

Yale French Studies

Moving Lips: Cinema as Ventriloquism

Author(s): Rick Altman

Source: *Yale French Studies*, No. 60, Cinema/Sound (1980), pp. 67-79

Published by: [Yale University Press](#)

Stable URL: <http://www.jstor.org/stable/2930005>

Accessed: 20/10/2011 10:42

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Yale University Press is collaborating with JSTOR to digitize, preserve and extend access to *Yale French Studies*.

<http://www.jstor.org>

Rick Altman

MOVING LIPS: Cinema as Ventriloquism

Conventional wisdom has it that the sound track in classical narrative films is by and large redundant. We see a door slam, we hear a door slam; the sound intensifies the sense of reality initially produced by the image, it anchors the visual reference to a slamming door in our auditory sense system, it precludes any distantiating which might possibly be produced by the sight of a slamming door unaccompanied by its familiar sound—but all in all nothing entirely new is contributed by mimetic sound, whence the acknowledged lack of independence of the sound track. Nearly every critic who has written about sound in film has made claims of this sort. In an effort to provide a new starting point for theoretical consideration of the sound track, I will challenge the most cherished assumptions of conventional sound analysis. First, I will show that the conventions of classical narrative have a strong tendency to make the *image* redundant. Next, I will demonstrate the inadequacy of the very notion of redundancy. Finally, I will postulate a new model for the conceptualization of sound-image relationships in the cinema: the sound track is a ventriloquist who, by moving his dummy (the image) in time with the words he secretly speaks, creates the illusion that the words are produced by the dummy/image whereas in fact the dummy/image is actually created in order to disguise the source of the sound. Far from being subservient to the image, the sound track uses the illusion of subservience to serve its own ends.

Redundancy of the Image

If we were to formulate descriptive rules defining the probability with which any given phenomenon might appear in a classical narrative film at any given moment, we would undoubtedly soon be

led to a rule such as this: an individual who speaks will in all probability be the object of the camera's, and thus of the audience's, gaze. In the political world, the right to free speech conveys a certain political power; in the narrative world, the right to speech invariably conveys narrational power, for by convention it carries with it a secondary right, the right to appear in the image. While there are many exceptions to this rule, even within the world of classical narrative cinema (e.g. Russell Rouse's *The Thief* /1952/, in which the protagonist never utters a word during the entire movie, or Jean Cocteau's *Les Parents terribles*, in which the listener rather than the speaker regularly appears on the screen in dialogue scenes), in general we may say that actors gain the right to a place in the image by virtue of having previously obtained a spot on the sound track. *I speak, therefore I am seen*. The most blatant application of this principle—and the heart of classical *découpage*—is the shot/reverse-shot sequence in which the camera is alternately pointed at each speaker in a dialogue situation.

Why, when so many theoreticians have called for a relationship of counterpoint between sound and image, should classical narrative films rely so heavily on a strategy of pointing the camera at the speaker? What advantage could possibly come from doubling the words with an image of the actor who produces them? In order to answer these questions it is necessary to reflect an instant on the general status of language in narrative cinema. Among sounds, language clearly reigns supreme. *The Jazz Singer* is commonly acknowledged as the first sound film not because it was the first to bring sound to film (*Don Juan* had done that a year earlier), but because it was the first to bring synchronized dialogue, i.e. language, to film. Before the experiments of the sixties in direct sound recording, hardly a word was uttered in a narrative sound film without a carefully planned simultaneous reduction in the level of all other sounds (music, background noise, synchronized effects). Sound editing is first and foremost word editing; its primary purpose is to assure the clarity of the film's dialogue. This emphasis on dialogue derives quite obviously from the role played by dialogue in the constitution of the narrative line. Without the dialogue the images

Rick Altman

are ambiguous, incomplete, underdetermined. But such a formulation is misleading, because it implies that the same images would be used with and without dialogue. This is, however, far from being the case. Perhaps the single most important difference between silent and sound narrative films lies in the latter's increased proportion of scenes devoted to people talking—devoted, that is, to moving lips.

The more we expose the importance of dialogue in narrative cinema, however, the less comprehensible becomes the practice of pointing the camera at the person talking. If it is the dialogue, the language, the words which count, then why show lips moving in time with the sound track? We can best answer this question by recognizing the effect of those moving lips: they transfer the origin of the words, as perceived by the spectator/auditor, from sound “track” and loudspeaker to a character within the film's diegesis. To put it another way, pointing the camera at the speaker disguises the source of the words, dissembling the work of production and technology. But such a recognition reveals, finally, the imprecision of the language I have been using. To say “pointing the camera at the speaker” is to have already been deceived by the ideology of synchronized sound. “Pointing the camera at the (loud)speaker” is precisely what does not happen in this case. Portraying moving lips on the screen convinces us that the individual thus portrayed—and not the loudspeaker—has spoken the words we have heard. The redundancy of the image—seeing the “speaker” while we hear “his” words—thus serves a double purpose. By creating a new myth of origins, it displaces our attention 1) from the technological, mechanical, and thus industrial status of the cinema, and 2) from the scandalous fact that sound films begin as language—the screen-writer's—and not as pure image.

Both of these repressed areas require more extensive commentary than I can provide here. In the case of the masking of technology, it is important to recognize how the dissimulation of sound technology complements and reinforces the better known masking of image production. When we say that *sound* is redundant, that it anchors the image, that it adds nothing, we are really saying that the major function of sound, considered from the standpoint of the image, is to

convince the viewer that the image exists independently of the technology which would mark it as a fiction. Similarly, when I say that the *image* of a “speaker” is redundant, I am really claiming that such an image, viewed from the standpoint of the sound track, serves to convince the auditor that sound exists independently of any technology which might mark it as a fabrication. The best way to convince a jury that a liar is telling the truth, as every Perry Mason fan knows, is to find another liar to corroborate the first liar’s testimony. Image and sound: as long as each reinforces the other’s lie then we will not hesitate to believe them both. This complementary ideology, whereby each track helps to hide the work of the other, must necessarily escape any theory which postulates the primacy of either one of the tracks.

The repression of the screenwriter has received less comment, but it derives from a no less important strain in the history of film theory. From the early days of sound, theoreticians have warned against the mixing of the cinema’s pure image orientation with the degraded language and practices of the theater. For the coming of sound represented the long-feared return of the cinema’s theatrical repressed. As the thirties progressed, however, it became increasingly clear all over the world that language, far from being anathema to the cinema experience, lay at its very heart. Unable to suppress language, cinema theory transferred its resentment to the source of that language, banning the screenwriter eternally from serious consideration. With the *auteur* “theory” the screenwriter was finally done away with all together, and the scandal of language’s dominance over and independence from the image was further repressed. As another in a long line of attempts to repress cinema’s theatrical origins, *auteur* criticism is a proper complement to the screen’s moving lips which so effectively disguise the role—indeed the very existence—of the screenwriter.

One further function must be attributed to classical narrative cinema’s tendency to “point the camera at the speaker.” Whereas early critics proposed to preserve the unity of the medium by simply neglecting, dismissing, or marginalizing the sound track, classical narrative cinema performs the same operation in a far more clever and

Rick Altman

convincing manner. Cinema, it is often claimed, is a mixed art, for it reaches its audience through two separate channels, audio and visual. Yet no one will claim that those humans who are watching/listening are “mixed” by virtue of their ability to employ simultaneously more than one sense. Classical narrative cinema takes advantage of this fact by using the model of human unity in order to bridge the image-sound gap. The moving lips which anchor the sound on the image track, and which appear to be producing the sounds we hear, simultaneously permit the cinema to constitute its own unity by identifying the two tracks of the cinematic apparatus with two well known aspects of human identity. Paradoxically, then, the spectator—who knows there is no contradiction between seeing and hearing—serves as cinema’s mirror, the *speculum* in which the film synchronizes its motor skills, establishes its own identity, and thus accedes to the Symbolic realm of language. This “reverse specularity” whereby cinema constitutes its own unity is a necessary part of the process whereby the spectator/auditor rehearses his own mirror-stage experience.¹ Of course the unity affirmed by this interchange is no more actual for cinema than it is for the divided auditor/spectator. If the human audience accepts the cinema’s unity, it is because it cannot affirm its own without admitting the cinema’s; conversely, the cinema appears to assent to the unity of the human subject only in order to establish its own unity. This collusion resembles the symbiotic relationship discussed earlier whereby image and sound count on each other to erase each other’s mode of production. Only when mirrored in the other does each side seem complete.

As common as the practice may be, “pointing the camera at the speaker” is only a special case of the more general pressure in classical narrative cinema to identify visually the source of a sound. Among the most basic of camera movements, defined as a function of the narrative, is the tendency to move the camera to a sound, to point it at the area from which the sound is coming (thus turning off-screen sound into on-screen sound). Now, as we well know from the

¹For the metaphor of specularity, as applied to the cinema, see Jean-Louis Baudry, “Ideological Effects of the Basic Cinematographic Apparatus,” trans. Alan Williams, *Film Quarterly*, 28 (Winter, 1974–75), 39–47.

preceding analysis, this formulation is entirely improper, because the camera does not actually point at the immediate source of the sound, but rather points at the character (machine, movement, etc.) which we are to consider as the source of the sound. In other words, the camera movement is here again at the service of the ideology which masks the film's image-sound split. At this point a few examples are in order.

Renoir's *La Règle du jeu* reveals the extent to which the search for a sound source can be central to shot composition in the work of a director renowned for his visual sense. At the very beginning of reel two (in 16 mm.) the servants of the château are seated at the downstairs table. As we watch the servants eat and listen to them talk about aristocratic conventions, the "Minute Waltz" fades in on the sound track. Is this music diegetic or not? It's certainly not on-screen music, but it could be wafting down from upstairs, where one of the guests might have just sat down to the piano. This suspicion seems confirmed when the camera follows Schumacher up the stairs—we seem to be using the waltz as a sound bridge to the next scene. But half-way up the stairs Schumacher encounters Marceau, whom we follow back down. As the scene continues, the camera proceeds to wander, as if continuing to seek out the source of the still unexplained waltz. Finally, it lights on a radio which up until then had remained outside the frame. There the camera holds for an instant and the sequence ends, as if this radio were the answer to the scene's sound enigma. (Indeed, if we have followed the live/mechanical performance dialectic in the film, we find strong reinforcement for this notion.)

A similar phenomenon takes place in a far less traditional manner in Rohmer's *Perceval*. At the beginning of the film Rohmer establishes with the camera two *loci*, one diegetic, the other the place where sound effects and music are produced. As the movie proceeds we often cut—or even pan—from the diegetic space to the "sound" space, as if to explain the source of the many aspects of the sound track unexplained by the diegesis. At one point we even cut to the "sound" space in order to witness the fabrication of the sound effects which we have previously read as "clanking of armor." Later on in

Rick Altman

the same film, a similar progression regulates the dialogue-image relationship. We often hear words spoken long before we know where they come from; these disembodied words are then associated with a specific character by means of a closer shot and lip synch.

What is it about sound that encourages its off-screen or disembodied use, and seems to call for rapid location of its source? Much has been written about the comparative phenomenology of sound and image, but no one, to my knowledge, has clearly delineated one important asymmetry between the two. At first, image-without-sound and sound-without-image would seem to be complementary and symmetrical situations. In fact, however, two considerations make these configurations quite different. First, an image without a sound differs from a sound without an image in that the former is a perfectly common situation in nature (a person standing quietly), while the latter is an impossibility (sounds are always produced by something imageable). Thus the completion of the former paradigm depends on the object within the image (the person may choose to say something), while the completion of the latter depends on the auditor (who must look around and *find* the source of the sound). Images call for no action on the part of the auditor. Or, to put it as Bresson has been reported to say, “A sound always evokes an image; an image never evokes a sound.”²

This fact stems from a purely physical difference between sound and light. Under normal circumstances light travels in straight lines only; only highly polished surfaces like mirrors will reflect light regularly enough to carry a recognizable image around a corner (the image consisting not of a single *point* of light but of an *area* of light). Sound, on the other hand, travels as a point rather than as an area, and thus provides fewer problems of reflection; even without special preparation, any surface will reflect rather faithfully any kind of sound. It is this difference which gives us the illusion of “hearing around corners” when we cannot see around corners. The consequent restriction of sight to those things present (and conversely, the definition of presence in terms of visibility) has an important

²Noël Burch, *Theory of Film Practice*, trans. Helen R. Lane (New York: Praeger, 1973), p. 90.

counterpart in the non-restriction of hearing to visibly present sources. The ramifications of sound's relative freedom as compared to the image are many; two in particular concern us here: 1) sound's ability to be heard around a corner makes it the ideal method of introducing the invisible, the mysterious, the supernatural (given that image = visible = real); 2) this very power of sound carries with it a concomitant danger—sound will always carry with it the tension of the unknown until it is anchored by sight.

As Béla Balázs has pointed out, there is “a considerable difference between perceiving a sound and identifying its source.”³ By virtue of its ability to remain sourceless, sound carries with it a natural tension. Whereas images rarely ask: “What sound did that image make?” every sound seems to ask, unless it has previously been categorized and located: “Where did that sound come from?” That is, “What is the source of that sound?” Far from ever being redundant, sound has a fundamental enigmatic quality which confers on the image the quality of a response, and thus a certain sense of finality (such as that sensed when the radio is located in the *Règle du jeu* sequence). *The image, in terms of sound, always has the basic nature of a question.* Fundamental to the cinema experience, therefore, is a process—which we might call the *sound hermeneutic*—whereby the sound asks *where?* and the image responds *here!*

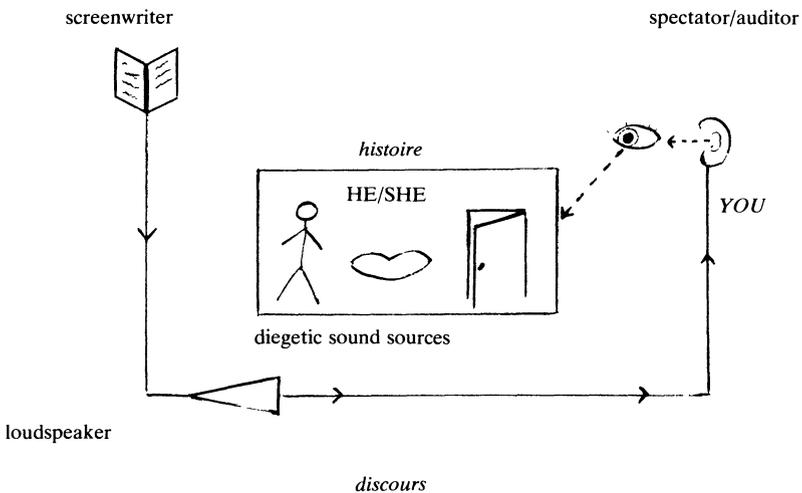
Once again, however, we must be reminded of the ideological play which underlies this hermeneutic. The loudspeaker makes a noise, thus producing tension, nervousness, frustration in the auditor/spectator. He looks around and sees no probable source for that sound except the image on the screen. Lacking any other source, and needing to anchor the sound at all costs, he accepts the notion that it comes from the image when in fact it comes from the loudspeaker. This rerouting of the sound from apparatus to diegesis is part of a fundamental progression in cinema whereby the *discours* connecting producers and consumers is masked by the *histoire* of the diegesis.⁴

³Béla Balázs, *Theory of the Film: Character and Growth of a New Art* (New York: Dover, 1970), p. 212.

⁴This process has been dealt with in a particularly clear and succinct fashion in Christian Metz, “Histoire/Discours (Note sur deux voyeurismes),” in *Le signifiant imaginaire* (Paris: Union générale d'éditions, 1977), pp. 111–20.

Rick Altman

The loudspeaker “speaks” directly to the audience; the sound track provides a sophisticated *discours* designed by the screenwriter and sound technicians to manipulate their audience, but this *discours* is recuperated to *histoire* when it is attributed to characters in a *diegesis*. Only because this diegesis seems so consummately unaware of the audience, however, does it succeed in keeping that audience from discovering the real source and purpose of this sound track and image made especially for their pleasure and fixation. In other words, the enigma initiated by the sound track is only a pseudo-enigma designed to short-circuit the logical progression whereby the spectator/auditor would continue his search until he found the technological source of the sound. By providing the audience with a more logical, simpler, and less threatening answer to the question “Where does the sound come from?” the image diverts attention from the sound’s true source, rerouting us instead through a visible, but fallacious, origin, as indicated in the following diagram:



This model radically questions the status of the so-called “ideology of the visible” as it relates to sound-image concerns. Far from being marginalized by the visible, far from being a slavishly redundant accompaniment to *His Majesty the Image*, sound now appears to be a far more clever Jacques than he at first seemed, for it is now apparent

that sound *uses* the visible to further its own cause. Asking all the questions, while the image is allowed to reveal no more than the answers, the sound track attributes its work to the image only in order to keep its own responsibility secret. Using the ideology of the visible as a front, the sound track remains free to carry on its own business.

The Sound Track as Ventriloquist

The notion of redundancy, as we have seen, is hardly a tenable one. Far from producing the same thing a second time, “redundancy” creates a functional supplement which invests its excess energy in the masking of cinema’s multi-media status. In other words, “redundancy” of image and sound is precisely similar to the “redundancy” of child and mirror image at the mirror stage—only by neglecting the important differences which exist between the two mirrored phenomena are the child (and cinema) able to constitute their identity, their unity. The question is thus no longer “Which is redundant.” but rather “What is the function of apparent redundancy?” “How and why does cinema constitute its unity?” and finally, “What are the conditions necessary for us to believe that sound comes from the image?”

These are not familiar questions; except for passages on synchronized sound in technical editing manuals, they are hardly addressed at all by the cinema community (though they clearly remain implicitly present in the technological literature on loudspeaker placement, sound-transmitting screens, and the use of stereo). We find much more help with these problems in a most unexpected place: the manuals and handbooks describing the fundamentals of ventriloquism. For the ventriloquist’s problem is exactly that of the sound track—how to retain control over the sound while attributing it to a carefully manipulated lifelike dummy with no independent life of his own. Indeed, the ventriloquist’s art depends on the very fact which we have found at the heart of sound film: we are so disconcerted by a sourceless sound that we would rather attribute the sound to a dummy or a shadow than face the mystery of its sourcelessness or the

Rick Altman

scandal of its production by a non-vocal (technological or “ventral”) apparatus.

In order to foster this pseudo-identification of the sound source, however, certain conditions must be met. The ventriloquist must avoid moving his own lips, thus disguising the true source of the sound (in the theater this function is fulfilled by dissimulation of the loudspeakers, often behind a sound-transmitting screen). He must also be able to move the dummy in time with his voice—not only lip movement but all other bodily motion as well must be rhythmically consistent with the sound (this is the obvious function of lip-synching in film, and the primary reasons why the camera must so often be pointed at the “speaker”). “Throwing” the voice thus refers to the process whereby a ventriloquist lures his audience into believing the (strong but mistaken) visual evidence rather than the (weak but correct) testimony of the ears. What is involved here is of course nothing more than the familiar “redundancy” which has proven of central importance all along. Unless the ventriloquist can produce a believable “redundant” lip movement in the dummy, he cannot induce us to transfer our allegiance from the aural to the visual witness within us. Paradoxically, then, a process which ultimately unifies the senses is based on their conflict and disagreement. Only by duping our powerful visual sense to be a false witness can the ventriloquist disguise his own responsibility for the sound emitted. As in cinema, the so-called ideology of the visible is so self-confident that it is easily led to perjure itself solely to gain the right to be seen in the witness’ stand.

But why would a ventriloquist want to perpetuate the illusion that not he but his dummy is speaking? Why give away one’s right to speech? Any confirmed fan of ventriloquism can easily provide the answer. In fact, ventriloquism manuals often spell it out clearly:

First of all, it is most important to analyze your own character carefully so that when you create the dummy’s character there is a good contrast. According to Freud, all of us have hidden desires that we suppress. The most successful ventriloquists have let these hidden desires be expressed in the personality of their dummies.

⁵Darryl Hutton, *Ventriloquism* (New York: Sterling, 1974), p. 28.

Perhaps the best known example of this relationship is provided by Edgar Bergen, whose technique is aptly described in these pointers for the aspiring ventriloquist:

You must remember, however, that **your partner will play the lead in your act and usually have the best lines.** *His voice and personality should be richer and stronger than yours.* That internationally famous ventriloquist, Edgar Bergen, is not so well known as his partner Charlie McCarthy. Bergen is, or seems, quite a shy person while Charlie has all the cheek in the world.⁶

This split helps us to understand why it is that the art in question has come to be known as *ventriloquism*. It is of course not the belly which speaks at all, yet ever since the Greeks the attribution of one's own voice to a synchronized dummy has been identified with the belly. Called *engastrimanteis* (belly-prophets), the ventriloquists of ancient Greece were taken to be prophets and were said to emit their prophetic voice from the belly: the head-voice may produce apparent truths, but the body-voice reveals hidden truth. This identification of the ventriloquist's voice with the belly, the *locus classicus* of that which is body (identified as it is with the bodily functions of eating, excretion, and sexuality), squares surprisingly well with the role traditionally allotted to the ventriloquist's disguised voice. Whereas the head-voice speaks the society's polite language, the body-voice speaks a more sincere, personal, and unguarded language, a language no longer watched over by the censorship of the conscious mind.

Returning to the cinematic model which led us to the art of ventriloquism, we find a curious reversal of our previous conclusions. **If the image is able to speak that which the sound track dare not say on its own, then we can rightly claim that the image represents the sound track's repressed,** the material which it has hidden from view until that material can be attributed to another source. But earlier I claimed quite the reverse—in view of silent cinema's uneasy birth out of theater and the fear of many early theoreticians that sound would simply plunge cinema back into the dread language-orientation of theatrical practice, I identified sound as the image's repressed, brought back to the surface with the coming of sound. By now,

⁶Douglas Houlden, *Ventriloquism for Beginners* (New York: A.S. Barnes, 1967), p. 24. Emphasis is Houlden's.

Rick Altman

however, such a reversal should no longer appear contradictory. Others have stressed the role which sound plays in completing and reinforcing the image; I have underlined the complementary relationship whereby sound uses the image to mask its own actions. Far from undermining each other, these two approaches are part of the overall strategy whereby, as we have seen, each track serves as mirror for the other and the spectator for the two together. Neither track accompanies the other, neither track is redundant; the two are locked in a dialectic where each is alternately master and slave to the other; this arrangement so suits both tracks that they studiously perpetuate the myth of cinema's unity—and thus that of the spectator—as if (and they are right) their very lives depended on it.

The fundamental scandal of sound film—and thus the proper starting point for a theory of sound film—is that sound and image are different phenomena, recorded by different methods, printed many frames apart on the film, and reproduced by an illusionistic technology. Voices are uttered by cardboard cones, by mechanical instruments, by machines designed to meet the challenge of a world in which cities are too populous to be addressed by a single unaided human voice. Cinema's ventriloquism is the product of an effort to overcome the sound-image gap, to mask the sound's technological origin, and to permit the film's production personnel to speak their sub-conscious mind—their belly—without fear of discovery.