

Branigan, Edward. « Sound, Epistemology, Film », dans Richard Allen et Murray Smith (direction), *Film Theory and Philosophy*. Oxford : Clarendon Press (1997), p. 95-125.

Le terme *épistémologie* doit être compris dans sa signification anglo-saxonne : Branigan analyse le rôle du son à l'intérieur d'une *théorie de la connaissance* cinématographique. L'auteur s'interroge sur la place accordée au son dans le processus spectatorial de construction d'une narration filmique (95). La compréhension de l'activité d'un sujet connaissant (le spectateur) passe donc par une analyse des différentes étapes menant à la composition d'un récit filmique. L'auteur base son étude sur les quatre principales étapes (niveaux) menant à l'élaboration d'un tissu narratif : 1) il décrit tout d'abord les différences et similitudes entre la perception des phénomènes sonore et visuel (« *Is sound less closely tied to the Kantian category of substance than vision?* ») ; 2) il explicite le cadre de référence qui déterminent les habitudes perceptives des spectateurs (« *What presuppositions about sound direct our search for knowledge from the visual features of film?* ») ; 3) il étudie l'influence de nos conceptions du temps sur notre perception des sons (« *May these presuppositions be altered to change our perception of the relationship between sound and light ?* ») ; 4) il analyse les processus cognitifs permettant de distribuer les données (*data*) dans l'espace filmique et de les inscrire dans un discours narratif. (*[H]ow do we expect sound to be of use to us in describing the world and in imagining a real world through the fictional depictions of a film ?*). L'ensemble de ces questionnements délimite les problèmes propres à l'épistémologie du son filmique (96) et présente en définitive les conditions *a priori* de l'écoute au cinéma.

#### *Son et lumière : asymétrie, matérialité, langage*

Le son et la lumière ont la même base physique ; ils sont des ondes en mouvement. Et pourtant nous les percevons de façons différentes. Il existe une asymétrie dans notre façon d'attribuer des propriétés visuelles et sonores aux objets. « It appears that objects are composed of visible (and other) properties while sound seems detachable and transient » (96). La couleur (résultat de la rencontre entre la lumière et une surface) est considérée comme une caractéristique de l'objet. Le son, lui, provient d'une source (une radio, une

porte, une botte), mais ne nous renseigne pas sur la substance, sur l'essence de l'objet. La couleur est une qualité *possédée* par l'objet ; le son est *produit* (et projeté) par une source (95). Par conséquent, la couleur est perçue comme une caractéristique permanente : nous croyons que la couverture du livre reste rouge même si la lumière s'éteint (ce qui est faux au niveau physique, mais vrai au niveau du sens commun). Le son, lui, est perçu comme un phénomène transitoire et contingent : le silence n'est troublé que lorsque les objets — fondamentalement inertes — bougent. En somme, lorsque le sujet connaissant tente de découvrir la substance et le positionnement des choses, la lumière apparaît comme une donnée permanente possédée par un objet alors que le son est une production contingente et éphémère tributaire de la mise en mouvement de cet objet (99). Au niveau perceptif, la permanence du visuel s'oppose à la contingence du sonore.

Mais pourquoi la lumière et la couleur semblent plus « attachées » aux objets matériels que le son et pourquoi ils présentent des repères plus efficaces afin de guider les mouvements de notre corps dans le monde ? (96) La prédominance du visuel sur le sonore s'explique par deux facteurs. Premièrement, la composition physique des ondes lumineuses permet une perception claire et distincte de l'espace. La lumière délimite avec une grande précision les contours des objets. Le son est plus flexible : il franchit les espaces mais donne des informations beaucoup plus diffuses sur les objets. Pour se déplacer et s'orienter, l'être humain a besoin de connaître l'emplacement exact des objets afin de construire mentalement une représentation de l'espace à parcourir. Le son n'offrant que des informations partielles et diffuses (puisque la réflexion des ondes altère les qualités acoustiques des sons), l'être humain a développé une façon d'évaluer la position des objets principalement à partir des données (*data*) visuelles (97). Deuxièmement, la propagation (le mouvement) des ondes sonores et visuelles diffèrent. Les premières sont produites mécaniquement par des objets et ont besoin d'un médium de transmission (l'air ou l'eau). La présence du son nécessite donc deux mouvements : la vibration d'une source et la vibration d'un médium (97). Le son est perçu à travers la modulation d'un médium élastique qui entoure le corps humain. Quant à la lumière, elle est une radiation électromagnétique qui n'a pas besoin de médium. Elle révèle ces propriétés à travers la réflexion des objets, alors que le son révèle ses propriétés indirectement à travers les jeux de tension et de détente d'un médium autour de nous (99). La perception *médiatisée* et toujours *en*

*mouvement* des occurrences sonores renforcent notre impression que le son est *produit et contingent*.

L'asymétrie perceptive entre les phénomènes sonores et visuels se répercute dans nos habitudes langagières. Suivant Metz, Branigan examine l'hypothèse selon laquelle les perceptions sont tributaires de la fonction nominale du langage (98). La perception d'un phénomène et le nom le désignant formeraient une entité indivisible, à la source de notre appréhension et de notre compréhension du monde (99). Ainsi, le son joue le rôle d'un *adjectif* alors que l'image est un *nom*. « When we see a « lamp » and can name it, the identification is complete and all that could be added would be merely adjectival – a « tall, reading » lamp. » (98). Or, quand nous entendons et nommons un son, l'identification demeure incomplète : « A 'whisling' sound still needs to be specified : the whisling *of* what? *From* where? The whisling *of* the wind *in* the trees *from* across the river » (98). L'écoute d'un son se double donc d'une investigation afin d'identifier la source d'émission, l'objet en mouvement. Le son fonctionne comme un adjectif parce qu'il décrit et caractérise une substance fondamentalement visuelle déjà désignée par un nom. Metz décalque sa théorie sur le perçu et le nommé de la structure « sujet-prédicat » des langues indo-européennes et de la distinction entre les qualités premières et secondaires issues de la tradition philosophique occidentale (98). La contingence et la dépendance du son envers sa source réduit celui-ci à un rôle de *prédicat* exprimant les *qualités secondaires* d'une substance principalement visuelle. Branigan conserve l'analyse de Metz mais rejette le déterminisme linguistique qui la soutient. La linguistique n'a pas à être le seul paradigme épistémologique dans la description de nos opérations perceptives ; le langage est seulement une catégorie d'opérations mentales (*mental computation*) (99). La différence fondamentale entre le sonore et le visuel se retrouve également dans la perception des phénomènes, comme nous l'avons vu, dans le dispositif cinématographique, dans l'acte de spectature et dans toutes activités cognitives.

Ces analyses phénoménales et linguistiques impliquent que la présence du son dans un film enrichit même les images les plus statiques puisqu'elle produit un mouvement continu dans la salle de projection. Le médium (l'air) traverse les espaces et touche autant le spectateur que les événements représentés à l'écran, ce qui crée une proximité (une intimité) entre le sujet connaissant et son objet de connaissance (99). La juxtaposition du son et de l'image au cinéma

permet de renforcer l'impression que l'espace représenté visuellement contient du mouvement et du volume. La vibration du son dans un espace en trois dimensions complexifie la représentation de l'espace visuel en deux dimensions. Le son est alors un indice de l'activité de sources sonores (le bruit des bottes marque la présence humaine) et des dimensions de l'espace (le bruit des bottes résonne dans un vaste amphithéâtre). Les qualités adverbiales, secondaires et transitoires n'empêchent donc pas le son de jouer un grand rôle dans la création d'un espace filmique unifié (100).

*Cadre de référence : le monde des perceptions moyennes*

Au cinéma, il arrive parfois des moments où le son conditionne notre perception des images. Par exemple, lorsque le son d'un nouveau lieu se fait entendre avant que l'image se transporte dans cet espace. À ce moment, le son précède la lumière dans notre compréhension du film : « [it] seems to be anticipating and creating space, addressing questions about the how and why of the objects seen » (100). Afin de comprendre ce renversement de la hiérarchie phénoménale visuel/sonore grâce au son asynchrone, Branigan étudie la relation entre le son, l'espace et le temps ainsi que le cadre de référence qui conditionne toute expérience possible.

Dans l'étude de la perception, « it is better to avoid characterizing sound in absolute terms in favour of relating it to a particular framework in which it may be known and to a particular set of questions one wishes to answer » (101). Le cadre de référence utilisé pour la présente étude est la *perception humaine* et le questionnement principal implique la relation entre la perception sonore *dans notre environnement quotidien* et la perception sonore *dans une salle de cinéma*. L'expérience humaine est confinée au « monde des dimensions moyennes » décrit par la mécanique newtonienne et ne comprend pas la structure infinie de l'univers décrite par Einstein, ni le monde microscopique de la mécanique quantique (101). Les mécanismes de la perception et de la pensée humaine sont adaptés au cadre temporel, spatial et causal du monde des dimensions moyennes. Ce monde ordinaire et quotidien est dominé par les objets solides. Pour survivre dans ce monde, l'être humain a besoin de percevoir des points de repère qui restent identiques dans le temps et l'espace. L'unité de base qui assure la stabilité et qui conditionne les actions humaines est ici l'objet solide. Il permet la *reconnaissance* à travers *l'identité* de *substances* ayant une série *d'attributs permanents*.

Dans ce monde, le son est désavantagé par rapport à l'image 1) parce qu'il n'a pas les mêmes fréquences que la lumière (ce qui nous informe de manière diffuse sur les caractéristiques des objets) et 2) parce qu'il est intimement lié au mouvement continu d'un médium élastique perçu comme dépendant de l'espace (« motion is interpreted as a movement from place to place of solid objects *within* a space ») (101). La lumière a préséance sur le son parce qu'elle nous révèle plus efficacement les propriétés des objets solides qui constituent l'unité de base du monde des dimensions moyennes. C'est ce qui explique que la lumière possède les propriétés des objets solides qu'elle illumine (permanence, clarté et distinction) ; elle révèle de manière directe (transparente, non-médiatisée) la substance de ces objets. C'est pourquoi le langage humain représente les éléments visuels sous la forme de noms. Le son, lui, apparaît comme éphémère et gouverné par le mouvement, « making it more likely to be represented in the form of a verb or adjective –as a predicate or attribute of some *other* object whose materiality is taken for granted. » (101). On comprend alors que l'ensemble de la réflexion sur le son est conditionné par ce cadre de référence, inspiré de la perception quotidienne. Par conséquent, un spectateur de cinéma applique ces habitudes perceptives forgées dans le monde des dimensions moyennes à l'expérience cinématographique : il cherchera à constituer un espace unitaire, il suivra l'évolution de la représentation des objets solides, il tentera d'identifier la source d'émission des sons.

#### *Conceptions du temps et perception des sons*

Le son asynchrone pose toujours problème : Branigan résout cette tension en supposant qu'il existe deux conceptions du temps dans le monde des perceptions moyennes. À chacune d'elle correspond une façon différente de percevoir les sons (102). La première conception est basée sur une perception directe des sons. Ceux-ci apparaissent, durent un certain temps, puis se fondent à nouveau dans une rumeur indifférenciée. Le son est alors une vibration, une modulation infinie. « If sound were to be frozen, there would be nothing to hear. It must always move *toward* something, or emanate *from* some (more permanent) object or state of being » (103). C'est sur cette observation empirique des phénomènes que se base Tom Levin pour affirmer que le son et l'image sont fondamentalement différents<sup>1</sup>. Ainsi, il est possible de les rendre commensurables ou de subordonner le sonore au visuel, mais seulement en

---

<sup>1</sup> Tom Levin, « The Acoustic dimension : notes on cinema sound », *Screen*, vol. 25, n° 3 (1984), p. 62-64.

invoquant des présupposés d'ordre idéologique. Cette conception est également compatible avec la perception du son en tant que verbe ou adjectif.

La deuxième conception est basée sur une perception indirecte des sons. Ces derniers s'inscrivent alors dans un contexte plus large (un discours ou une structure narrative). Dans ce paradigme, la répétition et l'organisation symétrique des occurrences sonores peuvent être perçus comme un arrêt, un retour, ou un ralentissement du temps. Par exemple, dans *Fist in the Pocket* (Bellocchio, 1966), lors de la mort de Lou Castel, le phonographe se dérègle et répète une note sur une image arrêtée (*freeze frame*). À l'instar du temps, la note est suspendue. Les cinéastes disposent donc, à l'aide d'une structure narrative, de différentes stratégies pour modifier notre perception temporelle. Le temps acquiert alors une nouvelle consistance et une extension dans l'espace.

Temporal *duration* is now being interpreted comparable to the *extend* of a solid object (e.g. It's been a *long* day [...]) and so may be conceived as something fixed and solid, not moving though still movable, capable of being shrunk, expanded, shifted or relocated ; capable even of being perfectly duplicated and repeated. In this view, time becomes a substance described by a noun or gerund (103).

Si le temps, caractérisé par le mouvement continu, peut, à l'intérieur d'une structure narrative, être désigné par un nom, le son, perçu à travers la vibration d'un médium, peut également acquérir des qualités nominales. C'est sur cette organisation narrative du temps que se base William Johnson pour affirmer que le son et l'image sont intrinsèquement égaux et idéologiquement neutres<sup>2</sup>. Ils sont soumis aux mêmes impératifs techniques et aux mêmes manipulations esthétiques. Par conséquent, toute inégalité ou domination de l'un sur l'autre est injustifiée<sup>3</sup>. En somme, « sound (and time) gains its equality when it is indirectly perceived [...] by being fixed within a context that is larger than identifying and locating objects in space. » (104).

---

<sup>2</sup> William Johnson, « The Liberation of echo : a new hearing for film sound », *Film Quarterly*, vol. 38, n° 4 (1985), p. 2-12.

—1989. « Sound and image : a further hearing ». ». *Film Quarterly*, vol. 43, no 1 (automne), pp. 24-35.

<sup>3</sup> Cf. le texte en langue française de Noël Burch. 1973. « Chapter 6 : On the structural use of sound ». Dans *Theory of film practice*. New York : Praeger. « It seems that the essential nature of the relationship between sound and image is due not to the difference between them, but rather to the similarity between them. »(p. 91).

À ces deux conceptions du temps correspondent deux types de perception du son et de la lumière au cinéma. Le premier type de perception est basée sur une acquisition de données (*data*) de façon directe et brève appelé *bottom-up*. Le spectateur organise immédiatement les données sonores afin de caractériser le son entendu (hauteur, timbre, forme, mouvement, allure) (105). Le processus *bottom-up* est sériel (il se répète à chaque nouveau son) et il est tributaire des qualités matérielles des sons. Il ne produit que des effets de courte durée puisqu'il varie selon le déploiement dans le temps de stimuli. Le deuxième type de perception, « based on acquired knowledge, memory and schemas (frames, scripts) [...], work primarily « top-down » on the data using a spectator's expectations and goals as principles of organization » (105). Le processus *top-down* est indirect parce qu'il organise les données dans un cadre de référence et des grilles d'analyses et de lecture indépendamment des premières stimulations phénoménales. Le processus *top-down* organise la perception empirique des sons (*bottom-up*) en suivant les règles causales, temporelles et spatiales propres au monde des dimensions moyennes. Le son au cinéma échappe donc à la stricte modulation empirique et s'inscrit dans l'activité cognitive d'un spectateur élaborant un tissu narratif. Il gagne alors une consistance nouvelle et aspire, à l'instar des formes visuelles, à la catégorie des objets solides. « Therefore, in cases where it is *knowledge* that is being measured and evaluated, and used to make predictions, sound is on its way toward equality. Sound achieves its equality when it is taken as a form of knowledge rather than taken as a state of the real world. » (106).

### *Son, narration, cognition*

Nous sommes maintenant en mesure de comprendre l'activité du spectateur lorsqu'il perçoit un son asynchrone. Comme le son filmique s'inscrit dans un contexte narratif qui le transcende, il acquiert une substance suffisante pour entrer en dialogue et même renverser la position dominante du visuel. Les différentes interactions entre le son et l'image doivent se comprendre à l'intérieur des balises épistémologiques créées par le dispositif filmique : ce dernier est un texte offrant différents niveaux narratifs (*levels of narration*) interprétés par un spectateur lui-même à l'intérieur d'une narration<sup>4</sup> (108). Les

---

<sup>4</sup> « 'Narration' refers not to what is told and happens in a story, but rather to the conditions of the telling – to the overall regulation and distribution of knowledge which determines when and how a reader acquires knowledge from a text [...] Note that the issue addressed by the concept of narration is not the actual production (recording, editing, mixing) of a sound (or an image) in

différentes données (*data*) proposées par la représentation audiovisuelle offrent une multiplicité d'interprétations. Le spectateur construit une narration qui relie plusieurs niveaux, confirme ou invalide différentes hypothèses nées au cours de la projection. Par exemple, le spectateur peut associer un son à différentes sources, ou nommer une source déjà perçue en tant que nouvel objet parce qu'un autre son se fait entendre (109). Ce processus de lecture est infini et peut toujours bifurquer lors de l'apparition de nouvelles données empiriques (*bottom-up*).

\*\*\*

En définitive, cette analyse montre comment notre perception est tributaire du degré de matérialité accordée aux phénomènes sonores. Le monde des perceptions moyennes a comme unité de base les objets solides. Le son fluctuant et éphémère est alors moins important que l'image dans la constitution des substances (tel que définit par Kant). Le cinéma bouleverse cette hiérarchisation puisque le son fixé 1) peut être répété, manipulé, travaillé comme un matériau solide 2) peut être un objet de connaissance s'inscrivant, au même titre que l'image, dans une structure narrative. Le son gagne alors une matérialité qui lui donne une importance dans le processus cognitif du spectateur. Cette nouvelle consistance est indirecte puisqu'elle ne concerne pas la nature du son, mais bien la *représentation* que nous avons du monde sonore : le son est *comme* un objet ; nos processus cognitifs le considèrent *comme* une donnée ; notre langage l'identifie *comme* un nom. Dans ce cadre de référence, la matérialité du son ne peut être affirmée qu'en niant ou en minimisant ses caractéristiques phénoménales : malgré que le son soit variable, médiatisé, contingent et produit, il est possible de se le représenter *comme s'il* était constant, immédiat, permanent et possédé. Il nous semble qu'en présence d'un paradigme qui tente de gommer les données fondamentales d'un phénomène sous une représentation, il ne faut pas modifier notre discours sur le son, mais bien changer de paradigme.

Frédéric Dallaire  
Mai 2008

---

a film, but its reception and interpretation by a spectator. » (123-124) Tout est dans la tête du spectateur.