Chapter 4: SOUND

In 1927, several minutes of crooning by Al Jolson in the otherwise silent *The Jazz Singer* signaled an end to the silent era. By 1929, the "talkies" were ascendant, and silents became almost nonexistent on the commercial film circuits.

The invention of an actual soundtrack both provided enormous practical advantages and changed the creative potentials of cinema. Theater operators no longer had to depend upon often capricious musicians, and a filmmaker could be certain that audiences would hear the same music he intended for the film. But sound strengthened the visuals as well by eliminating the need for titles. Jarring breaks in the middle of powerful scenes could be eliminated. The words on a heroine's lips could now be heard, and captions explaining a situation could be dropped. As the villain tied the heroine to the tracks, every "Help, save me!" no longer needed spelling out, nor did the explanation that the train was due at 5:30. The coming of the age of sound allowed the pace of visual images to increase manyfold and freed the visual image from the burdens of written language. The image could simply be an image.

Sound Terminology

The sounds contributing to motion pictures can be divided into three categories: spoken language, music, and ambient sounds (those contributing to the atmosphere or forming part of the natural surroundings). These sounds in turn are modified by **volume** (as when one person talks loudly and another softly), **pitch** (the "highness" or "lowness" of the sound), and **timbre** (the color or tone quality which enables sounds to harmonize--or be discordant--with other sounds). We often recognize voices by a combination of all three of these qualities.

Spoken language appears in two forms: dialogue and narration. If a character is visible (or assumed to be just beyond the limits of the frame) the words spoken are **dialogue**. Although written specifically for an audience, dialogue is spoken as if no public audience were present. The speaker aims his words at other characters or simply talks out loud in soliloquy. He or she appears unaware that a public audience overhears every word. Viewers become voyeurs overhearing the personal conversations of people.

Narration, on the other hand, is spoken language specifically directed to a theater audience. A voice addresses the audience, providing information not otherwise available. Technically, a narrating voice takes two forms: a **narrator**, who is either a character in a film or directly involved in its events, consequently holding a particular point of view; and a **commentator**, who is an omniscient source of supposedly unbiased information. The **narrator** appears most frequently in fiction films as one of the characters (often the main character), and the **commentator** (who usually remains unseen or, as in a football game, appears briefly on camera) provides the explanatory voice of most documentary films.

A **commentator** normally takes care to make her voice sound objective and trustworthy, while a **narrator**'s personal involvement appears to color his perception of events, frequently producing various tones of voice depending on his emotional state. Hence, a narrator is not always reliable or completely accurate since he has a personal interest in the events he describes. In some cases, a narrator will say one thing, while the visual images of the film will tell viewers that the truth is considerably different from what the narrator thinks it is. The result is dramatic irony. For example, a narrator might explain that the girl he is with loves him passionately, while the viewer sees that she obviously dislikes him.

Any spoken language not seeming to come from images on the screen is called **voice-over**. Sometimes, a narrator or commentator will appear on screen to talk directly and visibly to an audience, but most narration is voice-over. Occasionally, dialogue becomes voice-over by not synchronizing with the visual image it accompanies.

Music also has two forms: **local** and **background**. As with dialogue and narration, local and background music can be distinguished by their relationship to both scene and spectator. **Local music** originates in the scene itself and can be heard by the characters in the film as well as by the audience. An orchestra at an opera or a jukebox in a tavern provides local music. Local music often signals how we should feel about a character or situation. The music people listen to tells something about them, much as the visual images of their living quarters reveal personality. Local music also adds credibility to a scene. A

western bar needs western music, and a bar catering to young singles needs popular music.

Background music functions much like narration. It "tells" us how to respond to a film's visuals but has no source identifiable within the film. Background music heightens the dramatic impact of a film in several ways. In many films, a musical theme resurges throughout in order to create various attitudes in the viewer and to signal particular moods. Background music also provides a sense of locale, signals the passage of time, anticipates what is to come by echoing foreboding or romantic strains, and establishes and maintains moods. In older movies, background music even identifies character qualities--sinister music signals the villain, while sweet sounds accompany the hero and heroine.

Rather than using metronomes, filmmakers measure tempo in **clicks**. A **click** is the number of frames for each musical beat, so a 24 click would have 24 frames per beat. Since film moves at 24 frames per second, the click rate would be 60 beats per minute.

Ambient sound includes all nonverbal and nonmusical sounds contributing to a film. Ambient sounds are noises naturally accompanying a scene, such as frogs croaking at a lake, dogs barking at night, a plane's wheels touching ground, footfalls on a gravel walk, a door slamming shut, wind blowing through trees, or an owl hooting at midnight. The sources of ambient sounds need not be, and usually are not, seen. In life, the source of most sounds remains unseen, even though we usually know the source and could probably see it if we wished. The wind blowing through trees or a dog barking in the distance loses none of its immediacy by being heard and not seen.

Ambient sounds work much like background music in intensifying mood. A coyote howling at the moon brings feelings of loneliness, and the sudden roar of a car engine creates apprehension. Ambient noises strengthen the feeling of reality created by visual images and make the viewer mentally extend the limits of the frame.

Two elements of ambient sound deserve special mention: **artificial sound** and **silence**. **Artificial sound** is created synthetically, usually with electronic instruments or by techniques such as playing tapes backwards. It shows up most

frequently in experimental films or in films about science fiction or altered states of consciousness. **Silence**, on the other hand, is effective only as a hiatus in the presence of sound. When all sound ceases, silence can strongly affect mood. In a chase through a forest, for example, silence heightens fear and anticipation; in a love scene, it intensifies the romantic atmosphere. Anyone who has seen one of Hitchcock's thrillers knows how silence can arouse intense emotions.

Producing the final soundtrack of a film, called a **sound composite**, is a matter of patching together a variety of sounds in a process called the **mix.** Hundreds of tracks are recorded during both production and postproduction, and parts of these tracks must be combined for the composite. Spoken language, music, and ambient sound must all be coordinated with visual images.

Dubbing, the process of matching voice with the lip movements of an actor on the screen (also called **looping** or **lip sync)**, offers the filmmaker several advantages. Since all sound can be added later, the normal noise on a set can be ignored, allowing the filmmaker to work in faster, easier conditions. A breeze through trees or across a microphone, a cough, or a dropped glass will not interfere with filming and cause numerous retakes. Voices may even be substituted in certain cases, as when singing is required of an actor or actress with salty tonsils.

Technically, **dubbing** refers to replacing the entire dialogue track (as in dubbing a foreign film into English), while replacing dialogue originally recorded for a track is known as **looping.** Most **looping** is done by **ADR**, which originally meant "Additional Dialogue Recording" but now stands for "Automatic Dialogue Replacement." While loops could only be set up for a few seconds at a time, ADR allows an entire scene to be run though repeatedly. (Looping provides better lip sync, but the performer can work on the emotional flow of the scene during ADR.)

Other sound effects ("FX") get added by the sound editors. Some effects come from the sound library (such as applause), while others are produced by a **Foley artist.** During the Foley process, customized sounds are produced to match the visuals on the screen, such as the sounds of a dog walking on a sidewalk, a tree falling, a fist hitting a jaw, a person slogging through mud, or someone climbing a tree. The sound of ice cubes in movies might come from pencils in a glass but not from actual ice cubes.

Because sound provides information different from that coming from visuals, much effective sound is nonsynchronous. **Synchronous sound** simultaneously combines a visual image with the sounds apparently issuing from it. When characters speak on the screen and a viewer can tell by their lip movements that what they are saying is identical with what the viewer actually hears, the sound is **synchronous**. This is the most common use of sound.

Nonsynchronous sound, on the other hand, is not matched to the image nor recorded during shooting. Instead, it joins visuals from one source with sounds from another. A speaker's voice is often heard, for example, but the camera focuses on the person the speaker addresses, not showing speaker and words together. Hence, the viewer experiences both the speaker's comment, complete with voice inflections, and the reactions of the listener. Other applications of nonsynchronous sound occur when visual and audial images have no natural relationship, as when visuals show a bomber while the audio plays a Gregorian chant, or when visual scenes of a political convention are accompanied by the "baaing" sound of sheep. Nonsynchronous sound **counterpoints** visual components.

Critics also speak of **diegetic** and **nondiegetic sound**. Although these terms parallel synchronous and nonsynchronous sound, they are not the same. **Diegetic sound** consists of the actual sounds belonging in a scene (whether originally recorded live or later post-synchronized). Diegetic elements include actions and dialogue in their normal space and time (although these can rarely be shown fully in a film). These are the sounds we would be capable of hearing if we were in the space enveloping the story.

Nondiegetic sound does not come from the space enveloping the story. Background music is nondiegetic. So are the words of a commentator, or the noise of a crowd cheering as a climber neared the top of Mount Everest. Both diegetic and nondiegetic sound represent sound as a viewers perceive it, since they have no way to know whether sound is shot at the same time as visuals. The test for diegetic sound is the test of "could": could a participant in the scene recognize the sound as coming from the actual scene?

The term **MOS** refers to any segment of film taken without sound. The letters MOS come from early foreign directors who wanted pictures taken "mid out [without] sound."