Rick Altman

The Technology of the Voice Part. II.*

I

Missing from the three approaches roughly sketched in Part I. is a sense of the increasing specificity of microphone technology to cinema tasks —and thus of the increasing influence of cinema as such on microphone development. If the carbon transmitter was first developed for practical telephony and the condenser transmitter for laboratory sound calibration, the moving coil microphone is perfected quite specifically with sound cinema in mind (though on the basis of principles elaborated for other purposes); the ribbon and other directional microphones are the first to be created not only specifically for the cinema but also based on principles directly derived from cinema needs. Somewhere in our battery of approaches, we need a weapon capable of recognizing and accounting for this growing specificity.

On the whole, the extent to which microphone usage is medium-specific should serve as a warning that certain aspects of microphone development remain to be explained. It is instructive, for example, to contrast Green and Maxfield's article championing the carbon mike for public address usage in 1923 to Wente's 1936 overview of Bell research, in which he explains why «even a high-quality carbon microphone is unsuitable for sound picture recording »/1. Taking the «distant pick-up conditions » (p. 190) that prevail in sound cinema as a first cause, Wente never goes on to ask *why* it is that distant pick-up conditions continue to prevail when closer pick-up conditions would have provided better sound. He simply takes it for granted that new microphones are needed for the talkies (in this case his own redesi-

---

*Part. I. appeared in IRIS Vol. 3, nº 1.

/1. Green and Maxfield, « Public Address Systems » (1923) and Wente, « Contributions... » (1936), p. 190 (see notes 7 and 45 for exact references).

gned condenser mike)/2. The fortunes of the moving coil microphone involve similar seeming anomalies. Throughout the late twenties the call had been out for a microphone which would have lower impedance than the condenser mike (thus doing away with the heavy amplifier in the microphone case, increasingly problematic as microphone booms and single miking grew in popularity), which could be used farther from the sound source than the carbon mike, and which would have response characteristics better than either of the two existing systems. Bell's moving coil microphone met all of these requirements and more, yet the immediate popularity of the bidirectionnal ribbon mike kept the moving coil mike from winning the popularity that its qualities would seem to predict. We must conclude, I think, that more is at stake than the simple engineering standards of fidelity and versatility.

Similar conclusions are suggested by the treatment of reverberation within the Hollywood context. Sound is received by human subjects and microphones alike as a combination of direct and reflected impulses, the amount of the latter depending on the size of the room, the material of the walls, the number of people present, the distance of the receiver from the source, the frequency of the sound, and numerous other minor factors. Now, a long tradition within Western music-making, closely connected to the large romantic orchestra and consequently large concert hall, as well as to an even longer tradition of church music played in large resonant buildings, had by the 1930s created a preference for the reproduction of music with a high degree of reflected sound/3. (Characteristically, Thomas Edison alone went against this current, preferring the purity and intelligibility of direct sound even for music)/4. Recognizing the existence of this aesthetic judgment, engineers christened the quality that produces a high level

---

/2. Jones takes an exactly similar position : « In conclusion, it may be stated that the condenser and carbon types of microphone have been developed to a point where there is little to choose between them from the standpoint of quality of transmission... Although requiring less amplification than the condenser microphone the extent to which the carbon microphone is used at present is limited by the higher noise level obtained. The condenser type of microphone has therefore been adopted for most of the recording work in the sound picture field » (the high noise level being of course produced by distance from the sound source) ; from Jones, « Condenser and Carbon Microphones —Their Construction and Use », JSMPE 16 (jan. 1931), 19-20.

/3. See, for example, Hope Bagenal and Alexander Wood, *Planning for Good Acoustics* , New-York, Dutton, 1931, and M.C. Batsel, « Recording Music for Motion Pictures », JSMPE 25, Aug. 1935, 103-108.

/4. Read and Welch, 1976, p. 237, see note 43.

of reflected sound with a term of quasi-religious praise : liveness/5. The picture palaces of the silent era took full advantage of this quality in their construction : large, built out of highly reflective materials, and covered with decorative moldings that fragmented and perpetuated reverberations, these ornate barns often built up a three —or four— second reverberation time. Now, by the time the directional microphone appeared, it was already clear that quality sound cinema required both studios and theaters whose reverberation time would be significantly reduced/6, for what suited music apparently was deemed inappropriate for the combination of music, effects, and dialogue characteristic of the Hollywood product. Not surprisingly, the clash between the long reverb times required for music and the much shorter reverberation needed for dialogue intelligibility is quickly resolved in favor or the latter. Here again, it seems quite evident that the development of directional miking is in reaction to the specificity of dialogue requirements, over against the familiar needs of music reproduction. In order to understand the development of microphone technology, therefore, we must first understand why it is that sound cinema as a representational mode seems to require its own special sound reproduction system. Four separable but closely related characteristics of sound cinema reveal the extent to which microphone development depends on representational needs specific to this particular mode of narrative representation.

1. *Invisible miking.* Durable systems are built on strong, unquestioned assumptions. During the twenties and thirties one assumption about sound cinema stands out : mentioned everywhere, but never so much as discussed, let alone questioned, this is the notion that « the receiving device must be sufficiently sensitive to permit its being successfully concealed at a reasonable distance from the speaker or source of music to be photographed », as Lee Deforest insists as early

---

/5. On the notion of « liveness », see especially the work of J. P. Maxfield, including « Some of the Latest Developments in Sound Recording and Reproduction », *Technical Bulletin* (of the Academy of Motion Picture Arts and Sciences), 1935 ; « Wide-Range Reproduction in Theaters », JSMPE 26, jan. 1936, 67-78 ; and « An Acoustic Constant of Enclosed Spaces Correlatable with their Apparent Liveness », JASA 19, jan. 1947, 71-79.

/6. Once again, this current begins with Maxfield, « Methods of High Quality Recording... » (1926 — see in Part. I. note 27, IRIS Vol. 3, n° 1) ; it is rapidly picked up by, among others, Vern O. Knudsen, « The Hearing of Speech in Auditoriums », JASA 1, oct. 1929, 56-82 ; and Edward W. Kellogg, « Some New Aspects of Reverberation », JSMPE 14, jan. 1930, 96-107.

as 1923/7. The assumption that all sound-collection devices must be hidden from the camera is not only an important aspect of sound cinema, it is —along with the complementary notion that all image-collection noises (camera sounds, arc lamps, the director's voice, etc.) must be hidden from the sound track— the very founding gesture of the talkies. Consider, for example, Bell's 1926 demonstration film in which Lab Director Edward B. Craft lectures on the very Vitaphone system which the film demonstrates/8. In good Bell public address style, Craft stands in front of the improved double-button carbon mike praised by Green and Maxfield in their 1923 article. In any normal sense of the expression, this is « sound cinema », but the truth is that we don't really use those words (or any of the many equivalent expressions) in a normal sense. By « sound cinema » we mean instead something rather technical, something first satisfied not even by *Don Juan*, with its synchronized effects and music, but by the synchronized dialogue and invisible miking of *The Jazz Singer*. When we speak of the « coming of sound » we imply the triumph of a particular representational system which is narrative and not discursive, dialogue-oriented and not musical, and which thus has no room for revelation of the means of its fabrication. Unless the mike is hidden (or its presence justified by some overriding narrative situation), then the film doesn't represent in the desired manner. Before the technology, then, before any particular narrative, music, or dialogue, there exists a felt need to produce a particular kind of representational form, not just a narrative but this kind of narrative, not just sound cinema but this kind of « sound cinema », not just reproduced sound but sound reproduced in this particular way (and thus with this particular succession of microphones).

Now, practically, to make the mike invisible means to move it farther fom the sound source. Large close-ups do not present an enormous problem since the camera's field of vision is small relative

---

/7. Other pioneers of sound cinema felt just like Deforest, « The Phonofilm », JSMPE no 16 (may 1923), 62. Earl Sponable, for example, states quite clearly that « In recording, it is necessary, of course, that the microphone be either placed outside of the Movietone », JSMPE 11 (sept. 1927), 462. What's more, the dissimulation of the microphone is closely related to the changeover from an acoustic sound-collecting system (where musicians huddled in contrived formations in order to play specially built instruments into a collecting horn) to an electro-mechanical system (where recording conditions replicated concert conditions almost exactly); in the former the musicians constantly foreground the collecting system, while in the latter they are (nearly) free to ignore it. On this point see Maxfield, « Electro-Mechanical Sound Recording », *Bell Laboratories Record* 1, jan. 1926, 198ff; and « Methods of High Quality Recording... », 1926 —see note 27, 498ff.

/8. « The Vitaphone Tells Tales of Itself », *Bell Laboratories Record* 3, dec. 1926, 126-28.

to the close pick-up required by early mikes (2-5 feet for the carbon mike), but with medium shots and *a fortiori* long shots the problem becomes quite severe. In order to remain invisible, the mike must be placed either far in front of or well above the speaker ; in either case the ratio of reflected to direct sound will be unacceptably high, thus reducing intelligibility, falsifying the appropriate acoustic perspective, and increasing markedly the amount of noise produced by the carbon mike. The first problem of cinema microphone technology is thus a problem of distance from the sound source, which is caused by a felt need for invisible miking, in turn produced by the assumption that a particular type of representation and no other must be achieved.

2. Intelligibility of the voice. Right from the beginning —and even in their so-called pure research— Bell engineers understood what was at stake. In a 1928 Bell research report published in 1929, Harvey Fletcher explains one of the fundamental principles on which Bell technological developments were already being based /9. When speech frequencies below 500 cycles are filtered out, 60 % of the energy is removed, but only 2 % of the articulation is lost. Conversely, filtering out tones above 1.500 cycles removes only 10 % of the energy, but 35 % of the articulation. In other words, Fletcher concludes, the fundamental tone and first few harmonics carry the energy, while the higher frequencies permit understanding. In a move which will often be repeated by later sound technicians, Fletcher then identifies « energy » with the qualities which make speech sound « natural », thus setting up a fundamental opposition in sound reproduction between the understandable and the natural (p. 281), or as Carl Dreher would have it two years later, « intelligibility of dialog » versus « naturalness, or acoustic fidelity to the original rendition » /10. Within a few years this opposition largely disappears from the literature as the criterion of intelligibility is accepted as primary in determining the procedures used for sound collection and reproduction /11.

Theoretically, we may note the extent to which this stress on

---

/9. Harvey Fletcher, *Speech and Hearing*, New York, Van Nostrand, 1929, pp. 281ff.

/10. Carl Dreher, « Recording, Re-recording, and Editing of Sound », JSMPE 16, june, 1931, 756.

/11. The strongest statement on this subject is made by John G. Frayne in 1929 : « To insure high intelligibility in a sound-stage pickup it is customary practice to place the microphone as close to the actor as possible, the distance usually being limited only by the camera angle of the scene. For medium and long shots the microphone must necessarily be moved farther away from the actor than for close-ups, since in

intelligibility suggests that the implied referent is by no means an original scene (whether staged or documentary), but a narrative which lies behind, which authorizes and engenders that scene, and of which that scene itself is a signifier. This is why most music, many effects and even some dialogue can be allowed to remain below the level of intelligibility, but all sounds which subtend the narrative must be fully comprehensible. For it is the *narrative* that is being represented and not this or that character, this or that scene, this or that conversation. Two people may use different words, but we commonly say that they have told the same story; it is in this sense that we may say that the narrative is primary. (Such a claim is of course anathema to anyone for whom the cinema is primarily an art form, for whom the « same » story told by two different people creates two different works of art. My whole point here, however, is that Hollywood hardly saw things that way, and that we can understand its development of new technology and new techniques only by recognizing, indeed starting from, that fact).

On the practical side, the need for intelligibility in the context of invisible miking immediately poses the problem of reducing the pick-up of reflected sound. Since the ratio of reflected to direct sound increases with distance from the sound source, and since too much reflected sound radically compromises the intelligibility of dialogue, the only solution is either to reduce the distance between sound source and microphone or to limit the angle of collection. Only after unsatisfactory experiments with mikes hidden in flower pots and other « prop pick-ups »/12, did studio technicians finally perfect a mobile microphone boom which effectively reduced the distance between sound source and microphone. Sound concentrators and directional mikes were successively introduced to reduce the angle of sound collection. Later developments of similar effect include the directional loud speaker to reduce theatre reverberation/13, and dialogue

---

close-ups the camera covers a wider angle. However, no attempt is ordinarily made in practice to try to obtain the same acoustic as visual perspective of the scene. Although the latter would be desirable it is not feasible in practice », *Elements of Sound Recording*, New York, John Wiley, 1949, pp. 52-53. See note 31 regarding further information on this topic.

/12. This term is the contribution of Carl Dreher, in « Stage Technique in the Talkies », 1929 — see note 29, p. 2.

/13. On directional loud speakers, see especially Olson and Massa, 1934 — see note 32 in Part. I., op. cit; Maxfield and C. Flannagan, « Wide-Range Reproduction in Theaters », JSMPE 26, jan. 1936, 67-78; and C. Flannagan, R. Wolf, and W.C. Jones, « Modern Theater Loud Speakers and Their Development », JSMPE 28, march 1937, 246-64. For more recent surveys, see the appropriate chapters by Clark and Frayne, see notes 4 and 5 in Part. I., op. cit.

equalization/14, as well as the general tendency toward smaller, less live theatres/15.

3. Narrative focus. By itself, the insistence on intelligibility already implies a principle of choice. Some things must be intelligible, others have somewhat less importance. This notion of relative importance has a long history, to my knowledge never written, and certainly too long even to be sketched out here. Skipping over millennia of painting, sculpture, and other representational arts, however, we may usefully dwell momentarily on the history of the notion of focus since the invention of photography. It is a curious fact that the first photograph of Nicéphore Niépce, the father of photography, lacks the sharp shadows and clear outlines that we depend on for identifying objects. Like all early photographers, Niépce used an emulsion so slow that only a very long exposure (multiple hours) could assure fixation of the image. During this time, the sun continued to move, thus lengthening shadows and changing the aspect of the roofs at which Niépce's primitive camera was aimed. Throughout the mid-nineteenth century, photography remained an art of the still life; as emulsions became increasingly sensitive, sharp focus became a required photographic characteristic, requiring careful attention from photographer and subject alike. Or rather, I should say, the desire for sharp focus transformed the photograph from a rectangular space where all points are potentially of equal importance (where all are subjects), into a hierarchized space, a space with an «object», a focus. No longer do we see a picture, i.e. a work of art in which each part relates to the rest in a special way, we see instead a picture *of* something else, a representation, a piece in a potential narration. No longer is the photographer an artist; he has become a focuser, the lowest form of narrator. By the end of the nineteenth century technicians and photographers alike are rushing toward faster lenses and emulsions in order to be able to stop motion more effectively —in order, as it were, to reduce to object status an ever wider variety of mobile subjects. So codified will the notion of photographic focus become that in recent years we have seen the rise of the automatic-

---

/14. On automatic dialogue equalization and volume control, see Dreher (1931 — see note 30 in Part. I., op. cit.); «Current Developments in Production Methods in Hollywood», JSMPE 24, jan. 1935, 3-11; and the articles mentioned in note 39. in Part. I., op. cit. D.P. Loye and K.F. Morgan, «Sound Picture Recording and Reproducing Characteristics», JSMPE 32, june 1939, 634ff, provides a useful summary.

/15. The ease with which intelligibility of dialogue can be assured in a small theater was recognized as early as 1929, as reported by Harold B. Franklin, head of Fox West Coast Theaters, in JSMPE 14, march 1930, 302-308.

focus camera, based on conventions (of distance, placing, and size of object) that have come from over a century of photographic narrativizing. With this type of camera you cannot take a picture, you can only take a picture *of* something else.

It was of course not always that way. The original *camera obscura* had infinite focus thanks to the pinhole size of its «lens». The history of photography, however, is marked by a constant battle between a desire for better focus and a need for more light, just as cinema sound grows out of the conflicting desires for intelligibility and naturalness. The concept of focus, or relative importance, is thus passed on, as it were, from photography to sound, where it constitutes just as much a choice as in the visual field. The point here is not that someone must do the choosing in a system with a focus, thus opening the system to ideological effects (that point already having been made in the earlier ideological reading; the point is much more simply to recognize the power of the very notion of focus. What is it that motivates Dreher to divide the roundhouse scene from *Danger Lights* into dialogue, which must be fully comprehensible, and ambient «noises», which are reduced by the use of a parabolic concentrator to «a realistic background»/16? Why concentrate all the attention on a few banal remarks by the protagonists when so much fascinating sound-producing activity is taking place around them? Quite simply, because as RKO's chief sound engineer Dreher has a devotion to narrative which requires the notion of focus. In fact, it is instructive to note the language used by Dreher in an earlier article. Twice during the 1929 «Stage Technique in the Talkies» he points out the need for microphone placement to serve «the business of the play», revealing what is really at stake in the notion of focus/17. While Olson's reference to a narrative «center of gravity» (quoted earlier) reveals the technicians's attempt to naturalize the notion of focus, Dreher's off-hand comments say a great deal about the economic importance of narrative within the Hollywood enterprise. A picture or a sound —as long as it has a focus— is by definition more attractive, more desirable than a picture or a sound, however lovely, without a focus. For what Hollywood sells is stories, not films. It thus requires a microphone with a focus just as fully as it needs cameras that focus. Deep-focus mikes, soft-focus mikes, sharp-focus mikes —that can come later. The first requirement is for sound to follow in the path

/16. Carl Dreher, «Microphone Concentrators in Picture Production», JSMPE 16, jan. 1931, 26.

/17. Dreher, «Stage Technique in the Talkies», 1929 — see note 29, Part. I., op. cit., pp. 2 and 16.

already sketched out by photography, reaffirming technologically the principle of focus already present in the chosen representational form.

4. For-me-ness. Microphones that focus, directional microphones, thus exist in large part to control reverberation, thereby guaranteeing the intelligibility of the spoken word, itself ensuring the comprehensibility and coherence of the represented narrative. Here too, however, the process of overdetermination is actively at work. The desired sound —indeed, to a great extent the achieved sound— sports a high degree of direct sound with a severely restricted amount of reflected sound (chiefly present to satisfy the reverb-heavy reality code of music reproduction and to support an intermittent reality effect produced by acoustic perspective/18. The resultant sound, while on the one hand constituting a particularly recognizable sequence of words, on the other hand represents an especially familiar *type* of sound, one that we hear daily, that we implicitly compare to other possible sounds, and that we interpret in the context of this implicit commutation. I'm outside working in the garden. I hear someone calling, but I can't make out the words. I don't even bother to turn around because I can tell from the muffled nature of the sound (i.e. the high ratio of reflected to direct sound) that the caller is in some other yard and is not facing in my direction. A few minutes later another call comes. I pick up my head to reduce the reflection of sound coming from my own body. No, still too much reverberation; the call is not for me. Suddenly, a third voice pierces the air. I still can't understand the words but the ratio of direct to reflected sound is quite high; this caller is clearly facing in my direction, calling me and no other. I immediately straighten up and respond : « Yes, what do you want ? » When the reverb disappears, I know that the message is for me. I know that I have switched from the overhearing mode to the discursive mode, the mode where people look at me, talk to me, care what I think, want to make sure that I hear. For the standard of intelligibility, of limited reverberation, is ultimately a standard of for-me-ness, of implicit discursivity.

The directional microphone is thus the pick-up system that imitates the directness of public address (overt discursivity) in a situation which is on the surface narrative (thus excluding the auditor/viewer) but which by sleight of hand affords the auditor the same closeness, the same privileges characteristic of direct address. By providing (selected) privileged access to the conversations of others, directional microphones open the way to a sort of « voyeurism of the ear » which

/18. See notes 31 in Part. I., op. cit., and 11.

we might well dub the « eavesdropping syndrome ». We hear as if we were being directly addressed, yet never are we looked at, never do we share the space of the speaking characters, never do we recognize in them denizens of our own three-dimensional world. Coupled with electrical amplification, the directional microphone provides the perfect sound pendant to the perspectival image which seems made for me. I see the protagonists even when they are « lost » in a crowd ; so do I hear them in similar situations. When two people are talking or reacting to each other they are nevertheless both turned slightly toward me; in a similar way the sound track always gives me the sound of people talking toward me, never shielding the sound with their body. In front of the actors there is always a space cleared away, a space which seems to invite me to come into the image and to exercise a privileged vision ; whether by automatic or manual means, the sound mix always clears away an auditory space for the dialogue by reducing effects and music at the appropriate times. Just as the image-oriented aspects of classical narrative are wont to do, the directional miking prevalent from the mid-thirties on does a marvelously effective job of convincing me, the auditor, that this audio-spectacle is made for me, and what's more, that no one knows that I am listening.

II

In the previous pages I have sketched out a reading of the development of microphone technology which might properly be termed « representational » for it attributes the impulse to develop new technology to the choice of a particular representational mode. Now, as might readily be objected, the choice of a particular representational mode does not occur in a vacuum, and thus can hardly be given as a first cause. This choice is subject to pressures of various sorts, including the realist, ideological, and economic impulses implied by the various readings tentatively offered above. Indeed, it is this very multivalent nature of the representational mode that I wish to bring into play here in order to locate the important coordinating function fulfilled by the representational mode.

Let us return briefly to Carl Dreher's offhand characterization of a film's narrative focus as « the business of the play »/19. On the one hand, as we have seen, the notion of « business » corresponds to a

/19. Dreher, « Stage Technique in the Talkies », 1929 — see note 29 in Part. I., op. cit., pp. 2 and 16.

«realist» we mean this particular kind of representation. It is thus incomplete to suggest that microphone technology develops to increase realism, for this formulation leaves out an essential intermediate step. In order to increase a particular type of realism, a particular representational mode evolved ; it is in the service of this mode that microphone technology develops.

Now, realism is of course no first cause. Its very illusionism reveals the extent to which it serves ideological concerns. But how can we get at those concerns? Once again, we must pass through the basic representational mode, for it is this mode that provides the signifying system which permits ideology itself to signify. We have seen how important the manipulation of focus is to an ideological reading; by simultaneously choosing a focus and disguising the fact that a focus has been chosen, this particular brand of narrative text plies its wares. Yet the very notion of focus is a concomitant of this particular narrative mode. In other words, whether we are dealing with the realist, ideological, or economic reading of the history of microphone technology, we are necessarily dealing *first* with the place of miking within a particular representational system. If it often appears that the quarrels between proponents of opposing approaches reveal no more than acceptable alternate views of the same phenomenon, separable views which are nevertheless not entirely separate because they together constitute complementary pieces of a larger puzzle, it is largely because proponents of all three approaches have implicitly taken for granted the same limits, those of a particular representational mode. In doing so, they have necessarily agreed to share certain basic assumptions with other approaches which they deem antithetical to their own.

The first wave of technological criticism concentrated on the discourse of technicians and inventors themselves; this approach led to straight-forward technicist conclusions like that of the «engineering» reading presented here. Subsequent studies of technology often had the merit of locating another discourse contained by the technical; «realist», «ideological», or «economic», the individual readings generated by this approach were almost always seen as contradictory and mutually exclusive. While illustrating and recognizing the insights of the «engineering» and «revisionist» approaches, in this article I have been reaching for a third solution, one which attributes the generation of multiple (complementary) explanations out of a single set of phenomena to the multifarious (complementary) aspect of a particular representational mode. By paying closer attention to the relationship between technology and the representational purposes which it serves, perhaps we can open a path to a new and more versatile technological criticism.

L'histoire de la prise de son au cinéma -et tout particulièrement des diverses sortes de microphones introduits pendant les années trente- est racontée ici selon plusieurs hypothèses différentes. La première partie du texte montre comment la parole est d'abord aux techniciens, selon lesquels les micros directionnels auraient été introduits uniquement pour assurer une meilleure fidélité de la reproduction sonore. Trois autres points de vue possibles sont ensuite considérés : réaliste, idéologique, et économique. Cette deuxième partie tente une nouvelle approche de ces données, qu'on pourrait qualifier de « représentationnelle » de quatre règles tacites de la représentation sonore Hollywoodienne : invisibilité du micro, intelligibilité de la voix, focalisation narrative, et discursivité.*

* IRIS, Vol. 3, n° 1.