What Made Pistachio Nuts?
Henry Jenkins

Showstoppers: Busby Berkeley and the Tradition of Spectacle
Martin Rubin

Projections of War: Hollywood, American Culture, and World War II
Thomas Doherty

Laughing Screaming: Modern Hollywood Horror and Comedy
William Paul

Laughing Hysterically: American Screen Comedy of the 1950s
Ed Sikov

Primitive Passions: Visuality, Sexuality, Ethnography, and Contemporary Chinese Cinema
Rey Chow

The Cinema of Max Ophuls: Magisterial Vision and the Figure of Woman
Susan M. White

Black Women as Cultural Readers
Jacqueline Bobo

Picturing Japanese: Monumental Style, National Identity, Japanese Film
Darrell William Davis

Attack of the Leading Ladies: Gender, Sexuality, and Spectatorship in Classic Horror Cinema
Rhona J. Berenstein

This Mad Masquerade: Stardom and Masculinity in the Jazz Age
Gaylyn Studlar

Sexual Politics and Narrative Film: Hollywood and Beyond
Robin Wood

The Sounds of Commerce: Marketing Popular Film Music
Jeff Smith

Orson Welles, Shakespeare, and Popular Culture
Michael Anderegg

Pre-Code Hollywood: Sex, Immorality, and Insurrection in American Cinema, 1930–1934
Thomas Doherty

James Lastra

COLUMBIA UNIVERSITY PRESS
NEW YORK
EVERYTHING BUT THE KITCHEN SYNC
Sound and Image Before the Talkies

It is the purpose of the remaining chapters to examine the factors conditioning the emergence of "sound" within the American cinema. I will demonstrate not only how these events recapitulated several of the key features I associate with the emergence of the technical media more generally, but also how entirely new and unimagined possibilities may emerge out of a limited and concrete set of historical and material conditions. More specifically, in this chapter I examine how the entities "the cinema" and "sound" came to be defined in relation to one another historically, and how the very devices and institutions we associate with the Hollywood cinema were similarly defined and redefined over time. While the "coming of sound" may seem a particularly clear-cut instance of adapting a well-defined technology to a similarly stable aesthetic form, such is decidedly not the case. Not only were both technology and representational form in flux, but each helped to define or constitute the other in the process of their mutual interaction.

Moreover, the period before the putative coming of sound offers us a glimpse into how two new technologies with sometimes overlapping and sometimes quite distinct histories—namely, cinematography and phonography—could combine to form an integrated sensory experience that was neither audio nor visual, but distinctly audiovisual. However, the proper ratio between the senses—between hearing and seeing—was open to vigorous debate and competing models. Was the cinema an essentially aural medium, born of the spirit of publicly performed music, to which spectacles of various sorts might be appended, or was it a narrative visual form, to which sound could only be an "accompaniment"? During the early cinema period, especially, we can indirectly witness a confrontation between competing sensory regimes, each one adhering to its own dictates and coexisting at times unpeacefully. The dominant ideal of listening as attentive, sensitive, and receptively passive—as complete and self-contained—confronted the film industry’s emerging visual language, and each sense was forced to work with and against the other in a project of audiovisual collaboration. As the institution of cinema moved more and more firmly toward narrative, sound took its place as an integral, but also overdetermined, part of the sensory order characteristic of the classical cinema.

SOUND IN 1926

In retrospect, the sonic needs of Hollywood in 1926 seem almost obvious. Who, for example, could ever have doubted the primacy of intelligible, narratively important speech for the classical paradigm? The cinema of narrative integration, long since established, apparently mandated the practices of sound we have come to accept as the norm, as it simultaneously ruled out other approaches. David Bordwell, Janet Staiger, and Kristin Thompson, for instance, assert that "sound as sound... was inserted into the already-constituted system of the Classical Hollywood style," that "the centrality of speech became a guide for innovation in sound recording," and that the use of character-centered leitmotivs indicates how even musical accompaniment had become a fully integrated component of a narrative unity. Recorded sound simply fulfilled already existing functions, and indeed was their "functional equivalent." Change, then, was almost entirely technological.

However, the situation appears this tidy only in retrospect. It seems probable that the leitmotiv system was not universally adopted, or even dominant, nor was the function of sound accompaniment of a single type or exclusively narrative in intent or effect. Indeed, the very term sound accompaniment, which assumes the primacy of the image in all cases, obscures our view of the so-called pre-sound era and, furthermore, orients us toward the classical cinema in such a way that we may misconceive alternative traditions simply as fully integrated components of the classical style. Indeed, the perspective that does not expect to see discontinuities or complexities may not perceive particular deviations as deviations at all, encouraging us to notice those elements that led most directly to the ultimately triumphant set
of norms. The archive and the genealogy that I have produced in the first two chapters and in the present one encourage a different focus, and different conclusions.

As recent research confirms, the few years immediately preceding the "dawn of sound" are too narrow a historical slice by which to characterize phenomena as varied and complex as "film sound" and cannot adequately characterize the evaluative and practical frameworks within which sound was received. Although terms such as "cinema sound" appear to name objectively describable entities defined by the technologies that "produce" them, they are, in fact, not neutral descriptions. That is, the act of naming or describing determines the boundaries of the phenomena under discussion and effectively creates the contours of its object, and basic terms like synchronization, music, and effects may not always have meant what they have subsequently come to mean. The boundaries between media, and even between technologies, were and are open to dispute and renegotiation, as are the boundaries of function.

The title of this chapter evokes the heterogeneity of practices, functions, and effects of sound in the cinema, and hopes to open up calcified and often unexamined uses of the term "synchronization." To start, let me suggest that any fixed or purposeful relationship between sound and image legitimately may be thought of as synchronized. From this perspective, "lip-synch" is only one, rather banal, possibility. Consequently, I will avoid approaches to synchronization that focus exclusively on production, and thus ultimately on the relationship obtaining between sounds and sources in the everyday world, in order to ask if the represented sound and image correspond to the sound and image relationship obtaining in the "real world." Theories of sound representation that confine the profilmic and diegetic realms by reducing representation to inscription reinforce this attention, as do approaches that base their notion of synchronization on the naturalistic norms of later classical cinema. Our current view of synchronization and, more broadly, our current theories of film sound too often derive entirely from a consideration of sounds linked to pictured sources—their presumed real-world sources—as if representation were merely a matter of duplicating that world. As chapter 2 implies, however, it is better to focus on representational effects, putting the emphasis back on the film experience and moving it away from potentially irrelevant issues concerning, among other things, the arena of the profilmic. By expanding our sense of what it meant for sound and image to "go together," I hope to suggest how the historically shifting roles of sound have helped to define and redefine the mode of representation, the mode of production, and the pleasure characteristic of the Hollywood cinema at various historical moments.

In their seminal analysis of the classical Hollywood cinema, Bordwell, Staiger, and Thompson forcefully elaborate an account of the classical stylistic paradigm as a structured and structuring framework regulating the boundaries of aesthetic options and decisions. Arguing for a more or less "functionalist" approach to the analysis of textual production and systematicity, they discuss formal devices in terms of their specific roles within a historically and culturally defined set of textual possibilities. Thus, they argue, given the historically limited means (both technical and imaginative) available at any particular time, it is possible to reconstruct, through careful analysis and documentation, which formal devices could at any particular moment, and in any particular textual situation, substitute meaningfully for one another, the way a straight cut can substitute for a dissolve or a fade-out and fade-in. From this perspective it is easy to understand how and why certain textual forms persist despite possible alternatives, as well as how aesthetic options are ordered hierarchically in terms of probable use.

While powerful as an analytical tool, the functional equivalent, as a practical concept, is central to the organization of the Hollywood industry and its mode of textual production, which takes the continuity script as the basis for the realization of the cinematic text. This approach affirms the narrative and even verbal origin of all cinematic representation and structures the type and frequency of formal devices employed in the text. Unlike the earliest cinema exhibitions, which often revealed in the absolute specificity of each and every take, even encouraging multiple viewings in a single session, the mature cinema of narrative integration depends on the assumption that five filmed versions of the same narratively mandated event are virtually interchangeable. Since each shot or device is determined by the narrational function it is to fulfill, the field of representational possibilities is hierarchized in advance, and the possibility that two realizations of that function greatly differ is dramatically reduced. What the individual shot loses in phenomenal specificity, however, it gains in narrative saliency. The more regulated the relationship between script, realization, and final text, the more interchangeable the personnel can be as well—an economic benefit that cannot be overestimated.

The emergence of the classical cinema as a form of practice therefore
required nothing short of a radical redefinition of the cinematic image. From an essentially idiosyncratic form that relied on a "topographic" mode of spectatorship, which scanned the entire image, and a highly presentational mode of exhibition that emphasized performance, there emerged an essentially narrative pictorial form whose textual specificity was removed from the hands of the exhibitor and placed, as much as possible, within the mechanisms of the text itself. Despite this trend, however, until the sound era certain aspects of the film experience—like sound and music performance—remained beyond the producers' control.  

Like a modern printed text, whose typographic idiosyncrasies are of almost no importance in comparison with its signification, the filmic image became a relatively transparent signifier of a verbalizable event, rather than an "event" in and of itself. Whatever it thereby lost in complexity and "grain," it compensated with a gain in legibility and signifying stability. Particular filmic devices such as dissolves, fades, and wipes similarly abandoned whatever phenomenal appeal they might have had in favor of their ability to signify certain kinds of spatial and temporal articulations in a consistent manner. Within such a textual paradigm it seems reasonable to assume that, at least at a practical level, film producers evaluated devices in terms of their functional equivalence, and more or less ignored those aspects that were clearly nonequivalent.

When examining the state of film language at any specific historical moment, we are able, through the concept of functional equivalence, to analyze devices in terms of their paradigmatic substitutability, but it is less successful accounting for difference and contradiction. As a tool for explaining historical change, it can even be misleading. Bordwell, Staiger, and Thompson argue that Hollywood successfully innovates and adopts new representational technologies when they fulfill existing textual functions. This suggests a fairly uneventful sequence of problem and solution, with certain technologies succeeding because of their evident suitability to the classical narrative, others failing because they offered no functional gain. Certainly there is a great deal of truth in this position, especially given Hollywood's general conservatism in textual form and in investments in new machinery. All the major technological innovations, like sound, color, and widescreen processes, did indeed mesh with existing textual forms and functions. However, to think of these innovations as equivalently fulfilling certain functions robs our historical explanation of an important degree of complexity.

To the extent that, say, recorded sound did not precisely fulfill an existing function, it introduced an element of excess or contradiction that cannot simply be ignored or discounted. In fact, such contradiction seems generative of textual and even industrial transformation. To complicate matters further, given the multiplicity of possible textual functions for sound, to suggest that sound fell into a preordained slot, or even group of slots, is unnecessarily reductive. Even while appearing to fulfill a particular function (such as providing intelligible, narratively important dialogue), sound may simultaneously be performing other, nonnarrative or even nonrepresentational ones. The standards of sound representation eventually codified by Hollywood practice, although appearing to be yet another familiar set of techniques that aids the effortless assimilation of story information, and therefore simply another component of the classical narrative system, conceal a history of traditions of representation and exhibition that may be at odds with classical norms, and are therefore in direct conflict with them.

Sound technicians, and what might be called "sound practitioners," will be my direct focus in this and subsequent chapters since it is through them, their goals and standards and their systems of practices, that these varied traditions are realized. Rather than a simple matter of problem and solution, I see the innovation and adoption of new technologies as a process fraught with controversy and struggle—the stakes often involving the professional identity of the technicians involved.

In general, filmmaking followed a historical course that is by now familiar. After approximately 1907, narrative films dominated not only in sheer numbers but as the form around which the emerging industry grew. After that point as well, the heterogeneous and often boisterous field we call early cinema moved toward greater and greater formal and institutional unity. Along with the rise of narrative films came a more pervasive pressure toward narrative integration that encouraged producers and exhibitors to subordinate both filmic and extra-filmic elements to an overarching narrative effect, and thereby unify and regularize the cinema. Sound accompaniment was probably the element of the show most resistant to these pressures because its functions and effects were only occasionally narrative. In excess of their narrational functions, music, lecturing, and effects spoke directly to audiences in a celebration of their fleeting community. Whether through sing-alongs, virtuoso musical or vocal performance, ethnic or national music, conspiratorial mockery of the film, or simply their bodily presence, musicians and sound performers decisively mediated the experience of silent films in the nickelodeon and in the picture palaces. While few of us can recall the experience firsthand, it seems certain that sound and image coexisted in many more vari-
ations than in the talkie, and that sound might, at times, exceed its role as mere adjunct.

By briefly sketching the history of sound and image combination in the cinema, I will show how particular sound practices that apparently disappeared with the advent of the talkie continued to exist in the margins of the classical style. Many of these practices belong to a mode of cinema that might be characterized as more representational than representational, and therefore as belonging to a vision of the cinema—a cinema of attractions—alien to that shared by the sound technicians of the transition period, who were more uniformly immersed in the narrative norms of the classical cinema. In order to address these concerns, we need to interrogate our idea of accompaniment by asking, what was "sound," what was "synchronization," and what was understood as "a film"? As I will show, debates about these questions work to overcome conflicts between traditions of practice and institutional demands by bringing them into agreement, however tenuous that agreement might be.

SOUND’S DIRECT ADDRESS

That the silent cinema was never silent has become one of the great revisionist clichés of film studies. In point of fact, this is not quite true, for in many cases sound was an integral part of the nickelodeon program except at the time during which the film was projected. At many nickelodeons, music was played during the intermissions and with the obligatory illustrated song, but sound’s most important early role was as a means of attracting attention to the nickelodeon itself. Sound, in the form of phonograph, Barker, drummer, or piano, primarily served the function of hailing potential patrons. As nickelodeon historian Q. David Bowers notes,

Apart from numerous attempts to use phonographs to provide soundtracks for movies, the primary use of the record player was to attract patrons into the theatre. Thousands of nickelodeons had phonographs in the projection booth, with the horns projecting out through a wall into the street, playing through an opening usually above the ticket booth. The music thus provided was loud enough to attract passersby directly in front of the theatre but not sufficiently loud that merchants would complain up and down the street.

In 1910, Film Index’s music columnist Clyde Martin concurred, telling his readers that he was recently fired as a piano player “because they could not hear me on the street... That is the fault of the average exhibitor today; he doesn’t want a piano player, he wants a Bally-Hoo.” Martin’s rival at the Moving Picture World, Clarence Sinn, similarly recalls, “When music was first introduced to the picture theater, they ‘wooped’ ‘er up until the music could be heard out on the street.” Phonographs themselves were often called “barkers,” since they were so closely identified with this role. The hailing function of sound became so common, in fact, that several cities enacted legislation prohibiting the use of mechanical sound devices outside nickelodeons, and contemporary descriptions of theater design stress the importance of shutting out external noise, including barkers and the like.

However much “barking” may have dominated early film sound, other sources from the same period suggest that in some theaters music accompanied only the film. Whatever may have been the case, it is important to realize that cinema, as a cultural practice, was not dominated by a single approach. As Rick Altman argues,

Silence was one of a number of acceptable film exhibition approaches throughout the pre-1910 era. Indeed, in certain regions and in smaller theaters, silence during the film remained the rule even after 1910; in other situations, and for a limited number of years, silence was simply one of many operative strategies. In either case, audiences were not generally trained to expect continuous musical accompaniment of films until after the turn of the decade.” (“The Silence of the Silents,” 677)

Until the early ’teens, sound worked to solicit patrons and draw their attention—a function that was to support a more general strategy for directing spectators to performers in the theater, or to objects in the image. In fact, the direct address made possible by sound would shape most silent film sound in one way or another.

SOUND OVER IMAGE: THE STRUCTURAL PRIORITY OF SOUND

Beyond “barking,” which went on regardless of whether an image appeared on the screen or not, the best documented and most widespread audiovisual
form on the nickelodeon program was the "illustrated song," which came to prominence in the entertainment world in the 1880s and 1890s. The business flourished in the 1890s with the advent of the amateur photography boom. These songs, while a staple of nickelodeon programs, were not designed solely to entertain the patrons of the theater. Sheet music publishers routinely provided nickelodeon owners with free (but apparently inferior) slides in order to get them to perform and thereby promote certain new songs. While truly audiovisual, these were essentially sound performances that encouraged sing-alongs; the visual presentation was decidedly secondary. Song promotion and sheet music sales drove this practice and organized its formal properties, putting music ahead of images. In economic, conceptual, and structural terms, the song—sound—was clearly dominant over the image.

The illustrated song has an important and ultimately more influential parallel in the slide lecture. The slide lecture, often a travelogue, holds a privileged place in the history of cinema for several reasons. The collection of views typical of the travel lecture, which were united into a continuous whole by the mediation of the lecturer's account, and particularly by the fiction of a journey, offered an early and persuasive model for multishot narratives in the early cinema period. As Charles Musser has shown, the travel genre was one of the earliest proto-narrative forms of cinema because it offered a compelling logic that motivated and unified the succession of otherwise disconnected stills and films. In addition, by proposing an embodied observer located in the fictional world, this form paved the way for the subjectivation of space in the Hollywood classical narrative through devices such as the point-of-view shot. Significantly, travelogues were a sound-dominant form.

Subjectively organized spaces and identification with the traveler were implicit touchstones of many early accounts of travelogues (which began to include motion pictures as well as slides), and in "how-to" articles in the trades. Lecturers were prized for their ability to create a sense of vivid immediacy. A tour of Yellowstone, for instance, merited comment because it was presented by a stagecoach driver intimately familiar with the area, while a series of talks on Washington, D.C., New York, and Montreal likewise found favor because the lecturer had gone to these locations himself to acquire the views. John Stoddard and Burton Holmes, acknowledged masters of the genre, were singled out for the quality of their images, which "harmonized" with their lectures because they designed and shot their own slides to match the narratives, rather than employing stock images. So popular were these travelogues that, in 1908, the Keith theater chain was even using them as headliners.

Lecturing can tell us something about the role of music, too. Like the slide lecturer, the singer provided a sonic, and often narrative, pretext and continuity for a series of images, either stills or films, which illustrated the song. As in the practice of lecturing, images here accompany sound, not vice versa. The functional subservience of image to sound is emphatically declared by the widespread practice of illustrating songs with stock slides rather than with specially produced series, indicating that the motivation for the performance was aural, not visual. In fact, it was argued that the best lecturers never openly acknowledged the images but rather spoke continuously while an assistant changed the views at the appropriate time. Images were not irrelevant to the process, however, and both singers and lecturers met criticism when their slides did not live up to the sounds they were meant to accompany. Successful lecturers like Stoddard, Holmes, and Dwight Elmendorf stood above the competition by upgrading the images that accompanied their talks. Unlike many of their competitors, all three traveled to the locations they described in order to take photos that would therefore more precisely match the spoken descriptions. This insistence on better "matching" was later applied to song slides as well, and standards of matching offer a glimpse into early notions of synchronization.

Discussions of successful "matching" in lectures and illustrated songs stress a "harmony" between sound and image. Harmony, whether understood as illustrative accuracy or as thematic consistency, in fact describes another fairly stable tradition of synchronous relationships which has been submerged by our current limited perspective on sound and image possibilities. The praise associated with "harmony" was also accorded to other forms of synchronization, including ones that emphasized sound performance at the expense of the image. In lectures and songs, sound dominates and the image provides the momentary "shock" or "effect" that enhances the lecture. The lecture or the song would no doubt be most effective when sound and image "fit" or "synchronized" together, but as the use of stock slides and footage illustrates, it was the image that was fit or synchronized to the sound.

Beyond the lecture, promoting the sound component of the motion picture show seems to have been good business for exhibitors, even if manufacturers had already begun the drive toward narrative integration and image dominance. Writers from the 1907–1911 period repeatedly stress the differentiation, distinction, competitive edge, and more refined audiences available to those who added sounds, and especially music, to their programs.
prominence of sound at a “picture” show was no accident. Indeed, the idea of sound-track dominance influenced the very structure of at least one form of filmmaking. Of the early forms of “talking picture,” the ones that occasioned the most hope and, apparently, the most premature praise were the myriad phonograph and projector combinations. Going under the names Cameraphone, Chronophone, Theaterphone, Biographon, Cinophone Synchronizer, Phono-cinematograph, and Photophone, to name but a few, these variously constructed and variously successful portmanteau technologies were extravagantly praised as mechanically and aesthetically perfect, eliciting predictions of the imminent demise of the silent picture. Unfortunately, none of them adequately solved the problem of synchronization, although the American Cameraphone and the French Chronophone did enjoy a certain degree of success.

One early example, the Phono-cinematograph, was described in 1907 as a “thoroughly practical success,” although its synchronization system depended on the capabilities of a projectionist with a talent for on-the-spot adjustments as part of his performance. Significantly, the device required the performers to record the sound first and then film a performance that matched the sound track, much like our current practices of lip-synching. Other companies followed the same procedure, selecting a “stock” record from any record store and filming a series of images to “fit” it. The most common practice was to record a musical performance or an excerpt from a play and then “drill” the performers to “match” the record before filming the picture’s track. In the noted case of an early “Motographic Opera,” selections from Il Trovatore were recorded on a phonograph and images later matched to it. Advertising and critics’ responses make it clear that the overall form of the production stressed the musical résumé above all else, rendering the image incidental. Like many other sound practices, this sort of “synchronization” carries a distinctly performative flavor, but more important, these disc and film combinations all treated the sound recording as the structural basis of the entire representational amalgam.

Discursive, participatory, emphatically present, and lucrative for exhibitors, sound constantly threatened to usurp the image. The priority of sound over image became a governing structural feature of other types of films and in other exhibition practices as well. Often, in advertisements, producers would stress a film’s opportunity for “effects,” that is, for clever musical or percussive effects to accompany the image. While this may frequently have been limited to “being unduly noisy when a man falls from his horse,” in many cases the sound practices were more sophisticated. In an area of great theater concentration, like New York’s Union Square, at a time when film production could not keep up with demand, the success of a theater often depended on qualitative difference in presentation. Nickelodeons would compete by providing the most sophisticated, elaborate, or funny sound performances to fit a particular film. Articles suggested that middle-class “girls” who played “high class” music could draw high-class patronage as well. Others counseled that “props,” or sound effects, were the best way to improve bad pictures, while still others suggested that, measured in profit margin, sound effects were a sound investment. These strategies placed an unusual emphasis on performance, discursivity, and “liveness,” and both lecturers and “prop men” understood “synchronization” as a category of evaluation. Contemporary authors therefore repeatedly encouraged lecturers to “roadshow” their lectures to get both text and timing just right. The many articles praising Lyman Howe’s presentations use performance criteria in just this way, and ascribe his success to his troupe’s perfect performance.

Companies even began to produce films whose purpose was to motivate particular forms of sound accompaniment. A typical example is Biograph’s “Fights of Nations.” The film depicts a series of different racial and ethnic types who come into conflict, but who are reconciled in the end. The use of recognizable ethnic types provides a pretext for clever sound synchronization. As each type appears, the accompanists play a melody evocative of his or her homeland. As types enter into conflict, the themes compete for dominance and collide in discord. As the advertising suggests, the music underscores the conflict’s resolution into “harmony” by playing “America.” “Fights of Nations” is by no means an isolated instance. The trade press abounds with films whose chief purpose seems to be providing the opportunity for sound “effects,” which served not in support of the narrative but as attractions themselves.

Manufacturers even developed a short-lived generic form based upon the illustrated song—the song film—which included, according to Altman, “His First Success” (Pathé, August 1907) about a cellist; “Dot Leedle German Band” (Kalem, September 1907); “Our Band Goes to the Competition” (Pathé, September 1907); “The Irresistible Piano” (Gaumont, November 1907); “The Merry Widow” (Kalem, January 1908); “The Mad Musician” (Selig, March 1908); “The Merry Widow Waltz Crazy” (Edison, May 1908). Edison released the picture songs “Love and War” and “The Astor Tramp” using several songs in sequence as well as stereopticon slides and reading matter. The Edison catalog of 1906 announced: “We have at last succeeded in perfectly synchronizing music and moving pictures. The following scenes are
very carefully chosen to fit the words and the songs, which have been especially composed for the pictures. The film need not have been designed for effects, however, since writers often praised drummers for the sophistication of their "props" or "traps" and commended the tendency to look for every opportunity to supply effects. The emergence of sound practices that asserted their priority over the film, even if it had not been designed for sound accompaniment, quickly became a central concern of manufacturers, exhibitors, and critics.

Techniques of sound accompaniment were becoming so important to exhibitors, and therefore to the emerging industry, that commentators sought again and again to point to musical "effects" done properly, in order to encourage "appropriate" practice. To be sure, in the ‘teens and twenties such effects were often ridiculed as excessive, but for a time they represented one of the most sophisticated and lucrative approaches to exhibition. More importantly, the use of effects and piano introduced a performance-based aesthetic of sound to the average film and to the mainstream of cinematic practice. Like the Barker, but in concert with the screen, the singer, pianist, and drummer directly acknowledged their audiences, and even elicited their response through cleverness, ingenuity, and above all spot-on performances stressing synchronism. One measure of the sound practitioners’ success seemed to be exactness of his or her timing during performance. These and other sound practices stressed performance to such an extent that audience attention often seemed split between the world on the screen and the performers in the theater. This dual focus fostered the tendency to evaluate all sound, but particularly "synchronization," in terms of accurate performance.

**SOUND AGAINST THE IMAGE**

In addition to the stress sound effects accompaniment placed upon the performers, and therefore their direct and discursive relationship with the audience, it simultaneously and problematically altered the nature of the cinematic commodity. Although film producers had been generally successful in consolidating their control over the product, sound accompaniment, more and more a part of standard exhibitions after 1907, remained an arena wherein individual exhibitors could alter the public’s relation to the film. Perhaps inadvertently, effects accompaniment introduced a punctuated temporality of momentary audiovisual climaxes that could easily run counter to the temporality of the more purely visual narrative. Moreover, such techniques could destroy the intended spatial hierarchies of the image. Several comic stories survive that describe overactive drummers and effects men who, in an attempt to "enhance" a picture, go to absurd lengths to punctuate every image with a sound. One writer describes lecturing with a film only to be interrupted by the repeated sound of a bird whistle. Perplexed as to its supposed source within the image, he asks the effects man and is pointed to "a diminutive canary in a tiny wooden cage on a top shelf at a far corner of the room," whereupon he "wallops" the whistler and proceeds.

What makes this anecdote of more than passing interest is that it illustrates the conflict between two different conceptions of representation, and two modes of spectatorship. The drummer clearly searches the entirety of each image topographically in a quest for opportunities to display his craft, while the lecturer just as clearly adheres to what he feels is the self-evident hierarchy between narratively foregrounded elements and those left to the less important background. By adhering to a practice that was discursive and performative, the effects men created new hierarchies within the image, drawing spectator attention to incidental features because they could make noise. Under his gaze, the image ceased to signify in a predictable way, but became a pretext for virtuoso displays of sound. Consequently, the drummer could destroy the image’s narrative legibility, or create discontinuities between shots where the producers clearly sought continuities.

The drummer’s tendency to introduce potentially haphazard temporal and perceptual emphasis through sonic underlining presented producers with a dilemma: how do you control the type of sound performance in a theater in order to ensure the dominance of narrative over all other systems? The Edison Kinetogram, for one, began to publish musical suggestions for each film they advertised, as much to limit potentially transgressive forms of accompaniment as to suggest “appropriate” ones. The drive toward standardization moved beyond music to encompass effects, too. A writer from the *New York Dramatic Mirror* suggests:

Incidental music is claiming the intelligent attention of some of the picture manufacturers, notably the Edison and the Vitagraph. Some time ago the Edison Company commenced printing programmes of instrumental music suitable for Edison releases, and recently, the Vitagraph Company announced that it would introduce properly arranged piano scores with
each film of its manufacture. Now let some enterprising firm send along a prepared programme of sound effects to go with each subject, and another step forward will have been recorded.48

In general, the Kinetogran’s selections try to match narrative and musical mood, tempo, or theme. Falling within the category Bordwell calls “pleonastic” sound,49 these selections were chosen in such a way that they were subservient to and supported the image-narrative. Each was chosen on the basis of its ability to mirror some important aspect of the story, rather than any other criterion, such as opportunities for virtuoso performative display or for audience sing-alongs.

Regardless of such attempts at standardization, sound still functioned as a mediating device between the impersonality of the mass-produced film and the particular screening situation. Although it is doubtful that many sound practitioners or musicians consistently achieved the kind of truly interactive give-and-take necessary for direct audience engagement, sound was nevertheless one aspect of the exhibition that, being discursive and performative, addressed audiences directly and offered the structural conditions for shaping audience response. As the institutional or industrial film tended to shift from an aggregation of “attractions” to a form stressing narrative integration, and while the represented space of the film became more and more distinct from the space of the audience, sound retained and continued to develop a mode of sonic “attractions” maintaining a direct relationship of performer/audience copresence.50 Remaining one of the last arenas of exhibitor control, sound allowed various theaters not only to appeal to the religious, economic, and ethnic particularities of their audiences but offered the possibility of ridiculing the film, if it pleased the locals.

SOUND AND REALISM

Although performative sound implicitly threatened the dominance of the image, certain forms of sound practice could enhance both the exhibitor’s and the manufacturer’s prestige. Besides its highly touted benefits to attendance, sound apparently offered another type of symbolic capital that, itself, could be exploited to enhance revenues: realism.51 Realism connected the realm of sound performance with the demands of narrative integration by producing the effect that sounds emanated from sources on the screen. Under the rubric of “realism,” the drummer’s sound effects were seen as adjuncts to the image and confirmed its own realism by creating the impression that sound emanated from it, just as in the world outside the theater. Such “realistic” effects participated in the more general historical trend toward greater integration of sound within the diegesis created by the image. Lecturers lectured less and imitated voices more, and drummers reined in their tendencies to show off in favor of supplying narratively plausible effects.

The situation of “realistic” sound accompaniment was more complex than I have suggested so far, however. Competing with the drummer was an array of “technologies”—Humanovox, Actologue, and Humanoscope, to name a few—that combined film projection with troupes of performers who provided voices and effects from behind the screen. True to the general trend, hiding the performers tried to downplay the discursive aspects of sound performance by subsuming them within the diegetic world, but contemporary descriptions nevertheless always comment on the preparation, execution, and diversity of performance.52 The tension between realism and the flamboyant displays of skill it took to produce it profoundly shaped attitudes toward sound effects. Exhibitors such as Lyman Howe were cited for the care and practice with which they created their “illusion.” Howe’s career is exemplary of the importance of traveling companies to the history of sound in the cinema, as such groups were among the few who could repeat film material often enough to create a truly planned and practiced sound performance. A description of one of his shows is typical.

The sound effects were not only in time and always appropriate, but their novelty and range is surprising. . . . There were such details in the pictures given sound as would never be thought of by any but a specialist; for instance, the silence of a Venetian Gondola ride is broken not only by the sound of the dipping oars, but even the rhythmic knocking between the oarlocks is heard. . . . [With a building’s collapse] there was a single crashing, crackling sound equal in volume to the noise one would hear when watching the fire at such a distance as that at which the picture was taken.53

That the theater should be quiet enough to appreciate the dipping of oars is perhaps only as remarkable as the fact that such effects were attempted, and the reference to what might be called scale matching between image and sound is equally remarkable. As references to Howe—and especially his effects man, LeRoy Carleton—littering the trades confirm, it is consistently the quality of performance that was admired in his shows. Carleton, in particular, is praised as a consummate vocal performer, who “in the course of one
performance is called upon to make 115 changes of voice in his vocal mimicry of both humans and animals.” 54 One reviewer argued that

the ‘noise’ portion of the show—the use of stage effects to make the pictures more like real—is the best that has ever been used in Pittsburgh. Conversations of the subjects of the pictures, expressing every emotion as depicted on the faces of the pictures: the whir of machinery, rumble of railroad trains, swish of water in marine scenes, and various other things that help the onlooker to imagine that he is witnessing the real thing instead of a counterfeit presentment. 55

Such “realism” crucially depends upon precise performance, and is therefore always capable of usurping the film. Altman suggests, however, that using the voice in this way decisively tipped the balance away from the discursive tradition of the lecture and made speech “another sound effect, a method of further anchoring the film experience in reality” by solidifying the effect of simulated presence. 56

SOUND IN THE SERVICE OF THE IMAGE

Even while drawing attention to the performer’s skill, sound effects beyond the realistic use of the voice could still enhance the representation of narrative rather than deflect it. Indeed, sound accompaniment could become essential to the narrative film. One writer commenting in early 1908 gives an example of a film whose meaning was dramatically clarified by the intelligent use of effects. He enters a nickelodeon to view a film in which a couple, while crossing a field, suddenly pause and, inexplicably, bow their heads before proceeding. Seeing the film again in another theater, he notes that the previously puzzling pause is accompanied by the sound of a church bell, which clarifies the meaning of the gesture immeasurably. 57 In this case, sound produces the film’s legibility as a unified narrative, motivating what would otherwise be a perplexing discontinuity. This technique served the silent film long and well, and a similar effect is used early in Vitaphone’s A Woman of Affairs (1928), where several “silent” characters react to the sound of a horn offscreen. Had the film been projected without sound (as was likely in some unconverted theaters), the motivation of the characters’ response would have remained slightly obscure, but with a synchronized effect it is perfectly understood.

This sort of clarification became a more and more pressing issue between 1907 and 1910, as plots became longer and more complex and demanded an explicit narration. Some writers accurately pointed out that, for example, a version of Macbeth (1908) reduced to a single reel, or the various “Films d’Art” lately arriving in the United States, simply required the presence of a lecturer, lest they remain hopelessly opaque. 58 Earlier condensations like Edwin S. Porter’s Uncle Tom’s Cabin (1903) relied to a much greater extent on the spectator’s presumed familiarity with the story and consequent ability to fill in the gaps or discontinuities. Typical of the industry-wide trend toward self-contained modes of narration, however, these later films did not assume the same audience foreknowledge, and elliptical narratives became a problem rather than a strategy.

While some writers suggested that lectures simply helped the “less intelligent” understand the increasingly complex stories, 59 others like Van C. Lee championed the film lecturer, not as an attraction himself but as a necessary adjunct to the image, arguing that lectures can help everyone understand multishot films better. Like other writers promoting the “intelligent” use of sound effects, he cites Lyman Howe’s considerable success as an example of the lecture’s benefits. 60 One writer even argued that motion picture producers were so concerned with the problem of intelligibility that they meant exhibitors to read the film descriptions they included with the rented film aloud during screenings. This writer also noted that only a tiny minority of exhibitors ever did so. 61

In contrast to those who praised the promiscuous use of sound effects, other writers began to insist upon limiting them to “appropriate” effects. Even the drummer/piano pair praised for their tendency to “take every opportunity for effects,” and to “match all movements,” indicate that they, nevertheless, always play “with” the picture. 62 Other writers concur, arguing that, above all, musicians should avoid “interpreting” the picture wrongly. Thus, some argued even the ubiquitous coconut shell hoofbeats precisely synchronized to the image need to change their tone when the horse moves from pavement to dirt, lest the clip-clopping on a marble slab become comic. 63 Noted commentator W. Stephen Bush sums up the general tenor of the discussion, praising the voice that “runs with” the picture, explaining the figures and plot while bringing out “by sound and language the beauties that appear but darkly or not at all until the ear helps the eye.” 64

To a remarkable extent, Bush’s comments illustrate the growing tendency of one tradition of sound to absorb and integrate another that, otherwise, appeared fundamentally opposed to it. The “performative” tradition, allied
to a more topographic and less narrative approach to the image, tended to
treat the image as a pretext for the gratuitous production of sound. Deriving
its impetus from vaudeville, it stressed comic accentuation and an intermit-
tent, punctuated temporality, resulting in an antinarrative and antipsycholog-
ical form of humorous attention-grabbing. The later, and ever more domi-
nant, tradition stressed a rigid hierarchy in providing sound effects, separat-
ing the image into zones of importance and of unimportance. A gun wielded
by the villain was an appropriate cue for sound effects, but a tiny background
canary was not. Above all, synchronized sounds had to clarify or underscore
the story by matching mood, tempo, or character psychology, and simultane-
ously create and enforce the hierarchies of the image. Like so much else
discussed here, sound effects hierarchies were ultimately determined by the
necessities of the increasingly complex, psychologically motivated, multishtot
narrative film, which had emerged as the commodity upon which the emerg-
ing industry was erected.

As these trends indicate, one of the principal techniques that classical filmic
narration uses to create continuities out of the heterogeneity of formal ele-
ments is the motivation of discontinuities, such as produced by the material
interruptions of the multishtot image track. The proper “motivation” of these
evident discontinuities results in a unity at another level, that of the narrative
itself. A clearly hierarchical visual space where narratively essential elements
(say, human characters) are foregrounded in visual space against a generic
background allows for an easier link to succeeding shots, since the spectators
are already “primed” to attend to those characters. Images that lack such a
hierarchy (no distinction between “important” and “unimportant” elements,
say) require that each shot in a series be scanned in its entirety, topographi-
cally, and therefore with less emphasis on continuity across edits.

Therefore, any form of sound accompaniment that threatened to disrupt
the priorities of image space by redirecting attention to narratively inessen-
tial elements similarly threatened the norms of the hierarchical and lin-
erized images. Increasingly, all forms of sound work were required to
become, loosely speaking, pleonastic, underscoring and highlighting an
image hierarchy that was, in Hollywood, always mandated by the generative
narration. While it remained possible for a sound practitioner to display skill
in synchronizing effects, the range of permissible sources for sound was
ordered by the demands of the functions preordained for them by the narra-
tive. The villain’s gun could always make a noise, because it was always
important, but individual guns in a battle used as a backdrop only merited a
generic matching. However, the rolling surf in the background of a scene
could get a prominent sound accompaniment provided it did not interfere
with the legibility of the foreground narrative, or if the surf itself, for narra-
tive reasons, was being “foregrounded.” As I will argue below, much the same
series of accommodations were required of musical accompaniment.

Summarizing what would become the accepted wisdom of the post-1910 era,
S. L. “Roxy” Rothapfel counsels that “pianists must remember that they are
merely one of the cogs in the wheel that makes the picture theatre go round,
and as a rule, the people pay to see the picture, not to hear the pianist, so
therefore, play softly when occasion demands, and always remember the picture
comes first.”

“SYNCHRONIZING” THE PICTURE

The modes of audience address characteristic of film music were occasional-
ly at odds with the typical address of the image track. The richest, longest-
lasting, and most familiar tradition of synchronization (as it was called in the
silent era) involved “fitting” music to the picture. Influential how-to books
by Robert Beynon, Edith Lang and George West, and Erno Rapée discuss the
various methods of “synchronizing” music to a film. As implied by the
Edison Kinetogram’s earlier musical suggestions, these writers indicate that the
so-called compiled score, fashioned from existing selections chosen to match
the scenes in the film, dominated musical accompaniment. The two other
major approaches, the original score and the leitmotiv principle, while prac-
ticed by some, were comparatively rare. The reasons for this are several.
Among other things, compilations relied on well-known but not copyrighted
material, thus requiring less new learning and skill to play. Unless written in
advance by a composer, the improvisational use of leitmotivs in a manner
that did not simply repeat the same themes ad nauseam was beyond the
means of almost all musicians. As the true leitmotiv system could not,
frankly, be improvised in any case, compiled scores could still associate char-
acter or situation with a particular melody. Remember, too, that these were
often unschooled and vaudeville-trained musicians skilled in accompanying
hand springs with cymbal crashes and punch lines with rim-shots. Although
it is certain that something like the leitmotiv system was often used to under-
score narrative relationships, this tendency did not by any means dominate,
nor did its compatibility with the character-centered narratives of the emerg-
ing classical cinema somehow define, other musical practices as beyond the
classical paradigm.
Given the number of factors militating against the widespread adoption of the original score and/or leitmotiv system, it is surprising that other methods of “synchronizing” sound to a film have not been given more attention in the nonmusical histories of the era, especially since they help to explain some important aspects of film form. As the Edison films “Love and War” and “The Astor Tramp,” which chose images “to fit the words and the songs,” attest, it became common practice to match film scenes with musical selections whose title, style, or formal qualities were somehow related. Sometimes they matched in tempo, sometimes in mood, sometimes rhythm, sometimes in what they seemed to depict, but often they did so without regard for the connection between scenes. It has even been suggested that cases such as these had a precedent in the *opera ppastica*, whose libretto provided the context and the “excuse” for connecting an assortment of otherwise unrelated “greatest hits” from existing operas. As in a musical revue, the narrative exists for the purpose of motivating a series of show-stopping numbers exhibiting something like what Noel Burch calls “primitive autarky.”

Again, like some forms of sound effects accompaniment and like the earlier song-film, the image-based narrative is subservient to the direct-address of the performance, eliciting praise for the musicians rather than the film.

Musical accompaniment that emphasized discursive relations over submission to the film included the practice of “funning” a picture. “Funning” satirized scenes or entire films through musical puns, commenting on the picture through the title, lyrics, or melody of the accompanying music. A Moving Picture World columnist pointed, for example, to the use of the song “Oh You Kid” in a scene where the pharaoh’s daughter discovers the baby Moses. Another writer argued, “Bad judgement in the selection of music may ruin an exhibition as much as a good programme may help it. Imagine a pathetic scene showing a husband mourning his dead wife accompanied by the strains of ‘No Wedding Bells for Me!’ And yet this exact circumstance was noted by the writer recently.” Such techniques catered to particular audiences and their prejudices and/or drew attention to the musician’s cleverness or stupidity, but did so to the detriment of the film, whose uniformly coherent address was hopelessly fractured.

In conflict with the trend toward regularizing musical accompaniment and robbing it of any unsanctioned voice, many authorities, including Beynon, still considered a necessary component of good sound work to cater to the ethnic, class, and religious particularities of the audience. Funning fell as quickly out of favor with music columnists and large theater orchestras in the ’teens and twenties as did the unreflective use of musical clichés like “Home on the Range” and “Rock-a-Bye Baby” for the all-too-obvious scenes. But in small houses in ethnic enclaves or theaters with a single pianist or organist, these techniques continued to provide a nonnarrative or meta-narrative pleasure while implicitly cementing a bond between performer and audience. To the consternation of manufacturers, this community often came together at the expense of the film.

Theater owners learned quickly how to improve receipts by catering to local audiences through attention to particular ethnic or cultural tastes and traditions. Sometimes this attention amounted to little more than not programming a boxing match along with a Passion play. In deference to the more devout audience one could expect for the latter. In other cases it involved hiring conservatory-trained pianists to cater to Russian, Romanian, and Hungarian Jews who were thought to be “hard to please, musically.” Other writers, however, criticized sound practitioners who spoke in slang, with accents, or ungrammatically because these performers supposedly showed less respect for the audience. Patrons who were supposedly “deceived” into believing that they had attended a genuine (phonographic) talking picture had their illusions shattered by a writer for the Moving Picture World who heard nothing but “Dis, Dat, and Dem,” coming from behind the screen. Of course, this argument was a double-edged sword since one imagines that these mispronunciations unmasked the ruse because they were too much like those of the audience, and not the expected anonymous or “correct” type. Such grammar and pronunciation, and ethnically specific address in particular, were discouraged in the name of attracting a “better” crowd, yet what may have seemed “imperfections” in grammar from one point of view may, from another, have kept local audiences coming back.

Ultimately, the mass-produced and mass-consumed nature of the movies obviated such specific forms of address as it simultaneously sought to eradicate those forms of spectatorship that emphasized communal, collective, and local responses. Instead, they sought forms that encouraged an isolated, individual, and silent audience—that is, the norms of middle-class theatrical spectatorship. Despite these powerful trends, alternative modes of presentation and spectatorship persisted well into the late silent era, as scholars such as Mary Carbine have shown. Theaters in Chicago’s “Black Metropolis” employed jazz bands who, apparently, improvised along with Hollywood films, often adding “hot” numbers where Hollywood would have desired pseudoromantic mood music. One can easily imagine an orchestra with players like Louis Armstrong taking great delight in elaborate musical puns.
played at the film’s expense, and given the enormous improvising talents these bands must have included, all sorts of “funning” from sound effects jokes to race-specific “commentary” must have been common.²²

In contrast to such techniques, books like Rapée’s Motion Picture Moods (1924) and the various collections published by G. Schirmer and Sons encouraged pleonastic and supportive “isomorphisms” between narrative and music. Melody, with its necessary beginning, middle, and end, could mimic the structure and unity of the plot. The climax of the music, they advised, should match the climax of the narrative; fast-paced scenes should be paired with fast-paced music, battle scenes with battle-like program music, and love scenes with romantic melodies, thereby formalizing the growing trend toward redundancy of sensory address and information. Although the appropriate selections for each scene were still discrete and often tonally incompatible, these “how-tos” offered advice and examples illustrating the importance of modulation techniques that could provide continuous musical bridges between themes in different keys. These bridges functioned in a manner that imitated the image’s emphasis on creating a continuous whole out of discrete fragments. These essentially improvisational techniques allowed musicians to shorten or lengthen selections without obvious interruption, and to create a form of narrative integration between separate elements. In fact, some writers came to favor the accompaniment of the piano or organ over the orchestra precisely because it was easier for a single well-trained and flexible player to improvise connections between a piece in F and another in C minor.²³

In addition to mimicking the pluripunctual unities of the image—spatiotemporal wholes constructed across a series of discrete shots—Beynon, and Lang and West, counseled would-be pianists and organists to make themselves and their music subservient to the needs of the plot as well. While Beynon, for example, suggests that it is still appropriate to cater to the “religious, social or political leanings of a neighborhood.” or to make a pun,²⁴ these should be strictly limited since, in excess, they may distract from the picture. Beynon is adamant when he warns, “Never portray musically an emotion contrary to that depicted by the screen action.”²⁵ Lang and West provide a similar warning when they assert that “the player must be careful not to ‘italicize’ the situation so that it becomes distorted or burlesqued. Therefore he should avoid all excesses.”²⁶ Excess is defined in strictly narrative terms as what exceeds the demands of plot, emotional consistency, or tonal coherence. Both books argue that the ultimate goal is not to attract attention to the music, since this would “detract from the smoothness necessary in good picture accompaniment, and the leader should regard smoothness as of paramount importance.”²⁷ In addition, they argue against the frequent use of popular songs on the grounds that they may be overfamiliar or encourage audience members to whistle in the theaters. They likewise argue against folk songs, as they are often associated by the audience member with “some crisis in their careers,” and against all national anthems except as a form of narrative redundancy.²⁸

The general principle unifying these various suggestions encouraged musicians to minimize all those aspects of the presentation that could pull spectators out of their diegetic absorption, drawing attention toward the performers or other audience members and away from the image narrative. Likewise, the new attitudes implicitly argue against too-specific understandings of the audience and too-local means of addressing them and their concerns. All the advice offered by both books assumes the norms of middle-class spectatorship and specifically warns against encouraging any “deviant” spectatorship, whether through sing-alongs, burlesque humor, or parodies of the picture based on local attitudes toward the events or people depicted in the film. Ideally, the accompanist should be able to walk into a theater anywhere in the United States, play along with the picture, and please his (nonspecific) audience.

In what, then, did the vaunted “smoothness” consist? In essence, it seemed to require a kind of musical “narrative integration.” Like the images it accompanied, silent film music moved from a mode of intermittent “attractions,” or a mode stressing the autonomy of individual elements, to a form that stressed the absolute continuity of the music from one end of the picture to another.²⁹ Given the standard techniques and predictably variable quality of musical accompaniment, this was not always such an easy task. As the ‘teens progressed, writers like Clarence Sinn and Clyde Martin argued for the new norm of uninterrupted, continuous musical accompaniment, from the beginning of the picture to the end.³⁰ While Lang and West argue for an approach that starts from a memorable theme or two and builds the entire musical structure from improvisations around them, Beynon, writing only a year later, believes this technique encourages too much “key twiddling.” He argues instead for a more carefully prepared approach, centered not on improvisation but on the construction of a written score.³¹ Beynon summarizes the practices of music “synchronization” as follows:

The secret of synchrony lies not so much in careful timing of the selections as in the accurate judgment of the musical director. Music need not be cut
to fit the situation; but, if care be taken in the finishing of phrases, the
musical setting becomes cohesive—one complete whole that conveys to
the audience that sense of unity so essential to plot portrayal.\textsuperscript{92}

In other words, musical accompaniment should produce precisely the sense
of integrated and satisfying closure as the plot. Constructing such a musical
plot, however, was a fairly complicated process.

Beynon splits scores into three groups—the original, the compiled, and
the semi-original. Since original scores, while now the norm, were compar-
atively rare in the silent era (being confined, for the most part, to prestige
features in large urban houses), Beynon (and to a certain extent, Lang and West)
bases his general musical approach on the "semi-original" score, developed to
rectify "the faults of the compiled score."\textsuperscript{93} The table of contents of one of
Erno Rapée’s books \textit{Motion Picture Moods}, which collect bits of existing
music useful to a motion picture organism, can give a sense of the compiled
score. Rapée divides the material into categories like Aeroplanes, Battles,
Birds, Dances, Fire Fighting, Gruesome, Lullaby, Monotony, National,
Neutral, Oriental, Railroad, Rail, and Storm, and Wedding, using pace, mood,
emotion, nation, events, sources, and other categories as equivalently
descriptive.\textsuperscript{94} A score cobbled together out of a list such as this could easily
collapse into dozens of disconnected fragments, as the pianist shifted from
"Lullaby" when a baby appeared on the screen, to "Gruesome" as the villain
approached, to "Race" as the father raced home to the rescue. The potential
for discontinuity was, without a doubt, the peril that musicians were taught
to avoid most. Like the primitive "mode of attractions," or a "mode of dis-
continuity" that left individual textual units their autonomy, the compiled
score, even when it matched the individual shot brilliantly, could work against
the hard-won narrative continuity of the edited scene. Despite the obvious
necessity of the process, Rapée makes no mention of how these selections
were to be combined.

In general, these authors promote a specifically psychological and narrar-
tional musical aesthetic that requires the accompanist to "fit the predominant
emotion," discover the "significant nucleus around which the silent drama is
built," or locate "the dominant."\textsuperscript{95} He must be "emotional," needs "psy-
chological insight," should recognize emotions in the characters, and "above all,
learn to read facial expressions."\textsuperscript{96} As if teaching their readers to write scen-
arios, they clearly describe the Hollywood film as a character-based, psy-
chologically motivated, linear narrative form. Given that linearity, smooth-
ness, and coherence are the musical goals, as they are for the picture, we
might expect them to encounter some of the problems characteristic of the
multiunit image-narrative.

In contrast to Rapée, Beynon as well as Lang and West spend much of their
books teaching techniques of modulation, or improvised key-changing, in
order to connect the various selections into a more or less continuous whole.
Specifying their understanding of film continuity, Beynon stresses upper-voice
melody, and Lang and West a "solo-instrument-with-accompaniment" approach,
in order to achieve the musical equivalent of the foreground/background struc-
ture so essential to the narrative film image.\textsuperscript{97} Keeping the dominant melody
audible by placing it in an upper register or by giving it a different "voice" on the
organ is not so different from staging the action so that the characters are centered
and in focus, in order that we may follow them from shot to shot. Continuous melody,
therefore, assumes a role analogous to the image foreground and is susceptible to
its articulations. By this logic, the melody would change when scenes changed or
when locations shifted decisively.

The manner in which the authors handle cross-cutting and multishot
scenes involving inserts or, as they were known at the time, "flashbacks" is
very instructive for understanding how musical and filic continuity were
defined at the time. Given the implied link between melody and neoclassical
unities of time, place, and action, the flashback posed a particularly tricky
puzzle. As a result, each writer warns against mechanically mimicking the
material discontinuity of the image (the succession of shots) by changing the
accompanying music with it. Beynon argues that the musician should fit the
predominant emotion of a scene and advises against changing the theme
with each flashback or insert shot (allowing that the tempo might be
changed, if not the melody). This strategy attempts to avoid the effect of
choppiness in what, after all, are not separate scenes but only separate shots.
Lang and West similarly argue that during a "flash" (i.e., an insert, mental
image, another location, etc.) the music should not change its motive. Each
recognizes that the material discontinuity of the image is merely a change in
narrational level, "therefore the music should not change its character during
a flash-back, but it should be very much subdued . . . . The principles are: in
most cases not to disrupt the continuity of the music while the flash-back lasts,
but to change the intensity by playing the music, characteristic of the main
action, in a dynamic degree of loudness or softness which befits the second-
ary action."\textsuperscript{98} The authors add that in a cross-cut scene, however, a "competition"
between themes might very well be appropriate.\textsuperscript{99} Rapée makes iden-
tical recommendations regarding volume and melodic "competition" in his
later \textit{Encyclopedia of Music for Pictures} (1925).\textsuperscript{100}
These musical recommendations indicate a great deal of sensitivity to the intricacies of Hollywood narration, and are so flexible that they can be applied to a much later film like Old San Francisco (1927), where scenes between two musically delineated characters are accompanied by the theme of the one who dominates the scene. As in related suggestions requiring musicians to vary tempi according to the plot, or volume according to the proximity of the narrative climax, here the authors recapitulate many of the principles by which the Hollywood narrative is constructed, further submerging music within the demands of classical narration. Despite all the possibilities for audience address available to musicians, the address of the fictional image-narrative is absolutely sovereign for these writers, and would become almost inevitably so for all musicians in the sound era.

THE PERSISTENCE OF MARGINAL FORMS

Pre-talkie sound was, it is clear, a multiple and heterogeneous set of phenomena comprising a wide variety of functions. From “barking” to hailing, within and outside the text, sound emphasized its presence and its necessity by organizing and reorganizing the cultural practice of cinema. Sound often placed an enormous emphasis on the status of cinema as a kind of performance, whether through the elaborate practices of synchronous sound accompaniment, lecturing, or musical presentations of various kinds. Before 1910, sound very often dominated or threatened to dominate the image, in terms of conception, production, presentation, and even reception. Sound-based genres like the travel lecture and illustrated song, the visualized record, or the film that functioned as a pretext for musical or effects performance occupied a prominent place in the world of early cinema. As the image-based narrative began to assert its dominance, however, sound’s functions began to change. Sound could serve as a clarifying addition, as in the lecture, or it could provide important narrative motivation for on-screen events by supplying a missing sound cue. Or, musically, sound could aid in characterization; it could underline emotions, set moods, or emphasize tempi. On the other hand, music and commentary could work against narrative unity—by punning, by meta-commentary on the situations or characters, by irony, all of which might serve as a way to mediate between the film’s anonymous, mass-cultural address and its ethnic, economic, and social particularities.

While, in general, sound (especially music) did tend to become more and more subservient to the narrative, none of these traditions of sound disappeared entirely; they persisted on the margins of the classical style or in alternative modes such as the avant-garde. The role or roles that recorded sound was to fulfill were therefore anything but self-evident and would certainly be (as I hope I have suggested) the object of struggles for years to come. Simply because the transition to sound was relatively smooth does not mean that we should ignore the processes and principles of exclusion by which the classical style operates and defines itself and its boundaries.

Despite the growing symmetry between the form and address of the image-narrative and that of sound, even in studio-composed original scores, important elements of another less fully narrative mode persisted. As the American Cinematographer noted in 1928:

The screen drama of the immediate future will involve, not conflicting character, but competitive noise. The exposition of the rural drama will present the problem: “Who will emerge victor in a hog-calling contest and win the girl.” The development will proceed from one minor crisis and crash to another with steadily increasing volume until the heavy, with a portable boiler shop, threatens to drown out the hero's calliope. The denouement will come when the heroine blows up the field of oil tanks with her cigarette and scatters fragments of the heavy all over the pleasant meadow.

As sarcastic as this seems, the writer was fairly accurate. Vitaphone's 1927 Old San Francisco, otherwise a standard “silent” melodrama with silent film sound conventions based on character and ethnic difference, ends with a spectacular earthquake lasting several (ca. eight) minutes chock-full of synchronized sound effects. Just as the character O'Shaughnessy’s Irishness is motivated less by the inherent demands of the narrative than by the opportunity it offers to differentiate him by musical theme and written titles, so the earthquake’s spectacle is unwarranted by the film’s narrative economy. Earlier parts of the film are sporadically accompanied by diegetic sound effects often tied to dramatic narrative points. There is even an audio “point-of-view” shot showing, first, someone cocking an ear, then a close-up of horses’ hooves. The earthquake sequence, however, contains the film’s overwhelming majority of sound effects, and does so in an excessive manner.

The nonstop symphony of crumbling buildings, running victims, roaring fires, and Anna May Wong’s climactic scream offer a density of effects unheard in silent film, yet not entirely without precedent. Like some prede-
cessors, this spectralization of sound was not primarily narrative in motivation, nor essentially “realist.” Indeed, it stalls the plot, and its length is narratively unjustified. It does, however, foreground the technology of sound recording and sound reproduction, offering its perfect performance for appreciation and admiration. The sound does not support the narrative here, but narrative supports the gratuitous display of sound. Likewise, in The First Auto (1927), synchronized effects are not primarily narrative. From the opening drumbeat hooves and the shout of “C’mon Hank” (over an image of a crowd), which is followed by an intitule of the same phrase, through the sudden eruptions of a synchronized “Bob!” which is followed by conversation on intitules, to the gavel that starts the auction, sound is foregrounded not in order to present narratively important information (else it would not repeat “C’mon Hank”) but to address the audience directly—to hail them—to say, “Hey! Look and listen! This is important.” These visible sync-events serve to foreground the spectacle of technology while disregarding narrative.

Interestingly, the idea of sound, and especially synchronization, as performance carries over into the sound era, and into the realm of technology, as the ability of the device to ensure accurate synchronization seems foregrounded, itself, as a kind of performance.102 The earliest patrons of the Vitaphone certainly must have appreciated the new technology along these lines. Don Juan (1926), The First Auto, and Old San Francisco offered numerous opportunities for a spectacularization of the device’s performance through foregrounded sync-events. At points, the stories seem to exist purely to provide an excuse for a display of effects, in a manner that recalls the excuses for dramatic lighting effects in many DeMille productions.103 However important narrative may be in these films, at times it clearly takes a backseat to the display of pure technological marvel, and cinematic pleasure momentarily divorces itself from plot.

Nor did the “sound over image” tendency die with the silent cinema. The illustrated song, in particular, persists today in the form of music videos, for example, as it did in the thirties. A cartoon like Betty Boop’s Snow White (1933), featuring Cab Calloway, included songs owned by Paramount, its distribution company. Given the structural primacy of the sound track in these song-films, it is easy to see how they descend from the illustrated song. In a similar fashion, The Big Broadcast (1932) (also by Paramount, and featuring Cab Calloway) exploits several of the same songs as the Boop film, and even has an ingenious opening segment which mimics, with “real-world” images, the isomorphic techniques of the cartoon’s “mickey-mousing.”104 Again, although this unusual sound/image strategy gets a certain amount of narrative justification (it takes place in the clock-obsessed world of a radio studio), its chief motivation derives from the economic and structural importance of the sound track.

Although the “isomorphisms” of musical accompaniment and of the cartoon seem fully a part of the Hollywood narrational system, isomorphic techniques also supported one of the strongest nonclassical traditions of film sound. The basic principles of synchronization, while certainly amenable to the classical film and capable of clarifying its images, were by no means confmed to it. From Dziga Vertov’s Enthusiasm (1931) and Eisenstein’s Alexander Nevsky (1938), through mickey-moused cartoons, to Peter Kubelka’s Unscere Afrikareise (1961–1966), and Larry Gottheim’s Mouches Volantes (1976) and Four Shadows (1978), a marginal but powerful tradition of synchronization assumes the foreground. The sound and image relationships in these films are among the most carefully “synchronized” in the history of film, but they work in a manner that, while partially applicable to the Hollywood narrative, is generally quite opposed to it. In contrast to the more familiar sense of the word synchronization, in these films both sound and image “articulate” one another, ignoring the conventional divisions and hierarchies appropriate to narrative. As we have seen, however, traditonal, narrative synchronization may be a “special case” of this more general principle.

A Preordained Role for Sound?

As I’ve suggested, not only did Hollywood waver in its allegiance to representational modes during the transition, moving between discursive and diegetic forms, but the mode of representation understood to be characteristic of the very technology of sound was likewise up for grabs. Was sound an effect? Was it narration? Was it clarifying commentary? Should it function as an added form of omniscient or restricted narration? Was it realist or spectacular in nature? Even its technical nature was in dispute. Was it closer in form and purpose to the phonograph, the telephone, or the radio? Each device, while useful for grasping some aspects of the sound film phenomenon, validated different techniques and implied different representational norms. So, before “sound as sound” could be “inserted into . . . the Hollywood style,” it had to be determined just what sound “was” and what its appropriate functions were.

Recent histories have emphasized the ease with which sound was adopted into the classical paradigm, deafening and blinding us to the very real discon-
continuities of the transition and closing off the possibility of a history based on the tension between competing norms. Ideas of technological and representational self-evidence and self-identity and notions of functional equivalence derived from the end point of a historical transformation encourage us to hear history as a monologue. Concepts like jurisdictional struggle, technological redefinition, and functional near-equivalence provide the contradiction and conflict that can motivate historical change, encouraging us to listen to quieter voices from the past.

While the basic issues I have identified as arising from the emergence of the technical media are in some ways of quite general import, the manner in which any particular representational technology is normalized, regularized, and institutionalized is always quite specific. Just as the nature of Marey's and Helmholtz's experiments placed specific representational demands on contemporary sensory devices in order to render them adequate to the rigors of science, the emerging phonography and film industries asserted their own requirements for the proper implementation of both camera and sound recorder, shaping their characteristic uses in thoroughgoing ways. Contrary to what we might expect, explicitly theoretical debates about the general nature of technological representation were central to the development of Hollywood representational norms, and helped determine how sound technologies were understood and deployed. Particularly at moments of institutional crisis, reorganization, and transformation, these debates assumed a discursive importance that often shaped technological research and aesthetic experimentation, serving both as practical standards and regulative ideals.

In this chapter I will explore several influential manifestations of sound theory, examining them not only for their internal coherence and insight but also for the role they play in determining different forms of representational practice. As I have suggested in previous chapters, I am not only interested in how explicitly theoretical arguments illuminate our understanding of aes-