

Michel Chion

L'AUDIOVISIONE SUONO E IMMAGINE NEL CINEMA

Il contratto audiovisivo

1. Proiezioni del suono sull'immagine

Il cinema, arte dell'immagine, un'illusione, che si trova nel cuore della più importante tra le relazioni suono-immagine: quella del valore aggiunto.

Con l'espressione valore aggiunto designiamo il valore espressivo e informativo di cui un suono arricchisce un'immagine data, sino a far credere, nell'impressione immediata che se ne ha, che quell'informazione o quell'espressione derivino "naturalmente" da ciò che si vede, e siano già contenute nella semplice immagine.

Il fenomeno del valore aggiunto funziona soprattutto nel quadro del sincronismo suono/immagine, per il principio della sincresi, che permette di istituire una relazione immediata e necessaria tra qualcosa che si vede e qualcosa che si sente.

Il valore aggiunto è quello del testo all'immagine. Teorizzare che il suono nel cinema sia voco centrista, significa ricordare, che esso favorisce la voce. Si tratta della voce come supporto dell'espressione verbale. Il valore aggiunto del testo sull'immagine va ben al di là di un'opinione applicata su una visione e influisce sulla strutturazione stessa dell'immagine, inquadrandola rigorosamente.

Ci sono due modi per la musica di creare nel cinema una specifica emozione. Nel primo la musica esprima direttamente la propria partecipazione all'emozione della scena – musica empatica. Nell'altro, essa mostra al contrario una chiara indifferenza alla situazione – musica anempatica (musiche per pianola meccanica, celesta, carillon e orchestra da ballo).

L'effetto anempatico riguarda nella maggior parte dei casi la musica ma può anche essere utilizzato con i rumori. Per esempio in una scena violenta o dopo la morte di un personaggio, un evento qualunque continua a svolgersi come se niente fosse (*Psycho*, 1960).

Le percezioni sonore e visive sono molto diverse. Il suono presuppone, diversamente dal visivo, il movimento come prima condizione. Il suono implica necessariamente, per propria natura, uno spostamento, un'azione. L'orecchio analizza, lavora e sintetizza più in fretta dell'occhio. In un primo contatto con un messaggio audiovisivo, l'occhio è più abile spazialmente, l'orecchio temporalmente.

Uno dei più importanti aspetti del valore aggiunto è la percezione del tempo dell'immagine. Tale effetto di temporalizzazione ha tre aspetti:

1. animazione temporale dell'immagine, 2. linearizzazione, 3. vettorializzazione, o drammatizzazione dei piani.

La temporalizzazione dipende più dalla regolarità o dall'irregolarità del flusso sonoro che dal tempo nel senso musicale della parola.

Al suono sincrono dobbiamo attribuire l'aver fatto del cinema un'arte del tempo. Il tempo del film è divenuto no più un valore elastico, più o meno modificabile a seconda del ritmo di proiezione, ma un valore assoluto.

L'animazione temporale dell'immagine da parte del suono non è un fenomeno puramente fisico e meccanico: in essa svolgono il proprio ruolo anche codici cinematografici e culturali.

Il valore aggiunto è reciproco: se il suono rende l'immagine differente da come sarebbe senza di esso, dal canto suo l'immagine fa sentire il suono diverso da come risulterebbe se risuonasse nel buio. Il caso dei suoni orribili, sui quali l'immagine proietta per suggestione un senso che essi non comportano affatto di per se stessi, è l'esempio di tale reciprocità. In *Kiss me Deadly* di Robert Aldrich, quando la fuggitiva che Ralph Meeker ha preso a bordo in autostop viene catturata dai suoi inseguitori e sottoposta a tortura: di questa tortura, lo spettatore non vede altro che due gambe, e nel frattempo sente le grida. Questo è un tipico effetto sonoro.

2. I tre ascolti

Ci sono tre differenti disposizioni di ascolto: 1) l'ascolto causale, 2) l'ascolto semantico, 3) l'ascolto ridotto.

L'ascolto più diffuso è l'ascolto causale, che consiste nel servirsi del suono per informarsi, quanto più possibile su la sua causa. L'ascolto causale può avvenire a livelli diversi. Per esempio riconosciamo la causa precisa e individuale: la voce di una determinata persona, il suono di un oggetto unico. Possiamo anche, senza identificare la sorgente nel senso della natura dell'oggetto causale, seguire con precisione la storia causale del suono stesso.

Nel cinema l'ascolto causale è costantemente manipolato da ogni parte del contratto audiovisivo, tramite l'uso della sincresi.

L'ascolto semantico fa riferimento a un codice o a un linguaggio per interpretare un messaggio: il linguaggio parlato, ma anche i codici come il Morse. L'ascolto causale e semantico possono esercitarsi parallelamente e indipendentemente in una stessa catena sonora.

L'ascolto ridotto si rivolge alle forme proprie del suono, indipendentemente dalla sua causa e dal suo senso, e considera il suono – verbale, strumentale, aneddotico o qualunque – come oggetto di osservazione.

Tutti praticchiamo questo tipo di ascolto, quando individuiamo l'altezza di una nota o gli intervalli tra due suoni, praticchiamo l'ascolto ridotto senza saperlo, perché l'altezza è un carattere proprio del suono, indipendente della sua causa o del suo senso.

La questione dell'ascolto è inseparabile da quella dell'udire, come quella dello sguardo è legata al vedere.

3. Linee e punti

Il contrappunto audiovisivo s'incontra quotidianamente in televisione, per esempio durante l'emissione di certi eventi sportivi, in cui l'immagine segue la propria strada e il commento un'altra. Per un'immagine data, vi sono infatti centinaia di sonorizzazioni possibili, alcune producono il codice convenzionale, mentre altre ne fanno scivolare la percezione su un altro piano.

Per l'immagine è la fabbricazione tramite montaggio che ha creato l'unità specifica del cinema: il piano, una pratica per individuare una serie di punti di riferimento nel film. Il montaggio del suono nel cinema non ha creato un'unità specifica, le giunte sonore non sono evidenti all'ascolto.

Con il suono sono possibili due tipi di montaggio: udibile e non udibile. Jean-Luc Godard tagliava i suoni come le immagini, evidenziando le discontinuità e salti, e limitando al massimo il montaggio invisibile.

Quando parliamo di piano nel cinema, facciamo un nodo tra lo spazio del piano e la sua durata, tra la sua superficie spaziale e la sua dimensione temporale. Nei segmenti sonori invece, la dimensione temporale sembra predominare largamente, e la dimensione spaziale essere del tutto inesistente.

La colonna audio di un film non costituisce per ascolto un flusso senza tagli, perché distinguiamo comunque delle unità: frasi, rumori, temi musicali sono dello stesso tipo che nell'esperienza corrente, e vengono individuate in funzione di criteri specifici dei differenti tipi di suoni.

Il flusso del suono di un film si caratterizza per l'aspetto più o meno legato e concatenato dei differenti elementi sonori, o al contrario più o meno interrotto da bruschi salti. *Logica interna* del concatenamento audiovisivo è un modo di concatenamento delle immagini e dei suoni concepito per rispondere a un processo flessibile. *Logica esterna* risente degli effetti di discontinuità.

La funzione più diffusa del suono nel cinema consiste nell'unificare il flusso delle immagini:

- a livello temporale
- a livello spaziale
- tramite la presenza di una musica orchestrale che, sottraendosi alla nozione di tempo e spazio reali, inserisce le immagini in un stesso flusso.

La funzione di punteggiatura sta da molto tempo alla base della messa in scena teatrale. Nel cinema muto la punteggiatura era molteplice: gestuale, visiva e ritmica. Il suono sincrono ha

introdotto nel cinema non il principio della punteggiatura in se stessa, ma un mezzo più discreto e surrettizio di introdurla nelle scene senza demandarla alla recitazione degli attori o al découpage.

La partitura scritta da Steiner per il film di John Ford *Il traditore* si riferisce al principio che, in seguito, dominerà il novanta per cento della musica per il grande schermo: quello del *leitmotiv*, in virtù del quale ogni personaggio chiave od ogni idea-forza del racconto sono dotati di un tema che li caratterizza. Se è stata un'opera di Debussy a ispirare Steiner.

Gli elementi di sfondo sonoro contribuiscono a creare lo spazio di un film (l'abbaiare di un cane in distanza, lo squillo del telefono, la sirena dell'auto della polizia). Al di fuori del suo ruolo narrativo, l'elemento di sfondo sonoro può anche avere grazie al montaggio, funzione di punteggiatura.

Un movimento di macchina, un ritmo sonoro, l'evoluzione di uno degli attori mettono in moto nello spettatore un movimento di anticipazione. In *Figli di un Dio Minore* William Hurt lascia la sala di ballo, per poi girarsi e vedere Marlee Matlin che lo raggiunge, il suono della musica comincia a decrescere, lo spettatore anticipa il riunirsi dei due personaggi, sull'unione dei due innamorati c'è il silenzio.

È celebre aforismo di Bresson secondo cui il cinema sonoro ha introdotto il silenzio. L'impressione di silenzio in una scena non è il semplice effetto di un'assenza di suono, essa si produce soltanto quando viene raggiunta attraverso tutto un contesto e una preparazione. La quale consiste, nel caso più semplice, nel farla precedere da una sequenza rumorosa. Il silenzio non è mai un vuoto neutro: è il negativo di un suono. Gli altri rumori utilizzati nel cinema come sinonimo del silenzio sono: i versi di animali in lontananza, le pendole in appartamenti attigui, etc.

Un punto di sincronizzazione è in una catena audiovisiva, un momento saliente di incontro sincrono tra un momento sonoro e un momento visivo: un punto in cui l'effetto sincresi risulta più accentuato.

La sincresi (le parole sincronismo e sintesi) è la saldatura inevitabile e spontanea che si produce tra un fenomeno sonoro e un fenomeno visivo puntuale quando questi accadono contemporaneamente, e ciò indipendentemente da ogni logica razionale. Ci sono diversi passi di sincronismo che determinano un certo stile di film, in particolare a sincronizzazione labiale. I sincronismi stretti assoggettano i suoni ai movimenti della labbra, i sincronismi più larghi prendono in considerazione la totalità del corpo parlante.

4. La scena audiovisiva

L'immagine nel cinema non è il contenuto, ma il contenente, è il quadro. Una caratteristica propria del cinema è 'assenza di un luogo di immagini (in opposizioni alle videoinstallazioni, multivisioni). Al contrario, per il suono non esiste né un quadro né contenente preesistente.

Acusmatico significa "che si sente senza vedere la causa originaria del suono". La radio, il disco o il telefono sono media acusmatici. In un film un suono può compiere due tipi di percorso:

- esso viene prima visualizzato, poi acusmatizzato;
- esso comincia acusmatico, per essere visualizzato in seguito

Un suono o una voce mantenuti acusmatici creano in effetti mistero sull'aspetto della loro sorgente e sulla natura stessa. Il suono fuori campo nel cinema è il suono *acusmatico* rispetto a ciò che viene mostrato nel piano. Il suono *in* quello la cui sorgente appare nell'immagine. Il suono *off* quello la cui supposta sorgente è non soltanto assente dall'immagine, ma anche non diegetica.

Il suono d'ambiente avvolge una scena e abita il suo spazio senza sollevare la questione della localizzazione della sua sorgente (gli uccelli che cantano). *Il suono intero* corrisponde all'interno tanto fisico che mentale di un personaggio (respirazioni, battiti cardiaci). *Suoni in onda* (on the air) sono i suoni presenti in una scena, ma trasmessi elettricamente per radio, telefono, etc. Il suono della musica *on the air* (trasmessa o registrata) attraverserà più o meno le zone in/off/ fuori campo e si posizionerà per lo spettatore, come musica da schermo.

Suono e sorgente del suono dal punto di vista spaziale sono due fenomeni distinti. La musica che accompagna l'immagine da una posizione off, al di fuori del tempo e del luogo dell'azione è *la musica da buca*. *La musica da schermo* è quella che proviene da una sorgente situata direttamente nel luogo e nel tempo dell'azione, anche se tale sorgente è una radio.

Musica da buca/non-diegetica/commentativa/soggettiva.

Musica da schermo/diegetica/attuale/oggettiva.

L'effetto quinta si produce quando un suono legato a una causa suscettibile di entrare in campo, o che ne è appena uscita, si attarda su uno degli altoparlanti laterali posizionati all'esterno dello schermo: i passi di un personaggio che si avvicina o si allontana, il motore di un auto. Questi effetti del *fuori campo assoluto* sono possibili grazie al multipista.

Estensione dell'ambiente sonoro è lo spazio concreto più o meno aperto che i suoni evocano e fanno sentire intorno al campo. Il suono Dolby Stereo moltiplicando le possibilità di sovrapporre suoni in spazi, incoraggia il gioco dell'estensione.

La nozione di punto di ascolto è molto ambigua. Il concetto è stato modellato su quello di punta di vista. Punto di vista significa due cose:

- da dove vedo io spettatore,
- quale personaggio, nell'azione, si ritiene veda ciò che io vedo; è l'accezione soggettiva.

Comparativamente la nozione di punto di ascolto:

- un senso spaziale: da dove sento,
- un senso soggettivo: quale personaggio, a un dato momento dell'azione, si ritiene che senta ciò che sento io.

In certi casi è possibile attribuire all'ascolto una direzione. Per esempio c'è una differenza udibile tra la voce di faccia e la voce di schiena.

5. Il reale e la resa

La definizione di una registrazione sonora è sul piano tecnico, la sua finezza e la sua precisione nella resa dei dettagli, la funzione della sua banda di frequenza e la ricchezza di dinamica. Nel cinema, la definizione è per il suono una posta multipla e un importante mezzo espressivo.

Gli anni tra i '20 e i '40 conobbero la moda di una nozione oggi dimenticata, quella di fonogenia. La capacità che avrebbe permesso a certe voci di passare meglio nella registrazione. Questa nozione si riferiva alle condizioni tecniche dell'epoca. Il criterio era stato modellato su quello della fotogenia, molto in voga all'epoca delle star.

Il suono diretto nel cinema viene considerato non soltanto come l'unica soluzione ammissibile, ma anche come quella che semplifica tutto e dispense dal fare delle scelte. Nei film di Eric Rohmer nulla giunge a disturbare la concentrazione sui personaggi e sul loro testo. Per ottenere in *La Moglia dell'aviatore*, l'assenza di ambiente sonoro senza che nulla intervenga, il silenzio dello spazio che sta intorno ai personaggi è stato necessario a Rohmer scegliere delle ore precise per le riprese. La nozione del suono diretto, rispetto all'ambiente, non implica meno ricostruzione, foss'anche per mera sottrazione, di quella del suono postsincronizzato.

La questione del verosimile del reale sonoro è molto ambigua. Un'arte di registrazione come il cinema ha creato codici specifici di verità, legati alla sua natura tecnica. Per esempio, tra due reportage su una guerra, quello in cui l'immagine è mossa con difetti e sfocature, sembrerà più vero di quello con l'immagine impeccabile. Anche per il suono, l'impressione di realismo è spesso legato a una sensazione di disagio.

Nella questione della funzione realistica e narrativa dei suoni diegetici (voce, musica, rumori), la nozione di *resa* si contrappone a quella di riproduzione. Il suono viene riconosciuto come vero non se esso riproduce il suono prodotto nella realtà dallo stesso tipo di situazione, ma se rende le sensazioni associate a quella causa.

Una voce, dei rumori o una musica contengono sempre un certo tasso di indizi sonori materializzanti, da zero a un'infinità. In un suono gli indizi materializzanti sono quelli che ci rimandano all'intuizione della materialità della sorgente e al concreto processo dell'emissione

del suono. All'inizio di *Mio Zio*, nella scena di risveglio degli Arpel, i passi del ragazzino fanno sul cemento del giardino uno scricchiolio simpatico e concreto, mentre quelli di padre, un omone impacciato e misero.

6. L'audiovisione vuota

L'acusmetro è il personaggio acusmatico la cui posizione rispetto allo schermo si colloca in un'ambiguità, lo si potrebbe definire come né dentro né fuori. Possono essere descritti come "acusmetri" molti dei personaggi misteriosi o nascosti dietro la tenda (*Il Dottore Mabuse*). È proprio dell'acusmetro il poter essere privato all'istante di quei poteri misteriosi quando viene *deacusmatizzato*, cioè quando il volto da cui proviene la voce viene rivelato, e la voce si trova in tal modo, per sincronismo, attribuita a un corpo in cui essa è confinata e come rinchiusa.

La *deacusmatizzazione* è un dispositivo di svelamento drammatico, si trova nei film polizieschi o di mistero, quando finalmente il big boss di cui finora abbiamo sentito soltanto la voce, getta la maschera.

Si dà *sospensione* quando un suono naturalmente implicato dalla situazione, e in generale sentito in precedenza, viene a essere soppresso. Kurosawa ama le sospensioni, e le ha utilizzate spesso in *Ran*, per esempio nella scena dell'Inferno, un combattimento terrificante e muto accompagnato soltanto dalla musica.

Il ritmo è un elemento del vocabolario cinematografico che non è né sonoro né visivo. Dopo essere entrato nell'orecchio o nell'occhio, il fenomeno colpisce in noi una qualche area cerebrale connessa con le aree motorie, ed è soltanto a questo livello viene decodificato ritmicamente. La tesi sulla questione della percezione è d'altronde che ciò non vale soltanto per i ritmi, ma per altri percezioni di scrittura, di materia etc.

Al di là dei suoni e delle immagini

7. Degno del nome di sonoro

Alle origini del cinema sonoro la banda passante era ancora limitata. Il che imponeva, da un lato, di non mescolare troppi suoni per mantenerli udibili e, d'altra parte, di sceglierne uno principale. Più la banda passante del cinema si è allargata si perfezionavano nuove possibilità tecniche di miscelazioni dei suoni.

Il Dolby Stereo cambia sempre l'equilibrio dei suoni, in particolare introducendo un "surplus" percepibile nella riproduzione dei rumori.

Supercampo è il campo individuato, nel cinema multipista, dai suoni d'ambiente naturali, di rumori urbani, di musica etc., che circondano lo spazio visivo e possono provenire da altoparlanti situati al di fuori degli stretti limiti dello schermo.

Il multipista è stato definitivamente imposto da film musicali (*Woodstock* di Michael Wadleigh, *Tommy* di Ken Russel).

Con la nuova posizione occupata dai rumori, la parola non è più centrale nel film, e tende a essere reinserita entro un continuum sensoriale globale che la ingloba, e che occupa tanto lo spazio sonoro che lo spazio visivo.

8. Televisione, clip, video

Dire della televisione che essa è una radio illustrata, significa ricordare che in essa il suono, principalmente il suono della parola, è sempre primario, che esso non è mai fuori campo e che è sempre lì al suo posto, senza avere bisogno di immagine per localizzarsi.

Di tutti gli sport mostrati, il tennis è per eccellenza lo sport acustico. Nella trasmissione di una partita lo spazio acustico è naturalmente separato dallo spazio visivo.

Il video è sospeso tra un estremo e l'altro, tra un'immagine del tutto fissata in precedenza e un'altra che risulta da eventi in tempo reale; e tra l'unicità e la molteplicità degli schermi.

Una delle grandi differenze tra video e cinema, almeno nella sua forma sonora, i cambiamenti di velocità e fermo-immagine richiedono lavorazioni lunghe e costose, per ottenere ciò che in una regia video si ottiene immediatamente.

Lo sfarfallamento visivo che ritroviamo nei clip e nei video, interessa la rapidità dell'uditivo e del testo. L'immagine perde la propria qualità di superficie relativamente stabile, e sono i suoi cambiamenti di andatura o di aspetto a diventare significanti.

Se i clip funzionano, è sicuramente perché c'è un rapporto elementare tra colonna audio e colonna immagine, e le due non sono del tutto indipendenti. Tale rapporto si limita alla presenza puntuale di punti di sincronizzazione, nei quali l'immagine mima la produzione del suono. Per il resto del tempo, immagine e suono giocano l'una indipendentemente dall'altro.

9. Verso un audio-logo-visivo

Prima del sonoro, il cinema muto non era privo di linguaggio. Il linguaggio era presente a due livelli: esplicitamente nel testo dei cartelli, implicitamente con la maniera in cui le immagini erano organizzate.

Distinguiamo tre modalità di presenza della parola nel cinema: parola-teatro, parola-testo, parola-emanazione.

Nella parola-teatro, il caso più corrente, il dialogo sentito ha una funzione drammatica, psicologica, informativa e affettiva. Essa condiziona tutta la messa in scena del film.

La parola-testo è quella della voce-off e dei commenti, essa agisce sul corso delle immagini (voce di un narratore).

La parola-emanazione è quando la parola non viene necessariamente sentita e compresa per intero e soprattutto quando essa non è legata all'azione (il dialogo dei personaggi è intelligibile). La parola diventa una specie di emanazione dei personaggi, un loro aspetto.

10. Introduzione a un'analisi audiovisiva

L'analisi audiovisiva mira a cogliere la logica di un film o di una sequenza nella sua utilizzazione del suono combinata con quella dell'immagine. Per analizzare la struttura suono/immagine si può fare ricorso a un *metodo delle mascherature*. Esso consiste nel visionare una sequenza guardandola ora con suono e immagine, ora mascherando l'immagine, ora tagliando il suono.

Verrà caratterizzato l'andamento generale del suono, e in particolare la sua consistenza. La *consistenza* è funzione:

- di un equilibrio generale dei livelli
- della presenza più o meno consistente di un *riverbero*, che può smorzare i contorni sonori e creare una sorta di sostanza molle e unificante, in grado di legare i suoni gli uni agli altri
- dei fenomeni di mascheratura, legati alla coesistenza di diversi suoni nei medesimi registri di frequenza.

È interessante comparare il suono e l'immagine su una stessa questione di rappresentazione. Sul piano della velocità: il suono e l'immagine possono avere velocità diverse, e questa differenza può creare una sottile complementarietà di ritmo. Lo stesso vale per la questione della materia e della definizione.