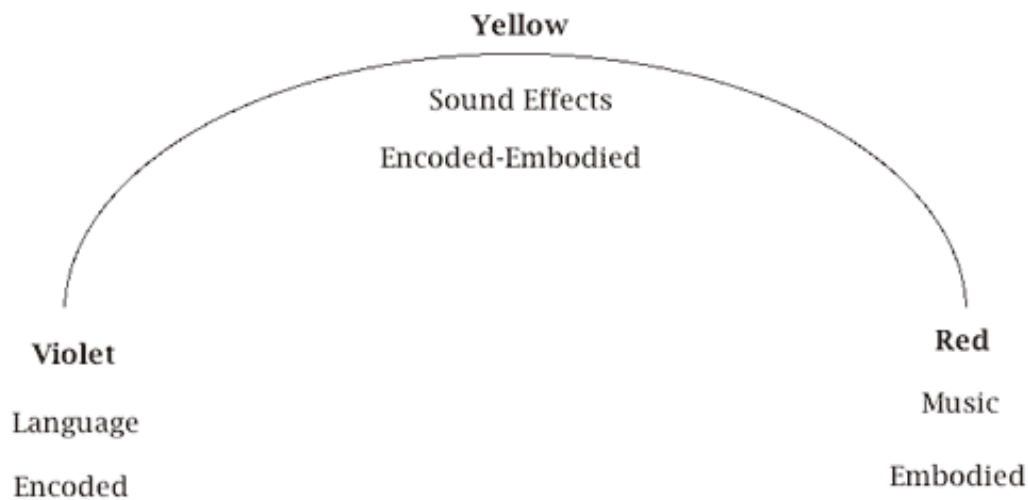
The clearest example of Embodied sound is music.

When you think about it, every language is basically a code, with its own particular set of rules. You have to understand those rules in order to break open the husk of language and extract whatever meaning is inside. Just because we usually do this automatically, without realizing it, doesn't mean it isn't happening. It happens every time someone speaks to you: the meaning of what they are saying is encoded in the words they use. Sound, in this case, is acting simply as a vehicle with which to deliver the code.

Music, however, is completely different: it is sound *experienced directly*, without any code intervening between you and it. Naked. Whatever meaning there is in a piece of music is 'embodied' in the sound itself. This is why music is sometimes called the Universal Language.

What lies between these outer limits? Just as every audible sound falls somewhere between the lower and upper limits of 20 and 20,000 cycles, so all sounds will be found somewhere on this conceptual spectrum from speech to music.

Most sound effects, for instance, fall mid-way: like 'sound-centaurs,' they are half language, half music. Since a sound effect usually refers to something specific - the steam engine of a train, the knocking at a door, the chirping of birds, the firing of a gun - it is not as 'pure' a sound as music. But on the other hand, the language of sound effects, if I may call it that, is more universally and immediately understood than any spoken language.



Green and Orange

But now I'm going to throw you a curve (you expected this, I'm sure) and say that in practice things are not quite as simple as I have just made them out to be. There are musical elements that make their way into almost all speech - think of *how* someone says something as a kind of music. For instance, you can usually tell if someone is angry or happy, even if you don't understand what they are saying, just by listening to the tone (the music) of their voice. We understand R2-D2 entirely through the music of his beeps and boops, not from his 'words' (only C-3PO and Luke Skywalker can do that). Stephen Hawking's computerized speech, on the other hand, is perfectly understandable, but monotonous - it has very little musical content - and so we have to listen carefully to what he says, not how he says it.

To the degree that speech has music in it, its 'color' will drift toward the warmer (musical) end of the spectrum. In this regard, R2D2 is warmer than Stephen Hawking, and Mr. Spock is cooler than Rambo.

By the same token, there are elements of code that underlie every piece of music. Just think of the difficulty of listening to Chinese Opera (unless you are Chinese!). If it seems strange to you, it is because you do not understand its code, its underlying assumptions. In fact, much of your taste in music is dependent on how many musical languages you have become familiar with, and how difficult those languages are. Rock and Roll has a simple underlying code (and a huge audience); modern European classical music has a complicated underlying code (and a smaller audience).

To the extent that this underlying code is an important element in the music, the 'color' of the music will drift toward the cooler (linguistic) end of the spectrum. Schoenberg is cooler than Santana.

And sound effects can mercurially slip away from their home base of yellow towards either edge, tinting themselves warmer and more 'musical,' or cooler and more 'linguistic' in the process. Sometimes a sound effect can be almost pure music. It doesn't declare itself openly *as music* because it is not melodic, but it can have a musical effect on you anyway: think of the dense ("orange") background sounds in *Eraserhead*. And sometimes a sound effect can deliver discrete packets of meaning that are almost like words. A door-knock, for instance, might be a "blue" micro-language that says: "Someone's here!" And certain kinds of footsteps say simply: "Step! Step! Step!"

Such distinctions have a basic function in helping you to classify - conceptually - the sounds for your film. Just as a well-balanced painting will have an interesting and proportioned spread of colors from complementary parts of the spectrum, so the sound-

track of a film will appear balanced and interesting if it is made up of a well-proportioned spread of elements from our spectrum of ‘sound-colors.’ I would like to emphasize, however, that these colors are completely independent of any emotional tone associated with “warmth” or “coolness.” Although I have put music at the red (warm) end of the spectrum, a piece of music can be emotionally cool, just as easily as a line of dialogue - at the cool end of the spectrum - can be emotionally hot.

In addition, there is a practical consideration to all this when it comes to the final mix: It seems that the combination of certain sounds will take on a correspondingly different character depending on which part of the spectrum they come from - some sounds will superimpose transparently and effectively, whereas others will tend to interfere destructively with each other and ‘block up,’ creating a muddy and unintelligible mix.

Before we get into the specifics of this, though, let me say a few words about the differences of superimposing images and sounds.

Harmonic and Non-Harmonic

When you look at a painting or a photograph, or the view outside your window, you see distinct areas of color - a yellow dress on a washing line, for instance, outlined against a blue sky. The dress and the sky occupy separate areas of the image. If they didn’t - if the foreground dress was transparent, the wavelengths of yellow and blue would *add together* and create a new color - green, in this case. This is just the nature of the way we perceive light.

You can superimpose *sounds*, though, and they still retain their original identity. The notes C, E, and G create something new: a harmonic C-major chord. But if you listen carefully you can still hear the original notes. It is as if, looking at something green, you still also could see the blue and the yellow that went into making it.

And it is a good thing that it works this way, because a film’s soundtrack (as well as music itself) is utterly dependent on the ability of different sounds (‘notes’) to superimpose transparently upon each other, creating new ‘chords,’ without themselves being transformed into something totally different.

Are there limits to how much superimposure can be achieved?

Well, it depends on what we mean by superimposure. Every note played by every instrument is actually a superimposure of a series of tones. A cello playing “A”, for instance, will vibrate strongly at that string’s fundamental frequency, say 110 per second. But the string also vibrates at exact multiples of that fundamental: 220, 330, 440, 550, 660, 770, 880, etc. These extra vibrations are called the *harmonic overtones*

of the fundamental frequency.

Harmonics, as the name indicates, are sounds whose wave forms are tightly linked - literally 'nested' together. In the example above, 220, 440, and 880 are all higher octaves of the fundamental note "A" (110). And the other harmonics - 330, 550, 660, and 770 - correspond to the notes E, Db, E, and G which, along with A, are the four notes of the A-major chord (A-Db-E-G-A). So when the note A is played on the violin (or piano, or any other instrument) what you actually hear is a chord. But because the harmonic linkage is so tight, and because the fundamental (110 in this case) is almost twice as loud as all of its overtones put together, we perceive the "A" as a single note, albeit a note with 'character.' This character - *or timbre* - is slightly different for each instrument, and that difference is what allows us to distinguish not only between types of instrument - clarinets from violins, for example - but also sometimes between individual instruments of the same type - a Stradivarius violin from a Guarneri.

This kind of *harmonic superimposure* has no practical limits to speak of. As long as the sounds are harmonically linked, you can superimpose as many elements as you want. Imagine an orchestra, with all the instruments playing octaves of the same note. Add an organ, playing more octaves. Then a chorus of 200, singing still more octaves. We are superimposing hundreds and hundreds of individual instruments and voices, but it will all still sound unified. If everyone started playing and singing whatever they felt like, however, that unity would immediately turn into chaos.

To give an example of non-musical harmonic superimposure: in *Apocalypse Now* we wanted to create the sound of a field of crickets for one of the beginning scenes (Willard alone in his hotel room at night), but for story reasons we wanted the crickets to have a hallucinatory degree of precision and focus. So rather than going out and simply recording a field of crickets, we decided to build the sound up layer by layer out of individually recorded crickets. We brought a few of them into our basement studio, recorded them one by one on a multitrack machine, and then kept adding track after track, recombining these tracks and then recording even more until we had finally many thousands of chirps superimposed. The end result sounded unified - *a field of crickets* - even though it had been built up out of many individual recordings, because the basic unit (the cricket's chirp) is so similar - each chirp sounds pretty much like the last. This was not music, but it would still qualify, in my mind, as an example of harmonic superimposure.

(Incidentally, you'll be happy to know that the crickets escaped and lived happily behind the walls of this basement studio for the next few years, chirping at the most inappropriate moments.)

Dagwood and Blondie

What happens, though, when the superimposure is *not* harmonic?

Technically, of course, you can superimpose as much as you want: you can create huge “Dagwood sandwiches” of sound - a layer of dialogue, two layers of traffic, a layer of automobile horns, of seagulls, of crowd hubbub, of footsteps, waves hitting the beach, foghorns, outboard motors, distant thunder, fireworks, and on, and on. All playing together at the same time. (For the purposes of this discussion, let’s define a *layer* as a conceptually-unified series of sounds which run more or less continuously, without any large gaps between individual sounds. A single seagull cry, for instance, does not make a layer.)

The problem, of course, is that sooner or later (mostly sooner) this kind of intense layering winds up sounding like the rush of sound between radio stations - *white noise* - which is where we began our discussion. The trouble with white *noise* is that, like white light, there is not a lot of information to be extracted from it. Or rather there is so much information tangled together that it is impossible for the mind to separate it back out. It is as indigestible as one of Dagwood’s sandwiches. You still *hear* everything, technically speaking, but it is impossible to listen to it - to appreciate or even truly distinguish any single element. So the filmmakers would have done all that work, put all those sounds together, for nothing. They could have just tuned between radio stations and gotten the same result.

I have here a short section from *Apocalypse Now* which I hope will show you what I mean. You will be seeing the same minute of film six times over, but you will be hearing different things each pass: one separate layer of sound after another, which should give you an almost geological feel for the sound landscape of this film. This particular scene runs for a minute or so, from Kilgore’s helicopters landing on the beach to the explosion of the helicopter and Kilgore saying “I want my men out!” But it is part of a much longer action sequence.

Originally, back in 1978, we organized the sound this way because we didn’t have enough playback machines - we *couldn’t* run everything together: there were over a hundred and seventy-five separate soundtracks for this section of film alone. It was my very own Dagwood sandwich. So I had to break the sound down into smaller, more manageable groups, called *premises*, of about 30 tracks each. But I still do the same thing today even though I may have eight times as many faders as I did back then.

The six premix layers were:

Apocalypse Now Premix Layers

1. Dialogue
2. Helicopters
3. Music (The Valkyries)
4. Small Arms Fire (AK47's and M16s)
5. Explosions (Mortars, Grenades, Heavy Artillery)
6. Footsteps and other foley-type sounds.

View These Clips Online at:

<http://www.transom.org/guests/review/200504.review.murch2.html>

Clips Courtesy of and Copyright American Zoetrope, 1979

These layers are listed in order of importance, in somewhat the same way that you might arrange the instrumental groups in an orchestra. Mural painters do somewhat the same thing when they grid a wall into squares and just deal with one square at a time. What murals and mixing and music all have in common is that in each of them the detail has to be so exactly proportioned to the immense scale of the work that it is easy to go wrong - either the details will overwhelm the eye (or ear) but give no sense of the whole, or the whole will be complete but without convincing details.



The human voice must be understood clearly in almost all circumstances, whether it is singing in an opera or dialogue in a film, so the first thing I did was mix the dialogue for this scene, isolated from any competing elements.

Then I asked myself: what is the next most dominant sound in the scene? In this case it happened to be the helicopters, so I mixed all the helicopter tracks together onto a separate roll of 35mm film, while listening to the playback of the dialogue, to make sure I didn't do anything with the helicopters to obscure the dialogue.



Then I progressed to the third most dominant sound, which was the "Ride of the Valkyries" as played through the amplifiers of Kilgore's helicopters. I mixed this to a third roll of film while monitoring the two previous premixes of helicopters and dialogue.

And so on, from #4 (small arms fire) through #5 (explosions) to #6 (footsteps and miscellaneous sounds). In the end, I had six premixes of film, each one a six-channel master (three channels behind the screen: left, center, and right; two channels in the back of the theater: left and right; and one channel for low frequency enhancement). Each premix was balanced against the others so that - theoretically, anyway - the final mix should simply have been a question of playing everything together at one set level.

What I found to my dismay, however, was that in the first rehearsal of the final mix everything seemed to collapse into that big ball of noise I was talking about earlier. Each of the sound-groups I had premixed was justified by what was happening on screen, but by some devilish alchemy they all melted into an unimpressive racket when they were played together.

The challenge seemed to be to somehow find a balance point where there were enough interesting sounds to add meaning and help tell the story, but not so many that they overwhelmed each other. The question was: where was that balance point?

Suddenly I remembered my experience ten years earlier with Robot Footsteps, and my first encounter with the mysterious Law of Two-and-a-Half.

Note About the Apocalypse Premix Clips: The picture for these six clips is black and white, which is what we mixed to except when checking the final, when we had a color print. The sound on these clips is mono, but in the premixes it was full six-track split-surround with boom (what we call 5.1 today).

Robots and Grapes

This had happened in 1969, on one of the first films I worked on: George Lucas's *THX-1138*. It was a low-budget film, but it was also science fiction, so my job was to produce an otherworldly soundtrack on a shoestring. The shoestring part was easy, because that was the only way I had worked up till then. The otherworldly part, though, meant that most of the sounds that automatically 'came with' the image (the sync sound) had to be replaced. A case in point: the footsteps of the policemen in the film, who were supposed to be robots made out of six hundred pounds of steel and chrome. During filming, of course, these robots were actors in costume who made the normal sound that anyone would make when they walked. But in the film we wanted them to sound massive, so I built some special metal shoes, fitted with springs and iron plates, and went to the Museum of Natural History in San Francisco at 2am, put them on and

recorded lots of separate ‘walk-bys’ in different sonic environments, stalking around like some kind of Frankenstein’s monster.

They sounded great, but I now had to sync all these footsteps up. We would do this differently today - the footsteps would be recorded on what is called a Foley stage, in sync with the picture right from the beginning. But I was young and idealistic - I wanted it to sound right! - and besides we didn’t have the money to go to Los Angeles and rent a Foley stage.

So there I was with my overflowing basket of footsteps, laying them in the film one at a time, like doing embroidery or something. It was going well, but too slowly, and I was afraid I wouldn’t finish in time for the mix. Luckily, one morning at 2am a good fairy came to my rescue in the form of a sudden and accidental realization: that if there was one robot, his footsteps had to be in sync; if there were two robots, also, their footsteps had to be in sync; but if there were three robots, *nothing* had to be in sync. Or rather, any sync point was as good as any other!

Robots from *THX 1138*



View This Clip Online at:

<http://www.transom.org/guests/review/200504.review.murch2.html>

Clip Courtesy of Warner Brothers

This discovery broke the logjam, and I was able to finish in time for the mix. But...

But why does something like this happen?

Somehow, it seems that our minds can keep track of one person’s footsteps, or even the footsteps of two people, but with three or more people our minds just give up - there are too many steps happening too quickly. As a result, each footprint is no longer evaluated individually, but rather the *group of footsteps* is evaluated as a single entity, like a musical chord. If the pace of the steps is roughly correct, and it seems as if they are on the right surface, this is apparently enough. In effect, the mind says “Yes, I see a *group* of people walking down a corridor and what I hear sounds like a *group* of people walking down a corridor.”

Sometime during the mid-19th century, one of Edouard Manet's students was painting a bunch of grapes, diligently outlining every single one, and Manet suddenly knocked the brush out of her hand and shouted: "Not like that! I don't give a damn about Every Single Grape! I want you to get the feel of the grapes, how they taste, their color, how the dust shapes them and softens them at the same time. "

Similarly, if you have gotten Every Single Footstep in sync but failed to capture the energy of the group, the space through which they are moving, the surface on which they are walking, and so on, you have made the same kind of mistake that Manet's student was making. You have paid too much attention to something that the mind is incapable of assimilating anyway, even if it wanted to.

Note About the THX-1138 Clip: *In the final mix, the music dominates, so it is hard to hear the multiple footstep effect except briefly at the beginning and the end. You will have to take my word that there was a full footsteps premix for this section, but it was in creating the sound for this scene, in 1970, that the idea of two-and-a-half layers first came to me. A clearer example of the multiple footstep effect can be found on the #6 Apocalypse Now premix.*



Trees and Forests

At any rate, after my robot experience I became sensitive to the transformation that appears to happen as soon as you have three of anything. On a practical level, it saved me a lot of work - I found many places where I didn't have to count the grapes, so to speak - but I began to see the same pattern occurring in other areas as well, and it had implications far beyond footsteps.

The clearest example of what I mean can be seen in the Chinese symbols for "tree" and "forest." In Chinese, the word "tree" actually looks like a tree - sort of a pine tree with drooping limbs. And the Chinese word for "forest" is three trees. Now, it was obviously up to the Chinese how many trees were needed to convey the idea of "forest," but two didn't seem to be enough, I guess, and sixteen, say, was way too many - it would

木 tree (mu)

林 woods (lin)

森 forest (sen)

have taken too long to write and would have just messed up the page. But *three* trees seems to be just right. So in evolving their writing system, the ancient Chinese came across the same fact that I blundered into with my robot footsteps: that three is the borderline where you cross over from “individual things” to “group.”

It turns out Bach also had some things to say about this phenomenon in music, relative to the maximum number of melodic lines a listener can appreciate simultaneously, which he believed was three. And I think it is the reason that Barnum’s circuses have three rings, not five, or two. Even in religion you can detect its influence when you compare Zoroastrian Duality to the mysterious ‘multiple singularity’ of the Christian Trinity. And the counting systems of many primitive tribes (and some animals) end at three, beyond which more is simply “many.”

So what began to interest me from a creative point of view was the point where I could see the forest *and* the trees - where there was simultaneously Clarity, which comes through a feeling for the individual elements (the notes), and Density, which comes through a feeling for the whole (the chord). And I found this balance point to occur most often when there were not quite three layers of something. I came to nickname this my “Law of Two-and-a-half.”

Left and Right

Now, a practical result of our earlier distinction between Encoded sound and Embodied sound seems to be that this law of two-and-a-half applies only to sounds of the same ‘color’ - sounds *from the same part of the conceptual spectrum*. (With sounds from *different parts* of the spectrum - different colored sounds - there seems to be more leeway.)

The robot footsteps, for instance, were all the same ‘green,’ so by the time there were three layers, they had congealed into a new singularity: *robots walking in a group*. Similarly, it is just possible to follow two ‘violet’ conversations simultaneously, but not three. Listen again to the scene in *The Godfather* where the family is sitting around wondering what to do if the Godfather (Marlon Brando) dies. Sonny is talking to Tom, and Clemenza is talking to Tessio - you can follow both conversations and also pay attention to Michael making a phone call to Luca Brasi (Michael on the phone is the “half” of the two-and-a-half), but only because the scene was carefully written and performed and recorded. Or think about two pieces of ‘red’ music playing simultaneously: a background radio and a thematic score. It can be pulled off, but it has to be done carefully.

But if you blend sounds from different parts of the spectrum, you get some extra latitude. Dialogue and Music can live together quite happily. Add some Sound Effects,

too, and everything still sounds transparent: two people talking, with an accompanying musical score, and some birds in the background, maybe some traffic. Very nice, even though we already have four layers.

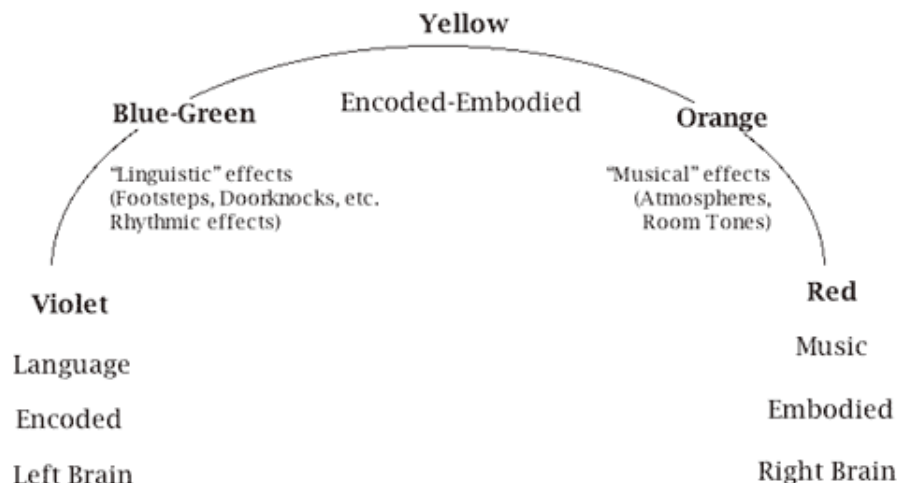
Why is this? Well, it probably has something to do with the areas of the brain in which this information is processed. It appears that Encoded sound (language) is dealt with mostly on the left side of the brain, and Embodied sound (music) is taken care of across the hall, on the right. There are exceptions, of course: for instance, it appears that the rhythmic elements of music are dealt with on the left, and the vowels of speech on the right. But generally speaking, the two departments seem to be able to operate simultaneously without getting in each other's way. What this means is that by dividing up the work they can deal with a total number of layers that would be impossible for either side individually.

Density and Clarity

In fact, it seems that the total number of layers, if the burden is evenly spread across the spectrum from Encoded to Embodied (from “violet” dialogue to “red” music) is double what it would be if the layers were stacked up in any one region (color) of the spectrum. In other words, you can manage five layers instead of two-and-a-half, thanks to the left-right duality of the human brain.

What this might mean, in practical terms, is:

1. One layer of “violet” dialogue;
2. One layer of “red” music;
3. One layer of “cool” (linguistic) effects (eg: footsteps);
4. One layer of “warm” (musical) effects (eg: atmospheric tonalities);
5. One layer of “yellow” (equally balanced ‘centaur’) effects.



What I am suggesting is that, at any one moment (for practical purposes, let's say that a 'moment' is any five-second section of film), five layers is the maximum that can be tolerated by an audience if you also want them to maintain a clear sense of the individual elements that are contributing to the mix. In other words, if you want the experience to be simultaneously *Dense and Clear*.

But the precondition for being able to sustain five layers is that the layers be spread evenly across the conceptual spectrum. If the sounds stack up in one region (one color), the limit shrinks to two-and-a-half. If you want to have two-and-a-half layers of dialogue, for instance, and you want people to understand every word, you had better eliminate the competition from any other sounds which might be running at the same time.

To highlight the differences in our perception of Encoded vs. Embodied sound, it is interesting to note the paradox that in almost all stereo films produced over the last twenty-five years the dialogue is always placed in the center no matter what the actual position of the actors on the screen: they could be on the far left, but their voices still come out of the center. And yet everyone (including us mixers) still believes the voices are 'coming from' the actors. This is a completely different treatment than is given sound effects of the "yellow" variety - car pass-bys, for instance - which are routinely (and almost obligatorily) moved around the screen with the action. And certainly different from "red" music, which is usually arranged so that it come out of all speakers in the theater (including the surrounds) simultaneously. Embodied "orange" sound effects (atmospheres, room tones) are also given a full stereo treatment. "Blue-green" sound effects like footsteps, however, are usually placed in the center like dialogue, unless the filmmakers want to call special attention to the steps, and then they will be placed and moved along with the action. But in this case the actors almost always have no dialogue.

As a general rule, then, the "warmer" the sound, the more it tends to be given a full stereo (multitrack) treatment, whereas the "cooler" the sound, the more it tends to be monophonically placed in the center. And yet we seem to have no problem with this incongruity - just the opposite, in fact. The early experiments (in the 1950's) which involved moving the dialogue around the screen were eventually abandoned as seeming "artificial."

Monophonic films have always been this way - that part is not new. What is new and peculiar, though, is that we are able to tolerate - even enjoy - the mixture of mono and stereo in the same film.

Why is this? I believe it has something to do with the way we decode language, and that when our brains are busy with Encoded sound, we willingly abandon any question of its origin to the visual, allowing the image to “steer” the source of the sound. When the sound is Embodied, however, and little linguistic decoding is going on, the location of the sound in space becomes increasingly important the less linguistic it is. In the terms of this lecture, the “warmer” it is. The fact that we can process both Encoded mono and Embodied stereo simultaneously seems to clearly demonstrate some of the differences in the way our two hemispheres operate.

Getting back to my problem on *Apocalypse*: it appeared to be caused by having six layers of sound, and six layers is essentially the same as sixteen, or sixty: I had passed a threshold beyond which the sounds congeal into a new singularity - dense noise in which a fragment or two can perhaps be distinguished, but not the developmental lines of the layers themselves.



With six layers, I had achieved *Density*, but at the expense of *Clarity*.

What I did as a result was to restrict the layers for that section of film to a maximum of five. By luck or by design, I could do this because my sounds were spread evenly across the conceptual spectrum.



1. Dialogue (violet)
2. Small arms fire (blue-green ‘words’ which say “Shot! Shot! Shot!”)
3. Explosions (yellow “kettle drums” with content)
4. Footsteps and miscellaneous (blue to orange)
5. Helicopters (orange music-like drones)
6. Valkyries Music (red)

Final Mixdown from *Apocalypse Now*

View This Clip Online at:

<http://www.transom.org/guests/review/200504.review.murch2.html>

Clips Courtesy of and Copyright by American Zoetrope 1979

Note About the Apocalypse Final Clip: *The soundtrack on this clip is stereo, but the final mix of the film is in 5.1, a format which we created specifically for Apocalypse Now in 1977-9 and which is currently the industry standard.*

If the layers had not been as evenly spread out, the limit would be less than five. And as I mentioned before, if they had all been concentrated in one “color zone” of the spectrum, (all violet or all red, for instance) the limit would shrink to two-and-a-half. It seems, then, that the more monochrome the palette, the fewer the layers that can be super-imposed; the more polychrome the palette, on the other hand, the more layers you get to play with.

So in this section of *Apocalypse*, I found I could build a “sandwich” with five layers to it. If I wanted to add something new, I had to take something else away. For instance, when the boy in the helicopter says “I’m not going, I’m not going!” I chose to remove all the music. On a certain logical level, that is not reasonable, because he is actually in the helicopter that is producing the music, so it should be louder there than anywhere else. But for story reasons we needed to hear his dialogue, of course, and I also wanted to emphasize the chaos outside - the AK47’s and mortar fire that he was resisting going into - and the helicopter sound that represented “safety,” as well as the voices of the other members of his unit. So for that brief section, here are the layers:

1. Dialogue (“I’m not going! I’m not going!”)
2. Other voices, shouts, etc.
3. Helicopters
4. AK-47’s and M-16s
5. Mortar fire.

Under the circumstances, music was the sacrificial victim. The miraculous thing is that you do not hear it go away - you believe that it is still playing even though, as I mentioned earlier, it should be louder here than anywhere else. And, in fact, as soon as this line of dialogue was over, we brought the music back in and sacrificed something else. Every moment in this section is similarly fluid, a kind of shell game where layers are disappearing and reappearing according to the dramatic focus of the moment. It is necessitated by the ‘five-layer’ law, but it is also one of the things that makes the soundtrack exciting to listen to.

But I should emphasize that this does not mean I always had five layers cooking. Conceptual density is something that should obey the same rules as loudness dynamics. Your mix, moment by moment, should be as dense (or as loud) as the story and events warrant. A monotonously dense soundtrack is just as wearing as a monotonously loud film. Just as a symphony would be unendurable if all the instruments played together all the time. But my point is that, under the most favorable

of circumstances, five layers is a threshold which should not be surpassed thoughtlessly, just as you should not thoughtlessly surpass loudness thresholds. Both thresholds seem to have some basis in our neurobiology.

The bottom line is that the audience is primarily involved in following the story: despite everything I have said, the right thing to do is ultimately whatever serves the storytelling, in the widest sense. When this helicopter landing scene is over, however, my hope was the lasting impression of everything happening at once - Density - yet everything heard distinctly - Clarity. In fact, as you can see, simultaneous Density and Clarity can only be achieved by a kind of subterfuge.

As I said at the beginning, it can be complicated to be simple and simple to be complicated.

But sometimes it is just complicated to be complicated.



Walter Murch mixing *Apocalypse Now*, 1979

Happy mixing!

A Conversation with Walter Murch

Nannette Drake Oldenbourg - *February 4, 2005 - #8*

Do you use left/right brain techniques in dealing with "the creator," yourself and colleagues?

Walter Murch - *February 5, 2005 - #9*

There does seem to be a double aspect to the creative process: a generative side that thinks up (or receives) the ideas; and a structural/editorial side that puts them in the right arrangement. It is kind of like those hypodermic epoxy glue dispensers that you get at the hardware store: one side is resin and the other side is hardener, and when you press the plunger, they are both squeezed out in equal proportions for the best result. If there is too much resin (generative creativity) and not enough hardener (editorial creativity) you get a sticky mess that never dries. If it is the other way around - too much hardener (editorial) and not enough resin (original ideas) -- the result hardens quickly to a thin and brittle consistency.

Sometimes when the film gets into a corner where it seems trapped, and no amount of logical thinking can pry it out, I will just subject the material to the "wisdom of the hands" -- a visceral and spontaneous reshaping where it seems that my logical brain is somewhere else and my hands are doing the work all by themselves.

To get to this point, though, there had to be (for me anyway) a lot of conscientious preparation -- viewing the raw material many times and taking in-depth notes.

Nannette Drake Oldenbourg - *February 5, 2005 - #10*

This may be a long shot, but back in Jay's introduction to you he wrote:

"Between films, he pursues interests in the science of human perception, cosmology and the history of science..." "

You might not have meant for anyone to take the comparison so literally, but I have to ask what you think about space in the universe and space in the mind, or at least our perception of sound in the mind.

Walter Murch -- February 6, 2005 - #13

From 600 BC Until the mid-1600's AD European astronomers believed that the stars were embedded in a black crystal sphere, and that beyond this sphere of stars was God. So in a sense the perceivable universe was contained, womb-like, within God. Which was (if you were a believer) a comforting thought.

When that black crystal finally began to crack apart, under the relentless prying of increasingly powerful telescopes, it left a profoundly disturbing aftertaste. Blaise Pascal, the French mathematician and philosopher, would now feel an unfamiliar shiver contemplating the nighttime sky: "Le silence éternel de ces espaces infinis m'effraye," he wrote in 1655. "The eternal silence of that infinite space terrifies me." Twelve years later, John Milton -- who visited the aged and imprisoned Galileo on a pilgrimage in 1638 -- would express similar thoughts in his poem *Paradise Lost*:

*Before his eyes in sudden view appear
The secrets of the hoary Deep -- a dark
Illimitable ocean, without Bound
Without dimension, where length, breadth, and highth,
And time and place are lost...*

When the child is in the womb, and sound is the only active sense, there can be no awareness of what we would call space: that discovery comes after birth, and it is a gradually unfolding one. I remember when I was four, living in Manhattan and looking out across the Hudson River and saying to myself about the distant (New Jersey) shore: "that is the other side of the world - that is China." The following year, when my sense of space had matured slightly, I decided that it wasn't China after all: it was France.

So in the last five hundred years or so we have been going through, culturally, a journey similar to the one that each of us experiences in the first ten years of our lives. In this case the womb-like enclosure of the sphere of stars has transformed into something else: is it the Cosmic equivalent of China? Or France? Or will we ultimately discover that it is New Jersey after all?

If you can find a copy of it, there is a wonderful book by Judith and Herbert Kohl called "The View from the Oak" (1977) which examines the difference senses of space that

each animal species has. But it is also true that our own sense of space changes as we mature, both individually and as a culture. Balinese space is not the same (I am guessing) as American space, and American space was different in 1705 than it is today.

So you are right, there probably is some connection between these very different pursuits in cinema and astronomy. Certainly at one level there is the pursuit of patterns: film editing can almost be defined as the search for patterns at deeper and deeper levels. Astronomy is probably the earliest human activity that we would recognize as science, and it also is certainly all about finding patterns.

What fascinates me particularly about planetary spacing is that there is clearly something regular going on there, but exactly what it is or what is causing this regularity has eluded us. We thought we knew the “what” of it in the early 19th century when Bode’s Law was generally believed to be true. But no one was ever able to explain the ‘why’ of Bode’s Law, and when Neptune was discovered, in 1846, it was also discovered not to fit Bode’s neat formulation and so the law fell into disrepute. Nonetheless, Pluto does fit where Neptune was supposed to have been. The curious thing, especially for me, is that the distance intervals predicted by Bode’s Law, expressed as numerical ratios, are also musical ratios. No one has observed this before. Investigation continues...

Jake Warga - February 7, 2005 - #15

"...the newly-born, silence must be intensely threatening, with its implications of cessation and death."

How do you treat...silence? Digital theatrical and home sound systems can do more with nothing than traditional optical release tracks. How has...silence...in its new absolute form tempted you, and why is it such a threat in contemporary culture?

Walter Murch - February 8, 2005 - #17

This is an answer to Jake Warga's question about silence. It is an extract from an article I wrote for Larry's Sider's School of Sound, which is held annually in London. The examples I was talking about were from Apocalypse Now.

This first section I'm going to play for you, from the Valkyries sequence in *Apocalypse Now*, is what I would call "locational silence." It's done for the purposes of demonstrating a shift in location but also for the visceral effect of a sudden transition from loudness to silence. This sudden silence -- cutting to a quiet schoolyard, also helps you share the point of view of the Vietnamese, who are shortly going to be overwhelmed with the noise and violence coming at them. Then, from the rear speakers only, you begin to hear the Valkyries and the sound of helicopters. The people on screen notice it, and then the sound moves rapidly forward and eventually encompasses the whole theater.

This cut to silence came about, partially, because of an idea of Francis's that I kicked against. While the helicopters were in flight, he wanted the tape that was playing the music to break. You'd be caught up in the excitement of all this Valkyries music and suddenly -- the music would stop! And then you'd see a soldier fiddling with the tape and it would start up again. I think I can understand what he was trying for, but, once you invoke this music, it's just so overwhelmingly powerful that I tried to find another way to get what he wanted but put it where I thought it would be more emotional and understandable: cutting to the peaceful schoolyard of the Vietnamese. All transitions to silence have a psychological component although, in this case, the geographical component is quite strong. We're with the helicopters and then -- cut - we are a dozen miles away and the music is gone. Nonetheless, there is a huge visceral effect on the audience when something that's so loud and sustained suddenly cuts off. This cliff of silence also gave us a chance to build up the level of the music again, which makes the moment when the helicopters actually hit the beach that much more powerful. If the music had been sustained and loud right from the beginning, without this

break, it would not be so effective. Your ears would get tired of it before the scene was over.

The next example is a different kind of silence, which has less geography and more psychology: the scene leading up to the tiger jumping out of the jungle. Chef (Fred Forrest) and Willard (Martin Sheen) have gotten off the boat and gone into the jungle to search for mangoes. Chef starts talking about himself and his ambitions: he wanted to be a cook so he joined the Navy ("heard they had better food") but they made him boil the meat and it was disgusting, etc. He's a Navy man and not comfortable in the jungle. Willard on the other hand is a jungle fighter and so he's first to become aware that something subtle is wrong with what he's hearing. It's like those scenes in Westerns where a cowboy says, "Ya hear that?" And the other cowboy says, "I don't hear anything." And the first guy says, "That's just the point." So, something changes in the sound of the forest, some little insect goes quiet, and Willard picks up on this, thinking, "Charlie. North Vietnamese out there gunning for us." And he moves deeper into the jungle, much to Chef's distress, because he's lost without the safety of his boat. Meanwhile the soundtrack is getting more and more minimal the closer they get to the source of the silence.

The important thing in scenes like this, where you are reducing the soundtrack incrementally, is the curve of the slope: how fast you bring the sound to silence. It's like those airplanes NASA uses when they're teaching astronauts about zero gravity. They fly to a high altitude and then arc back toward the ground following a specific curve where gravity disappears and they begin to float around the plane. Coming gradually to silence is a similar thing. If you come to it too quickly, it's as if you suddenly reached the bottom of a funfair ride--there's this jolt and you wonder, "What was that all about?" Very much what we just saw with the sudden cut to silence from the Valkyries. The reason it works in that case is that cut to silence accompanied a shift in location.

Here, with the two men in the jungle, the silence has to develop within the same environment. If it is effective, it makes you participate in the same psychological state as the characters, who are listening more and more intently to a precise point in space behind which they think "Charlie" is lurking. In fact, we never reach absolute silence in this scene. It feels silent, but it isn't. We've narrowed it down to a thin sound of a creature called a "glass insect" native to Southeast Asia; it makes a sound like a glass harmonica - when you rub your finger on the top of a wine glass, just that (sings note), which itself is a sound that puts you on edge. So we've done very much what Michel was calling "the silence of the orchestra around the single flute" -- or the single insect in this case. The trick here is to orchestrate the gradual elimination of "orchestral" elements of the jungle soundtrack: If it is done too fast, the audience is taken out of the moment; too slow, and the effect we were after, the tense moment before the tiger jumps, wouldn't have been as sharp.

I should mention that I spent many hours juggling the exact relationship between the move of the tiger and the roar of the tiger - ultimately the sound of the tiger is slightly delayed from its motion, so that you see leaves beginning to move and you think, "It's Charlie!" But that realization takes a little bit of time. What Michel was saying yesterday is that both sound and our appreciation of the sound are inevitably linked to time. You hear something but it takes you a certain amount of time to come to a realization about what it is. It can be fast, but it finally depends on our reaction time. So, just as the audience's reaction time is congealing around the belief that "it's Charlie," at that exact moment, the roar of the tiger happens. People are out of their chairs already because they think it's Charlie, and then they really jump because it's not Charlie, it's a tiger! And then the shooting starts.

The next clip demonstrates another kind of silence, even more psychological. What you saw with the tiger is

something that's not geographic displacement, but, nonetheless, it is something that could happen in reality. It's a possibility that the animals in the forest could all slowly get quiet like that. What I'm going to show you now is impossible, realistically speaking. This is what's called the Dolung Bridge sequence: the Americans are trying to build a bridge across the Dolung River, and the Vietnamese are bombing it to pieces every night. So it's an exercise in futility done for the needs of some military balance sheet back in Saigon. The reality of the situation is that nothing is being accomplished--people are dying and it's all hopeless.

The scene begins with the realistic sounds of bridge construction. You hear arc welders, you hear flares going off, machine guns and incoming artillery. As the scene goes on, though, you'll notice that the explosions and the machine guns are replaced by sounds of construction--the machine guns become rivet guns, for instance, so there's already a subtle warping of reality taking place. Francis called this scene "the fifth circle of Hell." Once the scene gets into the trench, the dilemma is explained: there's a Vietnamese soldier out there, a sniper taunting the Americans, and they're shooting wildly into the dark with an M-50 machine gun, but they just can't get him. Finally, out of frustration the machine gunner says, "Go get the Roach." "The Roach" turns out to be the nickname of a soldier who is a kind of human bat; he has precise echolocation instead of sight, so if he can hear the sound of the voice, he can then pinpoint his target, adjust his grenade launcher and, in the dark, shoot the sniper. As Roach approaches the camera, the rock music that has been echoing "in the air" of the scene, coming from all speakers in the theatre, concentrates itself in the center speaker only and narrows its frequency range, seeming to come from a transistor radio which Roach then clicks off, taking all the other sounds with it. After a brief rumble of distant artillery, there is now silence except for some kind of unexplained, slow metallic ticking. Visually you see the battle continuing--flashes of light, machine gun bursts,

flare guns--that should normally have sounds accompanying them -- but there is nothing else except the taunting voice of the sniper in the dark. You have entered into the skin of this human bat and are hearing the world the way he hears it. He aims, shoots, there's an explosion and then a moment of complete silence. Willard asks Roach if he knows who's in command, and Roach answers enigmatically: "yeah." Then the scene is over, we shift location and the world of sound comes flooding back in again.

There's a quote from Bresson about actors that could apply here: "An actor can be beautiful with all of the gestures that he could - but does not - make." You have to invoke the possibility of the sound. You can't simply be silent and say, "This silence is great"; instead you have to imagine the hundred musicians on stage in order for their silence to mean anything. You have to work with the psychic pressure exerted by the instruments or sounds that are not playing. This is the underpinning of what I try to do with sound, which is to evoke the individual imagination of each member of the audience.

In most films everything is "see it = hear it." The sounds may be impressive, but since they come from what you're looking at they seem to be the inevitable shadow of the thing itself. As a result, they don't excite that part of your mind that is capable of imagining far beyond what the filmmakers are capable of producing. Once you stray into metaphoric sound, which is simply sound that does not match what you're looking at, the human mind will look for deeper and deeper patterns. And it will continue to search--especially if the filmmakers have done their work--and find those patterns, if not at the geographic level, then at the natural level; if not at the natural level, then at the psychological level. And through the mysterious alchemy of film+audience, your understanding will re-project itself into the characters and situations themselves and you will assume that what you are feeling is coming from the characters and the situation rather than from something that is happening on the soundtrack. There's a process of

digestion and re-projection that goes on--it's in the nature of the way our minds work. John Huston used to say that the real projectors in the theater are the eyes and ears of the audience.

The ultimate metaphoric sound is silence. If you can get the film to a place with no sound where there should be sound, the audience will crowd that silence with sounds and feelings of their own making, and they will, individually, answer the question of, "Why is it quiet?" If the slope to silence is at the right angle, you will get the audience to a strange and wonderful place where the film becomes their own creation in a way that is deeper than any other.

Jeff Towne - *February 5, 2005 - #11*

I'm curious about the importance of foreground-background spacing, in terms of density and clarity and also in terms of "realism". I noticed in the clip we have here from *Apocalypse Now*, that the Wagner Valkyries music doesn't really sound as if it's moving, or changing distance from us, even though it logically is emanating from helicopters that are flying around at varying orientations to the action. But, of course, it works perfectly. Is it a similar thing to not needing stereo panning of dialog, that our perception of the scene overwhelms the need for any specific placement of a sound?

I know you sometimes find that sound of distance important, beyond simple volume changes. Can you talk about "airballing," when you think it, or similar spatial treatments are needed, and if you've found practical technical tools to make it easier than what you did on *American Graffiti*?

Walter Murch - *February 6, 2005 - #12*

The clips of the AN premixes on Transom are in mono, so there is no directional movement. The final mix clip, which is in stereo, is also restricted compared to the 'real' final mix which was in the 5.1 format, where the panning movement of the Valkyries music is more obvious, but we would need to listen to the film or the DVD in a 5.1 environment to hear this.

BUT, in this particular scene the music was presumably coming from many helicopters, all of which were landing and taking off simultaneously, so perceiving specific musical directionality for this scene is hard even in 5.1 -- especially with everything else that was going on in the sound track.

And then there is the creative balance between music as sound effect (source) and music as music (score). In this scene the music is functioning more toward the red end of the spectrum, as score, and in that case we would (and did) ease back on source-point directionality which would have emphasized the 'sound effect' nature of the music.

In this clip there are several layers of the music all playing at once, with different delays between each (because some helicopters are further away from us than others) and also each music track is being filtered differently, since a speaker at 500 feet distance has a different quality than one that is ten feet away. You can get some sense of this in the clips -- even in the mono version of the premix.

So -- to answer your question directly -- yes in this case we felt that being overly literal about the placement of the sound would distract from the emotional power of the music as such.

If you have the DVD of Apocalypse Redux, you can hear a full use of directional panning in the scene where Kilgore flies over the boat at night asking for his surfboard back. You would need to have a 5.1 setup to hear it the way we intended it.

"Airballing" or "worldizing" were nicknames we came up with in the early 70's for a technique of rerecording the sound in question in a real environment which duplicates or closely mimics the acoustics of the space you see on screen. Two machines are necessary: one to play the sound through a speaker, and another (some distance away) to re-record it once it has energized the space.

In the final mix this worldized track would be laid up alongside the original (studio quality) track and we would experiment until we found the correct balance between the two giving the right amount of power (studio track) and atmosphere (worldized track). We had to do this back then because there were no sophisticated reverb units that could achieve what we were after. Now, thanks to digital magic, there are dozens of very sophisticated devices (Lexicon, etc) on the market to achieve this kind of thing very easily in the studio. On the other hand, although it is labor intensive, there is something wonderful and 'handcrafted' about the worldizing technique because you can get exactly the environment that you want, and there is a random 'truth' to those tracks that electronic simulations can only approximate.

The reason for doing this atmospheric treatment at all, by whatever means, is that it gives you the acoustic equivalent of depth of field in photography: the sound is thrown

out of focus in an interesting way (to the degree that you want it to be out of focus) so that you can have something in the foreground (like a human voice) which is acoustically 'sharp' and in focus (ie: the balance of direct to reflected sound is tipped in favor of the direct) and this allows the audience to hear the voice clearly while also appreciating the atmospheric contributions of the other track. The comparison would be with portrait photography in a natural environment: the photographer would typically choose a long focal-length lens and keep the subject sharp while letting the foliage in the background go soft. The eye instantly knows it is supposed to look at the face, and yet it derives pleasure from the surrounding soft atmosphere. If everything were equally in focus, the eye would be confused. This is what happens in a mix if you simply lower the volume of a background track and perhaps filter it a bit: the background is lower in volume but it still has acoustically 'sharp' edges which attract the ear and distract your attention from the central focus which might be (and usually is) the human voice.

There is a lot of this technique, as it turned out, in Welles's *Touch of Evil* (1958), which I re-edited and mixed according to Welles's notes (he had been fired off the film in 1957).

The following is an excerpt from "Touch of Silence" -- a lecture I gave at the School of Sound in 1998. The excerpts in quotation marks are from the 58 page memo that Welles wrote after he had seen what the studio had done to his film.

=====

Let me read you one more page, which may give you some idea of what Universal was coping with. Welles is writing about the treatment of the music in the scenes with Susie (Janet Leigh) talking to Grandi (Akim Tamiroff.) "The music should have a low insistent beat with a lot of bass in it. This music is at its loudest in the street and once she enters the tiny lobby of The Ritz Hotel, it fades to extreme background. However, it does not disappear but continues and, eventually, there will be a segue to a Latin-type rhythm number, also very low in pitch, dark, and with a strong, even exaggerated, emphasis on the bass." Then his typewriter breaks into capital letters, and he centers the next paragraph right in the middle of the page:

"IT IS VERY IMPORTANT TO NOTE THAT IN THE RECORDING OF ALL THESE NUMBERS, WHICH ARE SUPPOSED TO BE HEARD THROUGH STREET LOUD SPEAKERS, THAT THE EFFECT SHOULD BE JUST THAT, JUST EXACTLY AS BAD AS THAT."

He continues: "The music itself should be skillfully played but it will not be enough, in doing the final sound mixing, to run this track through an echo chamber with a certain amount of filter. To get the effect we're looking for, it is absolutely vital that this music be played back through a cheap horn in the alley outside the sound building. After this is recorded, it can then be loused up even further in the process of re-recording. But a tinny exterior horn is absolutely necessary, and since it does not represent very much in the way of money, I feel justified in insisting upon this, as the result will be really worth it." Well, in that paragraph, if you are a Universal executive in 1958, are two completely incompatible concepts: After the music is recorded, "it can then be loused up even further;" and, "the result will really be worth it."

Put yourself in the shoes of Ed Muhl, the head of Universal at the time. Here is a director saying that he wants to take a well-recorded music track and louse it up by re-recording it in an alleyway and then maybe louse it up even further in the final mix, and to top it all off "the result will really be worth it." It wouldn't take much more to make you question everything that Welles was doing. And if he was taking a long time doing it, you might feel justified in removing the film from his control. When we were working on the recut, Rick (the producer of the recut) discovered that Mr. Muhl was still alive, 95 years old, and living in the San Fernando Valley, so we called him up to tell him what we were doing and see if he had anything to suggest. He was acerbic and unrepentant, felt that Welles had been a conceited poseur who never directed a film that made any money, and of course they were justified in taking Touch of Evil away from him.

Well, we now know exactly what he was up to with his alleyway recording - it was the analog forerunner of all the digital reverberation techniques that have blossomed over the last twenty-five years, allowing us to color a sound with an infinite variety of acoustic ambiances. For me, Welles's description of his technique had a

particularly strong impact because this is something -- what I then called "worldizing" -- I developed on my own in the late sixties. Or at least I thought I had until I read this part of the memo. Having heard Touch of Evil at film school I was probably subconsciously influenced by what Welles had done. Though I had always -- even when I was playing with my first tape recorder in the early 1950's -- been fascinated by the emotion that a spatial treatment of sound can give, and I was frustrated with the limited technical resources available to do that in the mid '60's.

Jonathan Mitchell - February 7, 2005 - #14

I think it's fascinating (as always) to hear your perspective on the way sound works. You have a refreshingly clear way of articulating ideas that are probably for most people very intuitive concepts... To what extent do you think of these ideas you're presenting as being principles that are fundamentally true? Or do you view them simply as tools you've come to rely on to get the job done?

I'm also interested to know about your experiences with the effect a picture has on one's perception of sound, and if the principles you've presented change when the picture is removed.

Jeff Towne - February 7, 2005 - #16

I join Jonathan in wondering about how some of these equations might change when there are no visuals to accompany. Do you think the density overload happens in the same way when we are only listening, or can we absorb more sounds? fewer? I suspect the same hot-cold balance is in effect, but I wonder if the thresholds are the same.

Walter Murch - February 11, 2005 - #21

As to whether these theories are actually true or not - we know so little (comparatively speaking) about the workings of the human brain and where the interface is between physical perception and conscious attention, that I think the best you could say is that

these theories are a kind of "acupuncture of sound" - or that they are similar to the understanding of architectural engineering during the middle ages. There is some truth in there - acupuncture 'works' and the cathedrals of the 14th century are still standing - but exactly where it is.... hmmm....

I know that these theories work for me, and that it helps me to have a theory behind what I do, whether it is 'true' or not.

In the 18th century many scientists (esp. Joseph Priestly) believed in something called phlogiston, which was the 'flammable' part of everything that could burn, and in burning this substance was released into the air. We know (thanks to Lavoisier) that this is not true, but however flawed the theory was it helped science to think deeply about combustion and the various ways that it works: it led to the (correct) understanding that rust is a kind of slow combustion, and that combustion of a sort takes place when we breathe. So even if a theory isn't 'true' it can still help move things forward.

I will leave you for the moment with the introduction to "In the Blink of an Eye" which deals with this topic:

=====

Igor Stravinsky loved expressing himself and wrote a good deal on interpretation. As he bore a volcano within him, he urged restraint. Those without even the vestige of a volcano within them nodded in agreement, raised their baton, and observed restraint, while Stravinsky himself conducted his own Apollon Musagète as if it were Tchaikovsky. We who had read him listened and were astonished.

"The Magic Lantern" by Ingmar Bergman

Most of us are searching -- consciously or unconsciously -- for a degree of internal balance and harmony between ourselves and the outside world, and if we happen to become aware -- like Stravinsky -- of a volcano within us, we will compensate by urging restraint. By the same token, someone who bore a glacier within him might urge passionate abandon. The danger is, as Bergman points out, that a glacial personality in need of passionate abandon may read Stravinsky and apply restraint instead.

Many of the thoughts that follow are therefore more truly cautionary notes to myself, working methods I have developed for coping with my own particular volcanoes and glaciers. As such, they are insights into one person's search for balance, and are perhaps interesting to others more for the glimpses of the search itself than for the specific methods that search has produced.

Walter Murch - *February 13, 2005 - #22*

When George Lucas and I were mixing the soundtrack for the film THX-1138 in 1970, our goal was to create a sonic 'alternate reality' for this imagined world of the indeterminate future. Also, because of budget restrictions at the time of shooting, the production design was not as elaborate as George would have wanted and we felt we might compensate by creating a denser world through denser sound.

So I built many tracks of what was going on next door and down the hall from THX's and LUH's cell. Everything in the film was sonically 'densified' so to speak, and when we mixed it all together, we were very happy. BUT we were mixing to a small fuzzy video image of a black and white 35mm dupe of the film, so the visual components of the total experience were extremely dim compared to the sound.

When we finally projected the mix with the color 35mm film, the experience was overwhelming, and not in a good sense. Francis Coppola said the experience was like having your head shoved in a garbage can and then someone banging on it for two hours.

What happened? Clearly the amount of sonic density that could be tolerated and even enjoyed when the image was small, dim, fuzzy, and black and white was not the same when that image was big, bright, crisp, and in Technicolor.

Chastened, we remixed the film, thinning out the track density and reducing the levels of auxiliary tracks. It is still a dense soundtrack, but tolerably so.

So as a result of this 'experiment' I suspect that a completely audio experience, pushed to the maximum, could cope with a couple of extra layers of sonic density.

The most helpful comparison probably would be with the experience that composers have had dealing with multiple melodic lines playing at the same time.

J.S.Bach wrote a series of fifteen Two-part Inventions and fifteen Three-part Sinfonias "...wherein lovers of the clavier, but especially those desirous of learning are shown a clear way not only (1) to learn to play cleanly in 2 voices, but also, after further progress (2) to deal correctly and well with 3 obbligate parts, simultaneously, furthermore, not merely to acquire good inventions (ideas), but to develop the same properly, and above all to arrive at a singing manner in playing, and at the same time to acquire a strong foretaste of composition.

- Johann Sebastian Bach
Kapellmeister to His Serene Highness the Prince of Anhalt Cothen
Anno Christi 1723

Each of these 'voices' is a melodic line, and the goal is to perceive each melody for itself and also to enjoy its interpenetration with the other(s). If the pieces are not carefully composed (and performed) the results are muddy and sterile.

In Bach's terms, this distinction between two and three voices is the equivalent of my law of $2 \sqrt[3]{2}$, which is the point where the individual elements can be perceived as well as how they all interact: density and clarity at the same time

Bach pushed against these limits -- he was one of the greatest composers, technically and artistically, in the history of music -- and responding to a challenge from king Frederick II wrote A Musical Offering which has one section for three voices and another for six voices.

According to Luc-André Marcel ("Bach" (1961) in the "Solfèges" series at the Editions du Seuil):

"Apparently during a visit to the palace, Frederick II asked Bach for a three-voice fugue, to which Bach complied immediately. Then the king asked for a six-voice fugue. Bach said he was unable to improvise it on the spot but that he would definitely look into it back in Leipzig. (The author adds "Ah! (Frederick II) would've gone up to twelve voices just to beat him!"). Apparently Bach took great fun in trying to write a work that would face the king as much as possible with his own technical limitations."

J. S. Bach himself did not specify any instrumentation, but his son Carl Philipp Emmanuel declared the fugue was intended as keyboard music. On one keyboard the voices cross a great deal and tend to lose their identities. This new arrangement divides the voices alternately between the keyboards; each keyboard ends up with what seems like a spare but beautiful 3-voice fugue, and yet they fit together in an astonishing six-voice texture which is one of the great achievements of the western musical tradition.

This last paragraph from the excellent “Inkpot” website of classical music reviews:
<http://www.inkpot.com/classical/>

In other spheres of human experience, we have three-ring circuses, but not six-ring ones.

This may be one reason why department stores and museums are so inexplicably tiring (especially to men!): there are many more than three ‘voices’ competing for your attention.

And most primitive tribes have counting systems which involve just the numbers One, Two, Three, Many.

An experiment: have someone toss an unknown number of coins on a table, and see how fast you can count them. Up to three coins, and your response will be instantaneous. When the number is four, a slight hesitation begins to creep in, and as the number of coins gets larger, the hesitation gets correspondingly longer. Apparently our brains are wired so that we can “see” three as a single unity grouping.

This also seems to be true for certain animals and birds that can count to three, but anything more simply becomes ‘many.’

Some autistic individuals have brain wiring which allows (or condemns) them to instantly see and count much larger groupings. Autistics are also painfully aware of much more going on simultaneously in their environment, events which the normal human brain sidelines before ever reaching consciousness. Autistics probably live in a world not of 2 1/2 but of 25 or 40.

jts - February 15, 2005 - #23

... While many docs now use foley and other sound effects to beef up their soundtracks, I think it seems to take away some of the realness (in observational scenes). Yet, it seems that audiences may now identify "real" sound with what they hear in movies. So, the question is ... In your opinion, what do audiences perceive as "real" sound? And, how much natural sound are they comfortable with?

Walter Murch - *February 18, 2005 - #27*

I remember reading a review, in one of the audio magazines of the 1920's, of the new 'electrical' 78 rpm machines, with the then-revolutionary tube amplifiers and speakers. Earlier in the century, the classic Edison and Berliner machines had been completely mechanical and acoustical: it was the power of the human lungs and muscle alone that caused the needle to vibrate and etch the sound waveforms into the wax master - and the reverse happened when you listened at home: the waveforms in the shellac record caused the steel needle to vibrate and move a diaphragm which was at the focus of a megaphone-like horn, producing a sound loud enough to fill a small room.

Anyway, the reviewer in the 1920's simply declared that with electrical recording and reproduction, ultimate sonic reality had been achieved: listening to these new electrical records, you would believe that the musicians were actually in the room with you, if you closed your eyes. No further improvement was necessary or conceivable

Discounting some hyperbole, it is probable that most people listening to these records for the first time in the 1920's did have some kind of a 'lifting-of-the-veil' experience. And it is that - the experience of the removal of a veil - which hits us powerfully and gives us the feeling that we are witnessing a higher level of 'reality.'

But this is all relative: Today, we listen to those same records and it is hard to credit the reactions of those innocent listeners in the 1920's.

Probably some similar fate awaits us, in the opinions of future audiophiles.

Walter Murch - *March 6, 2005 - #35*

This is additional response to JTS's questions on realism in sound.

When we record certain sounds for films, we use multiple recorders at various distances for different perspectives, and we also bring along a "bad" recorder, such as an old consumer-level video camera, to create a sound that has a degree of distortion in it. This is added to the final track, like a dash of bitters, to give the sound a 'realistic' quality it wouldn't otherwise have.

In other words, a too-perfect recording of certain sounds (munitions, car engines, rocket takeoffs, crashes, explosions, waterfalls) makes them sound unrealistic.

Ben Burt, when he was creating the soundtrack for an Imax film on the Space Shuttle, recorded the liftoff at Cape Kennedy with five different state-of-the-art recorders --

some very close and some as much as two miles away. In the end, the sound that he used for the main thrust was from a dictating microphone stuck out the window of a car traveling sixty miles an hour.

It is interesting that all of the sounds listed above have very high degrees of white noise in them.

When we were making the soundtrack for “The Unbearable Lightness of Being” we found that for the documentary section in the middle of the film (about the Russian invasion of Czechoslovakia in 1968) we had to add certain artifacts to make it sound real: microphone handling clatter and the sound that analogue tape recorders make when they are coming up to speed during an in-progress loud sound. These same artifacts, if we had put them in the main part of the film, would have broken the illusion of realism instead of adding to it.

The common element to both the distorted sounds and the artifacts of recording is that they subconsciously tell the listener that they shouldn’t really be hearing what they are hearing -- that the sound is in some way too dangerous to be heard ‘cleanly.’ Of course it is all relative: the tracks we call distorted today would have been considered amazingly clean back in the 1930’s.

Luke Andrews - *March 9, 2005 - #46*

... I think it's fairly obvious that most movie watchers (or movie listeners, if you will) don't consciously notice sound except occasionally to note a particularly effective musical score or masterful bit of dialogue...

Walter Murch - *March 12, 2005 - #50*

Yes, I think that’s true, but I would put the emphasis on the word “consciously” rather than the word “notice.” In fact, this is one of the real strengths of film sound -- that it flies under the radar of consciousness and can have an effect on an audience without them knowing why. A sound-inspired feeling is usually attributed to the acting, or the photography, or the sets, or anything other than the sound itself - and I don’t think I would change this even if the option was granted.

A paragraph from one of my other posts:

The mental effort of fusing image and sound in a film produces a "dimensionality" that the mind projects back *onto the image* as if it had come from the image in the first

place. The result is that we actually see something on the screen that exists only in our mind and is, in its finer details, unique to each member of the audience. We do not see and hear a film, we hear/see it. For a slightly different view of the same thing, see these articles by Randy Thom on designing sound for film.

<http://www.filmsound.org/randythom/confess.html>
<http://www.filmsound.org/randythom/confess2.html>
http://www.filmsound.org/articles/designing_for_sound.htm

Jackson Braider - *February 16, 2005* - #24

... One of the wonderful aspects of *A Touch of Evil* is that the score emanates from visible elements in the film -- a radio the viewer can see. What about the omnipresent score a la Williams -- is there negotiation between the two camps, is it paper/rocks/scissors, or do you all end up reading entrails?

Walter Murch - *February 20, 2005* - #28

About the overuse of music: the director John Huston compared it to riding a tiger -- fun while you're on, but tricky getting off. In other words, if the amount of music in a film's first thirty minutes is beyond some critical threshold, you create expectations in the audience that require you to keep going, adding more music, because now the decision not to have music becomes a statement in itself, and it may be more than the film can sustain.

There is no question that music can create emotion in a film, but the danger is that if you use music to create the emotion, rather than to modulate or channel emotion already created by the drama itself, you are using music the way athletes use steroids: to gain a short-term advantage at the expense of long-term health. This has never prevented films from using this technique -- in fact the list of the ten top-grossing films is made up almost exclusively of films in this category.

On the other hand, there are films that use almost no music at all, or only music that comes from sources (radios, etc) within the film itself. Clouzot's *Wages of Fear* (1953), the example that Jackson quoted -- Welles's *Touch of Evil* (1958), and Lucas's *American Graffiti* (1973). *Graffiti* is in a class by itself because there is no original music at all, and yet there is a chameleon-like use of the abundant source music (42 songs), constantly switching its function from atmospheric source to emotional score.

This is almost exactly what Welles did in *Touch of Evil*, though there were three moments where Henry Mancini wrote original music that Welles used as score. In our 1998 recut of *Touch of Evil*, we removed one of these at Welles's request (the title music over the credits) and replaced it with a montage of six or seven overlapping source cues coming from car radios, souvenir shops, and nightclubs.

A film that is emotionally engaging builds up a kind of electrostatic charge of emotion that well-placed music can then help to channel (to ground, in electrical terms). It lets the audience know what to do with the emotion that has already been invoked by the characters and the situations. In my view this is the real power and function of music in films. The holding off of this channeling, such as in *Wages of Fear*, creates an almost-unbearable buildup of tension and emotion that is an intrinsic part of that film.

On the other hand, a film that uses continuous music to tell its audience how to feel at every moment deprives itself of one of the most powerful techniques that cinema has - the modulation of this static charge of expectation and release. And the resulting sense that, like real life, you are on a roller coaster without a safety harness.

Ideally, the amount and placement of music in a film should be such that the filmmakers have the freedom to use music or not at any particular moment.

Jackson Braider -- *February 28, 2005 - #33*

What would you make of trying to give drawn objects the capacity to make noise? I wonder if one of the challenges of the medium is that the sound of animation is the truest element in its presentation. "Truest" and "life-like" are terms that might satisfy your philosophical bent -- does sound help cartoon viewers disengage disbelief?

Walter Murch -- *March 6, 2005 - #37*

As late as 1936, live-action films were being produced that added only 17 additional sound effects for the whole film (instead of the many thousands that we have today). But the possibilities were richly indicated by the imaginative sound work in Disney's animated film *"Steamboat Willie"* (1928). Certainly they were well established by the time of Spivack and Portman's ground-breaking work on the animated gorilla in *"King Kong"* (1933).

In fact, animation -- of both the *"Steamboat Willie"* and the *"King Kong"* varieties -- has probably played a more significant role in the evolution of creative sound than has been

acknowledged. In the beginning of the sound era, it was so astonishing to hear people speak and move and sing and shoot one another in sync that almost any sound was more than acceptable. But with animated characters this did not work: they are two-dimensional creatures who make no sound at all unless the illusion is created through sound from one reality transposed onto another. The most famous of these is the thin falsetto that Walt Disney himself gave to Mickey Mouse, but a close second is the roar that Murray Spivack provided King Kong.

For more on animated sound, go to this interview with former KPFA engineer Randy Thom, who just won the Oscar for best sound effects for "The Incredibles" and who received three other nominations this year for his work on the sound tracks of animated films.

http://mixonline.com/mag/audio_randy_thom/

Viki Merrick - *February 16, 2005 - #25*

I've been thinking a lot about this conflicting sensory concept, or tricking the senses if you will. Today I was taking a shower with this bar of soap that looks wonderfully like poured butterscotch. When my daughter brought it home I smelled it and said Oh I love that smell, like butterscotch.... GEEZE it's ORANGE she replied !

I'm not especially obtuse, no more than another but when I used that soap today, and apart from still wanting to bite on it, insistent on its visual butterscotch appeal, I was surprised at just how orange it smelled - wonderfully more orange than you'd expect. I suspect it has nothing to do with whatever ingredients they may have put in but more with this unexpectedness of sensory pairing. I think this because when I went to listen to the first of the helicopter audio layering examples, my computer was on mute. I was staring at the scene and it was SO silent, and I got really scared and also sad, in a more distant way. The impact of the descending helicopters, the outpouring of dust and men was made even more intense for me by the silence. Does the mismatching of sensory input heighten the individual sensory experience?

Walter Murch - *March 6, 2005 - #36*

About the mismatching of sensory input heightening the individual sensory experience: yes, absolutely. In fact that is a very good way of describing what I call "metaphoric sound" in the Womb Tone section.

The tension produced by the metaphoric distance between sound and image serves somewhat the same purpose as the perceptual tension generated by the two similar but slightly different two-dimensional (flat) images sent by our left and right eyes to the brain. The brain, not content with this dissonant duality, adds its own purely mental version of three-dimensionality to the two flat images, unifying them into a single image with depth added.

There really is, of course, a third dimension out there in the world: the depth we perceive is not a complete hallucination. But the way we perceive it -- its particular flavor -- is uniquely our own, unique not only to us as a species but in its finer details unique to each of us individually (a person whose eyes are close together perceives a flatter world than someone whose eyes are further apart). And in that sense our perception of depth is a kind of hallucination, because the brain does not alert us to what is actually going on. Instead, the dimensionality is fused into the image and made to seem as if it is coming from "out there" rather than "in here."

In much the same way, the mental effort of fusing image and sound in a film produces a "dimensionality" that the mind projects back onto the image as if it had come from the image in the first place. The result is that we actually see something on the screen that exists only in our mind and is, in its finer details, unique to each member of the audience. We do not see and hear a film, we hear/see it.

This metaphoric distance between the images of a film and the accompanying sounds is -- and should be -- continuously changing and flexible, and it often takes a fraction of a second (sometimes even several seconds) for the brain to make the right connections. The image of a light being turned on, for instance, accompanied by a simple click: this basic association is fused almost instantly and produces a relatively flat mental image.

Still fairly flat, but a level up in dimensionality: the image of a door closing accompanied by the right "slam" can indicate not only the material of the door and the space around it but also the emotional state of the person closing it. The sound for the door at the end of "The Godfather," for instance, needed to give the audience more than the correct physical cues about the door; it was even more important to get a firm, irrevocable closure that resonated with and underscored Michael's final line: "Never ask me about my business, Kay."

That door sound was related to a specific image, and as a result it was "fused" by the audience fairly quickly. Sounds, however, which do not relate to the visuals in a direct way function at an even higher level of dimensionality, and take proportionately longer to resolve. The rumbling and piercing metallic scream just before Michael Corleone kills Solozzo and McCluskey in a restaurant in "The Godfather" is not linked directly to anything seen on screen, and so the audience is made to wonder at least momentarily,

if perhaps only subconsciously, "What is this?" The screech is from an elevated train rounding a sharp turn, so it is presumably coming from somewhere in the neighborhood (the scene takes place in the Bronx).

But precisely because it is so detached from the image, the metallic scream works as a clue to the state of Michael's mind at the moment -- the critical moment before he commits his first murder and his life turns an irrevocable corner. It is all the more effective because Michael's face appears so calm and the sound is played so abnormally loud. This broadening tension between what we see and what we hear is brought to an abrupt end with the pistol shots that kill Solozzo and McCluskey: the distance between what we see and what we hear is suddenly collapsed at the moment that Michael's destiny is fixed.

THIS moment is mirrored and inverted at the end of "Godfather III." Instead of a calm face with a scream, we see a screaming face in silence. When Michael realizes that his daughter Mary has been shot, he tries several times to scream -- but no sound comes out. In fact, Al Pacino was actually screaming, but the sound was removed in the editing. We are dealing here with an absence of sound, yet a fertile tension is created between what we see and what we would expect to hear, given the image. Finally, the scream bursts through, the tension is released, and the film -- and the trilogy -- is over.

The elevated train in "The Godfather" was at least somewhere in the vicinity of the restaurant, even though it could not be seen. In the opening reel of "Apocalypse Now," the jungle sounds that fill Willard's hotel room come from nowhere on screen or in the "neighborhood," and the only way to resolve the great disparity between what we are seeing and hearing is to imagine that these sounds are in Willard's mind: that his body is in a hotel room in Saigon, but his mind is off in the jungle, where he dreams of returning. If the audience members can be brought to a point where they will bridge with their own imagination such an extreme distance between picture and sound, they will be rewarded with a correspondingly greater dimensionality of experience.

The risk, of course, is that the conceptual thread that connects image and sound can be stretched too far, and the dimensionality will collapse: the moment of greatest dimension is always the moment of greatest tension.

Viki Merrick - *February 16, 2005* - #26

Can you tell us a story about one of your suppressed volcanoes?

Walter Murch - *March 6, 2005 - #38*

One of the volcanoes that I have to keep under control is my love of dynamic range. When I started out, in the late sixties, the accepted dynamic range was 6 decibels: that is to say, the loudest sounds in the film could be six decibels over the level of average dialogue. This was simply the nature of what was known as the Academy standard optical track, and it had been that way since the late 1930's. In other words, the soundtracks of films of the late 1960's were technically the same as the soundtracks of the late 1930's.

There had been major improvements 'upstream' -- better microphones, magnetic sound recorders, etc. -- but as far as the ultimate delivery was concerned, the monophonic variable area optical track for *The Godfather*, say, was the same as the optical track of *Gone With The Wind*.

I wanted more, and I did what I could to 'trick' the projectionists to give my soundtracks an extra two or three decibels: I would mix the opening music a little bit low in the hopes that they would boost the level accordingly, and then later on in the film I would choose three places to completely saturate the optical track for maximum effect. I could think of doing this because there was no completely agreed-upon standard for sound levels of projection. There had been, back in the days when the studios ran their own theater chains up to the late 1940's, but when they were forced to divest themselves of control of exhibition, that link was broken and it was flapping in the breeze by the late 1960's.

This trick worked well up to the mix of *Godfather Part II*, where I went too far (here is the volcano part). The mix sounded fine in the relatively small mixing room at American Zoetrope, but when the sound was played in the big theaters, holding more than 1,000 people, my trick backfired. If the mix was played in those theaters at a level where the dialogue was loud enough, the loud sounds were too loud. And vice versa: if the film was played so that the loud sounds were at the right level, the dialogue was too low.

As a result, we had to remix the film to 'even out' the dynamic peaks and valleys (effectively, raising the level of the low dialogue by three or four decibels), and a whole new series of prints had to be made at considerable expense to Paramount. It was a searing, agonizing experience for me, one I never wanted to repeat.

All this happened in 1974. Within a few years, Dolby came along with their early "A-type" stereo optical tracks that lifted the dynamic ceiling by at least three decibels, bringing the first real change in the optical track since the mid-1930's. Dolby also brought around a standard "Dolby Level" which -- theoretically -- is the standard by which all films should be played back in a theater.

Since then, of course, there have been several major revolutions in the technical nature of theatrical reproduction of film sound, but I still keep that early lesson in mind: even though the sound is coming from a digital source far removed from those primitive Academy optical tracks, I never let the level of average dialogue fall below a certain level, and I never let the loud sounds get completely overwhelming.

Walter Murch - *March 7, 2005 - #39*

When I was working in England on "Julia" in 1977, the mixing theater where we were working was the same one that Stanley Kubrick used, and the man who ran the desk, Bill Rowe, was the man who mixed many of Kubrick's films.

One day I noticed a grease pencil mark on the master monitor level control and asked what it was. "That's Stanley's Level," Bill said. "It is three decibels below normal and he always makes us mix his films at that setting."

What this effectively did, Bill went on to explain, was to compress the dynamic range of Kubrick's films even more than the six decibels that was standard in 1976. Because the monitor level is lower, the dialogue is recorded at a higher level so that it sounds normal, but the loud sounds were still limited by the "brick wall" of optical saturation.

This trick limited the dynamic range of Kubrick's optical-track film to three decibels, which is roughly what the dynamic range of broadcast television is.

I suppose somewhere in Kubrick's past there had been a searing encounter with the devil of dynamic range similar to my experience on Godfather II. And in fact when we were preparing to mix Apocalypse Now we screened Dr. Strangelove, and sure enough, the machine gun fire in the film was only three decibels louder than average dialogue.

Jay Allison - *March 7, 2005 - #40*

In the theatre, the audience has to accept whatever playback level the projectionist chooses and whatever dynamic range you mixed. In radio, the audience controls the playback level, and they will correct your dynamic range if they don't like it. They'll turn your volume up and down and if they get tired of that, they turn you off. I suppose that's our devil of dynamic range.

Do you imagine some "average" movie theatre or do you mix for the best listening environment? I always wish I could mix for an audience in a quiet room, but that's not the radio audience. I tend to mix for car listening because that's where most of them are. This means some compression. I also tend to keep my playback volume quite low for the last pass. It seems to approximate the effect of road noise. You can tell really quickly if some section is very loud or quiet because it's very apparent out at low level. I suppose that's sort of what Kubrick's 3db-down trick did too.

Walter Murch - *March 8, 2005 - #41*

It used to be possible to imagine an average movie theater, but now with various-sized multiplexes and DVDs and in-flight films, etc. it is simply impossible to prefigure an ideal mix. As you can imagine, this is an area that is under considerable flux at the moment.

Out here at Lucasfilm (where I am editing "Jarhead") Gary Rydstrom just finished a 'near field remix' of the Pixar film Toy Story for a new DVD that is being issued. Amplifiers for DVD decks are now being fitted with settings that compress the dynamic range so that films with theatrical dynamics can be watched at low levels at home with no loss of information.

We do have something I call a "popcorn loop" which we run through the monitor when listening to the playback of the mix - a recording of a theater full of people, air-conditioning, rustling, low murmur - which serves the same purpose as your imagining the car motor. Indeed, Stanley Kubrick's low monitor setting and your low playback for the final pass are identical strategies. The popcorn loop achieves the same thing while allowing the monitor level to stay at the standard setting. The problem with listening at low monitor levels is the response of the human ear, which is non-linear - at lower levels there is a progressive dropoff in sensitivity to low and high frequencies (the Fletcher-Munson curve).

On the film "Ghost" I recorded the audience during a preview of the film, using directional microphones pointing away from the screen, and we would run this in the monitor chain when we were checking the final mix, to make sure that the level of dialogue was high enough during scenes where the audience was laughing loudly.

The devil is even more devilish these days because the theatrical digital formats (Dolby, DTS, SDDS) can give a dynamic range upwards of 25 decibels, compared to the 6 available in the 1960's. This is a tremendous amount of power - loud sounds can reach 110db, which is close to the threshold of pain - and putting it in the wrong hands is like giving a Ferrari to a teenager. In the old pre-Dolby days, mixing was like driving

Aunt Mabel's Dodge Saloon: there was a technical "Speed-limit" beyond which you couldn't go even if you wanted to (unless you were being particularly devious, as my experience on Godfather II).

One decibel is the smallest increment that the human ear can detect, and because the decibel scale is logarithmic, a sound 10 db louder than another has ten times the energy (wattage) of the lower sound. Since $25 \div 10 = 2.5$, and $10^{2.5} = 316$, this means the loudest sounds in films today can easily have more than three hundred times the energy of average dialogue. In pre-Dolby days, the loudest sounds in films could only have about four times the energy of average dialogue.

Jay Allison - March 8, 2005 - #42

Walter, do you wander into the Springfield Mall Cineplex in Wherever just to see how your work really sounds to the world? If so, what's that experience like? How different is it from the way you mixed it? I suppose, if your walla-walla popcorn track trick works, it should be pretty close.

Walter Murch - March 8, 2005 - #43

Yes, I do go out and listen to the mix in the 'real' world - it is sometimes a painful experience (left and right channels switched, no surrounds, no low frequencies, etc. etc.) When we did the Dolby Stereo re-release of American Graffiti in 1978, one theater had no sound coming from their surrounds, and when I talked to the manager about it he assured me that this was because the 'director wanted it that way.'

Mostly, though, things sound all right, which is a testimony to Dolby and THX, and the generally high standards of the average theater management.

But even so, every acoustical environment is subtly different, the way every violin sounds different, and as a result certain things the soundtrack are emphasized and certain others are less noticeable. In another theater, the opposite will happen. I file it all away somewhere in my subconscious and then put it to use in the next mix.

Walter Murch ~ *February 22, 2005 - #29*

Here is an issue common to radio and film:

Interview someone, or film someone in a real-life situation, and record their voice. Then transcribe those words to text and ask them to read it, and record this second version. Then compare the original recording with the first. Almost invariably, there is a wooden-ness to the second recording - sometimes more extreme than others.

The same phenomenon can be observed comparing teachers who read lectures from completely prepared text vs. teachers who give lectures 'spontaneously' from bare outlines.

Even trained actors have a hard time overcoming this: witness the tone of the presenters reading off a teleprompter at the Oscar ceremony.

How do members of the NPR news team train themselves to sound natural and spontaneous although they are frequently reading from prepared text?

What causes this wooden-ness? Is it the fact that the brain is doing two things at once: Reading and speaking? And the 'wood' is the tell-tale sign of this 'double think'?

Jay Allison - *February 22, 2005 - #30*

I think it's similar to writers who don't want to talk about their work until it's written. It saps the imperative to tell. Actors spend lifetimes trying to harness "the illusion of the first time." With reading, the important connection is to the brain. What makes you listen is hearing the sound of the brain at work, the rhythm of thought. It has little to do with the throat.

Jackson Braider - *March 9, 2005 - #45*

If I were a betting man, I would say that in the main, pubrad producers don't do volume well. We have been working so hard to achieve the ultimate intimacy between the lips at the mic and the ear at the speaker, "loud" seems kind of rude, crass... many of us envy the sheer number of decibels you wield. If we knew what to do with just volume -- thinking of volume as a force for good -- people would actually laugh at our jokes, rather than merely chuckle.

Walter Murch - *March 10, 2005 - #47*

... many of us envy the sheer number of decibels you wield ...

See the article by Richard Corliss on Jean Shepherd

<http://www.time.com/time/columnist/corliss/article/0,9565,168458,00.html>

"Shep" was my hero in the mid 1950's, when I was a teenager. He had a show on WOR in New York which began at midnight and went to 5am every night. Later he switched to Sunday nights 8 to midnight.

He could have single-handedly invented the idea of intimate radio. I don't know.

But he also had a trick which he would play every couple of weeks, called the "Hurling the Invective." He would explain the rules in a conspiratorial tone, and I would get sweaty palms thinking about what was going to happen...

He would go on dead air for about fifteen seconds. During that time you were to turn off the lights in your room, open the apartment window, take your table radio (what other kind were there) and put it on the windowsill, turn up the volume to maximum and then wait. I can still here the hiss of WOR and the room tone in Shep's studio when I had cranked it up all the way. Then, after what seemed a very long time, Shep would yell into the microphone, like the Drill Instructor for the Army of Greenwich Village: "ALL RIGHT YOU MEATHEADS! I KNOW WHAT YOU'RE UP TO - WHAT MADE YOU THINK YOU COULD GET AWAY WITH IT..." and so on for thirty seconds or so.

Then the air would go dead, and you had to reach around, turn the volume down, and get the radio back inside as quickly as possible.

If you were lucky, there would be five or six other listeners in your block who would have done the same thing, so the effect would be cumulatively catastrophic to the peace and quiet of Manhattan. Maybe that's where I developed my love of reverberation... Back in the dark of the room there would be a long pause, and a conspiratorial snicker, and Shep's 'normal' programming would continue.

The concept of these invectives were lifted by Paddy Chayevsky and put into the film "Network" where the Peter Finch character would ask his listeners to do the same thing with their television sets (a little improbable) and then yell "I'M MAD AS HELL AND I'M NOT GOING TO TAKE IT ANY MORE."

Talk about volume...

Walter Murch Links

FilmSound - Walter Murch Articles & Interviews:

www.filmsound.org/murch/murch.htm

***Time Magazine* - An Article about Jean Shepherd:**

www.time.com/time/columnist/corliss/article/0,9565,168458,00.html

Film Freak Central - A Conversation w/ Walter Murch:

www.filmfreakcentral.net/notes/wminterview.htm

About Transom

What We're Trying To Do



Here's the short form: Transom.org is an experiment in channeling new work and voices to public radio through the Internet, and for discussing that work, and encouraging more. We've designed Transom.org as a performance space, an open editorial session, an audition stage, a library, and a hangout. Our purpose is to create a worthy Internet site and make public radio better.

Submissions can be stories, essays, home recordings, sound portraits, interviews, found sound, non-fiction pieces, audio art, whatever, as long as it's good listening. Material may be submitted by anyone, anywhere - by citizens with stories to tell, by radio producers trying new styles, by writers and artists wanting to experiment with radio.

We contract with Special Guests to come write about work here. We like this idea, because it 1) keeps the perspective changing so we're not stuck in one way of hearing, 2) lets us in on the thoughts of creative minds, and 3) fosters a critical and editorial dialog about radio work, a rare thing.

Our Discussion Boards give us a place to talk it all over. Occasionally, we award a Transom.org t-shirt to especially helpful users, and/or invite them to become Special Guests.

Staff

Producer/Editor - Jay Allison

Web Director/Designer - Joshua T. Barlow

Editors – Sydney Lewis, Viki Merrick, Chelsea Merz, Jeff Towne, Helen Woodward

Web Developers - Josef Verbanac, Barrett Golding

Advisors

Scott Carrier, Nikki Silva, Davia Nelson, Ira Glass, Doug Mitchell, Larry Massett, Sara Vowell, Skip Pizzi, Susan Stamberg, Flawn Williams, Paul Tough, Bruce Drake, Bill McKibben, Bob Lyons, Tony Kahn, Ellin O'Leary, Marita Rivero, Alex Chadwick, Claire Holman, Larry Josephson, Dmae Roberts, Dave Isay, Stacy Abramson, Gregg McVicar, Ellen Weiss, Ellen McDonnell, Robin White, Joe Richman, Steve Rowland, Johanna Zorn, Elizabeth Meister



ATLANTIC PUBLIC MEDIA

Atlantic Public Media administers Transom.org. APM is a non-profit organization based in Woods Hole, Massachusetts which has as its mission "to serve public broadcasting through training and mentorship, and through support for creative and experimental approaches to program production and distribution." APM is also the founding group for *WCAI & WNAN*, a new public radio service for Cape Cod, Martha's Vineyard, and Nantucket under the management of *WGBH*-Boston.

This project has received lead funding from the Florence and John Schumann Foundation. We receive additional funding from The National Endowment for the Arts.