

La localisation auditive des sons dans l'espace

Georges Canévet

Laboratoire de Mécanique et d'Acoustique, CNRS
31, chemin Joseph-Aiguier
13402 Marseille cedex 20

Résumé : La perception auditive de l'espace résulte du décodage en temps réel ou faiblement différé d'indices acoustiques spécifiques contenus dans les sons. Ces indices sont essentiellement les différences interaurales de champ acoustique et des modifications spectrales causées par la diffraction des sons autour de l'observateur. L'analyse de ces indices montre toutefois qu'interviennent aussi des mécanismes perceptifs plus complexes, non encore élucidés. Ceci apparaît notamment en psychoacoustique des salles et dans les problèmes de simulation par écouteurs de l'espace sonore subjectif.

1. Introduction

La localisation des sources sonores est une des fonctions fondamentales du système auditif. Elle est assurée par une série de structures, qui vont de l'oreille externe aux aires auditives corticales. On en trouvera la présentation dans ce même ouvrage [Barone 1995] ou dans les monographies des éditions INSERM/SFA ([Dancer 1988] et [Rouiller 1992] par exemple). Ces structures opèrent un décodage particulier du champ acoustique qui atteint les tympans. Elles procèdent tout d'abord à un traitement approprié du signal acoustique, notamment la traduction des ondes mécaniques en activité électrique neurale. Puis, les réponses auditives ainsi engendrées sont intégrées aux autres informations sensorielles et cognitives disponibles, en vue de la construction de notre perception de l'espace sonore.

Les caractéristiques du champ acoustique pertinentes pour la localisation auditive sont souvent appelées *indices*. Les mieux connus sont les différences interaurales de niveau, de phase et de temps ; ils sont utilisés pour évaluer la position latérale des sons par rapport à l'observateur. Il en est d'autres, plus complexes et plus difficiles à cerner, qui apparaissent essentiellement comme des modifications spectrales spécifiques des signaux incidents. Ces indices spectraux servent en particulier à la localisation dans le plan vertical médian. Autrement dit, on peut admettre en première approximation que la localisation dans une direction quelconque de l'espace résulte de l'exploitation combinée de ces deux types d'indices. Ajoutons enfin que dans la perception auditive spatiale il intervient aussi des mécanismes dits "cognitifs" ; mais, à l'heure actuelle, on connaît peu de chose de ces mécanismes.

Dans cet article, nous présentons succinctement la question de la localisation d'une source unique, puis nous abordons celle de la localisation de sources multiples. La première catégorie de problèmes permettra d'analyser le détail des aspects

psychoacoustiques purs de la localisation auditive. La seconde catégorie sera abordée plutôt dans le but de présenter quelques faits avérés de la perception des environnements sonores réels.

2. Les différences interaurales du champ acoustique

Pour simplifier les choses, nos prédécesseurs ont tout d'abord examiné la question de la localisation dans le plan horizontal, c'est-à-dire de l'identification d'azimut d'une source acoustique unique. Ils ont établi que cette identification d'azimut se faisait par examen des différences de niveau créées par la source sur les deux oreilles, et des différences de phase ou de temps d'arrivée des signaux. Sans entrer dans le détail de l'analyse bibliographique, que l'on trouvera dans [Blauert 1983] ou [Canévet 1989], on se doit de citer les principaux auteurs. Les travaux les plus anciens [Venturi 1796] ont démontré le rôle joué par les différences interaurales d'intensité. Ensuite, Rayleigh, en 1877, a proposé une théorie plus élaborée [Rayleigh 1877], où il fait notamment remarquer que les différences d'intensité seules ne peuvent tout expliquer. Il faut faire selon lui appel à une autre grandeur physique : la différence interaurale de phase. Dans sa théorie, Rayleigh a précisé les intervalles de fréquence dans lesquels peuvent agir respectivement la phase et l'intensité. La phase est utile dans le grave, et jusqu'à environ 1500 Hz ; l'intensité prend le relais pour le médium et l'aigu, au-dessus de 1500 Hz. Les recherches qui ont suivi n'ont fait que confirmer et affiner les prévisions théoriques de Rayleigh.

En ce qui concerne les différences interaurales temporelles, il faut tout de même préciser qu'elles sont de deux types. Lorsqu'il s'agit de sons purs entretenus suffisamment longs, la différence de phase créée par diffraction autour de la tête permet l'identification de l'azimut de la source, dans les limites de fréquences indiquées ci-dessus. Mais lorsqu'il s'agit d'impulsions, ou de trains d'ondes brefs, cet indice disparaît. Le système auditif doit donc évaluer les différences instantanées des enveloppes des signaux à gauche et à droite, pour en déduire l'azimut. Cet indice temporel est valable sur tout le spectre audible.

Les valeurs exactes des différences interaurales de temps sont délicates à déterminer ; elles dépendent des caractéristiques temporelles précises des signaux acoustiques considérés [Kuhn 1977 et 1987]. Si l'on accepte la simplification de l'acoustique géométrique, on peut toutefois en déterminer une valeur théorique approchée. Le modèle géométrique le plus connu est celui de Woodworth (1938). Dans ce modèle, la tête de l'observateur est schématisée par une sphère, et la différence interaurale de temps d'arrivée à deux points diamétralement opposés se déduit de la différence de marche (dT) de deux rayons sonores provenant de l'infini avec une incidence α . La formule résultante s'écrit $dT = r/c(a + \sin \alpha)$, r étant le rayon de la sphère et c la vitesse du son. Par ailleurs, des résultats de mesures sur des impulsions ont été publiés par Feddersen et ses collaborateurs [Feddersen 1957]. Sur la figure 1, nous avons porté la courbe représentative de la fonction de Woodworth, ainsi que quelques-uns des points expérimentaux de

Feddersen. On notera que le maximum du retard interaural naturel est voisin de sept cents microsecondes. En introduisant artificiellement par écouteurs des retards supérieurs, on peut créer des effets perceptifs intéressants, qui n'ont pas encore été étudiés.

Les différences interaurales de phase ont été mesurées et publiées à de nombreuses reprises (voir [Mills 1958]). Leur évaluation théorique est également possible, en utilisant le modèle de la sphère, mais en prenant en compte les effets de diffraction (voir [Canévet 1985] ou [Canévet 1989]). La figure 2 présente les résultats théoriques obtenus (traits continus) ainsi qu'un certain nombre de points expérimentaux empruntés à Mills. La manipulation séparée de cet indice peut conduire à des situations "conflictuelles" du point de vue de la perception (voir aussi § 7 plus loin).

Quant aux différences interaurales de niveau, elle sont éminemment variables avec la fréquence et avec l'azimut. Des catalogues entiers ont été publiés, et nous recommanderons notamment celui d'Abbagnaro et ses coauteurs ([Abbagnaro 1975]). Les courbes donnant la surpression sur les oreilles, par rapport au champ incident, en fonction de la fréquence sont typiques. On reviendra sur ce point au § 4 (Fig. 4) plus loin. En général, la surpression est voisine de 0 dB en basse fréquence, c'est-à-dire que l'auditeur perturbe peu le champ acoustique, du fait que la longueur d'onde est grande par rapport aux dimensions de la tête humaine. La surpression augmente ensuite avec la fréquence jusqu'aux environs de 3 à 4 kHz, puis décroît pour atteindre un minimum d'antirésonance au voisinage de 8 kHz, et enfin remonte à nouveau. Dans l'extrême aigu, les variations de niveau deviennent en fait très complexes, à cause des effets de diffraction, notamment dus au pavillon.

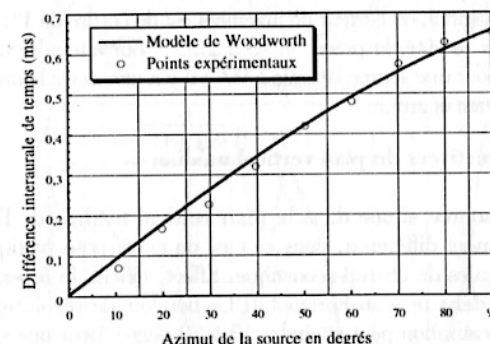


Figure 1 : Différences interaurales de temps d'arrivée d'impulsions aux oreilles, mesurées [Feddersen 1957] et calculées à partir du modèle de Woodworth [Canévet 1989], en fonction de la position latérale de la source.

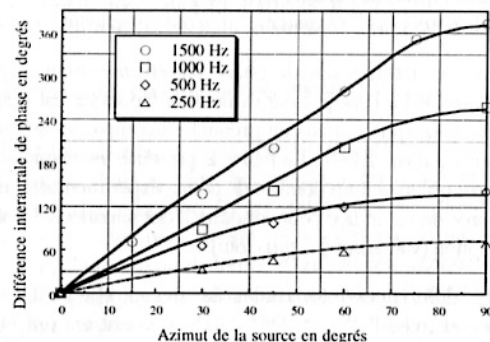


Figure 2 : Différences interaurales de phases en fonction de l'azimut de la source. Valeurs calculées (traits pleins, [Canévet 1985]) et valeurs mesurées (symboles, [Mills 1958]).

Muni de ces indices de niveau, de phase et de temps, l'auditeur peut déterminer la direction des sons, grâce à des structures nerveuses spécifiques du système auditif [Aschoff 1992], qu'il utilise à la manière de corrélateurs. La détermination se fait avec une bonne précision. Toutes les recherches effectuées sur cette question des capacités de localisation dans le plan horizontal l'attestent. La plus récente est peut-être celle de Middlebrooks ([Makous 1990], [Middlebrooks 1992]). On peut résumer les faits de la manière suivante : la précision de la localisation est maximale pour une source de face ou proche du plan sagittal, où l'erreur de jugement est de l'ordre de 1° ; à mesure que la source s'éloigne sur le côté, la précision se dégrade pour atteindre une erreur d'une dizaine de degrés pour une source latérale à 90° ; il y a une assez bonne symétrie entre les performances avant et arrière.

3. Les bandes directives du plan vertical médian

Pour une source située dans le plan vertical médian de l'observateur, la situation est totalement différente ; dans ce cas, on ne dispose pratiquement plus de différences interaurales du champ acoustique. Effectivement, le repérage spatial y est plus difficile que dans le plan horizontal. Le pouvoir de résolution est faible, et l'incertitude à la localisation peut atteindre 15 à 20 degrés pour une source au-dessus de la tête ([Damaske 1969]).

En fait on s'est aperçu que, dans le plan vertical, la direction perçue ne correspondait pas forcément à la direction d'incidence réelle des sons. Plus précisément, on a remarqué que le système auditif semblait localiser la source dans une direction qui est imposée par le spectre de son signal. La démonstration en a été faite par Blauert

(1969/70). Selon les réponses des auditeurs qu'il a testés, les sons semblent presque toujours provenir de l'avant si leur spectre appartient aux bandes 250-500 Hz ou 2-4 kHz. Ils semblent au contraire provenir de l'arrière si leur spectre se répartit dans les

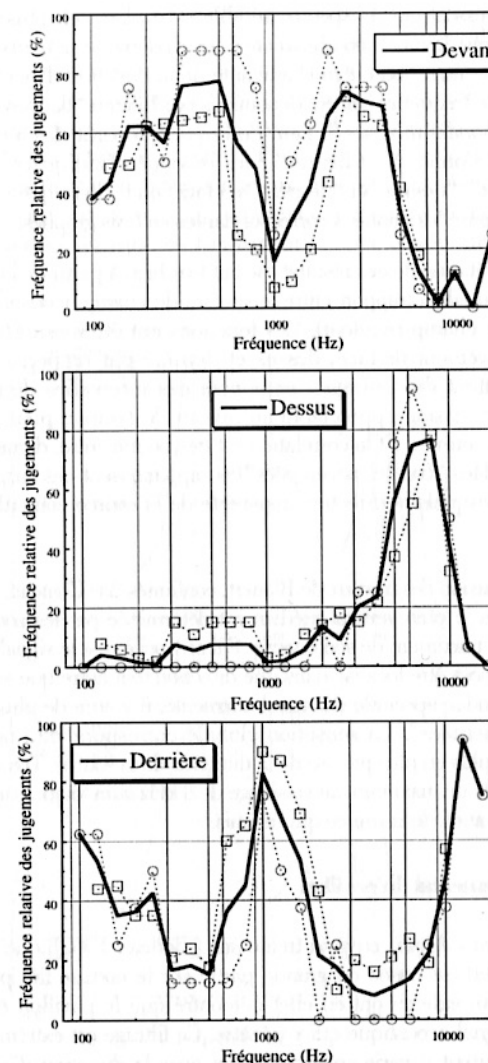


Figure 3 : Direction apparente de sons présentés dans le plan vertical médian. Les points de mesure sont empruntés à [Blauert 1969/70] pour les symboles carrés et à [Chateau 1995a] pour les symboles ronds. Les traits continus représentent les moyennes de deux études.

bandes 700-2000 Hz et 10-15 kHz. Enfin, seule la région des 8 kHz peut procurer l'impression qu'un son vient de dessus. Ce partage en trois zones est arbitraire, et on doit le considérer comme une première approximation. Des mesures récentes de Chateau ([Chateau 1995 a et b]) ont toutefois confirmé la teneur de ces résultats et les ont même précisés. Il en ressort que le spectre audible est partagé en plusieurs domaines, qui chacun conditionne la direction apparente d'un son lorsque sa fréquence appartient à ce domaine ; et cette direction est indépendante de la direction d'incidence réelle du son. Ces intervalles de fréquence ont été dénommés par Blauert "bandes déterminantes de la direction", ou "*bandes directives*". Leur description, à partir des résultats de Blauert et de Chateau, est donnée à la figure 3 : les ordonnées indiquent le pourcentage des réponses "devant", "dessus" et "derrière" en fonction de la fréquence ; pour des raisons de clarté on a séparé les données correspondantes en trois graphes.

L'interprétation de ces résultats se fait très bien à partir de la notion de fonctions de transfert des oreilles (rapport entre le spectre du champ acoustique sur les tympans et le spectre du champ incident). Ces fonctions ont été mesurées pour des sources situées respectivement de face, dessus et derrière par rapport à l'observateur. Les mesures permettent de distinguer nettement des intervalles de fréquences où l'on observe une surpression appréciable au niveau du tympan pour certaines positions particulières des sources. Et la corrélation est très bonne entre, d'une part les intervalles de fréquences (dites "bandes accentuées") où apparaissent ces surpressions, et d'autre part ceux pour lesquels la direction apparente de la source coïncide avec sa direction réelle..

La conclusion des travaux de Blauert, confirmés par Chateau, est que la sensation de direction dans le plan vertical médian est déterminée par les zones spectrales où le signal possède le maximum de son énergie. S'il se trouve que le signal est un son pur ou à spectre étroit, il peut être localisé dans une direction tout autre que sa direction réelle. A mesure que l'étendue spectrale du signal augmente, il y aura de plus en plus de bandes directives "alimentées". La sensation globale correspondante permettra alors une identification de plus en plus précise de la direction de la source. Toutefois, un signal dont le spectre présente un maximum au voisinage de 8 kHz aura toujours tendance à créer une impression d'élévation de la source qui l'a émis.

4. Le rôle fondamental du pavillon

On a admis depuis environ trente ans ([Batteau 1967]) que la localisation hors du plan horizontal est basée en grande partie sur le codage fait par les pavillons de l'oreille. Plusieurs auteurs ont en effet démontré que le pavillon exerçait un effet de filtrage sur le signal acoustique qui y pénètre. Ce filtrage est extrêmement complexe et particulier, et surtout il varie spécifiquement avec la direction d'incidence du son. Il intervient si la fréquence du signal est suffisamment élevée, disons à partir de 2 kHz, lorsque la longueur d'onde est suffisamment petite en comparaison des dimensions du

pavillon. Ceci a été parfaitement démontré par les mesures de Shaw et Teranishi ([Shaw 1968]). La figure 4 donne une présentation schématisée des résultats de ces auteurs, pour deux incidences latérales différentes : l'une à l'horizontale, l'autre à la verticale.

Comment le système auditif interprète-t-il ce filtrage ? On peut adopter deux approches. L'une consiste à considérer, comme Shaw et Teranishi ([Shaw 1968]), le pavillon et le conduit auditif comme un ensemble de résonateurs. On voit bien, d'après leurs résultats (cf. Fig. 4), que l'oreille externe est le siège de résonances multiples dont la première fréquence propre apparaît aux environs de 3 kHz. Les suivantes se situent entre 7 et 15 kHz, suivant la direction de la source. Par conséquent, le dosage spécifique des différentes résonances produit une résultante qui dépend de la direction du son. Ce dosage des résonances, variable avec l'incidence, constituerait une sorte de clef permettant la localisation (l'analogie est facile entre la forme des spectres de la Fig. 4 et celle d'une clef).

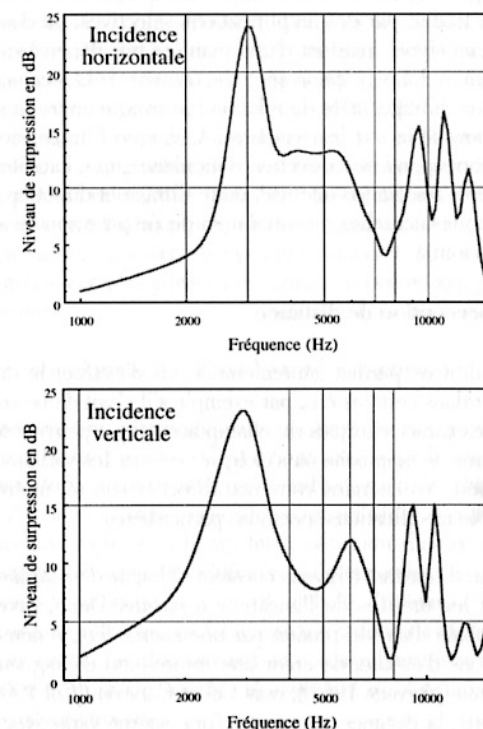


Figure 4 : Fonctions de transfert en amplitude du pavillon de l'oreille, pour une source latérale située dans le plan horizontal (graphe supérieur) ou dans le plan vertical. Représentation schématisée, déduite des données de Shaw et Teranishi ([Shaw 1968]).

L'autre approche a été proposée d'abord par Batteau en 1967, et reprise par Wright et al. ([Wright 1974]) ou par Watkins (1978). Cette seconde approche considère le pavillon comme un système de réflecteurs multiples. Selon ce point de vue, les différents plis du pavillon renvoient une partie du signal incident vers le conduit auditif. Chacune de ces réflexions vient s'ajouter au signal primaire incident avec un certain retard. Or ces retards sont fonction de l'incidence (entre 2 et 80 microsecondes pour l'azimut, entre 100 et 300 microsecondes pour la hauteur). La combinaison des ondes directes et réfléchies serait caractéristique de la valeur du retard, donc de l'incidence. Lorsque les retards sont brefs, le signal retardé introduit une atténuation des fréquences aiguës. A mesure que les retards s'allongent le spectre du signal incident se trouve modulé de plusieurs maximums et minimums (comme par filtrage en peigne). Il se trouve que ces déformations spectrales ne sont pas perçues comme des modifications de timbre, mais comme des changements de direction d'un même signal.

En résumé, l'effet du pavillon sur un signal acoustique est équivalent à celui d'un filtre. Cet effet se traduit par des amplifications sélectives, ou des atténuations spectrales, appliquées au signal incident d'une manière qui dépend spécifiquement de la direction de la source. La perception spatiale, au sens de la localisation, dépend alors des variations de ce filtrage, sorte de relation biunivoque entre la position de la source et le champ acoustique sur les tympans. A chaque filtrage, donc à chaque fonction de transfert, correspond une direction d'incidence, que l'auditeur a appris à reconnaître. Et c'est cette association acquise, entre filtrage et direction, qui est mise à profit dans les techniques actuelles de simulation de l'espace sonore subjectif sous écouteurs (voir § 8 plus loin).

5. Les indices de la perception de distance

Le système auditif ne permet généralement pas d'évaluer la distance d'une source. Il peut y arriver dans certains cas, par exemple s'il s'agit d'une source connue, ou "familière". Certaines caractéristiques du champ acoustique peuvent toutefois créer la sensation que la source se rapproche ou s'éloigne : ce sont les variations d'intensité, les variations du rapport des énergies entre son direct et son réverbéré lorsqu'il est présent, et enfin certaines modifications spectrales particulières.

Lorsqu'une source émettant à niveau constant s'éloigne d'un auditeur, l'intensité acoustique atteignant les oreilles de l'auditeur diminue. Donc, inversement, en réduisant l'intensité globale d'un son produit par une source fixe, il doit être possible, dans certaines conditions d'écoute, de créer une impression d'éloignement de cette source. Stevens et Guirao ([Stevens 1962]), puis Col et Canévet ([Col 1990]) ont étudié cette question. Selon eux, la distance apparente d'une source varie avec le niveau du son comme une fonction puissance à exposant négatif ; autrement dit, le logarithme de la distance apparente décroît linéairement lorsque le niveau augmente. Mais ces études seraient à approfondir.

Le second indice relève de l'expérience quotidienne : lorsque dans une salle un locuteur s'éloigne, le rapport du son direct au son réverbéré décroît. Par conséquent, on a coutume d'associer cette décroissance relative du son direct à un éloignement de la source. De ce fait, elle devient un indice pour l'évaluation de la distance. Cet effet est connu depuis longtemps, mais il a surtout été examiné récemment par Mershon et King ([Mershon 1975]), qui voulaient en particulier faire la part de la réverbération et de la diminution d'intensité du son direct. Leurs résultats peuvent se résumer de la manière suivante : les différences d'intensité, prises isolément ne permettent pas d'estimer la distance, mais elles sont utilisables pour en apprécier assez finement les variations. Au contraire, la réverbération semble avoir une qualité d'indice absolu, qui permet le positionnement d'une source surtout dans l'espace proche de l'individu. Plus la réverbération est élevée, plus la distance perçue est grande.

Enfin, chacun sait que la densité spectrale d'un signal acoustique varie au cours de sa propagation, par absorption inégale des graves et des aiguës. Ainsi, un sujet est capable, après un certain entraînement, d'évaluer les distances relatives d'émission avec une précision acceptable. Inversement on peut, par filtrage approprié du signal, créer la sensation d'éloignement ou de rapprochement d'une source immobile. C'est ce qu'ont fait notamment Butler et al. ([Butler 1980]). En manipulant la composition spectrale de signaux, ils ont réussi à donner à des auditeurs l'impression que leur distance apparente variait ; notamment, les bruits filtrés passe-bas semblent beaucoup plus éloignés que les bruits filtrés passe-haut. Il est intéressant de noter également dans leurs expériences que l'azimut perçu variait aussi un peu avec les caractéristiques spectrales. Les signaux à spectre grave ont tendance à être localisés vers l'arrière, et d'autant plus qu'ils sont plus graves. Les signaux à spectre aigu par contre sont localisés plutôt correctement.

6. Effet stéréophonique, effet d'antériorité, adaptation

La présentation effectuée dans les paragraphes précédents concernait la localisation d'une source unique. C'est à notre avis la meilleure approche pour décrire les mécanismes de base de la perception auditive spatiale. Dans la réalité, il est plus fréquent que l'on ait à localiser une source en présence d'autres sources, ou dans le bruit, ou dans une salle. Le champ total, résultant du rayonnement de plusieurs sources, va aussi engendrer des différences interaurales de temps et d'intensité, et donc créer une sensation de localisation globale. Toutefois, l'azimut apparent de la source résultante ne sera pas, sauf exception, confondu avec celui d'une des sources réelles qui l'ont créée. Le plus souvent, le son semblera provenir d'une direction de l'espace où il n'y a pas de source, c'est-à-dire qu'il proviendra de ce que nous appelons une source virtuelle.

L'effet stéréophonique est l'exemple le mieux connu de sources acoustiques virtuelles. On sait, depuis de nombreuses années, déplacer l'image sonore créée par deux

haut-parleurs en agissant sur les différences de niveau des signaux. Une présentation générale de ces techniques et de leurs effets a été faite par Gardner (1968). En ce qui concerne les différences de niveau, une étude complète a été faite par Leakey (1959), utilisant deux haut-parleurs situés symétriquement par rapport à l'observateur, dans la position habituelle de l'écoute stéréophonique. Il a montré que l'azimut apparent passe continuellement du haut-parleur de gauche à celui de droite lorsque la différence de leurs niveaux passe environ de -15 dB (gauche plus fort) à +15 dB (droite plus fort). Une différence de phase entre les haut-parleurs peut aussi déplacer l'image résultante. En particulier, par manipulation correcte des alimentations des sources, on peut donner l'impression de sources en mouvement ([Chowning 1971]). Une étude théorique de ces effets, basés sur des calculs de diffraction par une sphère, a été faite par Canévet (1985).

Lorsque les sources réelles génératrices de sources virtuelles ne fonctionnent pas simultanément, des phénomènes spéciaux peuvent se produire, et notamment un effet dit d'antériorité (parfois appelé effet de Haas, voir [Haas 1951]), qui est l'un des effets fondamentaux de l'écoute stéréophonique. Si l'on écoute un enregistrement stéréophonique en se plaçant à égale distance des haut-parleurs, on a la sensation que le son provient d'une source centrée entre les haut-parleurs. Par contre, si l'on se rapproche de l'un des haut-parleurs, le signal arrive plus tôt aux oreilles dans sa direction que dans l'autre, et on a la sensation que le son provient uniquement de ce haut-parleur proche. Ceci illustre ce que l'on appelle l'effet d'antériorité. Le principe de l'effet d'antériorité est donc que, lorsqu'un même signal est émis successivement par deux ou plusieurs sources, avec un léger décalage temporel entre les émissions, l'oreille fusionne l'ensemble, en un signal qu'elle localise généralement dans la direction de la source alimentée la première.

L'effet d'antériorité n'apparaît toutefois que sur les transitoires (percussions, instruments à cordes et parole par exemple). Il est inexistant avec des sons purs continus, excepté à la mise en route. De plus, il dépend du décalage temporel entre les sources. En règle générale, pour un couple de sources situées symétriquement par rapport à un observateur, la source virtuelle passe du centre au haut-parleur de côté lorsque le décalage varie de 0 à 1 ms environ. Elle reste confondue avec ce haut-parleur pour des retards entre signaux de 1 à 50 ms, suivant le type de signaux. Au-delà de 50 ms la séparation se fait de plus en plus nette. Elle est totale pour un retard de 70 à 100 ms, c'est-à-dire que l'on perçoit alors deux sons successifs, provenant des haut-parleurs respectifs.

Un autre phénomène intéressant apparaît, dans certaines conditions, pour des retards encore plus longs : l'adaptation à moyen terme ([Canévet 1995]). L'adaptation modifie la direction apparente d'un son, lorsque l'auditeur a été préalablement exposé à un son de même caractéristique spectrale mais de direction différente. Les conditions d'apparition de ce phénomène ne sont pas encore totalement déterminées. Néanmoins, on peut rapporter ici les conditions expérimentales où il a été découvert. Il s'agit d'une

expérience en chambre sourde, où le sujet est face à un haut-parleur émettant un son continu ; un court instant après le début du son continu, un second son est émis par un haut-parleur latéral (15° à droite). Suivant le décalage temporel du son de face et du son latéral, l'azimut apparent de ce second peut varier considérablement (voir Fig. 5). Au départ (sons simultanés) la position de l'image dépend de la relation de phase entre les sons (voir [Canévet 1985]). Après quelques millisecondes, par effet d'antériorité, le second son est attiré vers le centre. Au-delà de la zone temporelle d'action de l'effet d'antériorité, l'azimut se déplace vers le côté, droit ou gauche, suivant la position initiale du couple de sons simultanés ; on notera (Fig. 5) que le déplacement peut atteindre 60 à 70°.

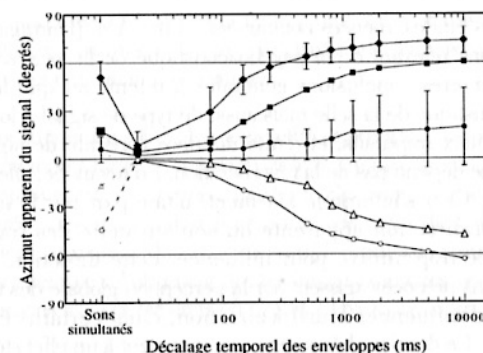


Figure 5 : Azimut apparent d'un son bref à 4 kHz, émis à 15° sur la droite d'un auditeur, en chambre sourde, après exposition au même son venant d'une source de face. L'abscisse indique le décalage entre l'apparition du son de face et celle du son de côté. (Adapté de [Canévet 1995]).

La zone d'intervention de l'antériorité est de l'ordre de la centaine de millisecondes, comme on l'a indiqué précédemment ; c'est également ce qu'avait trouvé Rakerd (1986) dans des mesures (totalement indépendantes) de localisation dans les salles (voir § 7 ci-après). L'effet qui prend la suite, que nous considérons comme une forme d'adaptation à moyen terme, est certainement dû à une variation lente de la sensibilité auditive : à mesure que cette sensibilité varie, les différences interaurales du champ induisent une différence d'activité neurale variable, et donc une perception variable de l'espace sonore.

En ce qui concerne l'effet d'antériorité, les mécanismes de base ne sont pas encore connus avec certitude. Zurek a émis l'hypothèse, à la suite d'expériences appropriées qu'il serait trop long d'exposer ici [Zurek 1980], que l'effet d'antériorité pourrait s'expliquer par une réduction temporaire de la sensibilité différentielle interaurale ; réduction équivalente dans ses effets à une inhibition.

7. Localisation dans les salles

Quelles que soient les bases physiologiques de l'effet d'antériorité, on connaît son utilité en acoustique des salles, où chaque son primaire est suivi d'une série d'échos, dus aux réflexions sur les parois. Pour des réflexions précoces, on a tendance à percevoir essentiellement le son direct, ce qui est plutôt bénéfique à la réception des messages. Des retards plus longs sont perçus comme de la réverbération, ou de véritables échos. Si le signal est un son de parole, l'intelligibilité va alors se dégrader, en raison du masquage exercé par les signaux réfléchis sur le signal primaire. Cet aspect de l'effet d'antériorité a fait l'objet d'une étude détaillée de Santon (1974, 1976).

De nombreuses études, menées notamment à l'IRCAM ([Hartmann 1983]), ont approfondi les questions relevant de la psychoacoustique de la localisation dans les salles. L'une des premières conclusions générales à retenir est que la localisation dépend certes de l'acoustique de la salle mais aussi du type de signal à localiser. Ainsi, la localisation de signaux impulsifs, et d'une manière générale de signaux à riche fourniture transitoire, ne dépend pas de la réverbération. En revanche, elle dépend de la géométrie de la salle. Ceci s'interprète aisément, d'une part par le fait que l'effet d'antériorité ancre la direction apparente du son sur celle des transitoires; la réverbération est donc trop tardive pour influencer cette direction. D'autre part cependant, les réflexions précoces agissent sur la perception globale des transitoires, et sont donc capables d'influencer leur localisation, dans certains cas même de "délocaliser" la source. Ce dernier phénomène est équivalent à un effet stéréophonique, où la source virtuelle finalement perçue résulte de la combinaison de la source primaire réelle et des sources d'images d'où proviennent les réflexions. On comprend donc l'importance de l'architecture de la salle dans l'organisation de la séquence des réflexions, et par conséquent dans la possibilité d'exploitation de l'effet d'antériorité pour la qualité des messages sonores.

Dans le cas de signaux entretenus, par exemple de signaux de musique, les interférences créées par l'ensemble de l'onde directe et des ondes réfléchies sont généralement complexes. Il se peut alors que le champ résultant produise sur les oreilles des signaux dont les différences instantanées, d'intensité et de temps par exemple, sont incohérentes du point de vue de la perception. On dit alors que les indices du champ acoustique sont conflictuels ([Canévet 1985]). La localisation, dans ce cas, est en principe impossible. En fait, Hartmann a observé, à plusieurs reprises ([Rakerd 1985], [Hartmann 1989]), que les auditeurs semblent effectuer une analyse séparée des différents indices de directivité, puis qu'ils appliquent à ces indices une sorte de "pondération perceptuelle" pour décider de la direction probable du son. C'est une vue des choses que Hartmann appelle "hypothèse de plausibilité", qui sous-entend l'intervention de mécanismes cognitifs, évidemment indéterminés, dans l'évaluation des différents indices physiques disponibles, y compris d'éventuels indices visuels. Les ventriloques pourraient en témoigner.

D'autres facteurs psychoacoustiques importants influencent la perception spatiale dans les salles. On sait depuis les travaux de Stevens (1975) que les sons peuvent procurer une impression de "volume". Les sons aigus sont perçus comme plutôt concentrés ou ponctuels, les sons graves apparaissent plutôt larges et diffus. Cet attribut élémentaire des sons est particulièrement important dans la perception du paysage sonore. On l'utilise parfois au cinéma pour créer des ambiances ("sensurround" ou autres) propices à faire participer subjectivement le spectateur. Une autre caractéristique acoustique intervient aussi dans l'impression spatiale : la cohérence interaurale. Il est facile de démontrer par exemple que des bruits présentés par écouteurs vont créer l'impression d'une image très concentrée (vers le centre de la tête) si les deux bruits sont parfaitement cohérents. Inversement, s'ils sont totalement incohérents, l'impression obtenue sera celle d'une image diffuse, étalée dans l'espace auditif intracrânien. Pour une cohérence intermédiaire, l'image aura une forme plus ou moins complexe, et plus ou moins confinée.

En acoustique des salles, le même effet se produit, c'est-à-dire que l'impression d'espace plus ou moins large est assurée par le degré de cohérence entre le son direct et les réflexions provenant des parois. Le degré de cohérence des signaux binauraux contribue à créer une impression plus ou moins nette de "plénitude" spatiale, ou d'encombrement, qui peut dépasser le simple contour géométrique de la source (l'orchestre par exemple). Autrement dit, les images auditives occupent un volume plus vaste que si les sons avaient été émis en champ libre, toutes choses égales par ailleurs.

De nombreux facteurs supplémentaires contribuent évidemment à la qualité acoustique d'une salle. Et il faut souligner à ce sujet l'importance des études développées, notamment à l'IRCAM de Paris, pour la définition des paramètres psychoacoustiques qui gouvernent le jugement sur une salle (voir par exemple [Warusfel 1987 et 1992], [Jullien 1989] et [Lavandier 1988]). Une série de tests perceptifs a permis à ces auteurs de discerner les critères objectifs utilisés pour caractériser la qualité acoustique d'une salle. On a même pu en déduire des algorithmes pour la simulation de certaines de ces caractéristiques par traitement du signal ([Jot 1991 et 1992]). Ce type d'études est fondamental pour la définition des futurs lieux d'écoute. Car elles permettent la caractérisation, la mesure puis la simulation de "fonctions de transfert psychophysiques"; ces fonctions relient une grandeur d'entrée physique (le signal acoustique) et une grandeur de sortie sensorielle ou perceptive. Le même problème se pose aux compositeurs de musique produite par ordinateur : comment construire des signaux, et surtout comment fixer leurs caractéristiques physiques, pour qu'ils induisent telle ou telle qualité subjective, référentielle ou émotionnelle ?

8. Simulation de l'espace sonore subjectif

En parallèle avec le développement de la notion d'environnement virtuel, on assiste à de nombreuses tentatives pour recréer un paysage sonore virtuel, ou simulé, à

l'aide d'un nombre limité de sources. Il y avait déjà eu dans le passé la vénérable époque de la quadriphonie ou de "l'ambiophonie". Mais faute de support théorique et surtout psychoacoustique convenable, le succès avait été limité, et les quelques effets intéressants obtenus tenaient davantage de la distribution spatiale du son que de la reproduction contrôlée de l'espace sonore subjectif. La première approche sérieuse semble avoir été celle de Jessel (1973). A la base de ses travaux, il y avait le principe de Huygens, selon lequel on peut recréer n'importe quel champ acoustique à partir d'une collection bien choisie de sources secondaires. Le rôle de ces sources est de reconstruire à l'identique le champ primaire. Alors, par analogie avec la célèbre technique de l'électromagnétisme, Jessel a introduit le terme "holophonie" pour désigner sa technique de reconstruction d'images auditives. Dans le sens opposé, c'est-à-dire en inversant le rayonnement des sources secondaires, on obtient le non moins célèbre anti-bruit actif. La difficulté de cette approche était que, en toute rigueur, le nombre des sources secondaires devait être infini pour obtenir une reconstruction exacte du champ souhaité. Les réalisations pratiques du système Jessel ont été poursuivies dans le domaine de l'anti-bruit, mais pas dans celui de la simulation d'espace.

Quelques années plus tard, une idée différente faisait son apparition. Avant d'être transmis aux haut-parleurs, le signal électrique devait être traité de façon à prendre en compte les "fonctions de transfert de la tête" (voisines des fonctions représentées à la Fig. 4). Sakamoto et ses collaborateurs [Sakamoto 1981] ont tenté cette technique, développant ainsi un projet voisin de M. Schröder et B. Atal [Schröder 1963]. Les résultats sont excellents, mais ne valent en général, et par principe, que pour une seule position de l'observateur, ou en tout cas pour une région limitée de l'espace.

Enfin, les travaux actuels se concentrent sur la restitution des impressions spatiales à l'aide d'une simple paire d'écouteurs. Techniquement, la simulation de l'espace par écouteurs nécessite de créer, sur les tympans d'un auditeur, un champ acoustique identique à celui que produirait une source réelle de l'environnement. Cela est théoriquement possible, par convolution d'un signal monophonique quelconque avec les fonctions de transfert des oreilles de l'auditeur (du champ libre au tympan). Le résultat de ce traitement est que le signal, transformé par convolution, est entaché de toutes les altérations qu'aurait subi un signal réel par propagation jusqu'au tympan. Il contient donc toutes les informations spatiales d'un signal réel. Le premier auteur à réaliser ce type de simulation semble avoir été Pösselt (voir [Pösselt 1986]). Actuellement les résultats des recherches contribuent à affiner la technique (voir notamment [Wightman 1989 a et b]). Mais on peut faire état de quelques réalisations concrètes intéressantes, notamment de Pelliex et ses coauteurs [Pelliex 1992 et 1993] ainsi que de toutes celles qui sont décrites dans la présente publication. Il faut également faire mention de l'application qui est faite de cette technique à la conception de salles assistée par ordinateur. Dans ce cas on introduit en plus, dans la convolution, les réponses impulsionnelles de la salle projetée. De nombreux problèmes freinent cependant le développement de ces travaux. Les plus importants sont, d'une part la difficulté de créer par écouteur l'impression

d'extériorisation des sons, qui tendent à rester subjectivement au voisinage de la tête, et d'autre part les fréquentes confusions avant/arrière qui reflètent une maîtrise et une connaissance encore insuffisantes de tous les mécanismes psychoacoustiques impliqués dans la localisation. Les recherches fondamentales se poursuivent donc activement dans ce domaine, (voir par exemple [Chateau, 1995 a et b], [Hammershoi 1994], [Hartung 1993], et [Wightman 1994]).

9. Conclusion

Il est important de bien garder à l'esprit la distinction qui doit être faite entre les notions voisines de localisation, perception spatiale et spatialisation auditives. Dans cet article, nous n'avons exposé que les questions relatives à la localisation auditive, dans les limites de ce qui est connu avec une certitude raisonnable. La localisation, au sens du positionnement des sources dans l'espace, n'est cependant qu'une partie, essentiellement physique et sensorielle, de la perception de l'espace sonore au sens large. Les problèmes purement perceptifs, qui prennent le relais de la localisation, sont complexes, multidisciplinaires, multidimensionnels car relevant presque de l'expérience individuelle, et de toute façon les recherches dans ce domaine n'en sont qu'à leur début. La spatialisation en revanche est une technique florissante, qui servira certainement beaucoup la création musicale et artistique, et même les recherches de perception.

10. Bibliographie

- [Abbagnaro 1975], L. A. Abbagnaro, B. B. Bauer, E. L. Torick, *Measurements of diffraction and interaural delay of a progressive sound wave caused by the human head*, J. Acoust. Soc. Amer., 1975, 58, 693-700.
- [Aschoff 1992], A. Aschoff, *Complexe de l'olive supérieure et lemnisque latéral*, in : "Le système auditif central - Anatomie et physiologie", Editions INSERM/SFA, Paris, 1992, 109-150.
- [Barone 1995], P. Barone, *Localisation spatiale d'une source sonore : bases anatomiques et neurophysiologiques*, In : "Le son et l'espace", Rencontres Musicales Pluridisciplinaires Informatique et Musique, mars 1995, Lyon.
- [Batteau 1967], D. W. Batteau, *The role of the pinna in human localization*, Proc. Roy. Soc. (London), 1967, série B, 168, 158-180.
- [Blauert 1969/70], J. Blauert, *Sound localization in the median plane*, Acustica, 1969/70, 22, 205-213.
- [Blauert 1983], J. Blauert, *Spatial Hearing. The psychophysics of human sound localization*, The MIT Press, 1983, (version anglaise actualisée de l'ouvrage publié en 1974 : Räumliches Hören, Hirzel, Stuttgart).
- [Butler 1980], R. A. Butler, E. T. Levy, W. D. Neff, *Apparent distance of sounds recorded in echoic and anechoic chambers*, J. Exp. Psychol., Human Perception and Performance, 1980, 6, 745-750.
- [Canévet 1985], G. Canévet, *Aspects physiques de la détection et de la localisation masquées*, Acustica, 1985, 57, 122-132.