

Psychoacoustique musicale

Cours 9 : La localisation de la source sonore

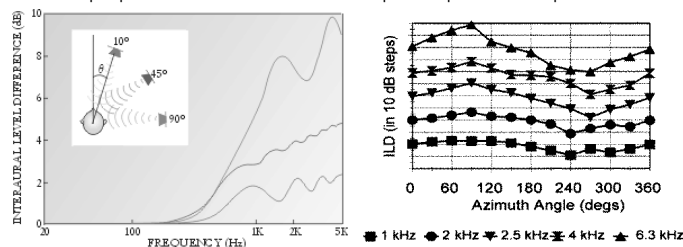
16 mars 2006

Calcul de la fréquence critique (correspondant aux dimensions d'une tête) :
diamètre de la tête = 23 cm = 0,23 m $\rightarrow f = c / \lambda = 344 / 0,23 \approx 1500$ Hz

À 500 Hz, la longueur d'onde d'un son est de 68 cm, ce qui représente plus de 3 à 4 fois le diamètre d'une tête moyenne. À cause de la diffraction, la DNI est donc petite pour des fréquences inférieures à 500 Hz, si la source est au moins à un mètre de distance. Cela veut donc dire que **la DNI est un indice de localisation surtout utilisé pour les hautes fréquences** (au-delà de 4000 Hz).

Figure de gauche : DNI en fonction de la fréquence **calculées** pour une source dans le plan azimutal (comprenant les deux oreilles et le nez). La source est placée à un azimut θ de 10°, 45° ou 90° par rapport à la direction frontale (résultats d'un **modèle sphérique** où les oreilles seraient situées à deux poles opposés).

Figure de droite : DNI en fonction de l'angle d'azimut **mesurées** pour différentes fréquences. On remarque que la différence de niveau s'accroît pour des plus hautes fréquences.



1. L'audition binaurculaire

La perception des sons dans l'espace est essentiellement reliée à la perception binaurculaire, c'est-à-dire au fait d'avoir deux oreilles. La localisation des sons dans l'espace est complémentaire des autres sens, notamment la vision et la perception de son corps dans l'espace (proprio-perception).

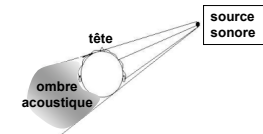
La localisation comprend trois dimensions spatiales, soit l'**azimut** (gauche-droite), le **plan vertical** (haut-bas), et la **distance**.

2. Localisation à l'azimut (gauche - droite)

Les indices pertinents sont les **différences de niveau interauriculaires** (DNI) et les **différences de temps interauriculaires** (DTI). Ces deux indices sont indépendants.

2.1. Les Différences de Niveau Interauriculaires (DNI)

Il y a 120 ans, Lord Rayleigh avait déjà compris l'importance pour la localisation de la **différence de niveaux d'intensité arrivant aux deux oreilles**. À cause du phénomène de la **diffraction**, un son venant de la droite sera plus fort à l'oreille droite qu'à l'oreille gauche, et ce d'autant plus que la fréquence est élevée. En effet, la **tête** constitue un **obstacle à franchir** pour les ondes sonores et une **"ombre acoustique"** apparaît, d'autant plus marquée que la fréquence est élevée.



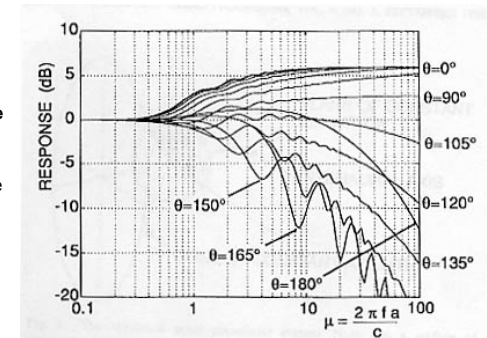
Le modèle sphérique de Rayleigh

Lord Rayleigh a modélisé la tête par une sphère et a résolu les **équations de propagation d'une onde autour d'une sphère rigide**. Cette approximation permet d'expliquer certaines observations en rapport avec les DNI. Il a ainsi obtenu la réponse en fréquence de ce modèle sphérique, illustrée sur la figure ci-dessous, en fonction de la fréquence normalisée ($\mu = a / c$) où a est le diamètre de la tête et c la vitesse du son.

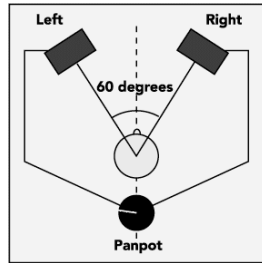
Le modèle est tracé pour plusieurs valeurs de θ , angle correspondant à la **différence entre l'azimut de la source et l'emplacement de l'oreille**.

On peut remarquer que pour des angles $\theta > 100^\circ$ (point dans l'hémisphère opposée à celle de l'oreille), le modèle se comporte comme un filtre passe-bas.

(Ref.: Brown & Duda 1998).



La **plus petite différence de niveau perceptible** serait approximativement de **0.5 dB**, à toutes les fréquences !!! Les DNI sont donc des indices de localisation potentiels à n'importe quelle fréquence, à condition que la différence soit supérieure à 1 dB. Ceci explique l'efficacité de la **stéréo panoramique**.



La figure illustre la configuration stéréo standard avec les haut-parleurs à 60 degrés. Si l'angle est plus élevé, un 'trou' apparaît au centre du champ sonore.

Les potentiomètres traditionnels n'opèrent que sur le volume sonore si bien que dans cet exemple, l'auditeur entend un son provenant de sa gauche parce que le volume sonore du haut-parleur gauche est plus élevé.

Conclusion : dans une situation "naturelle" (localisation d'une source unique ponctuelle), les **DNI** ne sont utilisées que **pour la localisation des hautes fréquences**.

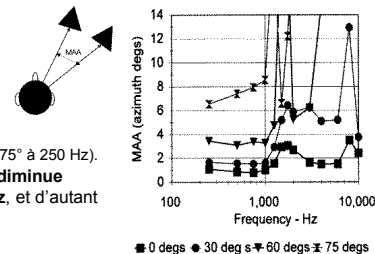
Alors comment expliquer que l'on puisse quand même localiser des sources émettant des basses fréquences ?

Des expériences ont déterminé que l'on peut localiser un son pur à 500 Hz avec beaucoup de **précision**. À l'avant ($\theta = 0^\circ$), la sensibilité $\Delta\theta$ serait de **1° à 2°**.

Une différence d'angle d'azimut de 1° correspond à une DTI de 13 µs !!!

Le système nerveux ayant des délais synaptiques de l'ordre d'une milliseconde, comment expliquer qu'il puisse encoder une si petite différence de temps ? La réponse à cette énigme semble être dans le **traitement parallèle de l'information** par le cerveau : l'information temporelle est transmise à grand nombre de neurones à la fois. Un centre de traitement du cerveau moyen (l'olive supérieure) serait capable d'effectuer une **opération de corrélation croisée** sur les signaux provenant des deux oreilles, dont l'information de déphasage est extraite.

Comme le montre le graphique ci-contre, l'angle minimal audible (*minimal audible angle* - MAA) pour des sons purs est faible à l'avant aux basses fréquences ($\Delta\theta = 1^\circ$ pour un azimut $\theta = 0^\circ$ à 250 Hz), mais le MAA augmente avec l'angle d'azimut ($\Delta\theta = 6,5^\circ$ pour un azimut $\theta = 75^\circ$ à 250 Hz). On remarque aussi que la **sensibilité diminue drastiquement entre 2000 et 4000 Hz**, et d'autant plus que l'angle d'azimut est élevé. Comment peut-on expliquer cela ?



2.2. Les Différences de Temps Interauriculaires (DTI)

Pour les basses fréquences (grandes longueurs d'ondes $\rightarrow \lambda > \text{tête}$), il semble que notre système auditif binaurculaire utilise la **différence de phase ($\Delta\phi$)** ou **différence de temps d'arrivée (Δt)** entre les signaux arrivant aux deux oreilles (les signaux étant repérés par les maxima ou les passages à zéro).

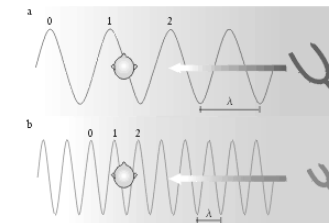
Une différence de phase $\Delta\phi$ correspond à une DTI $\Delta t = \Delta\phi / (2\pi f)$. Pour calculer une approximation de la DTI, on considère la demi-circonférence correspondant au parcours de l'onde autour de la tête pour passer d'une oreille à l'autre : $d = \pi R$ où R est le rayon de la tête. La DTI vaut alors $\Delta t = d / v$ où v est la vitesse du son dans l'air.

→ Rayon de la tête = 10 cm = **0,1 m**
 Demi-circonférence = $0,1 \pi =$ **0,314 m**
 $DTI = \Delta t = 0,314 / 340$
 $=$ **924 µs = 0,924 ms** qui est la plus grande valeur possible (dans le cas où la source est en face d'une des deux oreilles).

Pour les grandes longueurs d'onde, on peut utiliser la formule pour la **diffraction par une sphère** : $\Delta t = (3R / v) \sin \theta$ où θ est l'angle azimutal (gauche-droite). À 90° (source en face d'une des deux oreilles), $\Delta t = (3 \times 0,1 / 340) =$ **0,882 ms**.

Le problème de l'ambiguïté de phase avec les DTI

Un problème d'ambiguïté de phase survient quand la longueur d'onde est comparable à la distance entre les mesures (par les deux oreilles respectivement).



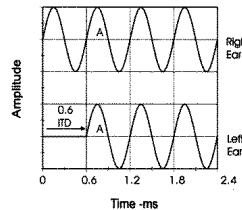
Considérons un signal venant de la droite. **Sur la figure (a)**, l'onde est telle que le maximum "1" arrive à l'oreille droite avant d'arriver à l'oreille gauche. Parce que la longueur d'onde est plus grande que 2 x le diamètre de la tête, il n'y a pas de confusion causée par les autres maxima de l'onde ("0" ou "2").

Sur la figure (b), la longueur d'onde est plus courte que 2 x le diamètre de la tête. Chaque point du 2^{ème} cycle arrivant à l'oreille droite est immédiatement précédé par le point correspondant du 1^{er} cycle à l'oreille gauche. L'auditeur en conclut que la source vient de la gauche !

Appelons Δt le délai entre les deux oreilles.

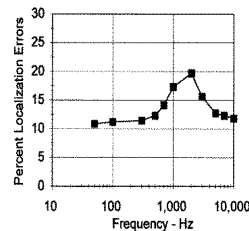
- **période $> (2 \Delta t)$** : pas d'ambiguïté
- **période = $(2 \Delta t)$** : ambiguïté totale
- **$(\Delta t) < \text{période} < (2 \Delta t)$** : localisation apparente de la source du côté opposé

De même, l'ambiguïté de phase peut faire qu'un son pur à 1666 Hz venant de la droite pourrait être perçu comme venant de l'avant.



Imaginons que la DTI vaut 0,6 ms quand le son est présenté à droite (c'est-à-dire que l'onde atteint l'oreille droite 0,6 ms avant l'oreille gauche). Or, la période d'un son pur à 1666 Hz vaut précisément $1/1666 = 0,0006 \text{ s} = 0,6 \text{ ms}$.

Comme l'illustre la figure ci-contre, des ondes à 1666 Hz arrivent donc **en phase** aux deux oreilles ce qui donnent l'impression que le son est situé devant plutôt qu'à droite !



Ce phénomène fut confirmé par une expérience consistant à compatibiliser les erreurs de jugement pour la localisation d'une source sinusoïdale. Le graphique ci-contre donne le pourcentage d'erreur en fonction de la fréquence. On remarque que le taux d'erreur augmente autour de 2000 Hz.

Localisation latérale exceptionnelle chez les chefs d'orchestre

Aussitôt que la source n'est plus en face, notre localisation latérale devient très mauvaise.

Or, comment les chefs d'orchestre, tête bien en avant, font-ils pour savoir que c'est le violoncelle 2 (et pas le 3 ou le 1) qui joue faux (le violoncelle étant normalement à la droite du chef) ???

Une étude a montré que les chefs d'orchestre ont développé une capacité exceptionnelle pour localiser latéralement. C'est le résultat de la **plasticité cérébrale** qui fait que leur expertise (via leur expérience) leur permet d'être nettement meilleurs que la population générale.

Ainsi, la pratique d'une activité permet de développer de nouvelles connexions au cerveau, et ainsi d'augmenter les ressources dédiées à cette activité. On l'a démontré beaucoup pour des aspects liés à l'expertise musicale, entre autres.

Mais la nature est bien faite. Le système auditif binaurculaire semble être beaucoup moins sensible aux DTI pour des fréquences entre 1000 et 1500 Hz, fréquences pour lesquelles la différence de phase est ambiguë.

Pour des sons purs, il semble notre système n'est pas sensible à des DTI pour des fréquences supérieures à 1500 Hz.

Cependant, pour des sons complexes modulés, des DTI pour des hautes fréquences (4000 Hz) peuvent être détectées. Il faut préciser que les fréquences de modulation doivent être inférieures à 1000 Hz. Notre système de localisation basé sur les DTI est donc inapte à suivre des variations rapides.

On devine l'idée que nous utilisons deux types d'indices (les DNI et les DTI) pour la localisation d'abord à Lord Rayleigh, et plus tard à Stevens et Newman.

C'est ce que l'on appelle la **théorie duplex de la localisation**.

La localisation azimuthale : en résumé...

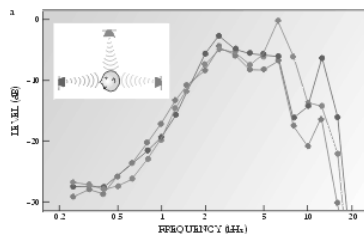
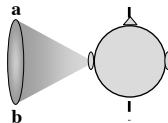
- Notre système binaurculaire est sensible à des DNI à toutes les fréquences, pour des ondes incidentes planes, bien que des indices de DNI n'existent physiquement que pour des fréquences supérieures à 500 Hz.
- Les DNI deviennent des indices de localisation très efficaces au-delà de 3000 Hz.
- Les DTI sont des indices de localisation pour les plus basses fréquences, inférieures à 1500 Hz.
- Pour des sons purs ou à bande étroite autour de 2000 Hz, ces deux indices (DNI et DTI) sont inefficaces et la localisation est imprécise dans cette région.

3. La perception avant/arrière et haut/bas

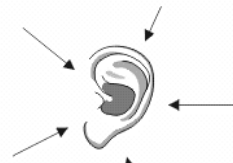
Limitations des indices de localisation basés sur des différences (DNI, DTI)

Considérant le modèle de la tête sphérique, tous les points compris dans un plan vertical perpendiculaire à l'axe passant par les deux oreilles, correspondent à des signaux arrivant aux deux oreilles avec la même intensité et la même phase. Les indices de DNI et DTI ne permettent donc pas de distinguer le haut, le bas, l'avant et l'arrière, dans le plan vertical.

D'autre part, des sources sonores situées en tout point de la surface d'un cône, appelé **cône de confusion**, causeront des DTI identiques. Sans le support visuel, le son peut être perçu comme provenant du côté opposé sur le cône, par rapport à son point d'origine réel.



Malgré tout, la localisation avant/arrière et haut/bas est possible grâce à l'interprétation par le cerveau du filtrage acoustique du son par la **Fonction de Transfert Anatomique** (FTA), correspondant au filtrage par le pavillon, la tête, les épaules et le haut du tronc. La FTA est ici illustrée pour trois positions dans le plan vertical.



Les sons venant de directions différentes vont être affectés de manières différentes, à cause de la forme du pavillon, notamment.

En fonction de la direction, certaines fréquences seront atténuées et d'autres amplifiées.

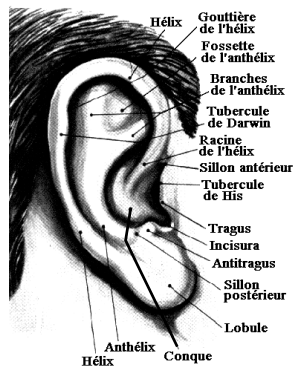
En particulier, les ondes sonores **provenant de l'arrière** sont **amplifiées** autour de **1000 Hz**, et celles **provenant de l'avant** sont **amplifiées** autour de **3000 Hz**.

Au-delà de 4000 Hz, la longueur d'onde est inférieure à 10 cm et donc inférieure aux dimensions des détails de la tête et en particulier, du pavillon de l'oreille.

Le **pavillon** devient alors un obstacle significatif pour les hautes fréquences et agit donc comme un **filtre passe-bas**. Un **pic à 7000 Hz** semble être une caractéristique du filtrage d'une source sonore située **au-dessus de la tête**.

Pour raffiner les modèles (celui de Rayleigh notamment) et les réponses en fréquence correspondantes, on peut effectuer des mesures à l'aide d'un mannequin tel que le KEMAR (Knowless Electronics Manikin for Acoustic Research) qu'on appelle aussi "Dummy head" en anglais.

Les FTA illustrées sur le graphique de la page précédente ont été obtenues à l'aide de ce dispositif.



Délais et résonances au niveau du pavillon

Les circonvolutions du pavillon provoquent des échos du son direct atteignant le tympan.

Ces délais courts agissent comme un **filtre en peigne** sur le signal original. Des petits mouvements de la tête correspondent à des changements importants dans la forme du filtrage.

Aussi, la **conque**, cavité profonde du pavillon, est le lieu de résonances dont l'amplitude dépend de la direction des ondes incidentes. (La conque tire son nom du coquillage servant de trompette aux dieux de la mer dans la mythologie).

Nous effectuons inconsciemment des **petits mouvements de la tête** pour obtenir différentes informations de filtrage d'une source donnée. En les combinant, la localisation devient plus précise.

Remarque : Les Fonctions de Transfert Anatomiques sont traduites en anglais par "**Head-Related Transfer Function**" (HRTF).

Illustration du filtrage anatomique (Head-Related Transfer Functions) pour différentes positions dans le plan vertical latéral

(plan vertical passant par les deux oreilles)

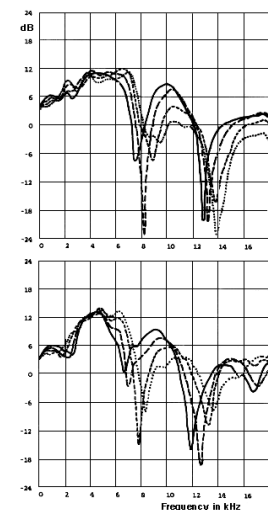
Les graphiques ci-contre donnent les HRTF pour deux sujets différents.

La source sonore est située à 2 m à gauche du sujet et est déplacée du niveau de l'oreille (0°) jusqu'à un angle de 30° au-dessus du niveau de l'oreille, dans le plan latéral.

Légende :

Ligne continue : 0°
Tirets longs : 10°
Tirets courts : 20°
Pointillés : 30°

Réf: G. Kendall & W. Martens, "Simulating the cues of spatial hearing in natural environments," Proceedings of the 1984 *International Computer Music Conference*.



Problème posé par la localisation par interprétation du filtrage anatomique

Si on ne connaît pas d'avance les caractéristiques timbrales de la source, on ne peut pas facilement déterminer si un profil spectral particulier (une amplification autour de 7000 Hz par exemple), est dû à un filtrage acoustique lié à l'emplacement de la source, ou fait simplement partie des caractéristiques de la source elle-même.

→ la connaissance préalable de la source est nécessaire.

Cette confusion est très marquée pour des sons purs ou à bandes étroites (de quelques demi-tons). **L'auditeur va localiser la source en fonction de sa fréquence seulement, et non pas en fonction de son emplacement réel !!!**

Les fréquences graves sont plutôt perçues provenant d'en bas, et les fréquences aiguës d'en haut. Par exemple, la fréquence 8000 Hz est toujours perçue comme provenant du haut, indépendamment de son emplacement réel.



Conclusion des points 1., 2. et 3.

Le système auditif binaurculaire utilise trois indices différents pour localiser les sources sonores. Chacun des indices est plus efficace dans une gamme de fréquences particulière (basses, moyennes et hautes fréquences respectivement).

Pour les fréquences inférieures à 1000 Hz, la longueur d'onde est relativement grande par rapport aux dimensions de la tête, et les différences de phase interauriculaires sont traduites en **différences de temps interauriculaires (DTI)**. Cet indice est utilisé pour localiser les sons dans le plan horizontal (localisation gauche-droite).

Pour les fréquences entre 1500 et 5000 Hz, la longueur d'onde est plus petite que la tête. La tête représente un obstacle pour l'onde sonore et projette une ombre acoustique. Il en résulte des **différences de niveau interauriculaires (DNI)**. C'est donc l'indice de localisation prédominant pour les fréquences moyennes.

Pour les fréquences supérieures à 5000 Hz, la longueur d'onde est plus petite que les dimensions des détails de la tête, et en particulier du pavillon de l'oreille. L'onde sonore va subir un filtrage acoustique, dépendant de sa direction d'origine.

La transition d'un type d'indice de localisation à un autre est relativement continue et comme dans un environnement naturel, les sons présentent généralement un spectre large, le système auditif va combiner les trois indices afin de former une impression subjective de l'emplacement de la source.

4. La perception de la distance de la source

a) **En champ libre**: l'**analyse spectrale** par le système auditif est le principal indice pour la perception de la distance (si la référence connue). En effet, la répartition des graves et des aigus change avec la distance (dispersion des graves, directivité des aigus, absorption des aigus par l'air). Cette faculté est plus précise si la source est en mouvement.

b) **Dans un espace réverbérant**: l'indice principal est la **proportion son direct/son réverbéré**. La précision dépend moins de la connaissance de la source. Cet indice permet d'évaluer la distance mais plus rarement la localisation.

Malgré tout, notre cerveau analyse l'ensemble de l'image sonore et est capable de séparer le son directement perçu depuis la source et ses multiples réflexions, grâce à l'analyse de l'altération des signaux réfléchis et des informations sur leur direction. Il reconstruit alors une image sonore virtuelle.

La **vision** a une importance capitale : le cerveau ajuste l'image sonore reconstituée pour la faire coïncider aux sources sonores situées dans son champ visuel. Cette correction influence aussi la localisation des sources qu'il ne peut pas voir.

Finalement, notre perception de la distance peut aussi être influencée par notre expérience avec un type de source particulier tel que la **voix**. Dans une expérience, on avait présenté des voix criées, parlées, et chuchotées à des pressions sonores identiques. Les voix criées étaient systématiquement évaluées comme étant plus distantes, suivies des voix parlées, puis des voix chuchotées.

5. L'effet de précedence (ou effet de Haas)

Lorsqu'un son arrive multiplié aux oreilles (à cause de réflexions sur des obstacles environnants), **l'information sur la localisation n'est retenue que pour la première manifestation de ce son.**

Le système auditif **ne tiendrait pas compte des signaux succédant le premier signal, et ce pendant les 40 premières millisecondes**, supposant que les signaux arrivant plus tard sont des échos ou de la réverbération, signaux qui ne possèdent pas d'information spécifique sur la localisation de la source.

Remarque :

Lorsque les délais entre l'onde directe et les ondes réfléchies s'étendent jusqu'à des valeurs **entre 50 ms et 500 ms**, on perçoit de la **réverbération**.

La réverbération varie en fonction des délais, qui dépendent de la pièce et du type de source. Normalement il n'y a pas de difficulté à localiser, mais on a l'impression d'être entouré de sons.

Lorsque le délai entre l'onde directe et les ondes réfléchies se situe **entre 500 ms et 1 seconde**, deux attaques distinctes sont entendues. On entend distinctement que le son est répété. C'est ce que l'on appelle l'**écho**.

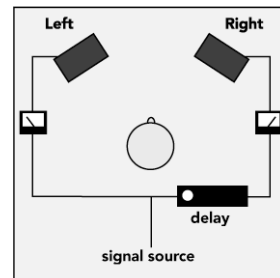
Helmut Haas a montré que les délais pour une certaine gamme de fréquences peuvent être compensés par des changements d'intensité.

Par exemple, un délai de 10 ms requiert un ajustement de 10 dB pour replacer l'image auditive au centre. Cette zone de fusion ou d'intégration s'étend jusqu'à des délais entre 30 et 35 ms. Des délais plus importants seront perçus comme des échos distincts.

Application : stéréo avec délai

→ prendre un signal, le diviser en deux, retarder l'un des deux et envoyer les signaux résultants respectivement à gauche et à droite.

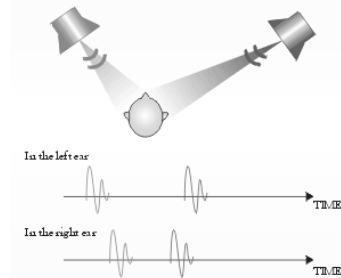
Cependant, pour que le volume apparaisse identique des deux côtés, **le niveau du signal retardé doit être plus élevé pour compenser l'effet de Haas** qui tente de nous faire croire que ce n'est qu'un écho.



L'**effet de précedence** est démontré facilement à l'aide d'un système stéréo reproduisant un enregistrement monophonique (même signal envoyé aux deux haut-parleurs).

Si on se place juste à **distance égale** des deux haut-parleurs, on entend le son comme **provenant de l'avant au centre**. Si maintenant on se déplace vers le haut-parleur gauche par exemple, le son **paraît provenir entièrement de ce haut-parleur**. Les indices de DNI et de DTI sont en compétition et la compétition est gagnée par le premier son qui parvient à l'auditeur, selon le principe de l'effet de précedence.

Bien que le son paraît provenir du haut-parleur gauche, le **haut-parleur droit continue à contribuer à l'intensité du son et à un sentiment d'étendue spatiale**. Ceci peut être vérifié en coupant soudainement le haut-parleur droit. La différence est très marquée.



L'effet de précedence ne dépend pas seulement des différences interauriculaires, mais aussi des différences spectrales dues au filtrage anatomique.

6. Latéralisation (interne) vs localisation (externe)

La **localisation** se produit en champ libre tandis que la **latéralisation** se produit quand l'auditeur entend la source via des écouteurs sur les oreilles, les différences entre les sons entendus à gauche et à droite (DNI et DTI) étant manipulées de façon à simuler la situation en champ libre. Dans ce cas, la source n'est pas entendue comme provenant de l'espace extérieur, mais bien comme **provenant de l'intérieur de la tête**.

En effet, ce sont les pavillons (via le filtrage par les fonctions de transfert HRTF) qui donnent la perception qu'un son vient de l'extérieur plutôt que de l'intérieur de la tête. Le casque d'écoute qui « écrase » les pavillons sur la tête empêche à ces derniers d'exercer leur effet filtrant.

Voici une expérience démontrant l'importance des pavillons pour la localisation externe de la source :

Batteau (1967) a fabriqué des pavillons artificiels à partir de moules de vrais pavillons, et a inséré des microphones dans chaque pavillon. Les pavillons étaient installés sur une tige métallique. Le son recueilli par les microphones était ensuite transmis par écouteurs à des auditeurs. C'est comme si les auditeurs avaient des pavillons réels, mais éloignés (puisque dans une situation sous écouteurs, le pavillon n'intervient pas dans la perception du son). Les auditeurs ont rapporté que le son semblait provenir de l'espace extérieur, et non de la tête comme c'est le cas normalement avec des écouteurs.

7. Démasquage binaural

Le démasquage binaural est une mesure de l'amélioration dans la détectabilité d'un signal qui se produit dans une condition d'écoute binaurale. Le seuil de masquage d'un signal peut parfois s'abaisser quand on écoute avec deux oreilles plutôt qu'avec une seule. Ça dépend de plusieurs facteurs.

Le démasquage est donc **toute amélioration du seuil masqué obtenu lorsque deux oreilles plutôt qu'une sont utilisées**. Le démasquage binaural est exprimé en dB, en fonction de la fréquence du son masqué.

Le démasquage binaural démontre que **le système auditif est sensible à la phase**. En effet, dans certaines conditions, il y aura une diminution du masquage exercé par un type de stimulus lorsque le signal et le masque ont des relations de phase différentes.

Le démasquage binaural n'est pas un phénomène limité aux sons purs.

Des effets similaires ont été observés pour des sons complexes, la parole, des clics. En fait, il semble qu'à chaque fois que la phase ou le niveau du signal sont différents aux deux oreilles, notre habileté à détecter et à identifier les signaux s'améliore. Or dans la vraie vie, des différences de phase et de niveau pour chaque oreille se produisent seulement quand les sons proviennent d'endroits différents dans l'espace. Une des implications du démasquage binaural est donc que **notre détection de signaux sonores, incluant la parole, va s'améliorer quand le masque et le signal à détecter ne seront pas situés au même endroit dans l'espace**.

8. L'effet "cocktail party"

C'est la combinaison de la vue et de l'ouïe (via le démasquage binaural) qui permet de **focaliser l'attention auditive** sur quelque chose qui nous intéresse particulièrement dans un milieu bruyant. Un microphone placé au même endroit ne renverra que l'image sonore bruyante du milieu d'origine (d'où les difficultés rencontrées par les personnes portant un appareil auditif dans un milieu bruyant tel qu'un restaurant).

La localisation des sources sonores est donc un facteur important dans notre capacité d'analyse des scènes auditives et de séparation perceptuelle de ces sources (en combinaison avec la cohérence spectrale et la cohérence dynamique).

9. Le rôle de la vision

Lorsque les indices visuels et auditifs sont en conflit, **ce sont les indices visuels qui l'emportent généralement sur l'audition**. Pensons au cinéma : le son vient des haut-parleurs situés de part et d'autre de l'écran et sur les murs latéraux de la salle (même chose pour la télévision : le haut-parleur n'est pas dans l'écran), et pourtant nous percevons clairement que le son provient directement de la bouche des acteurs.

L'**effet McGurk** est aussi un exemple de situation où les indices sont en conflit et où la vision vient perturber ce que nous entendons (un « Ba » est entendu « Da » s'il est visuellement articulé comme un « Ga »). Ainsi, le cerveau tient compte de tous les indices et tente de se faire une idée (la plus plausible possible) sur la source du son.

Quelques définitions :

Monodique ou monaural (*monotic*) : stimuli présentés à une oreille seulement

Diotique ou binaural (*diotic*) : stimuli identiques présentés aux deux oreilles

Dichotique (*dichotic*) : stimuli différents présentés aux deux oreilles

Homophasique : lorsque la phase relative du signal aux deux oreilles est la même que la phase relative du masque.

Antiphasique : lorsque les phases relatives du signal et du masque aux deux oreilles sont opposées (une est inversée, l'autre pas).

10. Réorganisation cérébrale chez les personnes aveugles

La vision est le sens qui crée l'espace extra-personnel, et qui l'emporte sur l'audition lorsque les deux sont en conflit. Cependant, les aveugles de naissance semblent être capables de recréer un espace extra-personnel sans vision. Franco Leporé, professeur de psychologie de l'Université de Montréal qui travaille sur la réorganisation cérébrale chez des personnes aveugles, a démontré scientifiquement que **les aveugles localisent mieux que les voyants en situation monaurale**. Cette recherche a été publiée dans la revue *Nature*.

Franco Leporé a testé trois groupes de personnes: (1) des voyants (avec yeux bandés lors de l'expérimentation), (2) des aveugles de naissance, et (3) des aveugles avec une vision résiduelle (qui peuvent distinguer la lumière mais pas les formes). Les sujets avaient devant eux un périmètre de localisation, c'est-à-dire une série de haut-parleurs placés en demi-cercle, à hauteur des yeux. Des sons (bruits à bande large, contenant hautes et basses fréquences) étaient présentés aléatoirement et les sujets devaient pointer la direction de provenance. Avec leurs deux oreilles, tous les sujets étaient capables de pointer assez correctement. Par contre, en monaural, donc avec une seule oreille, seuls les sujets aveugles de naissance (sans vision résiduelle) étaient capables de localiser assez correctement à l'azimut. Les voyants et les autres aveugles (avec vision résiduelle) avaient des performances catastrophiques.

Des expériences ultérieures ont démontré que pour localiser, les aveugles de naissance utilisent les indices spectraux, en particulier l'information apportée par les hautes fréquences, qui interagissent fortement avec les pavillons d'oreille. Cette réorganisation cérébrale se ferait dans des zones postérieures du cerveau, probablement des zones qui auraient été dédiées à la vision si cette modalité sensorielle avait été présente chez eux.