

# CLEAR DENSITY, DENSE CLARITY

## by Walter Murch

### Volcanos and Glaciers

#### A cautionary note

Film director Ingmar Bergman had read Stravinsky's words of advice to young conductors: "Restraint! Always restraint!" But then Bergman watched with dismay as friends of his, young conductors "without even the vestige of a volcano within them," nodded in agreement, raised their batons and... observed restraint.

When Bergman finally got to see Stravinsky conduct in person, however, he was astonished to find him swaying wildly on the podium, thrashing at the orchestra and using his baton like a whiphandle.

Most of us are searching for some 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 carried 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.

So most of what we are going to be discussing tonight are cautionary tales: methods and observations that I have developed over the years for coping with my own particular glaciers and volcanos. They work for me, but I suggest that you take them with the appropriate grains of salt - their usefulness to you may be more for the insights they offer into one person's search for balance rather than for any of the specific methods that the search itself has produced.

### 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.

The general level of complexity, though, has been steadily

increasing over the seven 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. Sixty 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 tonight 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 tonight 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 right (where we had violet); and something else I'll call *Embodied* sound, which I'll put over on the left (red).

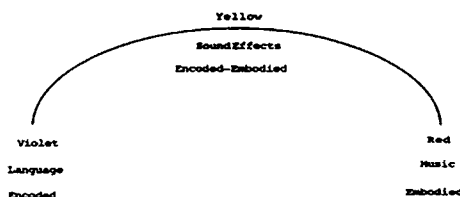
The clearest example of Encoded sound is *speech*.  
The clearest example of Embodied sound is *music*.

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 is happening right now, as you listen to this lecture. The meaning of what I am saying is 'encoded' in the words I am using. 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 lie 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 monotonously even - 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. Schonberg 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.'

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

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 waveforms 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's sandwich

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.

The problem, of course, is that sooner or later (mostly sooner) this 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 *premixes*, of about 30 tracks each. But I do the same thing today even though I may have three times as many faders as I did back then.

The six premix layers were:

1. Dialogue
2. Helicopters
3. Music (The Valkyries)
4. Small Arms Fire (AK47's and M16s)
5. Explosions (Mortars, Grenades, Heavy Artillery)
6. Miscellaneous (Footsteps, Winds, Fires, Shrapnel, etc)

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 have in common is that in each art 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 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. I mixed this music 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 (miscellaneous). 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 onscreen, 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.

## 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!

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 footstep 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.

## 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 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 a little 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 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.

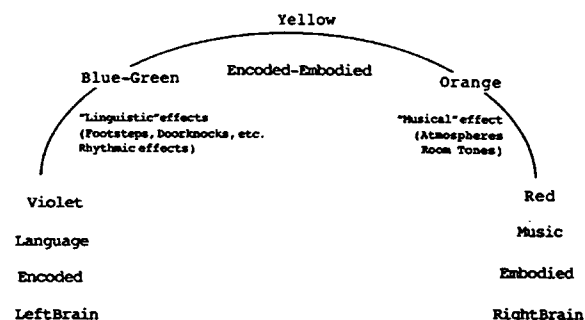
But if you blend sounds from different parts of the spectrum, you get some extra latitude. Dialogue and Music can live together quite happily. Add some Sound Effects, too, and everything still sounds transparent: two people talking, with an accompanying musical score, and some birds in the background, 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 achieve 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 contribute 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. And yet the audience still believes the sound is coming from the actors, even though they may be moving around freely. 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. Even "orange" sound effects (atmospheres, room tones) are given a full stereo treatment. "Blue-green" sound effects like footsteps, however, are usually given the same "center" placement as dialogue, unless the filmmakers want to call special attention to the steps, and then they will be placed and moved along with the action.

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 origin to the visual, allowing the image to "steer" the source of the sound. When the sound is Embodied, however, and relatively 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 do both simultaneously clearly demonstrates the differences between our two hemispheres.

Getting back to my problem on *Apocalypse*: it appeared to be caused by the fact that I had 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 over 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) Miscellaneous (sounds ranging from blue to orange)
- 5) Helicopters (orange music-like drones)
- 6) Valkyries Music (red)

This does not mean I *always* had five layers cooking. Conceptual density is something that should obey the same

laws 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 that should not be surpassed thoughtlessly..

So if I had five layers going and 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 the music should actually be louder there than anywhere else. But for story reasons I wanted to hear the 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 here are those five layers:

- 1) Dialogue
- 2) Other voices, shouts, etc.
- 3) Helicopter
- 4) AK-47's and M-16s
- 5) Mortar fire

Under the circumstances, the music was the layer that had to be sacrificed. 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.

The bottom line is that the audience is primarily involved in following the story: the right thing to do is ultimately whatever serves that endeavor, in the widest sense. When this helicopter landing scene is over the impression should be that everything happened at once - Density - yet that everything was heard - 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.

Good evening, thanks for listening, and happy mixing!

February 22, 1998