

Jullier, Laurent. *Les Sons au cinéma et à la télévision. Précis d'analyse de la bande-son*. Coll. « Cinéma et audiovisuel ». Paris : Armand Collin, 1995.

L'analyse de la bande sonore dans un cadre cognitiviste implique quatre présupposés :

1) « [L]a perception n'est pas un simple enregistrement, mais plutôt une *réorganisation* du monde qui nous entoure, à laquelle il est impossible de se soustraire » (16). Les données sensorielles sont donc soumises à des « significations imposées » qui ont pour but « de nous rendre le monde intelligible » (16). Fidèle à ces racines kantienne, le cognitiviste postule que « le monde tel qu'il est » (le noumène) est inaccessible. Notre appréhension du monde est plutôt le résultat d'une réorganisation des perceptions selon des « modèles cognitifs pré-existants » : « [n]ous percevons moins « ce qui est » que « ce qui *doit* être » en fonction de ce qu'il nous est déjà donné de savoir sur le monde » (16). Par conséquent, l'analyste doit décrire les catégories qui commandent notre perception et comprendre les processus d'organisation des phénomènes menant à notre représentation du monde.

2) Le son filmique est un objet de connaissance parce qu'il nous informe sur les caractéristiques d'un monde fictionnel. « [Les] produits audio-visuels assignent aux sons la tâche de transporter des informations » (9). Par conséquent, l'analyste doit centrer ses recherches sur « ce pouvoir qu'ont les sons d'*apprendre* quelque chose à leurs auditeurs. » (9).

3) Cet apprentissage est conditionné par les films et par les institutions. Ceux-ci forgent des « consignes de lecture » (10) qui permettent au spectateur d'organiser l'information véhiculée par les sons. La perception et la compréhension des sons dépendent de ces consignes. C'est à l'aide de ces règles que le spectateur se forgera des buts et des attentes. Par conséquent, l'analyste doit « *expliquer* rationnellement ce que le spectateur fait des données sonores d'une œuvre audio-visuelle, et [se] *donner des outils* qui [lui] permettront [...] de reproduire cette explication en l'adaptant à la forme et au contenu d'une œuvre donnée » (9). Le cognitiviste travaille donc à l'élaboration d'un schéma unique d'opérations mentales qui, s'il est assez sophistiqué et

complexe, pourra expliquer les conditions de toutes expériences cognitives possibles (tous les contextes, tous les contenus, toutes les formes)¹.

4) L'étude des processus cognitifs nécessite la représentation d'un spectateur modèle : a) cet individu est le « spectateur coopérant », c'est-à-dire « celui qui s'efforce de *comprendre* ce que l'œuvre audio-visuelle « dit » » (10)² ; b) ce spectateur écoute du cinéma narratif, parce que sa quête de sens implique un questionnement : « quelle information utile [...] du point de vue du récit [...] est [véhiculé] par ce son-là ? » (11) ; c) ce « spectateur ne dispose que d'une seule et unique façon de percevoir le monde », « la même pour le film et la vie quotidienne », « celle que cent cinquante millions d'années d'évolution des mammifères lui ont fournies. » (16 et 57). L'analyste élabore ses outils selon ce triple postulat : a) une théorie cognitive de l'audiovisuel doit expliquer la compréhension d'un spectateur coopérant ayant des buts et des attentes ; b) une théorie cognitive de l'audiovisuel a comme *corpus* les œuvres qui agencent des sons et des images afin d'élaborer un récit ; c) une théorie cognitive de l'audiovisuel se base sur les études de la perception quotidienne parce que les mêmes processus se retrouvent dans la perception esthétique. En somme, une théorie cognitive de l'audiovisuel doit, pour être scientifiquement valable, s'ériger à l'intérieur d'une théorie générale de la connaissance qui considère : que la valeur cognitive de l'art repose dans sa capacité à raconter des histoires, que la perception repose sur des schémas mentaux invariants et que l'organisation des informations est dictée par des consignes de lecture extérieures aux sujets.

C'est à partir de ces présupposés que Jullier construit son « précis d'analyse de la bande-son ». Le son a une valeur informative ; il « porte en lui

¹ Jullier considère que cette approche est avantageuse parce que sa méthode d'analyse cognitive peut « s'appliquer indifféremment à tous les types de sons ». De plus, « la liberté dans le choix des sons se retrouve dans celui des œuvres » (11). Or, cette démarche tend à surévaluer la fonction et la valeur informative des sons alors qu'elle n'est qu'un élément parmi d'autres dans les films de plusieurs cinéastes (Bresson, Godard, Lynch, Denis, Grandrieux, Van Sant). Jullier admet ce désavantage (11) mais tient tout de même à conserver cette approche scientifique dans la démarche, universelle dans l'application, versatile dans le choix du corpus et égalitaire dans l'évaluation esthétique (il y a du bon dans tout).

² « Parmi tous les buts que le spectateur peut se fixer lorsqu'il décide de s'asseoir pour regarder un produit – se distraire, s'instruire... - on a choisi ici d'étudier celui de la *compréhension* des données de l'écran » (95).

des renseignements de deux natures, les uns concernant la *source* dont il émane et les autres concernant l'*espace* qui sépare et entoure les deux points en question » (14). Pour le cognitiviste, l'écoute d'un son à des fins informatives se résume à l'*identification* d'une source (la question « quoi? »), à la *localisation* de la source dans l'espace représenté (la question « où? ») et à l'*intégration* de l'amalgame son/source dans la structure du récit (la question « pourquoi? »).

L'identification des sources : écoute causale, révolution électrique, continuum

Comme l'être humain n'a qu'une seule façon de percevoir les sons, une étude de la perception esthétique doit tout d'abord décrire les processus de la perception quotidienne. Dans notre vie de tous les jours, l'identification d'un son suppose une « [h]abitude profondément inscrite en nous » que l'on nomme l'*écoute causale*. Celle-ci « consiste à deviner quelle source est la cause du son entendu » (15). Cette propension à connecter un son à une image visuelle est une question de survie pour la plupart des animaux : le prédateur doit détecter la présence d'une proie s'il ne veut pas mourir de faim et la proie doit sentir la présence du prédateur s'il veut éviter la mort (19). Depuis des millénaires, ce type d'écoute domine notre activité auditive. Selon la loi de la causalité, un bruit est considéré comme « l'indice d'une nouvelle situation » (19), le présage d'une modification de notre environnement immédiat. L'écoute causale démontre donc notre capacité à distinguer et grouper les sons : nous ne restons pas « passivement à ouïr globalement, en s'y laissant baigner, le bruissement du monde » (19). Nous distinguons les sons et nous identifions leurs sources d'émission.

Le cinéma narratif fonde une grande partie de son pouvoir illusionniste — de sa capacité à construire une représentation cohérente du monde — sur son aptitude à faire croire que le lien entre le son et l'image est, à l'instar de la vie quotidienne, un lien causal. Il convoque alors un type d'écoute maîtrisé par le spectateur coopérant, et qui permet l'ancrage d'un son à sa (supposée) source visuelle. Comme « le système auditif semble prioritairement fait pour répondre aux demandes de l'écoute causale » (19), il est évident que le cinéma narratif, qui tente de représenter un monde de manière efficace, privilégie ce type d'écoute au détriment de l'écoute réduite qui nécessite une déconstruction de nos habitudes perceptives. La compréhension d'un récit audiovisuel dépend de

ces « compétences premières », de ce type d'écoute partagée et pratiquée quotidiennement par tous les humains (30)³.

Malgré tout, l'écoute d'une œuvre audiovisuelle reste différente de l'écoute quotidienne des sons parce que les haut-parleurs procurent une écoute indirecte des sons. Depuis la *révolution électrique*, les sons peuvent se séparer de leur contexte d'émission (32). Alors que l'écoute directe met en jeu une « chaîne de communication » unique (de la *source* vers l'*auditeur*, en passant par l'*espace* de propagation des ondes), l'écoute indirecte crée plutôt trois chaînes de communication distinctes :

- la *fixation* représente le mouvement de la source naturelle vers le microphone ;
- les *conversions* représentent les multiples manipulations des sons fixés sur un support ;
- la *diffusion* représente le mouvement du haut-parleur vers l'oreille de l'auditeur.

³ « La plupart des œuvres audio-visuelles [...] sont d'abord comprises parce qu'elles font appel à des mécanismes mentaux ancrés [dans le spectateur] : reconnaissance de formes, dénominations, comparaisons avec des images mentales stockées... » (29). Afin de décrire ce processus d'identification, le cognicien de l'audiovisuel utilise un modèle général composé de trois moments :

1-Le corps doit premièrement convertir le stimulus en sensation. À ce niveau neuro-sensoriel, les « modules agissent comme des automates et sont encapsulés, c'est-à-dire indépendant des connaissances du sujet » (17). Les modules transforment la modulation de l'air en image sonore. Cette dernière est une « représentation psychologiques d'entités sonores présentant un comportement acoustique cohérent » (17).

2-Le cerveau traite ensuite les sensations. À ce niveau perceptif, les modules effectuent des opérations de groupement et de stockage immédiat. Ces processus perceptifs sont « automatiques ; ils correspondent à une perception gouvernée par les *données de l'écran* (data-driven) » (processus ascendants (*bottom-up*)) (17).

3-Le cerveau met en relation les sensations traitées. À ce niveau cognitif, les modules « traitent les *données de l'écran* comme des échantillons inductifs destinés à être projetés et testés dans le cadre des attentes et des buts du sujet » (processus descendants (*bottom-up*))(17).

L'identification en trois temps aboutit à des représentations conscientes des sources « sur lesquelles pourront s'appuyer des raisonnements, des prises de décisions et des projets d'action » (18).

Cette fragmentation a modifié la nature informative du son. L'auditeur ne peut plus seulement considérer un bruit comme l'indice de la présence d'une source active, il doit maintenant se demander s'il provient d'une source directe ou indirecte. Ainsi, « les mécanismes de l'écoute causale en situation acousmatique [sont] remis en question » (34). Le spectateur ne perd pas son habileté chèrement acquise au cours de milliers d'années d'évolution, mais il développe une « écoute causale de second niveau » (31). Ce niveau d'écoute se forgera au fil des deux époques de la révolution électrique. Chacune de ces époques est marquée « par l'avènement de deux technologies qui ont profondément modifié les habitudes d'écoute et de fabrication des sons diffusés » (32). Lors de *l'époque des découvertes, l'avènement des sources-relais* favorise le transport des sons d'un espace à l'autre et *l'avènement du stockage des sons* permet le transport des sons d'un espace-temps à l'autre. Lors de *l'époque moderne, l'avènement de la synthèse des sons* supprime l'espace de départ et *l'avènement des prothèses nomades* (baladeur) supprime l'espace d'arrivée⁴ (32). L'écoute causale de second niveau tente toujours de relier les sons à leur source, mais cette source peut être médiatisé par l'entremise de ces inventions technologiques.

L'écoute indirecte a deux conséquences pour le cinéma. Premièrement, « en situation d'écoute indirecte, c'est le résultat qui importe, puisque le processus de fabrication demeure inaccessible à l'auditeur » (34). La bande-son ne proposera plus « une sorte de stylisation imitative d'un son, comme chez les bruiteurs du muet, [elle fera croire] que le bruit « a été », [qu'il est la trace d'un véritable événement] » (34). À l'instar de la photographie, le son peut avoir la qualité indicielle d'une trace. Et puisque l'auditeur n'a pas accès à la chaîne de fixation, le concepteur sonore a une grande liberté afin de créer l'illusion que les sons sont produits par des sources visuelles. Deuxièmement,

les sons qui émanent de sources-relais [ont] une variabilité plus grande que celle de leurs homologues naturels [...] ; alors que depuis l'aube des temps le tonnerre, par exemple, a toujours été perçu par les hommes comme une manifestation sonore à la dynamique exceptionnellement importante et générant des fréquences très basses, on peut maintenant entendre de « minuscules » coup de tonnerre émaner du haut-parleur d'une petite radio portable» (35).

Ce peu de fidélité entre les sons fixés et leur source d'émission dilue la notion d'échelle forgée lors de l'écoute causale de premier niveau. De plus, « le

⁴ Les prothèses nomades placent le son au centre de la tête de l'auditeur.

répertoire mental de couples son/source se construit sur une base de traits pertinents invariants beaucoup plus réduite qu'avant l'avènement de l'écoute indirecte » (36). Par exemple, le son du tonnerre n'a plus besoin d'avoir une dynamique et des basses fréquences imposantes pour être reconnu comme tel.

Cette altération des qualités acoustiques est due à la réduction de la bande passante. « Des origines à la fin des années soixante-dix, le cinéma sonore a écarté les sons au-dessus de 8000 Hz et au-dessous de 45 Hz » (36). Évidemment, la catégorie sonore qui a souffert des imperfections de transmission est celle des bruits, puisque ceux-ci couvrent habituellement toutes les fréquences audibles. « Les bruits arrivant [...] aux oreilles du spectateur passablement abîmés, on ne pouvait pas les charger d'un rôle important dans la narration filmique » (37)⁵.

En définitive, les processus menant à l'identification des sources facilitent le classement, par le spectateur, des différents sons. L'organisation de la matière sonore en catégories est à prime abord difficile puisque « la bande-son ne se présente pas comme un objet *découpé* » (7). La bande-son est perçue comme un continuum, « un magma ininterrompu » qui se déploie verticalement (simultanéité des sons) et horizontalement (déroulement ininterrompu de sons sans coupure apparente) (7). Autant les créateurs, les analystes que les spectateurs ont privilégié pour l'édification d'un récit les sons qui utilisent des

⁵ Cette citation nous permet de soulever un nouveau présupposé du cognitivisme de l'audiovisuel : un son d'une bonne fidélité acoustique donne plus d'information sur le monde qu'un son de mauvaise qualité. Par conséquent, l'auditeur peut reconnaître plus facilement un son fidèle parce qu'il possède une représentation mentale tirée de son expérience quotidienne. Le propre d'une œuvre étant d'être comprise, le cinéma narratif tend vers une construction « naturaliste » (18), qui crée ultimement une adéquation entre la représentation auditive et le monde réel. Voici quatre extraits qui nous aident à comprendre le lien nécessaire entre naturalisme, narration, récit et causalité : « Au cinéma et à la TV, [le naturalisme est] une façon de présenter les *donnée de l'écran* la plus proche possible, compte tenu des seules limites technologiques du dispositif, de la perception telle qu'elle est définie par les Gestaltistes. » (195). « Dès l'enfance on nous habitue à un naturalisme audio-visuel qui présuppose un son pour chaque chose et une cause (visible) pour chaque son. C'est pourquoi tout le cinéma narratif classique, dont une partie du pouvoir évocateur repose sur l'effet de réel, est très attentif à ne faire entendre que des sons à l'identification aisée, et dont la source sera mentalement rangée par le spectateur à un endroit précis et prévu » (18). « [Dans une œuvre narrative], le *récit* vise à faire comprendre au spectateur l'histoire racontée » (194). « [Le récit est une] organisation des *données de l'écran* qui se rapportent aux liens causals entre les événements de l'histoire » (196).

systèmes notables préexistants. En effet, la *voix* sert de support aux mots et la *musique* relève du système d'écriture polyphonique occidental. Étant difficilement notables, les bruits sont exclus des recherches. En croisant les axes verbal et musical, il est possible de se représenter le continuum et ses catégories (49) :

Tonal\verbal	_____	chant	_____	Sprechgesang	_____	parole	_____
Verbal\non tonal	_____	chant	_____	parole	_____	vocalises, rires, cris,	_____
						musique tonale	
						instrumentale, bruits	
Tonal\non verbal	_____	mus.	_____	tonale	_____	instrum.	_____
non tonal	_____					mus.	_____
						sérielle	_____
						bruits	_____
						Non	
						verbal\non tonal	

Cette partition du monde sonore reprend en fait les compétences et les connaissances des analystes et des spectateurs.

La localisation des sources dans l'espace représenté : processus imaginaire, consignes de lecture, mondes narratifs

Dans le cinéma monophonique, le positionnement des sons dans l'espace visuel est un « processus imaginaire » (57). Lorsque la source est visible à l'écran, « l'ancrage consiste [...] à confondre mentalement dans une même position de l'espace (l'espace tri-dimensionnel dont la base est l'image « dépliée ») source et espace imaginés avec source et espace représentés » (62-63). Ce processus illusionniste d'ajustement de la perception dans la salle de projection et de la représentation à l'écran est « automatique » (57). Il fait appel au processus cognitif ascendant (*bottom-up*). Lorsque la source n'est pas visible à l'écran, le spectateur place la source à une certaine distance de l'espace visuel représenté (hors-champ contigu, absolu, fosse d'orchestre) (63). Dans ce cas, le positionnement n'est pas automatique, il est tributaire des conventions cinématographiques et fait appel au processus cognitif descendant (top down).

En résumé, on peut considérer que les opérations mentales qui concernent l'ancrage des sons filmiques oscillent entre deux modèles, le « circuit court », où les automatismes qui interviennent dans la perception du monde peuvent fonctionner (caractère « naturaliste » du dispositif cinématographique, faible importance des conventions, illusion imposée); et le « circuit long », où les automatismes qui interviennent dans la perception du monde ne peuvent plus fonctionner (caractère non-naturaliste du dispositif cinématographique, grande importance des conventions) (67).

Jullier a déjà décrit les processus automatiques de type *bottom-up*. Il s'attarde maintenant au circuit long, c'est-à-dire aux consignes de lecture qui permettent au spectateur de situer les sources non-visibles dans l'espace représenté. Les *consignes internes* sont données par le film et les *consignes externes* sont données par les institutions (96). Le spectateur tente de comprendre les données de l'écran à l'aide de ces consignes. Et puisque « [c]omprendre, c'est construire une représentation » (Denhière 1990 : 70), le spectateur bâtit des « mondes possibles narratifs » (96). Ces derniers ont comme « base invariable » le monde quotidien. « En l'absence d'indications sur les « propriétés alternatives » d'un monde, c'est le monde réel qui sert de référence, dans le doute, aux inférences du spectateur. Tout au long de la projection d'une œuvre narrative de fiction, le spectateur est amené à faire à la fois des *prédictions* et des *vérifications* (conformité à ce qu'il connaît déjà du monde ; en cas de non-conformité, il faut redéfinir le monde, en modifier les lois) » (96-97). La délimitation par le spectateur des différents mondes du film est un travail régi par les consignes de lecture. « Il ne s'agit plus de s'occuper du spectateur « tel qu'il est », mais du spectateur tel que le film et l'institution l'*incitent* à être » (103). Historiquement, l'institution cinématographique a mené le spectateur à distinguer le monde-narratif, les hétéro-univers (univers de croyance d'un personnage), la fosse et le monde de la fabrication et de la projection (103). Le spectateur ne place pas les sons selon la dichotomie diégétique/extra-diégétique ; il situe les sons le long de cet axe où les mondes ont différents degrés d'extériorité par rapport au récit.

La théorie cognitive de l'audiovisuel met en place une véritable « pensée de l'adéquation ». Elle tente de décrire des schémas mentaux invariants, en adéquation avec un monde quotidien invariant et avec des œuvres constantes dans leurs visées narratives. La perception esthétique doit égaler la perception quotidienne. L'opération de base du spectateur est toujours la même : il mesure le degré d'adéquation du monde représenté au monde réel. Les descriptions de l'analyste seront également jugées selon leur degré de conformité avec les habitudes du spectateur coopérant. Ce dernier est en fait une représentation idéale d'un sujet invariant qui obéit aux consignes de l'institution et qui adopte les postulats cognitif et narratif du chercheur, à savoir qu'une œuvre transmet des informations afin de faire comprendre une histoire. Dans ce paradigme,

l'acte créateur rejoue ce même mouvement : le cinéaste a l'intention de « dire » quelque chose et, pour le faire efficacement, il s'imagine comment le spectateur coopérant réagira à son montage d'images et de sons. Pour réussir à raconter son récit, le créateur doit donc faire correspondre ces intentions avec les processus cognitifs du spectateur. L'œuvre est le reflet de cette intentionnalité : le lieu d'une rencontre entre l'émetteur du message et son récepteur. En somme, cette pensée vise ultimement à une adéquation entre l'analyse du cognitiviste, la compréhension du spectateur et les visées du créateur.

Frédéric Dallaire
Juin 2008