

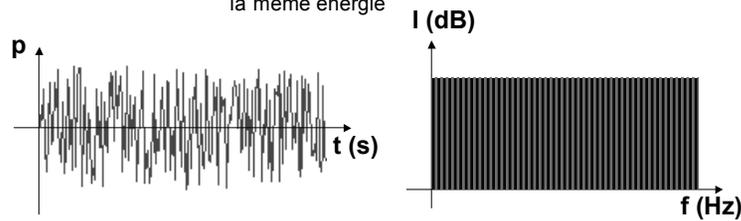
Psychoacoustique musicale

Cours 5 : La sonie des sons complexes

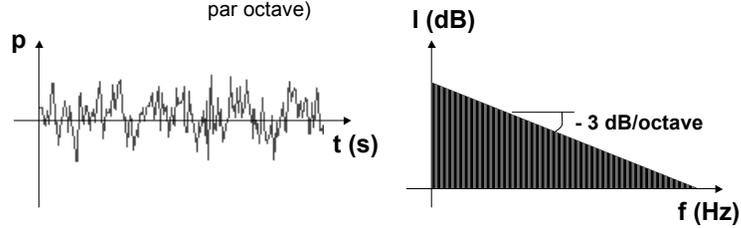
9 février 2006

Notions préalables :
les filtres et le bruit

Bruit blanc : toutes les fréquences reçoivent statistiquement la même énergie



Bruit rose : bruit blanc pondéré à -3dB/octave (énergie constante par octave)



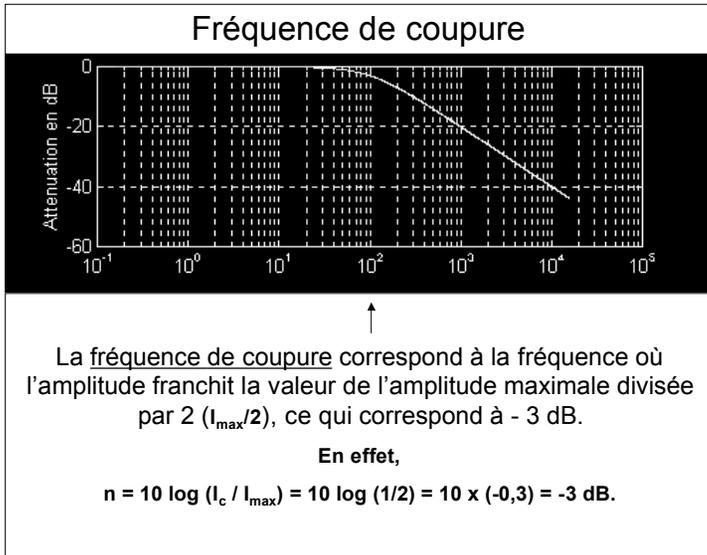
Les filtres

Un filtre est défini par

- 1- sa bande passante
- 2- sa ou ses fréquence(s) de coupure
- 3- sa constante de temps

La **bande passante** est l'ensemble des fréquences qui pourront traverser le filtre. Les fréquences limites de la bande passante sont les **fréquences de coupure**.

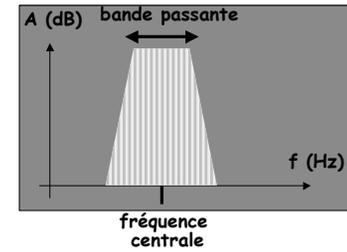
La **constante de temps** est le temps que met un filtre pour atteindre 90% de son régime permanent. Elle est inversement proportionnelle à la largeur de bande, c'est-à-dire que le filtre prend d'autant plus de temps à réagir que la bande passante est étroite.



Les analyseurs à filtres passe-bande

Il y a différents types d'analyseurs basés sur une banque de filtres passe-bande :

- La fréquence centrale des filtres passe-bande peut être fixe ou variable (balayage du spectre)
- La bande passante peut être plus ou moins étroite :

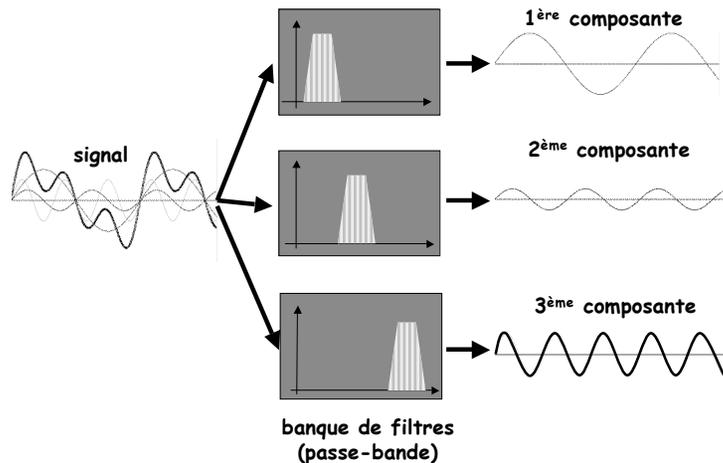


octave, 1/3 d'octave, 10%, 3% et 1% (%^{lage} de la fréquence centrale)

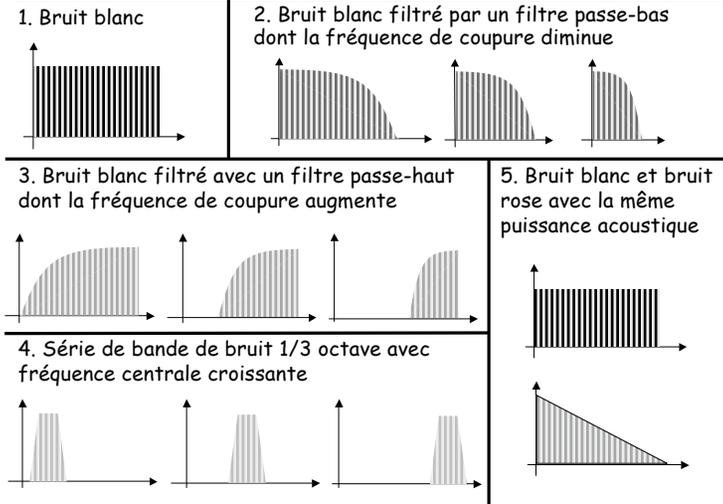
Plus la bande passante est étroite, plus l'analyse est fine.

- L'axe des fréquences peut être linéaire (si la bande passante est constante) ou logarithmique (si la bande passante dépend de la fréquence centrale = racine carrée du produit des fréquences de coupure inférieure et supérieure)

Analyse par une banque de filtres



Bruit blanc filtré ("Auditory Demonstrations" no 5)



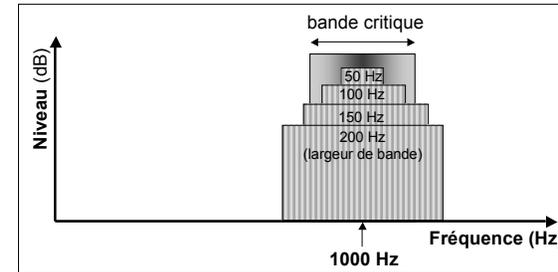
L'oreille effectue-t-elle une analyse spectrale du son ou décrit-elle la forme d'onde du signal sonore?

La notion de **bandes critiques** met en évidence les principales caractéristiques de cet "analyseur auditif" qu'est la cochlée.

Mise en évidence de l'existence de bandes critiques : expérience de Feldtkeller et Zwicker (1956)

- **Son de référence** : son pur à 1000 Hz
- **Son de comparaison** : bande de bruit dont la largeur de bande est augmentée progressivement tout en conservant l'énergie totale dans la bande (largeur de bande x niveau de spectre = constante)

Tant que la largeur de bande est inférieure à une largeur dite « critique », la sonie demeure constante (elle est toujours associée à la même intensité du son pur de référence à 1000 Hz). Au-delà, la sonie augmente.



La bande critique

La bande critique est la largeur de bande des « filtres auditifs » de l'oreille interne. Elle correspond à la largeur de bande d'un bruit en-deçà de laquelle la sonie reste constante si la pression sonore est maintenue constante.

Les psychoacousticiens modernes utilisent la formule suivante pour estimer la largeur d'une bande critique (*Equivalent Rectangular Bandwidth*) :

$$ERB = 24,7 (4,37 F + 1) = 108 F + 24,7$$

où **F** est la fréquence centrale en kHz du filtre auditif correspondant.

(B. C. J. Moore and B. R. Glasberg, « A revision of Zwicker's loudness model », 1996).

Exemple: à 1000 Hz, $ERB = 24,7 (4,37 \times 1) + 1 = 132,6$ Hz

fréq.	ERB	(ERB/f)	fréq.	ERB	(ERB/f)	fréq.	ERB	(ERB/f)
100 Hz	35,5 Hz	(35 %)	500 Hz	78,7 Hz	(16 %)	1300 Hz	165 Hz	(12,7 %)
150 Hz	41 Hz	(27 %)	600 Hz	89,5 Hz	(15 %)	1500 Hz	186,6 Hz	(12,5 %)
200 Hz	46,3 Hz	(23 %)	700 Hz	100,3 Hz	(14 %)	1700 Hz	208,2 Hz	(12,2 %)
250 Hz	51,7 Hz	(21 %)	800 Hz	111 Hz	(13,8 %)	2000 Hz	240,6 Hz	(12 %)
300 Hz	57 Hz	(19 %)	900 Hz	121,8 Hz	(13,5 %)	2250 Hz	267,6 Hz	(11,9 %)
350 Hz	62,5 Hz	(18 %)	1000 Hz	132,6 Hz	(13,2 %)	2500 Hz	294,5 Hz	(11,8 %)
400 Hz	68 Hz	(17 %)	1150 Hz	148,9 Hz	(13 %)	3000 Hz	348,5 Hz	(11,6 %)

Répartition des 24 bandes critiques de l'échelle de Bark

E. Zwicker and H. Fasl, Psychoacoustics, Facts and Models, 1990.

Bande critique	Fréquence (Hz)			
	Minimum	Maximum	Largeur	
0	0	100	100	} largeur constante (accroissement linéaire)
1	100	200	100	
2	200	300	100	
3	300	400	100	
4	400	510	110	} largeur augmente logarithmiquement (à peu près 20 % de la fréquence minimum → tierce mineure)
5	510	630	120	
6	630	770	140	
7	770	920	150	
8	920	1080	160	
9	1080	1270	190	
10	1270	1480	210	
11	1480	1720	240	
12	1720	2000	280	

Répartition des 24 bandes critiques de l'échelle de Bark (suite)

Bande critique	Fréquence (Hz)		
	Minimum	Maximum	Largeur
13	2000	2320	320
14	2320	2700	380
15	2700	3150	450
16	3150	3700	550
17	3700	4400	700
18	4400	5300	900
19	5300	6400	1100
20	6400	7700	1300
21	7700	9500	1800
22	9500	12000	2500
23	12000	15500	3500
24	15500	22050	6550

Sommation d'énergie ou de sonie

Si la largeur de bande de bruit est **inférieure** à la bande critique, il y a une **sommation de l'énergie**.

D'après Stevens,

$$\text{sonie totale} = k (I_1 + I_2 + I_3 + \dots)^{0.3}$$

où I_1, I_2, \dots sont les intensités des composantes fréquentielles du bruit

Si la largeur de bande de bruit est **supérieure** à la bande critique, il y a une **sommation de sonie**.

$$\begin{aligned} \text{sonie totale} &= \text{sonie}_1 + \text{sonie}_2 + \text{sonie}_3 + \dots \\ &= k I_1^{0.3} + k I_2^{0.3} + k I_3^{0.3} + \dots \end{aligned}$$

(Remarque: sonie_n = sonie dans la bande critique n)

Dans ce cas, la sonie totale obtenue est plus élevée.

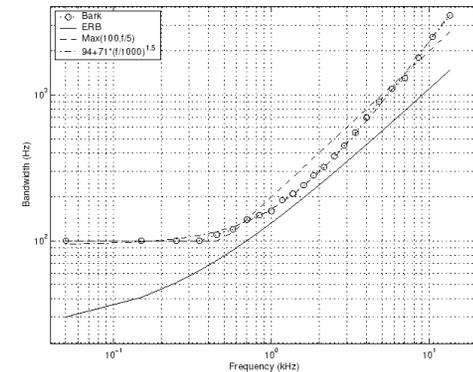
Exemple: $(4 + 5 + 6)^{0.3} = 2,25 < (4^{0.3} + 5^{0.3} + 6^{0.3}) = 4,85$

La sommation de la sonie s'applique dans le cas d'un son complexe harmonique par exemple, du moins pour les composantes dans les basses fréquences qui se retrouvent dans des bandes critiques distinctes.

Comparaison de l'échelle ERB et de l'échelle de Bark

L'ERB (*Equivalent Rectangular Bandwidth*) est liée à la bande critique exprimée en Bark mais elle est mesurée à l'aide de la méthode du bruit à bande réjectée (*notched-noise*) plutôt que via les expériences classiques de masquage impliquant un bruit masquant à bande étroite et un son pur masqué (cf. pages suivantes).

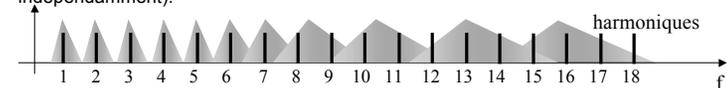
Les différentes méthodes utilisées pour la mesure de la bande critique (par Zwicker, Moore, ...) expliquent l'obtention de résultats expérimentaux menant à l'établissement de plusieurs modèles qui diffèrent légèrement, surtout dans les basses fréquences.



Les bandes critiques interviennent dans deux contextes musicaux différents:

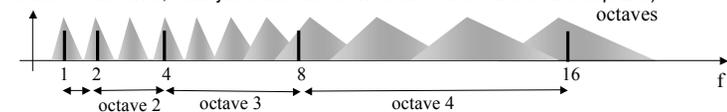
- la résolution des harmoniques d'un son complexe:

dans ce cas, c'est la **largeur absolue** de la bande critique qui compte (**ERB**). La largeur absolue des bandes critiques **augmente** avec la fréquence et de ce fait, les bandes critiques comprennent **de plus en plus d'harmoniques** vers les hautes fréquences. Les premières harmoniques sont donc facilement discriminées, mais les harmoniques d'ordre plus élevé fusionnent (on dit qu'elles ne sont pas « résolues » indépendamment).



- la distance fréquentielle minimale entre les voix (ex: opéra)

dans ce cas, c'est la **largeur relative** de la bande critique qui compte (**ERB / f** = largeur de bande sur fréquence centrale). La largeur relative des bandes critiques **diminue** avec la fréquence et de ce fait, les bandes critiques correspondent à des **intervalles musicaux de plus en plus petits** vers les hautes fréquences. Ceci explique pourquoi, pour bien distinguer les voix dans le registre grave, elles doivent être séparées davantage que les voix dans le registre aigu (ex: une octave entre une basse et un tenor, mais juste une tierce entre une contre-alto et une soprano).



Le masquage : définitions

L'effet de masque ou le masquage désigne le phénomène par lequel la présence d'un son peut empêcher la perception d'un autre son qui autrement serait audible. Autrement dit, le seuil d'audibilité d'un son - appelé "signal" - est élevé par la présence d'un autre son - appelé "sons masquant" ou "masque".

Cet effet est important dans le contexte de l'analyse de la scène auditive, du mixage, de l'écriture musicale, des techniques de compression (minidisc, MP3, ...), etc... On peut trouver des similitudes avec le masquage visuel.

Son masquant : son dont la présence rendra plus difficile la perception du "signal". Appelé "bruit" en communication → rapport signal-sur-bruit (SNR : *Signal-Noise Ratio*)

Son masqué : son qu'on cherche à percevoir et dont on mesure le seuil en présence du "son masquant". Le "son test" pour une expérience de masquage est généralement un son pur ou une bande étroite de bruit.

Courbe d'effet de masque ou **seuil masqué** : courbe similaire à la courbe de seuil d'audition (courbe de plus bas niveau parmi les courbes d'isotonie). Elle s'obtient de la même manière. Représente le niveau d'intensité tout juste perceptible (en fonction de la fréquence) d'un son en présence d'un son masquant.

Bruit blanc : bruit dont l'énergie est statistiquement égale à toutes les fréquences (ne présente aucune périodicité).

Bandes de bruit : bruit blanc filtré avec des filtres passe-bandes.

Les courbes d'effet de masque : l'étude de Wegel et Lane (1924)

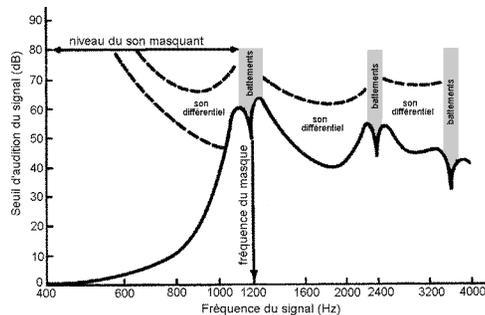
Cette étude classique concerne également l'interaction entre deux sons purs présentés simultanément. Le **son masquant** (f_1) est présenté à un **niveau fixe** (en dB SPL) et le **seuil masqué** pour le signal à détecter (f_2) est déterminé pour différentes fréquences. La courbe obtenue est une **courbe d'effet de masque** qui représente la quantité de masquage produit par un son masquant donné en fonction de la fréquence du signal masqué.

La courbe du seuil masqué ci-contre met en évidence la présence des **non-linéarités** de l'oreille lorsque le son présenté est à fort niveau (le masque est ici à 80 dB SL).

Des **battements** sont aussi ressentis pour des fréquences harmoniques du signal ($2f_2, 3f_2, \dots$)

Les **sons différentiels** sont des sons dont la fréquence correspond à la différence des deux sons en présence ($f_2 - f_1$).

On remarque une forte **asymétrie** de l'effet de masque, qui est plus efficace pour des **fréquences supérieures** à la fréquence du masque (c'est pourquoi il faut plus de femmes que d'hommes dans les chœurs).

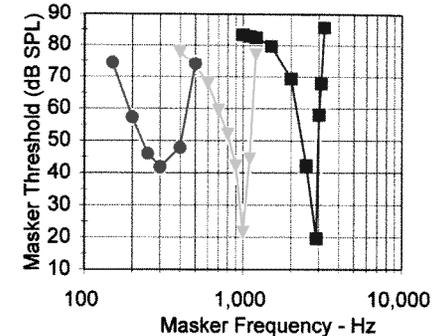


Le masquage tonal et les courbes d'accord psychophysique

Le masquage tonal consiste à masquer un son pur par un autre son pur. Plus le masque doit être fort pour couvrir le signal, moins le masque est efficace. Plus le masque est faible, plus il est efficace.

Ci-contre, trois **courbes d'accord psychophysique** pour le masquage simultané d'un signal à :
 - 300 Hz (●),
 - 1000 Hz (▲),
 - 3000 Hz (■).

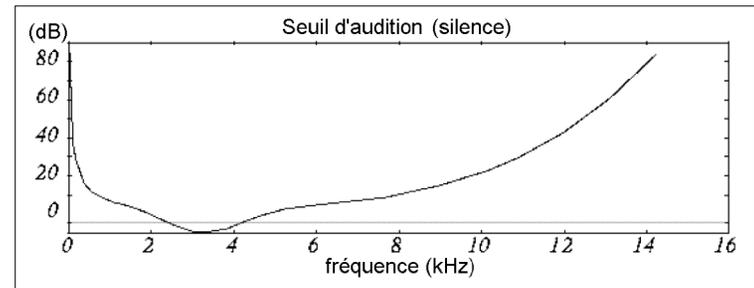
Ces courbes présentent, en fonction de la fréquence, le **niveau d'intensité du son masquant** nécessaire pour masquer le **signal** (qui est un son pur à une fréquence fixe).



Plus on s'éloigne de la fréquence à masquer, plus le niveau du son masquant doit être élevé pour obtenir le même effet de masquage. Ces **courbes d'accord psychophysique** sont similaires aux **courbes d'accord neural** mesurée chez les animaux.

Courbes d'effet de masque : seuil d'audition

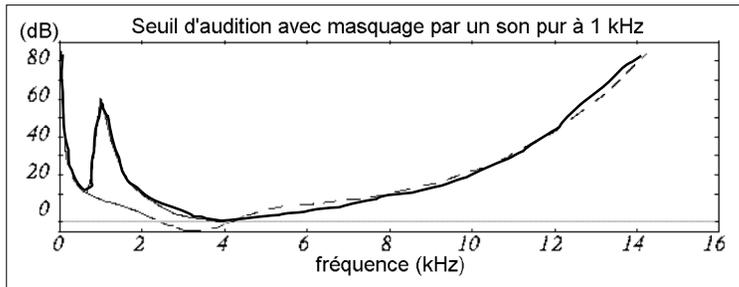
Niveau d'intensité du seuil de perception dans le silence en fonction de la fréquence.



Cette courbe correspond à la première courbe d'isotonie à 4 phones.

Courbes d'effet de masque : masquage par un son pur

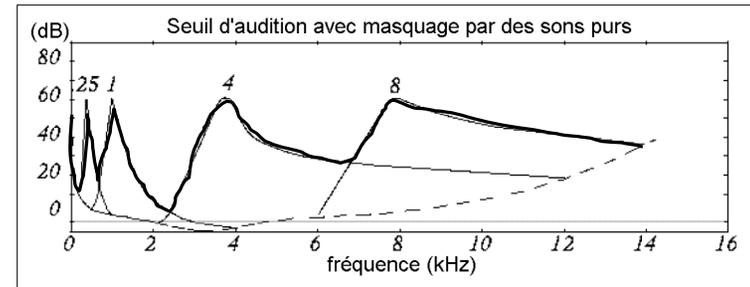
Niveau d'intensité du seuil de perception en présence d'un son masquant à 1000 Hz en fonction de la fréquence.



Le seuil d'audition est déformé (vers le haut) dans la région du son masquant.

Courbes d'effet de masque : masquage par des sons purs

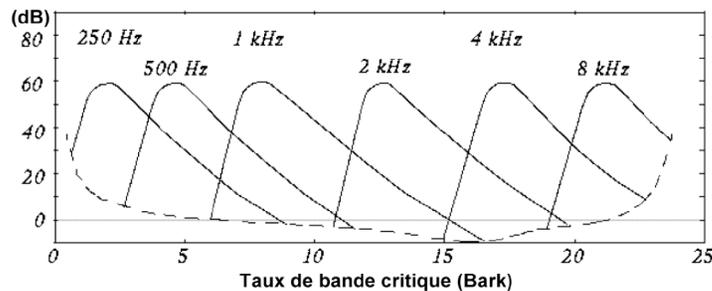
Niveau d'intensité du seuil de perception du son masqué en fonction de la fréquence.



Si un son en cache un autre qui autrement serait perceptible, c'est qu'il excite la même région de la membrane basilaire que celui qui est caché.

Courbes d'effet de masque par des sons purs, représentées avec une échelle de fréquence en "Bark"

1 Bark = largeur d'une bande critique

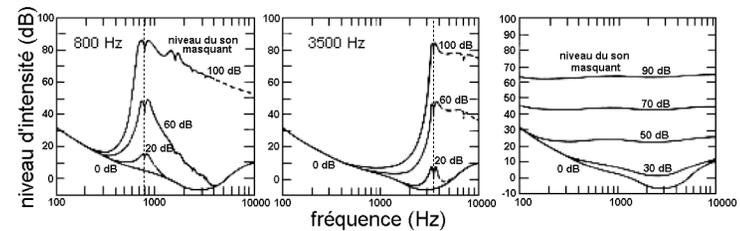


Les différentes courbes de masquage ont maintenant la même largeur.

Courbes d'effet de masque en fonction de l'intensité du son masquant

Seuil de masquage pour différentes intensités d'un masque :

(a) son pur à 800 Hz (b) un son pur à 3500 Hz (c) bruit blanc



de Miller, Language and Communication, McGraw-Hill, N.Y., 1963, pp. 54-55.

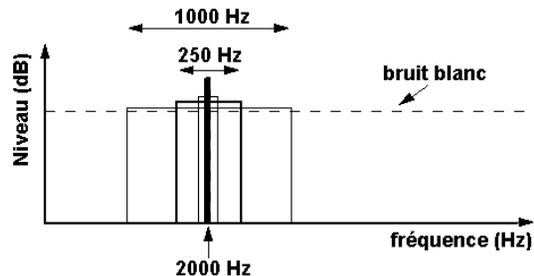
Dans le cas d'un masque "son pur" (a, b) : on remarque l'asymétrie du masquage (*upward spread of masking* : étalement du masquage vers le haut)

Dans le cas d'un masque "bruit blanc" (c) : plus on augmente le niveau du bruit, moins le seuil du signal dépend de la fréquence (la courbe d'effet de masque s'aplatit).

Le masquage par le bruit

Masque d'un son pur par une bande de bruit de largeur croissante

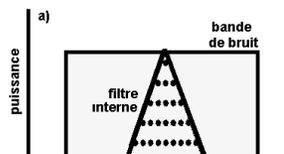
Démonstration (basée sur une expérience de Fletcher dans les années '40) : Vous allez entendre un son pur à 2000 Hz présenté en 10 intensités décroissantes par paliers de 5 dB. Ce son va être masqué successivement par un bruit blanc, puis des bandes de bruit de largeur 1000, 250 et 10 Hz respectivement, toutes centrées sur 2000 Hz, la fréquence du son masqué. Comptez le nombre de paliers que vous entendez dans chaque cas.



En résumé :

Seule l'énergie contenue dans une bande critique a un effet masquant.

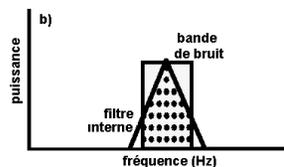
Donc, si la largeur de bande du son masquant s'élargit, seule la partie qui se situe à l'intérieur de la bande critique masque le son masqué.



La figure ci-contre représente le « filtre interne » (triangle) en présence

- a) d'un bruit à large bande
- b) d'un bruit à bande étroite

Le bruit à large bande (a) produit une puissance maximale à la sortie du filtre interne.



Par contre, la largeur de bande du bruit à bande étroite (b) est inférieure à la largeur de bande du filtre interne (l'aire commune est plus petite).

L'effet de masque sera donc **plus important** pour un signal dont la fréquence est au centre du filtre interne dans le cas du **bruit à large bande** (car la totalité du filtre est sollicitée).

Discussion :

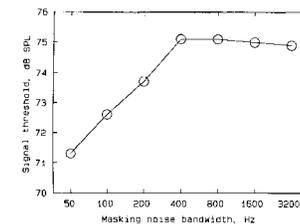
En l'absence du son masquant, on peut entendre les 10 paliers.

En présence du bruit blanc et des bandes de bruit larges, on peut seulement entendre 5 paliers environ. Il n'y a pas de changement lors du passage à des largeurs de 1000 et 250 Hz, alors que le son masquant est clairement perçu moins intense (à énergie totale égale, la sonie d'un son complexe à large bande est beaucoup plus grande !).

Lors du passage à 10 Hz de largeur de bande, on peut entendre plus de paliers. Ceci s'explique par le fait que la bande de bruit est alors plus étroite que le filtre auditif centré à la fréquence du son masqué, et il y a donc moins d'énergie qui vient perturber le son à 2000 Hz dans cette zone.

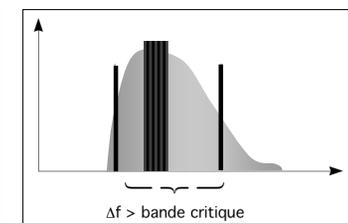
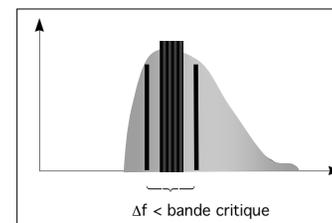
La largeur du filtre auditif correspond à la "bande critique".

Dans le cas de l'expérience décrite ici (largeur de bande décroissante), la bande critique est la largeur de bande de bruit à partir de laquelle l'effet de masque commence à décroître.



Masque d'une bande étroite de bruit par deux sons purs

On augmente progressivement la distance entre les deux fréquences des sons purs, masquant une **bande de bruit étroite**.



Observation : le seuil masqué varie peu tant que la distance entre les deux sons masquants reste inférieure à la valeur "critique". Au delà de celui-ci, le seuil décroît rapidement. Si les sons masquants sont assez éloignés, ils ne fournissent plus suffisamment d'énergie au filtre auditif centré sur le son masqué, et de ce fait, ne participent plus au masquage de ce dernier.

(Cette expérience peut servir à la mesure des bandes critiques.)

Application des phénomènes de masquage

la compression audio (MP3, ...)

- 1.) On utilise des **filtres** pour diviser le signal audio (échantillonné à 48 kHz, par exemple) en **sous-bandes** correspondant approximativement aux 32 bandes critiques.
- 2.) On détermine le niveau de masquage pour chaque bande causé par des bandes voisines utilisant un modèle psychoacoustique.
- 3.) Si la puissance d'une bande est sous le niveau de masquage, on ne l'encode pas.
- 4.) Sinon, on détermine le nombre de bits nécessaire pour représenter le coefficient, de sorte à ce que le bruit introduit par la quantification est sous le niveau de masquage.
(Rappel : 1 bit de quantification introduit 6 dB de bruit).
- 5.) On forme le flux de bits (bitstream)

Le masquage temporel

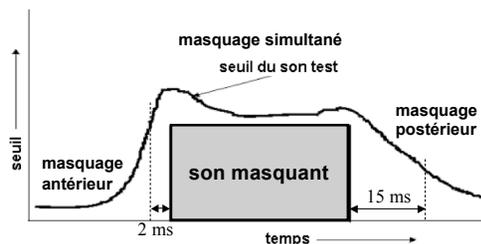


Diagramme schématisant le changement relatif du **seuil d'audition** du signal en fonction de la position temporelle du signal test par rapport au son masquant.

Le masquage temporel

Masque postérieur

Phénomène par lequel un son masquant cache un son masqué entendu postérieurement:

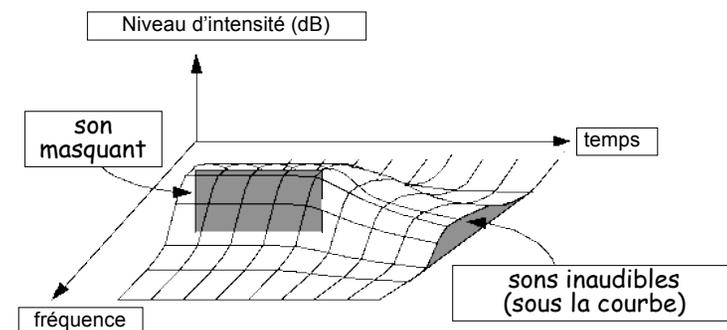
- **De 0 à 3 ms** : la courbe d'effet de masque demeure inchangée.
- **De 3 à 200 ms** : celle-ci diminue progressivement.
- **Au-delà de 200 ms** : l'effet de masque a disparu.

Application en musique : les notes de musique ont une durée moyenne de 100 ms. Les attaques sont perçues plus nettes quand la durée entre deux notes est plus grandes.

Masque antérieur

Phénomène par lequel un son masquant cache un son masqué entendu antérieurement (surprenant à première vue !). En effet l'oreille a besoin d'un certain temps pour réagir. Si le son antérieur est faible par rapport au son masquant, la sensation n'aura pas le temps de se former et son énergie sera intégrée à celle du stimulus masquant.

Le masquage simultané (selon la fréquence) et temporel (selon le temps)



Induction auditive

Ce phénomène a d'abord été découvert dans le domaine de la perception de la parole en présence de bruits, puis a été généralisé.

Dans le cas de la parole, si la phrase a du sens, on "entend" les syllabes qui sont pourtant cachées par du bruit.

S'il y a, de par le contexte, une probabilité qu'un son puisse être présent à un instant donné, et si certaines des unités périphériques excitées par un son plus fort incluent celles qui le seraient par le son plus faible anticipé, alors ce son plus faible peut être entendu comme s'il était effectivement présent.

Influence de la durée sur la sonie

La sonie d'un son ne dépend pas QUE de son niveau. Sa durée a aussi une influence. Pour les seuils de détection et les échelles vus précédemment, les sons avaient une durée d'au moins 250 ms. En effet, **pour des sons de 250 ms et plus, la sonie est indépendante de la durée**. Par contre, **pour des sons de plus courte durée, la sonie est proportionnelle à la durée** : elle décroît lorsque la durée décroît.

Ainsi, à intensité équivalente, un son de 20 ms est jugé moins fort qu'un son de 40 ms. On en déduit qu'une impulsion de 20 ms doit être de 10 dB plus intense (x 10) qu'une impulsion de 200 ms pour lui être isosonique. L'**énergie** (puissance x durée) doit donc être maintenue **constante** pour une sonie constante, dans le cas des sons très courts.



La valeur de durée limite au-delà de laquelle la sonie n'augmente plus avec l'augmentation de la durée, qu'on appelle la **constante d'intégration de l'oreille**, a été établie à **250 ms** par Yost, mais d'autres auteurs donnent les valeurs de **500 ms** ou encore **100 ms**.

Pour des sons très courts, la sonie dépend donc de la **durée** mais aussi de la **fréquence de répétition** des sons (à ne pas confondre avec la fréquence du son lui-même).

Démonstrations sonores :

- **continuité apparente (n° 28) → illustration du seuil de pulsation**



- **continuité perceptuelle d'un glissando interrompu (n° 29)**



- **l'effet "barrière" pour le cas de la parole (n° 31)**



Un son entendu au moins 200 fois par seconde (200 Hz) sera considéré isosonique à un son continu.

Par contre, pour des fréquences de répétition comprises entre 20 et 200 Hz, l'impulsion sera **isosonique** à la référence **si le produit de son intensité et de sa fréquence ($I \times f$) est constant**. Pour une fréquence 10 x plus petite, l'intensité devra être multipliée par 10 (+ 10 dB). Ainsi, les deux sons ci-dessous ont la même sonie :



En-deçà d'une durée de 100 ms et d'une fréquence de répétition de 200 Hz, la sonie est proportionnelle à la durée et à la fréquence de répétition. La sonie sera conservée constante **si le produit de la durée et la fréquence de répétition ($d \times f$) est constant**. Pour une fréquence 2 x plus grande, la durée devra être 2 x plus petite pour maintenir la sonie constante. Ainsi, les trois sons ci-dessous ont la même sonie :

