

# The sound of sound

A Brief History of the Reproduction of Sound in Movie Theaters

*by Rick Altman*

---

---

## **What should sound sound like?**

When you stand in the stereo showroom, or when you move your speakers around the family room, how do you know when the sound sounds right? When THX creator Tomlinson Holman designs crossover circuitry or specifies speaker type and placement, how does he know when he has it right? Discussing the home THX system, Holman stresses the importance of making films sound in the home just as they do on the dubbing stage or in the theater.[\[1\]](#) But that only begs the question--How do the people who design theater acoustics know when the sound sounds right? In fact, how do any of us decide whether a sound reproducing system represents the original sounds properly?

Clearly, it's not simply a question of fidelity to the original sound source. How many of us have actually heard Toscanini at La Scala or the final mix of Star Wars on the dubbing stage? Yet even though we've never heard the original, we have very clear ideas of how the copy should sound. In fact, depending on our hearing experience, we harbor quite divergent ideas about how Toscanini--and everything else--should sound. When Aesop's Country Mouse paid a visit to his city-dwelling cousin, he found the urban soundscape not to his liking at all. Indeed, if the City Mouse were to drop in on his country cousin, he would probably conclude that there is something 'unnatural' about a nocturnal soundtrack featuring no more than the sound of crickets. For we learn to hear by hearing, and in doing so we form quite specific notions about how sound should sound.

To an extent as yet unrecognized, cinema sound depends heavily on the very same process. Though it is typically studied as an independent phenomenon, the history of film sound cannot be

properly understood unless it is correlated with the major sound practices of each era. By listening to available sound, each generation learns just what constitutes acceptable sound. But since the sound available to each generation changes with transformations of taste and technology, it stands to reason that the standards by which cinema sound is judged must vary from decade to decade.

These changes are reflected throughout the history of cinema through such developments as increased frequency and dynamic range, modifications in the role accorded to music, shifts in the relationship between sound scale and image scale, and innovations like stereo or surround sound. Changing notions of how sound should sound are thus readable through the history of various cinema sound practices. In particular, a fascinating record of varying spectator expectations regarding sound is encapsulated in decisions as simple as the placement of speakers in the theater.

Unless they happen to be sitting next to a misbehaved surround speaker or watching a stereo film with a dead channel, most people pay little attention to the location of speakers. Indeed, theater designers have systematically followed Hollywood's tendency to dissimulate technology inside the theater (even though it is often touted on the marquee outside), so much so that most spectators have literally never even seen a cinema speaker (other than the surrounds). Yet the location of speakers is remarkably indicative of contemporary presuppositions about sound. Indeed, the history of cinema sound may conveniently be divided into five periods, each featuring a different speaker configuration designed to match cinema sound to current standards of how sound should sound.

During the latter half of the 1900s, the cinema industry entered into a profound crisis. With the rise of the nickelodeon, the number of theaters exhibiting films had grown so rapidly that producers were unable to meet demand. Forced to show the same film as the competitor down the block, theater owners looked to sound practices to differentiate their products.

Where previous films had been only intermittently accompanied by a vaudeville orchestra, a lone untrained pianist, or not at all, exhibitors now sought to raise the tone of their establishments through sound. Eschewing popular music and ragtime, theaters

instead increasingly featured light classical accompaniment performed by competent musicians.

Before continuous musical accompaniment became the rule, however, enterprising exhibitors labored mightily to make films sound like live theater. From 1908 to the early Teens, the human voice commonly accompanied film projections. During the late Aughts, films were often supplemented by carefully rehearsed actors speaking lines in sync with the image. Indeed, there were enough "talking picture" troupes (calling themselves Humanovo, Actologue, Ta-Mo-Pic, and the like) to support a New York academy dedicated to training behind-the-screen actors. For theaters unable to afford the full troupe, a live narrator was often used to secure the narrative coherence of films longer on spectacle than clarity.

The real attention-getters, however, were the dozens of experiments with sound-on-disc synchronization. The first of these systems to achieve a modicum of success was Cameraphone, an avowed attempt to can vaudeville performances--image and sound --for inexpensive distribution to the hinterlands. With one hundred locations by the end of 1908, and continued expansion in 1909, Cameraphone was soon joined by a bevy of imitators: Vivaphone, Electrograph, Phoneidograph, Picturephone, Phonoscope, Gaumont's Chronophone, the British Cine-phone, and many others, culminating in 1913 with Edison's ill-fated Kinetophone. Every one of these systems, it should be noted, aims not at providing synchronized musical accompaniment, but at reproducing the human voice (in keeping with the current generic term for the phonograph: "talking machine" ). It is thus hardly surprising that, after many experiments locating the loud speaker near the projector (the simplest solution) or to the side of the screen (the traditional arrangement for combined slide and phonograph presentations), virtually every early synchronized sound system settled on a speaker location behind the screen (fig. 1), where the resultant sound could most easily be assimilated to the body of the characters observed on the screen.

Primarily the province of undercapitalized, independent enterprises, sound-on-disc fell prey by the early Teens to a systematic producer campaign to feature continuous musical accompaniment and

narrative sound effects in preference to the human voice. By the mid-Twenties, light classical orchestral or organ accompaniment had become so pervasive as to relegate speech entirely to the written form of inter-titles.

It is thus not so much the technology that changes with the Vitaphone system that precipitated Hollywood's conversion to sound in the late Twenties. After all, even though it benefits from Bell Laboratories' advances in electric recording and amplification, Vitaphone is still nothing more than an improved version of the dozens of sound-on-disc systems popular around 1910. Important changes had come not in technology alone, but also in audience expectations regarding sound. No longer was speech the film accompaniment of choice. Instead, discs were called upon to provide the expected musical support for films that continued to carry speech on intertitles.

When the Vitaphone system was first exploited commercially in 1926, we thus find an entirely new speaker configuration, again reflecting current assumptions about what kind of sound merits reproducing. While one speaker is maintained behind the screen--in order to reproduce infrequent speeches, like Will Hays's introduction to the initial Vitaphone shorts program--the other is located in the orchestra pit, pointing upwards, simulating the sound of the orchestra it has displaced (fig. 2). Pity the poor projectionist, frantically switching back and forth from one speaker to the other, according to the type of sound reproduced.

From the films produced for the Vita-phone system during its first year of operation, it is clear that Warners thought of synchronized sound as serving alternately one of two purposes: either to replicate music or to serve as a public address system (hardly surprising, since the Bell Labs research leading to Vitaphone had included the development of a new public address system and a new phonograph, the Orthophonic Victrola). The first Vitaphone shorts systematically stress musical uses, while the first year's features range from Don Juan's ninety- nine percent musical accompaniment (August 1926) to tentative experiments with what we might call "megaphone speech" in *The First Auto* (June 1927). While the latter film uses intertitles for all normal conversation, the Vitaphone system is pressed into service each time a character shouts or calls

out to another character, thus taking advantage of the public address-like amplification provided by the behind-the-screen speaker quite properly identified as loud.

Designed primarily for sounds made to be amplified, sounds that their makers seek to project to a larger public, the Vitaphone system nevertheless proved unable to determine its own fate, for technologies depend as much on their use as vice versa. Starting with *The Jazz Singer* in October 1927, audiences were increasingly exposed to a new kind of sound--not the theatrical kind meant to be projected to a larger public, but a new more intimate sound that is presented as private, and thus can only be overheard. When Jolson sings to the crowd in *Coffee Dan's*, like generations of vaudeville and theatrical performers before him he is purposely projecting his voice to a large audience; but when he sings and talks privately to his mother, an entirely new kind of relationship is established between the performer and the amplification system. At *Coffee Dan's*, performer and technology are aligned, the amplifying potential of the one overtly serving the other's amplificatory purpose; in the privacy of the family living room, however, the amplifying technology operates in spite of and against Jolson's quiet demeanor, thus changing us spectators from the destined audience of a self-conscious performer to a group of auditory voyeurs intent on hearing sounds that are not meant for us.

The new function of the antiquated sound-on-disc technology spawned by this important change in filmmaking style is reflected as of 1929 by a revised loudspeaker configuration. No longer present to replace the orchestra, the sound now abandons the pit to settle fully behind the screen. Whereas 1926 sound practice recognized the pit orchestra as the source of all music (typically thought of as accompaniment), the many musical films of the 1927-29 period increasingly locate the source of music on the screen. As revealed in a 1929 Western Electric ad (fig. 3), this new standard is recognized in theaters by henceforth placing both speakers behind the screen, so that all sound can once again be identified with the activity presented on that screen.

Note that there is nothing particularly logical about this change. Why should the voice of Fox's Movietone News announcer come from behind the screen? It would make more sense to identify him

with the projection of the film by locating his speaker near the projector, or to recognize his off-screen status by placing his speaker next to the screen. Locating his voice behind the screen creates a spurious identification between the announcer and the images he presents. And of course it is precisely this identification that the new arrangement seeks to establish. Increasingly, during Hollywood's heyday, the screen displaces all other aspects of the film experience, to the point where generations of film theorists have assumed that the whole of the cinema may be reduced to the screen alone, thus missing the point that the speakers of Hollywood's classical period are dissimulated behind the screen on purpose, in order to hide the real source of the sound by attributing it to the image.

Ironically, the turn away from the classical tendency to dissimulate sound sources occurs as a side effect of a movement designed to increase identification between sound and image. Not content with a generalized correspondence between screen image and behind-the-screen sound, technicians caught up in the high-fidelity movement sought to enhance the spatial correspondences between cinema sound and image. Following up on the 1933 Bell Labs experiments with broadcast stereo, in 1940 Western Electric demonstrated a four-track stereo system (left-center-right-control) aimed instead at the recording industry. Before stereo records began to flood American markets in the late Fifties, however, stereo had been adopted by the cinema industry under the most confused of circumstances. First introduced in Cinerama's early Fifties travelogue extravaganzas, cinema stereo was given the double task of meeting the needs both of fidelity (accurate spatialization) and of spectacle (rapid, energetic movement). Only the familiar ping-pong sound of early stereo records and films could simultaneously capture these two standards, yet the panning of dialog across a wide screen and back ran directly counter to the expectations of both cinema spectators (who had been trained to expect single-source sound by classical Hollywood films and speaker placement) and home high-fidelity listeners (who had been trained to regard monaural reproduction as the norm).

When Fox tried to impose magnetic stereo on all CinemaScope users, four-track for 35mm (left-center-right-surround) and six-track for 70mm (adding half-left and half-right channels), they thus

found themselves bucking both economic and representational objections. While the fully panned dialog championed in the mid-Fifties by Fox and Todd-AO offered gains in a certain sort of fidelity, it failed to match current (monaural) notions of high fidelity. The surround speakers created the inverse problem. Used only intermittently, usually to reinforce spectacular visual effects, surround sound worked directly against the ideal of spatial fidelity applied to the three directional front speakers. So contradictory did this system appear that most studios simply refused to follow Fox's lead. As John Belton reports, M-G-M, Warners, Columbia, and Universal refused to ping-pong dialog, reproducing it instead in mono, while most studios shied away from the surrounds, with Columbia never using the fourth channel at all.[\[2\]](#)

The parallel development of stereo sound for music and cinema over the past forty years offers a fascinating view of the way in which technological systems may be retrofitted to existing standards. To make a long story short, the difficulty of matching Fifties cinema stereo to current monaural standards led to virtual abandonment of stereo as a narrative tool during the Sixties and early Seventies, with only music regularly receiving stereo treatment (in keeping with stereo's conquest of the home music market during this period). Surround channels were so seldom used that surround speakers fell into disrepair, offering more static than anything else.

However, the late Seventies application of the new Dolby optical stereo variable area matrixing with improved noise reduction to *Star Wars*, *Close Encounters of the Third Kind*, and other fantasy blockbusters initiated a new era in speaker usage (fig. 4). At first, a new generation of sound specialists labored mightily to employ the surround speakers to enhance spatial fidelity. Having failed to learn a lesson from the mistakes of Fifties stereo technicians, the sound designers of the post-*Star Wars* era regularly placed spatially faithful narrative information in the surround channel. Recalling the 3-D craze in the mid-Fifties, for a few years every menace, every attack, every emotional scene seemed to begin or end behind the spectators. Finally, it seemed, the surround channel had become an integral part of the film's fundamental narrative fiber.

But not for long. Listening to theatrical reproduction of the sound

he had designed for Star Wars and its sequel, The Empire Strikes Back, Ben Burtt discovered that due to poor equipment and managerial disinterest the narrative sound events he had carefully placed on the surround channel were simply not being properly played in the theaters.<sup>[3]</sup> Starting in 1983 with the third film in the series, The Return of the Jedi, Burtt initiated a new strategy, soon emulated by other sound designers. All narrative information would henceforth emanate from the front speakers, with the surrounds used for spectacular (but nonessential) enhancements. Thus freed from any responsibility to present narrative events or even spatial fidelity, the surrounds began a new career (especially in fantasy or horror films) as purveyors of spectacular effects. Not since the antics of the vaudeville-trained drummer accompanying silent comedy had cinema accorded such a place of independence and honor to sound effects.

While the surrounds were being liberated from the demands of spatial fidelity or narrative relevance, a similar transformation was taking place with the front speakers. Since channels two and four of all six-channel 70mm prints (feeding the half-left and half-right speakers) had long since been simply extrapolated from a four-track master, they offered no new information. Beginning with Star Wars, a new function was assigned to these speakers: to provide a boost for available low frequency sound. Corresponding with Hollywood's renewed attempt to attract the youth market through concentration on sci-fi, adventure, horror, and musical superproductions, the creation of two "baby boom" channels realigned cinema sound with a new and unexpected model, the rock concert with its characteristic overamplification and earth-shaking bass.

Whereas Thirties film practice fostered unconscious visual and psychological spectator identification with characters who appear as a perfect amalgam of image and sound, the Eighties ushered in a new kind of visceral identification, dependent on the sound system's overt ability, through bone-rattling bass and unexpected surround effects, to cause spectators to vibrate-- quite literally--with the entire narrative space. It is thus no longer the eyes, the ears, and the brain that alone initiate identification and maintain contact with a sonic source; instead, it is the whole body that establishes a relationship, marching to the beat of a different woofer. Where

sound was once hidden behind the image in order to allow more complete identification with that image, now the sound source is flaunted, fostering a separate sonic identification contesting the limited, rational draw of the image and its visible characters.

By the time the "baby boom" speakers and the surrounds had been liberated from narrative responsibilities, the center channel had already become specialized in dialog reproduction. So deep-rooted is Hollywood's dedication to dialog intelligibility (we mustn't forget that the conversion to sound was initiated by the ultimate purveyors of dialog: the telephone company and its subsidiaries), that nothing but perfectly understandable dialog could possibly satisfy spectator expectations. Given Hollywood's establishment during the Thirties of a clear preference for clarity of dialog over careful matching of sound and image scales, it is hardly surprising that stereo imaging would eventually be reserved primarily for music, with dialog being routed uniquely through the center speaker.

What we see taking place over the past forty years is thus a systematic dismantling of the unified classical Hollywood system whereby all sounds would be fused into a single, unified soundtrack and funneled through a single cluster of speakers behind the screen. Creating the fiction that all sound derives from and serves the image (the familiar myth that has led to such a high level of disregard for cinema sound in general), this classical framework has been done away with by broad dissemination, over the past decade, of a new system of discrete parts. Whereas the soundtracks of the Thirties and Forties were marked by their ability to share a single invisible loudspeaker (or a cluster of speakers all reproducing the same sound at the same time), the new approach offers four virtually independent sound outlets, each separately engineered and visibly located to serve a specific need and to correspond to a different set of sound standards.

The new configuration and its purposes are most obvious in the many proprietary home audio/video systems (including the home version of THX) that use Dolby Pro Logic encoding to emulate the cinema theater situation. Receivers featuring Yamaha's Digital Sound Field Processing, for example, offer six speaker outputs (digitally processed from the four tracks on Dolby-encoded laserdiscs): left-center-right-left surround-right surround-subwoofer

(fig. 5). In 1929, these six channels would have made no sense whatever, but when considered in terms of the multiple and varying requirements enforced by our soundscape and our listening experience, they openly reveal their source and function.

The left and right speakers offer standard stereo. Over the last quarter-century, stereo has become increasingly specialized in the reproduction of music (records, tapes, CDs, FM multiplexes, most uses of TV stereo), while narrative uses of the very same media (particularly radio and television) have remained in the monaural mode. The left and right channels of homevideo systems are thus primarily dedicated to the reproduction of music. In fact, all Pro Logic receivers offer the option of returning the system to a traditional home stereo mode, routing music from nonvideo sources solely through the left and right channels, while dosing down all other channels.

The center speaker offers a separate monaural channel, to which all dialog is shunted. Listening to the center channel is like listening to a telephone during a music concert, simultaneously satisfying our expectations for music reproduction (large room with high levels of long, slow reverberation and a wide frequency range) along with the standards that we have learned to apply to dialog transmission (spacelessness and no reverb, with a relatively narrow frequency range).

By virtue of its physical separation from the screen and because it carries no sound events of crucial narrative importance, the surround channel (or two channels in the case of THX, Yamaha, and certain other processors) is released from the standards we apply to the front channels (directional fidelity for the stereo left/right combination; equal intelligibility throughout the theater for the center). Seeking "effects that are out of this world" (as a recent Adcom ad suggests), contemporary films commonly create domains in which any sound effect, however farfetched, will be deemed acceptable. Not just the fantasy worlds of outer space and Transylvania, but also the apparently realistic realms of heavy military machinery and undersea exploration create atmospheres in which synthesized or digitally massaged sounds coming through the surround speakers can add to our pleasure, in spite of--or rather because of--our inability to judge whether the sounds we are hearing

have any correspondence to reality.

Note how different this logic is from the standards applied to the limited number of effects fed through the left and right frontspeakers, which are judged by altogether different notions of spatial fidelity.

Derived from the baby boom speakers in 70mm theaters, the subwoofer reproduces all low frequency sounds. In addition to extending the bass response of speakers with insufficient bass extension, the subwoofer's floor-shaking capacity offers the possibility of representing cinema as a more participatory event. Yamaha's ad says that "Cinema DSP blurs the line between watching a movie and actually being in one." It might well have said that subwoofers blur the line between listening to film music and actually being present at a rock concert, thus radically modifying the identificatory relationship between the audience and the film.

Just as all modern music speakers involve a combination of woofers, midrange, and tweeters, each serving a specific purpose and range governed by a network of crossovers, so current theatrical and home configurations involve a series of quite different speakers, each dedicated to a different purpose, connected by Dolby Pro Logic and the twin needs of narrative and spectacle. While the logic is the same as it was in 1909, with the success of the technology depending in large part on its ability to conform to contemporary notions of what kind of sound deserves reproduction, and how that sound should sound, today's results are far removed from those of the beginning or even the middle of the century. Instead of alternately satisfying divergent sound needs through differing sound systems and speaker configurations, we have entered into an era where careful manipulation of technology and representation alike have made it increasingly possible to satisfy a large number of contradictory needs simultaneously.

As sound technology becomes increasingly microminiaturized--moving first from theater to home and now to multimedia computer workstation--it is tempting to speculate about future developments. Will CD-ROM- equipped computers need center speakers if they are to be used for talking books or voice-illustrated encyclopedias?

Will they have built- in subwoofers next year, so as to provide the bass response needed for certain styles of music? Will they feature FM connections to surround speakers, so that video games will feel truly wrap-around? We live in exciting times, which only become more fascinating when we apply to them the logic systematically applied to past developments in sound: in order to succeed, each new sound technology must satisfy the needs created by the other sound practices to which potential consumers are accustomed.

-----

Footnotes:

1 Tomlinson Holman, "Home THX," Stereo Review, April 1994, pp. 54- 60.

2 John Belton, Widescreen Cinema (Cambridge: Harvard University Press, 1992), p. 205ff.

3 Larry Blake, Film Sound Today (Hollywood: Reveille Press, 1984), p. 45ff.

Altman, Rick, The sound of sound..., Vol. 21, Cineaste, 01-01-1995, pp 68.