



La perspective sonore

Jean Chatauret

DANS **COMMUNICATIONS** 2009/2 n° 85 , PAGES 139 À 154
ÉDITIONS **LE SEUIL**

ISSN 0588-8018

ISBN 9782021002423

DOI 10.3917/commu.085.0139

Date de mise en ligne : 01/01/2012

Article disponible en ligne à l'adresse

<https://shs.cairn.info/revue-communications-2009-2-page-139?lang=fr>



Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...
Scannez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



Distribution électronique Cairn.info pour Le Seuil.

Vous avez l'autorisation de reproduire cet article dans les limites des conditions d'utilisation de Cairn.info ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Détails et conditions sur [Cairn.info/copyright](http:// Cairn.info/copyright).

Sauf dispositions légales contraires, les usages numériques à des fins pédagogiques des présentes ressources sont soumises à l'autorisation de l'Éditeur ou, le cas échéant, de l'organisme de gestion collective habilité à cet effet. Il en est ainsi notamment en France avec le CFC qui est l'organisme agréé en la matière.

Jean Chatauret

La perspective sonore

C'est à la perception visuelle et aux techniques graphiques que le terme de « perspective », au sens premier, s'applique. Cependant, notre système auditif nous permet de percevoir des plans, d'évaluer les distances relatives de différentes sources sonores. Et les techniques de captation et de reproduction du son visent, surtout depuis l'avènement de la stéréophonie, à la réalisation d'une « image sonore ».

Au cours des années 1950 et surtout 1960, l'essor du « microsillon » et de la « haute-fidélité », puis l'apparition des technologies à transistor permirent à un public de plus en plus large de découvrir le son « stéréophonique ». Au début de la chaîne de production les techniques de captation évoluaient, et l'on vit des équipes d'ingénieurs du son, d'électroniciens et d'acousticiens mettre au point des dispositifs de prise de son stéréophonique. Ces systèmes se devaient d'être polyvalents. Qu'il s'agisse d'enregistrer un concert de musique classique, un discours ou une ambiance en extérieur, un dispositif microphonique stéréo avait pour fonction de restituer à l'auditeur une image sonore réaliste, exempte de déformations, de distorsions, respectant les proportions des sources, leur placement frontal et distal dans l'espace sonore¹. L'analogie avec l'image semble évidente : comme le photographe qui, en plaçant judicieusement l'objectif de son appareil, donne à voir une scène depuis un point de vue qu'il a choisi, l'ingénieur du son, en déterminant le placement d'un système unique de prise de son stéréophonique, se devait de donner à entendre à l'auditeur une scène sonore réaliste, en relief, si vraie qu'on l'appellerait « stéréophonique », du grec *στερεος* : « solide »².

Et pourtant, qu'il s'agisse des phénomènes physiques ou perceptifs, des fonctions sémantiques du son et de l'image, de la façon dont nous réagissons aux informations visuelles et sonores, rien entre audition et vision n'est identique ni même sérieusement comparable. Pourtant aussi, perceptions sonore et visuelle sont si intimement liées que l'une ne vient

jamais seule sans que l'autre ne nous manque. L'absence de son (est-ce le silence ?) nous fait voir autrement, tout comme l'obscurité ou le hors-champ confèrent aussitôt au son un caractère particulier. Enfin, on évoque souvent le son avec les mots du visuel : on parle d'une « voix claire », d'une « tonalité sombre », ainsi que d'image et de perspective « sonores ».

Qu'est-ce que la perspective sonore ?

J'en présenterai deux facettes. L'une concerne les phénomènes physiques et perceptifs qui nous permettent de localiser des sources sonores dans l'espace. L'autre est d'ordre esthétique : du naturalisme des débuts de la stéréophonie aux mixages contemporains que CD, télévision, cinéma nous donnent à entendre, on se demandera quel est le chemin parcouru. J'évoquerai ces questions avec les mots et l'expérience d'un ingénieur du son, avec le « point de vue » particulier de celui qui chaque jour réalise des « images » sonores.

LE SON DANS L'ESPACE ET LE TEMPS

La perception des distances.

Nous promenant au bord d'un lac, nous observons, à quelques mètres, un enfant qui joue en lançant des pierres dans l'eau : nous entendons alors le son que fait une pierre qu'un enfant jette dans l'eau, autrement dit le son de l'eau qu'une pierre frappe, jetée par un enfant³. Continuant notre promenade, nous percevons, mais maintenant dans le lointain, le son des pierres que l'enfant jette. Ce n'est pourtant plus le même objet sonore. En quoi diffère-t-il du premier ? Et maintenant supposons que l'on puisse entendre ces mêmes sons dans un autre contexte : dans un appartement, par exemple. Les identifierions-nous comme « le son d'une pierre jetée dans l'eau », dans le premier cas proche, dans le second plus lointain⁴ ? Probablement pas. Notre interprétation des sons est intimement liée à un faisceau contextuel⁵. Considérons d'abord le domaine qui nous vient logiquement à l'esprit, s'agissant de perspective : l'espace.

Le son est une onde produite par la vibration mécanique d'un support fluide ou solide et propagée grâce à l'élasticité du milieu environnant sous forme d'« ondes longitudinales ». Pour être perçue par l'homme, cette vibration doit avoir une fréquence comprise entre 20 et 20 000 hertz. En outre, son énergie doit se situer entre le seuil (minimum) d'audibilité et le seuil maximum, dit seuil de la douleur. Ces seuils sont variables suivant la fréquence. Le son se propage dans l'air à une vitesse d'environ

340 mètres/seconde. Mais cette vitesse varie en fonction de la pression, de la température et de l'hygrométrie de l'air.

L'onde sonore, en parcourant une distance, est atténuée : d'une part parce que, dans le cas d'un rayonnement sphérique, son énergie se disperse dans l'espace – d'où une atténuation théorique de 6 décibels chaque fois que la distance est doublée –, d'autre part en raison de la viscosité de l'air. Cette atténuation n'est pas la même à toutes les fréquences et l'atténuation perçue est plus importante pour les basses et les élevées (courbes d'isotonie de Fletcher et Munson).

Voilà donc quelques éléments permettant de caractériser la perception de la distance d'une source sonore : atténuation de l'énergie sonore, modification du spectre fréquentiel – donc d'un élément constitutif de ce que l'on appelle en musique le « timbre » –, mais aussi modification des transitoires qui caractérisent tout autant que le timbre un objet sonore. Ainsi, si le son de la pierre jetée dans l'eau est tout à fait identifiable à quelques mètres et dans le contexte d'une promenade au bord d'un lac, il n'est pas certain que ce même son, nous parvenant à quelques centaines de mètres de son origine, serait si facilement reconnu.

La présence.

À l'impression d'éloignement, on oppose la proximité. On parle de présence. On confond souvent présence d'un son et puissance – ou force – d'un son. Parce que amplifier la force d'un son est une des façons les plus simples de donner une impression de proximité ; c'est le b.a.-ba du mixage – et parfois d'un mauvais mixage. On pousse un son pour « mieux » l'entendre, du coup, un autre élément sonore paraît plus « lointain », on le renforce, et commence une inflation vite incontrôlable du plus fort et du plus près jusqu'à la saturation des circuits électroniques. En agissant sur le spectre des fréquences, on peut renforcer l'impression de proximité sans pour autant augmenter le niveau d'un son. Certains microphones sont caractérisés par une « bosse de présence », renforcement des fréquences aiguës dans une zone sensible de notre audition, aux alentours de 4 000 hertz.

Absence, silence.

Si la présence d'un élément sonore signifie proximité mais aussi puissance, alors il faut évoquer l'éloignement spatial d'une source, mais aussi son éventuelle absence. Le silence absolu n'existe pas sur terre. Il serait

plutôt ce que nous ressentons lorsqu'un son nous manque. Nous ne pouvons ressentir ce silence que parce que le son qui manque a existé, nous est connu et que nous pensons qu'il pourrait réapparaître. Quand et où un objet sonore a-t-il disparu ? Pour combien de temps ? Est-il caché ou masqué par un autre son ?

Le silence, c'est aussi une façon de ressentir le temps qui passe. Et nous verrons que, s'agissant du son et de la perspective sonore, temps et espace sont indissociablement liés. On ne s'arrête pas devant une image sonore pour la contempler comme on le ferait devant un tableau. On n'arrête pas le son qui passe.

La réverbération.

L'onde sonore ne se propage pas selon une ligne droite et unique entre sa source et l'oreille de l'auditeur. Elle se disperse de façon sphérique, rencontre des obstacles aux caractéristiques acoustiques variables, se diffracte et se réfléchit partiellement, dans des directions qui dépendent d'une part des fréquences, d'autre part de la nature et de la forme de ces obstacles. Tant et si bien que le son perçu de notre pierre jetée dans l'eau, dès l'origine fort complexe, n'est pas unique : nous percevons son bruit directement perçu, mais aussi le même bruit réfléchi par l'eau du lac, peut-être aussi par des arbres alentour : plusieurs sons du même événement sonore nous parviennent, chacun différent, modifié et retardé par son parcours spécifique.

Il s'agit là d'un phénomène essentiel pour la perception des distances : en fonction du lieu où nous nous trouvons et de la distance entre nous et la source sonore, nous percevons plus ou moins de son direct (l'onde qui nous parvient en premier, directement de la source⁶) ou de son réfléchi (toutes les autres ondes). On parle de « champ direct », de « champ diffus » et de « distance critique » – distance à laquelle les énergies directes et réfléchies seraient égales. Mais on comprend aisément que les ondes réfléchies sont de natures très diverses, que le moindre groupe d'ondes émis dans un espace clos génère aussitôt un ensemble extrêmement complexe de sons réfléchis, de résonances et d'ondes stationnaires.

L'importance déterminante de la réverbération dans la perception de la perspective sonore ne peut pas être évaluée uniquement de façon quantitative. On perçoit plus ou moins de réverbération, mais c'est par sa durée, par la forme de sa décroissance dans le temps, par sa densité à des instants plus ou moins rapprochés du son direct, que la réverbération détermine la sensation de la perspective.

Un ingénieur du son, voulant placer « artificiellement » un son au lointain dans une image sonore, pourra avoir recours à quelques artifices : baisse du niveau sonore, atténuation des basses et hautes fréquences (les graves et les aigus), retard du son pour simuler un parcours plus long, un peu de réverbération artificielle. Facile à dire, pas si facile à réaliser, car si notre oreille est parfois fort tolérante et peut se laisser duper, elle sait aussi être d'une extrême sensibilité. Cela dépend du contexte temporel (ce que nous avons entendu auparavant et ce que nous nous attendons à entendre), des informations visuelles qui nous parviennent simultanément, et de la fonction de ce son-là dans l'écriture sonore, qui détermine la sélectivité de notre attention et, pour une part importante, de l'imaginaire de l'auditeur. Jusque-là, point n'est besoin de stéréophonie pour restituer le relief sonore : une seule enceinte acoustique devrait y suffire.

L'effet de masque.

Mais un autre phénomène vient appauvrir notre capacité à distinguer toutes les informations sonores disponibles : un son faible, audible dans un environnement sonore calme, n'est plus perceptible dans un milieu bruyant. Un son est d'autant plus masquant que son spectre de fréquences se rapproche de celles des sons masqués. Plus le son masquant sera fort, plus le masquage sera efficace.

Les sons de faible intensité absolue (c'est souvent le cas de la réverbération) et relative (atténuation d'un son due à la distance) sont un facteur déterminant de la perspective sonore. En conséquence, le masquage partiel ou total de certains sons faibles altère d'abord la perception de la perspective. Il suffit d'écouter un programme de musique classique dans l'environnement bruyant d'une voiture : au-delà d'un certain niveau de bruit de fond, l'espace et la profondeur originellement présents dans la prise de son disparaissent, puis l'altération du contenu va jusqu'au masquage de certains éléments harmoniques et mélodiques.

La représentation stéréophonique.

Nous pouvons percevoir les sons provenant de tout l'espace qui nous entoure. Notre système auditif binaural est capable d'en déterminer la direction, avec plus ou moins de précision suivant que ces sons nous parviennent frontalement ou de l'arrière, par exemple. C'est dans un cône d'attention frontal que notre capacité à déterminer précisément la pro-

venance d'une information sonore est la plus grande (c'est également – avec des nuances – notre zone de vision)⁷.

Comme nous l'avons dit plus haut, le son émis par une source sonore nous parvient d'abord sous la forme d'un « son direct », qui nous permet de déterminer la direction de la source et d'un ensemble de sons réfléchis qui peuvent, eux, nous parvenir de toutes les directions, modifiés et retardés au fil de leur parcours. Alors que notre champ auditif couvre 360 degrés dans les trois dimensions, la stéréophonie, dont l'angle d'écoute est limité par deux enceintes acoustiques, ne nous propose qu'une scène sonore d'une ouverture de 60 degrés sur un plan horizontal⁸. La stéréophonie ne peut donc pas proposer une restitution fidèle de la scène sonore, mais une représentation de celle-ci.

Cependant, en ouvrant sur 60 degrés le champ d'audition, la stéréophonie permet à l'ingénieur du son de placer différemment dans l'image sonore des éléments susceptibles de se masquer l'un l'autre. Cela favorise une écoute plus fine, et une meilleure perception de la perspective.

Ainsi le « réalisme » éventuel de l'image sonore stéréophonique dépend-il autant de l'intention de celui qui la réalise que des qualités intrinsèques d'un dispositif microphonique : il s'agit bien de faire des choix, de sélectionner lesquels des éléments sonores perceptibles dans la réalité par notre système auditif doivent être conservés, puisque tous ne peuvent être reproduits sans masquage dans l'espace limité de l'image stéréophonique. « À la volonté de cumul du réel, faut-il opposer la nécessité de la perte⁹ ? »

Un tel choix est évidemment interprétatif. L'ingénieur du son conservera les éléments qui lui semblent devoir faire sens, éliminant ou limitant les sons qu'il juge inutiles, gênants, masquants. Il invente alors une nouvelle mise en perspective sonore, dans un sens plus large que celui évoqué plus haut, puisqu'il s'agit non seulement de positionner des sources sonores dans l'espace stéréophonique, mais encore de les sélectionner.

Dans des conditions d'écoute naturelles, ce ne sont pas forcément les éléments sonores les plus forts ou les plus présents, ou ceux perçus frontalement, qui sont les plus signifiants. Ce sont ceux que notre système perceptif sélectionne, choisit de suivre, même s'ils sont lointains, postérieurs, partiellement masqués.

Le son multicanal.

Le son multicanal, par exemple le format 5.1 proposé aujourd'hui pour les dispositifs domestiques, puisqu'il permet une diffusion sur 360 degrés,

n'offre-t-il pas enfin la possibilité d'une reproduction « fidèle » d'une réalité sonore ?

Notons tout d'abord que ce standard ne prévoit pas l'installation d'enceintes en hauteur – et limite donc *a priori* la reproduction au plan horizontal¹⁰, sauf à avoir recours à des traitements numériques de type spatialiseur¹¹. Ajoutons que ce standard, issu des dispositifs destinés au cinéma, prévoit une répartition des cinq enceintes principales d'une façon non homogène autour de l'auditeur.

Cependant, la perspective spatiale sera perçue comme beaucoup plus naturelle, et l'agrandissement à 360 degrés (dans le plan horizontal) du champ de l'image sonore donne à l'auditeur une incontestable facilité d'écoute : il est plus facile de concentrer son attention sur une source sonore frontale dont les ondes réfléchies sont perçues sur l'arrière – ce que permet le 5.1 – que de devoir chercher le « son utile » dans la superposition frontale du son direct et du son réverbéré, comme c'est le cas en stéréophonie. Pour les mêmes raisons, l'ingénieur du son pourra, s'il le souhaite, être moins sélectif dans le choix des sons qu'il donne à entendre, plus nuancé et plus subtil dans ses mises en perspective, plus raffiné dans ses traitements sonores (par exemple son action sur les timbres), car l'élargissement du champ d'écoute atténue les phénomènes de masquage.

Perspective sonore et cinéma.

Si le cinéma avait été « parlant » en 1895, les frères Lumière nous auraient peut-être montré quelques secondes plus tôt la voie vide de la gare de La Ciotat, avant que la locomotive n'apparaisse. Le son lointain de la machine à vapeur, agrémenté (c'est incontournable) de quelques sifflements stimulant notre imaginaire par le son déjà connu d'une locomotive au loin bien qu'encore invisible, aurait renforcé l'effet de perspective visuelle et de rapprochement progressif. Nous laissons au lecteur le soin d'imaginer, en son multicanal, le passage de la locomotive, puis, après sa disparition sur la gauche de l'écran, son arrêt hors champ (crissements de freins, halètements de la machine...) en très gros plan sonore, tandis que la foule sur le quai se déplace et que les premiers voyageurs descendent des voitures. Que se disent-ils ? Qu'en laisserait-on entendre au milieu de ce vacarme ?

Le cinéma, puis le montage au cinéma ont révolutionné la notion de perspective. La perception de l'espace, du positionnement de ce qui est donné à voir et de celui qui voit, ne dépend plus d'une image, mais de l'enchaînement de plusieurs plans, d'axes et de focales différents. Les plans sonores, en général, ne suivent pas le montage images plan par plan

à l'intérieur d'une séquence. Sauf cas très particulier, que se passe-t-il lorsque les valeurs de plans sonores changent en synchronisme et en apparente corrélation avec les plans visuels ? On ressent ces « sautes » comme une rupture spatiale et/ou temporelle. S'agissant de films musicaux, du fait de la continuité musicale (donc temporelle) de l'œuvre, le spectateur-auditeur ressentira un déplacement du point d'audition, alors qu'une succession de plans visuels différents sans changement du son n'est pas ressentie comme une suite de ruptures.

Que percevons-nous lorsque, dans un concert filmé, un gros plan sur un instrument survient ? Sans que rien n'ait été changé dans la balance sonore, nous percevons plus clairement la ligne dudit instrument¹². Il y a en quelque sorte un effet de démasquage, d'éclairage de la source sonore. Est-ce pour autant une modification de la perspective ressentie ? Non, car le son global, dont la balance n'a pas été modifiée, assure la pérennité du « point d'audition ». Ce phénomène de « soulignement sonore » par l'image sans modification de la perspective globale est à rapprocher de l'effet d'« aimantation » décrit par Michel Chion. C'est un outil passionnant, en particulier dans le domaine musical.

S'agissant d'opéra filmé, le réalisateur et l'ingénieur du son jouissent d'une liberté de « mise en scène sonore » qui serait incongrue en musique instrumentale. Perspectives sonores et visuelles, sons *in* et *off* se combinent. Les possibilités sont nombreuses, limitées cependant (et heureusement, car l'effet gratuit n'est jamais loin) par la continuité de l'œuvre, et par le respect nécessaire du discours musical.

Vers la fin du « naturalisme » ?

Des dispositifs de diffusion sonore comportant de très nombreuses enceintes sont aujourd'hui utilisés au théâtre, en sonorisation de concert, dans des salles de cinéma à écran géant. Le concept de « scénographie sonore » concrétise-t-il l'aboutissement et en même temps la fin d'une période de l'histoire de la représentation sonore : celle de l'intention utopique de reproduire le réel dans un autre temps et dans un autre lieu ?

Finalement, seul notre imaginaire, stimulé par ce qu'il perçoit et en se référant à des sons déjà entendus et stockés dans notre mémoire, peut nous faire ressentir l'impression d'« une » réalité. Et comme l'écrit Claude Bailblé à propos de l'image cinématographique, « le spectateur ne vient chercher de l'imaginaire qu'à condition de se défaire d'un trop-de-réel ». De mieux en mieux la technologie offre au créateur la possibilité d'imaginer des écritures sonores sophistiquées, donnant ainsi toute sa place à l'enveloppement, au relief, autorisant sans masquage une multitude

d'événements sonores simultanés proposés à l'écoute de l'auditeur. L'auditeur, quant à lui, retrouve non pas l'écoute du réel, mais la liberté d'écoute qui caractérise le monde réel : « Ce qui est primordial pour l'écoute, c'est la conservation de la liberté de choisir ce que l'on a envie d'entendre dans ce qui nous est offert. Et de pouvoir attribuer des temps d'attention différents selon l'intérêt porté à ce que l'on a choisi¹³. » N'est-ce pas une proposition dynamique et séduisante pour évoquer la perspective sonore ? Tout comme la perspective graphique, elle instaure le point de vue – dirait-on le « point d'écoute » ? –, dont le concept est à la fois spatial et temporel.

ESTHÉTIQUE ET PERSPECTIVE SONORE

Perspective et musique.

Avant que la technologie n'en permît l'enregistrement, la musique ne prenait vie que dans le lieu de son interprétation. Elle était parfois composée pour un lieu déterminé. On imagine mal le chant grégorien chanté dans un gymnase, pas plus que des danses folkloriques interprétées dans une abbaye, une grande symphonie romantique jouée en plein air sans sonorisation. Tout cela se fait cependant de nos jours, avec plus ou moins de bonheur, le plaisir de l'auditeur-spectateur dépendant de ce qu'il attend de l'expérience. La littérature musicale foisonne d'effets sonores de mise en perspective, par des échos, des réponses, des contrastes. Souvent, le compositeur indique une scénographie sonore : trompettes de coulisse, cloches au loin, « banda » d'opéra défilant en fond de scène¹⁴...

Un enregistrement discographique du *Parsifal* de Wagner me revient à l'esprit. La fin du troisième acte se situe dans le château des chevaliers du Graal. Wagner a donné dans la partition même des précisions sur la salle du Graal, mentionnant l'existence d'une coupole. Il précise le positionnement de chacun des trois chœurs. Celui des chevaliers, « en bas », entourant Parsifal, un autre chœur « à mi-hauteur », le troisième tout en haut de la coupole, très loin.

L'ingénieur du son et le directeur artistique d'un enregistrement se doivent-ils de proposer une représentation sonore réaliste de la salle du Graal telle que décrite par Wagner, alors qu'un metteur en scène d'opéra moderne s'en affranchira pour donner une vision personnelle de l'œuvre ? Dans quelle mesure et comment faut-il, avec le disque – donc sans éléments visuels – mettre en perspective les trois chœurs ? Quelques éléments de réponse :

Le premier est qu'il le faut, absolument, et pour des raisons musicales. Tout comme l'interprète, l'ingénieur du son est lié à l'œuvre écrite par un contrat moral. Ce contrat, c'est la partition. Il dit ce qui peut être fait et ce qui ne peut pas l'être, avec des imprécisions, des lacunes, des manques et des libertés. L'important est de comprendre et de respecter le sens général du contrat. Wagner souhaite évidemment une mise en perspective des trois chœurs et un effet de hauteur pour les deux derniers. Le chœur d'en haut doit, en quelque sorte, tomber du ciel, comme la lumière – Wagner le précise. Il est vraisemblable que ce chœur ne soit pas visible du public. À l'ingénieur du son de traduire cela.

Le deuxième élément de réponse est que la salle du Graal n'existe pas – et même si elle existait, il n'est pas indispensable d'en restituer au disque l'acoustique exacte pour respecter l'œuvre¹⁵. On peut déduire que les indications du compositeur concernent autant l'apparence du décor que les questions musicales.

Ma troisième remarque est que, si Wagner précise la disposition des sources sonores, il ne dit rien, et pour cause, sur les positions de l'ingénieur du son, des micros, de l'auditeur final. S'il avait vécu aujourd'hui, il aurait vraisemblablement exprimé des désirs sur le rendu final en stéréo, en multicanal, et peut-être même aurait-il demandé que l'on disposât les enceintes de façon particulière.

L'enregistrement de concertos est souvent l'occasion d'un débat – intérieur ou public – sur la position sonore du soliste par rapport à l'orchestre. Il n'est pas si facile de séparer les impératifs musicaux des questions esthétiques propres à l'image sonore. Telle œuvre, par son écriture, requiert une séparation acoustique marquée entre soliste et orchestre. Cette séparation peut s'exprimer de plusieurs façons : présence (proximité) du soliste, puissance, « clarté » du timbre.

Je pense à une œuvre singulière parce qu'elle n'est ni concerto ni à proprement parler symphonie : *Harold en Italie*, de Berlioz. Cette œuvre écrite pour alto et orchestre place l'alto solo dans un rôle musical original : il est le héros romantique, seul parmi la foule, témoin rêveur et inspiré, mais son discours est comme autonome, indépendant de l'orchestre. Il n'est pas déclamatoire, il est ailleurs et en même temps là. Berlioz l'a voulu ainsi et l'alto, qui n'a pas la brillance du violon, ne trouve pas toujours naturellement la présence requise. On dit que Berlioz indiqua que l'altiste devait se situer plusieurs mètres en avant de l'orchestre. Mais si l'on suit l'indication au pied de la lettre, le rendu sonore peut être déroutant... L'œuvre fut créée dans la petite salle du Conservatoire de Paris – aujourd'hui Conservatoire d'art dramatique – et on voit mal comment, physiquement, cette disposition a pu être réalisée. À l'équipe

d'enregistrement d'imaginer, d'interpréter, et de respecter l'esprit des indications musicales de Wagner ou de Berlioz.

Interprétation musicale et prise de son.

L'évolution esthétique de la prise de son classique au cours des quarante dernières années montre comment l'on peut – dans les cas réussis – mettre en valeur et respecter le discours musical sans pour autant coller à une esthétique sonore naturaliste qui consisterait à installer l'auditeur à la « meilleure » place d'une salle de concert.

Je me souviens avoir assisté à un concert en l'église Saint-Eustache de Paris où Mstislav Rostropovitch interprétait les *Suites pour violoncelle seul* de J.S. Bach. Malgré l'installation d'une conque pour focaliser le son, il était extrêmement difficile de percevoir clairement l'instrument et les nuances du jeu au-delà du cinquième rang de chaises (l'église était cependant pleine, et le public enthousiaste). Quelques années plus tard, ce grand musicien enregistra les *Suites* pour le CD et le DVD à l'abbaye de Vézelay. Ce lieu était son choix, il en aimait l'atmosphère, la beauté, le calme. Dans les deux cas, il s'agit de lieux peu adaptés au jeu d'un violoncelle seul, et aucune donnée musicologique ne peut inciter au choix d'une grande église.

La prise de son nous propose un violoncelle (pas n'importe lequel, car ce Stradivarius, le « Duport », est exceptionnel) présent, défini, placé dans l'acoustique généreuse de Vézelay : c'est une mise en perspective réussie d'un superbe instrument dont le son assez présent nous procure un plaisir véritablement sensuel et d'un lieu inspirant et mystique. D'autres interprètes, plus proches de l'esthétique baroque, rejetteraient ce choix acoustique, n'y trouvant ni un support à leur interprétation ni une pertinence musicologique.

Pop music et record-business.

Rendons à César ce qui lui revient : c'est probablement à la pop music, à la formidable inventivité des groupes et des producteurs des années 1960 (on pense bien sûr aux Beatles), que la prise de son classique doit le courage iconoclaste d'avoir dépassé l'image sonore naturaliste... Avec des moyens techniques fort limités en comparaison des outils actuels, une nouvelle génération inventa le plaisir du son de proximité, l'efficacité musicale d'un effet sonore bien réalisé, l'intérêt esthétique des changements de plan sonore au gré des arrangements instrumentaux.

Mais si les plus brillants inventèrent cette nouvelle perspective sonore variable, dynamique, aux axes multiples, d'autres, moins inspirés, se contentèrent de placer au premier plan toutes les sources sonores et d'enrober l'ensemble d'un fond de sauce réverbérant. Écoutez ou réécoutez certains tubes des années 1970 et 1980 !

Il ne faut pas non plus négliger le pouvoir économique de telle ou telle star – celle qui fait vendre – qui, assistant au mixage, exige quelques décibels supplémentaires sur sa voix ou son instrument. Certains acteurs de cinéma spécifient par contrat, paraît-il, les détails des plans où ils apparaissent, imposent leur bon profil... Il en est de même, tacitement ou pas, de certains musiciens. La perspective sonore prend alors un relief inattendu. Mais cela n'est pas vraiment nouveau : faut-il rappeler que la crise iconoclaste qui déchira le monde chrétien ne se nourrissait pas que de motivations théologiques ? L'histoire de la perspective s'en trouva bouleversée.

La question du « prêt-à-entendre ».

Ainsi, les progrès technologiques de l'après-guerre apportèrent une remarquable amélioration de la qualité des systèmes de reproduction sonore, et la possibilité d'écouter des programmes dans tous les lieux de vie : dans le salon et en voiture, en plein air et dans l'ascenseur. Un nouveau marché de masse était né. Et avec lui de nouvelles façons de produire le son.

Paradoxalement, la « hi-fi » permettait aux professionnels de proposer une perspective sonore riche et raffinée, tandis que les conditions de « consommation » du son (multiplicité des lieux et des contextes) incitaient à offrir un son consommable en toute circonstance et, par conséquent, simplifié, appauvri, afin que le son considéré comme « utile » reste toujours audible. Ainsi, lorsque j'ai débuté – c'étaient les dernières années du microsillon et de la musicassette –, on réalisait deux bandes master différentes pour le disque – produit noble destiné à l'écoute domestique – et pour la cassette – essentiellement destinée à l'écoute nomade : voiture, promenade, etc. Aujourd'hui, la situation est encore plus paradoxale : on peut écouter un CD ou visionner un DVD dans un home-theater ou sur un coin de table, choisir de n'utiliser qu'Internet et un laptop et télécharger musique et films. Les conditions d'écoute sont innombrables, les formats de diffusion multiples.

Les formats de diffusion sonore.

Considérons la grande variété des caractéristiques techniques des formats et supports de diffusion actuels : mono (configuration plutôt accidentelle aujourd'hui), stéréo (c'est le « prêt-à-diffuser » le plus courant), multicanal 5.1 (l'avenir très proche), la haute définition (SACD, DVD audio), le son à dynamique compressée (radio et TV), les compressions de débit (MP3, dolby digital, DTS, Internet en général) : comment réaliser une perspective sonore, indispensable musicalement, de telle sorte qu'elle puisse « fonctionner » dans tous les formats de diffusion et dans la plupart des situations d'écoute ?

Nous avons expliqué plus haut comment notre alto solo, toujours intelligible en reproduction 5.1, risque de subir des masquages dans une restitution réduite en stéréophonie. Il faudrait donc mixer l'œuvre différemment – et même la capter différemment – pour obtenir l'intelligibilité souhaitée du soliste, et un positionnement perspectif satisfaisant. Et il ne s'agit pas alors de réaliser une image sonore réaliste. Notre problème est d'ordre musical. Mais peu de labels ou de producteurs audiovisuels peuvent aujourd'hui financer une double captation microphonique et/ou un double travail de mixage. Une pratique courante consiste à réaliser un « down-mix » stéréo de la version 5.1, méthode qui donne des résultats acceptables dans la plupart des cas. Mais « acceptable » est-il acceptable, s'agissant d'art et d'artistes ? Pire, il est maintenant courant de fabriquer du 5.1 – *via* quelques algorithmes astucieux – à partir d'une version stéréo : on appelle cela l'« up-mix »... C'est de cette façon que la plupart des chaînes de télévision se préparent à gérer les « flux » sonores en régie finale. Les chaînes nous donneront donc à entendre de la mono up mixée, de la stéréo up mixée, et de temps en temps peut-être du 5.1 original... Pauvre Berlioz, pauvre Wagner ! Et pauvre perspective sonore !

Dynamique sonore et spécifications des chaînes de télévision.

La « dynamique sonore » désigne le rapport entre le niveau le plus fort d'un événement sonore et son niveau le plus faible. Pour un son enregistré, il pourra s'agir de niveaux électriques ou bien, dans le domaine du son numérique, de quantification (nombre de « bits »). L'ouïe accepte, dans la zone fréquentielle la plus sensible (aux alentours de 2 000 hertz), des écarts absolus de 120 décibels. Numérisé (quantifié), un tel écart requiert environ 21 bits théoriques.

Les télévisions, afin d'uniformiser les niveaux sonores des différents programmes et dans le souci de garantir aux auditeurs des conditions d'écoute confortables, même dans un environnement bruyant, exigent des producteurs de programmes une dynamique maximale de l'ordre de 20 décibels. Si, afin de satisfaire à cette demande, on ne conservait que les 20 décibels supérieurs dans une œuvre symphonique classique, il y aurait de longs moments de silence ! Mais on peut aussi décider de « raboter » les niveaux les plus forts, de pousser les niveaux les plus faibles, pour que le tout rentre dans la boîte. C'est un peu tout cela que l'on fait, avec, au final, une sérieuse altération de la dynamique musicale d'une œuvre. Et c'est précisément la perspective sonore qui s'en trouve le plus altérée. Revenons à nos trois chœurs de *Parsifal* : on les entendra, après compression, pratiquement tous au même niveau. La réverbération, artificiellement renforcée par la compression, noiera le chœur le plus lointain dans une boue épaisse et masquante.

Le pire, c'est que, loin d'obtenir par ces spécifications de réduction de dynamique une homogénéisation des niveaux sonores, le spectateur-auditeur doit constamment ajuster son niveau d'écoute, de la publicité aux films, des infos aux programmes musicaux. On assiste à une guerre des niveaux subjectifs. Certains programmes, et notamment la publicité, vont bien au-delà des spécifications des chaînes, en écrasant la dynamique sonore (à peine quelques décibels entre niveaux minimal et maximal), en éliminant tous les sons qui risqueraient d'être masquants, en ne conservant que le son « essentiel ». Plus aucune perspective sonore, au mieux deux plans : parole devant, musique ou bruitage derrière ! C'est prêt à consommer.

Par contraste, le son plus subtil d'un concert classique ou d'un film semblera lointain et terne. On ressent qu'il faut faire un effort d'attention pour percevoir certaines phrases ou certains sons *off*, car tous les éléments sonores ne sont pas au même niveau. On pousse le volume du téléviseur. Rien n'y fait... Un « zapping » tonitruant sur les chaînes voisines achèvera-t-il de dissuader l'auditeur ?

SON UNIQUE, PENSÉE UNIQUE... ADIEU LA PERSPECTIVE ?

Tel est, en résumé, l'état des choses – ne dit-on pas « état de l'art » ?

D'un côté, une technologie de plus en plus sophistiquée qui permet de créer des espaces sonores complexes et complets, si riches que l'auditeur retrouve enfin cette caractéristique de l'audition humaine dans le monde réel : pouvoir choisir parmi tous les sons proposés à l'écoute ceux qui font

sens pour lui à un moment donné, laisser des associations se former, singulières, personnelles à chacun. De l'autre côté, la multiplicité des supports de diffusion et des conditions d'écoute, les enjeux de communication considérables dont l'audiovisuel est devenu le vecteur incontournable, qui tendent à appauvrir dramatiquement le contenu sonore délivré par les médias. *In fine*, c'est l'esthétique du son, quels que soient les supports et les conditions d'écoute, qui en pâtit. Entre un son simplifié à l'extrême et un son complexe et riche de multiples sensations potentielles, comment l'auditeur fait-il son choix ? Évolution, révolution de la perspective sonore ? Certainement. Disparition ? Je ne le crois pas. Le travail sur la perspective sonore reste une démarche esthétique en constante évolution, dont la définition même est sujette à débat, mais qui, parce qu'elle fait sens, participe du principe de plaisir.

Jean CHATAURET
jean.chatauret@soundways.fr
ingénieur du son
enseignant à l'École nationale supérieure Louis-Lumière

NOTES

1. André Charlin inventa dès 1954 sa « tête artificielle », et diverses radios européennes contribuèrent à l'avènement des couples dits « XY » et « A-B ». L'un des plus célèbres dispositifs est le « couple ORTF », constitué de deux capsules cardioïdes (directives) espacées de dix-sept centimètres et formant un angle de cent dix degrés.

2. En géométrie, « solide » qualifie les volumes. L'intention de reproduire le « relief » sonore n'est pas liée uniquement à la stéréophonie : toutes les techniques de prise de son, de la monophonie des débuts au son « 5.1 » surround actuel, s'en préoccupent. Mais le « réalisme de la scène sonore » était un critère majeur pour les pionniers de la stéréophonie.

3. Il ne s'agit pas ici de jouer avec les mots : comment nommer un son ? On a généralement recours à une définition « causale » qui, hors contexte, ne permet pas d'imaginer précisément le son en question.

4. Son que nous évoquerons vraisemblablement, dans les deux cas, par l'onomatopée « plouf ! ».

5. En fait, il n'est pas possible que ce son, capté au bord de l'eau, ne soit que celui de la pierre heurtant l'eau. Il y a inévitablement un ensemble de micro-informations qui évoquent un contexte : clapotis de l'eau, bruissements de feuilles, chants d'oiseaux, pas sur l'herbe, cris d'enfants...

6. Cette onde directe nous parvient toujours la première, car elle suit le trajet le plus court. Son antériorité joue un rôle déterminant dans notre capacité à localiser une source sonore. On appelle cela l'« effet Haas », ou « effet de précedence ».

7. Ces phénomènes perceptifs, que nous abordons ici de façon superficielle, ont été étudiés par des chercheurs comme Michel Chion (*Le Son*, Paris, Armand Colin, 2004) et Claude Bailblé (*La Perception et l'Attention modifiées par le dispositif cinéma*, thèse de doctorat, Paris VIII, 1999 ; « Le son documenté », *La Revue documentaires*, n° 21, 2007).

8. La norme prévoit en effet que l'auditeur soit placé à l'un des sommets d'un triangle équilatéral, les deux enceintes étant situées aux deux autres sommets.

9. Daniel Deshays, *Pour une écriture du son*, Paris, Klincksieck, 2006, p. 79-82.

10. Certains labels, comme Dabringhaus & Grimm, éditent des DVD audio prévus pour une disposition en hauteur des enceintes acoustiques.

11. Les techniques de spatialisation permettent aujourd'hui d'étendre le champ d'écoute au-delà des enceintes acoustiques, aussi bien horizontalement que verticalement.

12. Ce phénomène n'est pas systématique. S'agissant par exemple d'une harpe, le parfait synchronisme entre les doigts de l'instrumentiste et les sons produits par l'instrument rendent le phénomène évident. S'agissant d'un cor, dont les sons ne sont pas corrélés strictement, pour le profane, aux gestes visibles du corniste, l'effet sera peut-être moins clair. La connaissance que l'auditeur-spectateur possède des timbres des instruments est un facteur complémentaire qui renforce l'effet d'intelligibilité.

13. Daniel Deshays, *Pour une écriture du son*, *op. cit.*, p. 95.

14. L'essentiel de la musique que nous écoutons est fondé sur la répétition, le développement, la reprise transformée, inversée, transposée, d'une cellule musicale – cellule mélodique, harmonique, rythmique... Il est donc là encore question de mise en perspective. Mais cela n'est pas à proprement parler notre sujet, qui porte sur la représentation sonore, plus que sur la musique elle-même.

15. Il est à noter cependant que, dans le cas précis de *Parsifal*, Wagner avait, dans son testament, exigé qu'il ne fût interprété qu'à Bayreuth. Cette volonté fut respectée pendant un certain nombre d'années après la mort du compositeur, puis transgressée.

RÉSUMÉ

La perspective sonore, le « point d'écoute », ou « point d'ouïe », sont pour le preneur de son des concepts familiers, comme le « point de vue » l'est au photographe, et à l'« honnête homme »...

Partant de l'expérience perceptive de ce phénomène, le texte tente de dégager, en s'appuyant sur le vécu d'un ingénieur du son spécialisé dans la musique classique, les tendances esthétiques actuelles de l'« art » de capter et de reproduire le sonore.