

L'adaptation au local du système de reproduction par la méthode Physiotemporelle™

OU L'INFLUENCE DU LOCAL SUR L'ÉCOUTE

par Patrick Thévenot

L'objectivité d'une écoute est déterminée non seulement par les caractéristiques des enceintes (courbes de réponse et de phase dans chaque direction, distorsion, rapidité d'écoulement d'énergie / au temps, ...), mais aussi par le local d'écoute, ainsi que par le *couplage physique et psychoacoustique* de ceux-ci.

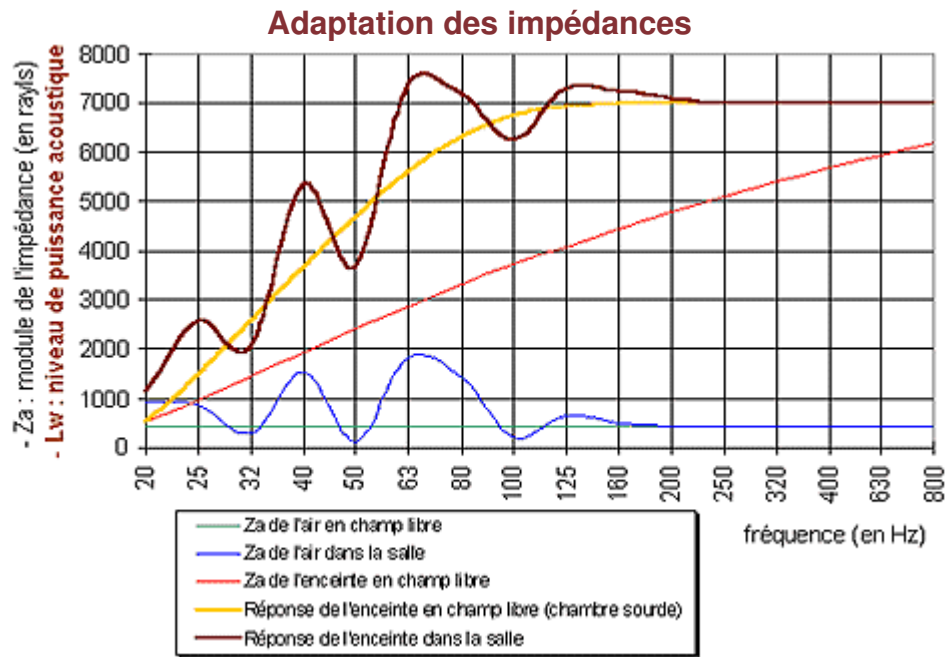
Il faut dissocier *l'influence dans le grave* (< 200 Hz) du local d'écoute que nous qualifierons de couplage physique, puisque l'énergie que restitue l'enceinte à l'emplacement d'écoute est *réellement modifiée* et *l'influence au-delà de ces fréquences* (de 200 Hz à 20 kHz), où le son direct fourni par l'enceinte reste le même, alors que le local nous le fait *percevoir différemment*. Il s'agit alors d'un couplage psychoacoustique.

COUPLAGE PHYSIQUE AU LOCAL DANS LE GRAVE (< 200 Hz)

Dans cette partie du spectre, et compte tenu des longueurs d'onde grandes devant la distance d'écoute, nous sommes incapables de dissocier en timbre le son provenant du local et celui provenant de l'enceinte.

Par conséquent, la réponse globale dans cette plage doit être la plus régulière possible. Or, si le local n'est pas homogène, les ondes stationnaires qui se forment, déterminent des impédances acoustiques très variables en fonction de la fréquence et du lieu. L'adaptation d'impédance acoustique entre l'enceinte (générateur) et l'air du local (récepteur) varie donc beaucoup et la courbe de réponse de l'enceinte devient très irrégulière (même si celle-ci est "droite" en champ libre !).

La régularité de la réponse est ainsi déterminée par les qualités acoustiques du local et de l'enceinte, ainsi que de leur couplage (voir courbe ci-après).



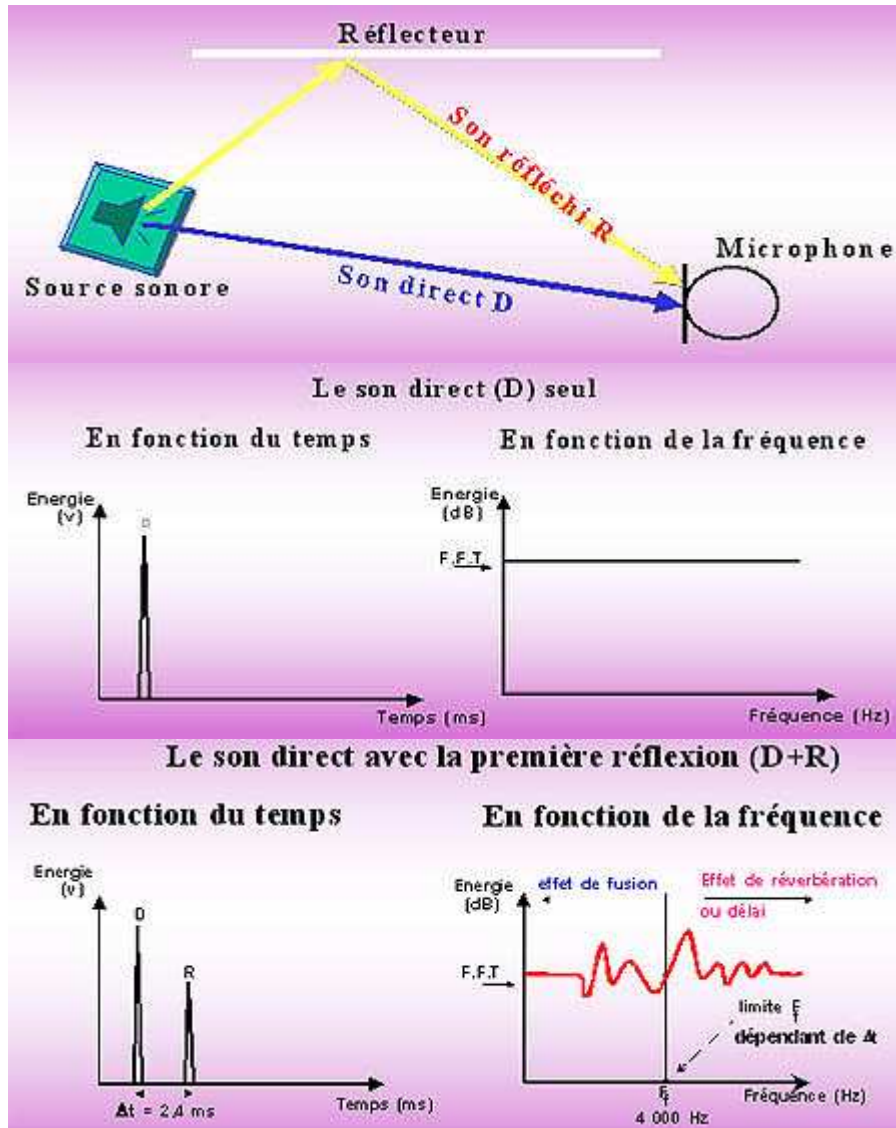
A défaut de trouver des solutions acoustiques (suppression des ondes stationnaires), *une égalisation classique mais prudente* peut être envisagée à l'endroit d'écoute, en prenant pour règle qu'un excès d'énergie (bosse) peut toujours être 'raboté', mais qu'un manque d'énergie (trou) ne peut souvent pas être compensé électriquement puisque le traitement n'aura pour d'autre effet que d'exciter encore plus l'onde stationnaire et donc, d'augmenter dangereusement la distorsion.

Une solution souvent intéressante consiste à encastrer les caissons de grave, de manière à éviter des filtres en peigne dus aux réflexions à l'arrière des enceintes. De là, deux choix sont possibles :

- soit toute l'enceinte est encastrée, et on doit s'assurer de la parfaite **neutralité du bafflage** prolongeant la partie médium-aigu, afin d'éviter les irrégularités de rayonnement dans les premières millisecondes pour ces fréquences.
- soit seul le grave est encastré et on doit retarder le satellite médium-aigu, afin de conserver **une réponse impulsionnelle et une phase correctes**.

COUPLAGE PSYCHOACOUSTIQUE DU BAS-MEDIUM A L'AIGU (> 200 Hz)

L'influence des premières réflexions (avant 20 ms)



On constate, aux mesures, des déformations de la courbe de réponse dues au retard entre le son réfléchi R et le son direct D.

Le problème est de savoir *si nous percevons ou pas ces déformations* et de quelle manière nous les interprétons.

Une étude psycho-acoustique portant sur 150 personnes a été menée à l'INA dans le cadre des stages de la formation professionnelle. Celle-ci a montré que seule la partie de la courbe située en dessous d'une fréquence particulière appelée *fréquence de fusion (F_f)* est réellement significative de notre perception.

Dans cette plage, nous ne pouvons pas dissocier le son réfléchi du son direct : *il y a fusion* entre les deux.

Au delà de cette plage, nous percevons distinctement les deux et la perturbation mesurée n'est alors pas perçue. Il y a l'effet bien connu de *réverbération*. Celle-ci modifie certains paramètres, comme la transparence et la localisation en profondeur de la source, mais pas le timbre.

La fréquence F_f varie inversement avec le retard D_t entre le son direct et le son réfléchi (Ceci est du à la constante de temps de l'oreille qui varie avec la fréquence).

Taylor Made Systems a établi une courbe spécifique permettant de quantifier le phénomène et donc de le corriger.

Grâce à cette courbe, on peut définir *objectivement la neutralité d'une écoute* à condition de tenir compte de cet effet de fusion par une *égalisation méthodique dite Physiotemporelle® effectuée seulement et uniquement dans la zone de fusion*.

Par conséquent, *ni l'enceinte, ni le local d'écoute ne doivent ajouter quoi que ce soit dans cet intervalle de temps* (surtout aux fréquences élevées), *sinon il y aura dénaturation de la source sonore enregistrée* et donc mauvaise analyse de l'enregistrement (*même si l'effet de ce défaut paraît agréable, voir ajoutant une précision plausible !*!).

Les réflexions qui ne peuvent être supprimées par traitement acoustique (comme par exemple l'énergie réfléchie par la console) ne peuvent être égalisées que par méthode Physiotemporelle®, sélective en fonction du temps et de la fréquence, afin de tenir compte de cet effet de fusion (et surtout pas avec un bruit rose ou par FFT classique !).

L'influence du champ réfléchi diffus (après 20 ms)

Cette énergie (entre 200 Hz et 20000 Hz) est perçue comme provenant à part entière du local et provoque un *effet de masque* d'autant plus gênant que celle-ci sera importante et se situera plus haut dans le spectre.

Néanmoins, elle doit demeurer suffisante afin d'assurer un confort acoustique ambiant naturel, nécessaire à une compatibilité du mixage à l'écoute en milieu domestique (chez le client).

Un écart de niveau moyen de 5 à 10 dB (suivant la fréquence) est acceptable entre l'énergie avant 20 ms et l'énergie après 20 ms, pour que celle-ci ne soit pas gênante. Cette mesure peut être matérialisée par l'indice de clarté C20, dont la formule est :

$$C20 = 10 \log \left(\int_0^{20ms} p(t).dt / \int_{20ms}^{\infty} p(t).dt \right) .$$

Ce rapport d'énergie doit être stable avec la fréquence, sinon il peut en résulter une *modification de l'équilibre tonal subjectif* de la source. En effet, un effet de masque plus important à certaines fréquences provoque une imprécision et donc une impression de diminution de niveau à ces mêmes fréquences. Inversement, un effet de masque moins important engendre un son plus précis et " incisif " qui sera donc subjectivement perçu comme plus puissant. Une variation de cet effet de masque en fonction de la fréquence peut donc engendrer une variation du timbre.

L'étude de la courbe du *C20 en fonction de la fréquence* permet de mettre en évidence ce phénomène.

Avec des *enceintes à directivité croissante et contrôlée*, on peut corriger cet effet en "jouant" sur l'orientation de l'enceinte par rapport à l'auditeur. On peut alors changer l'équilibre tonal du son direct indépendamment du son réfléchi (lié à la puissance acoustique restituée par le couple "enceinte-local", qui ne dépend que très peu de l'orientation). Dans ce cas, on modifie réellement la courbe C20 en fonction de la fréquence.

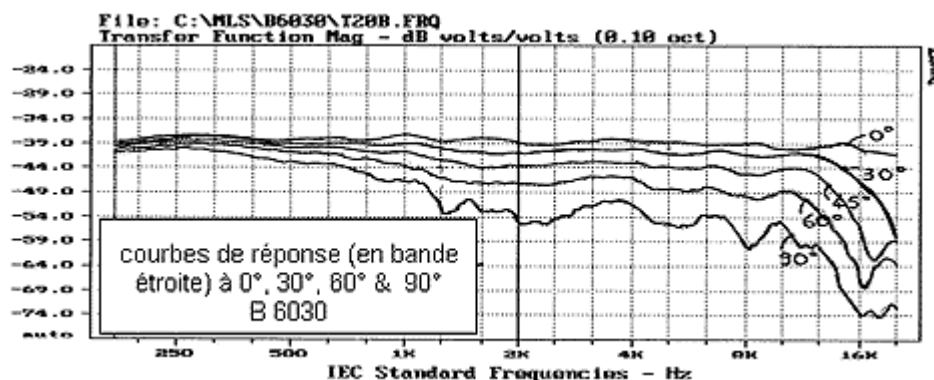
Avec des *enceintes à directivité constante*, on ne peut compenser cet effet que par une égalisation globale, en modifiant la puissance acoustique que l'enceinte fournit au local. Dans ce cas, le C20 reste identique et on essaye de palier au déséquilibre 'subjectif' par un déséquilibre 'objectif', ce qui ne n'est pas rationnel, car pas de la même nature du point de vue de l'audition.

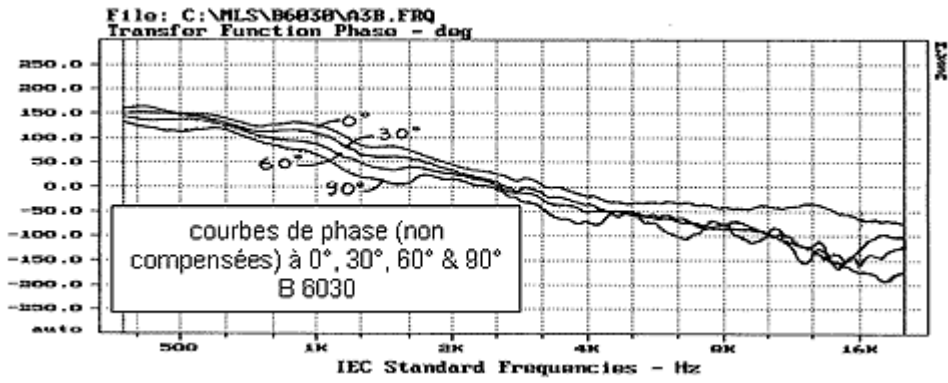
Ces 2 solutions ne sont pas équivalentes.

Pour répondre à ces exigences, une solution pour les enceintes consiste à posséder des courbes de réponse non accidentées et présentant une *directivité régulièrement croissante* en fonction de la fréquence (ce qui assure une phase de rayonnement stable sur 360°). De plus, le bafflage doit permettre *un écoulement parfait de l'énergie de 0 à 10 ms* (condition nécessaire à l'analyse objective des premières réflexions... de l'enregistrement. Les mesures types ci-après illustrent ces performances.

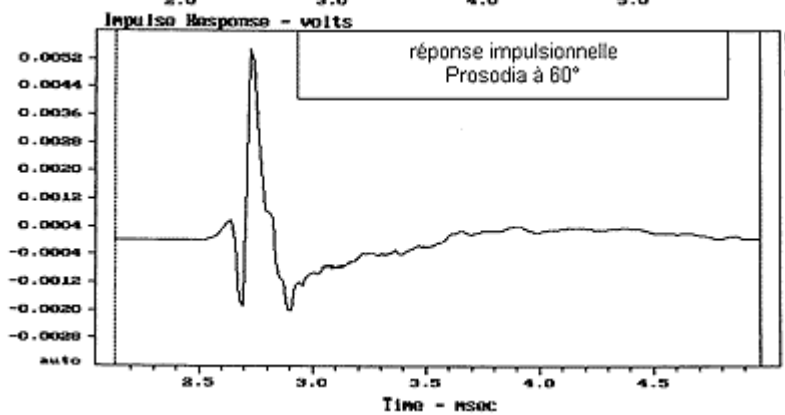
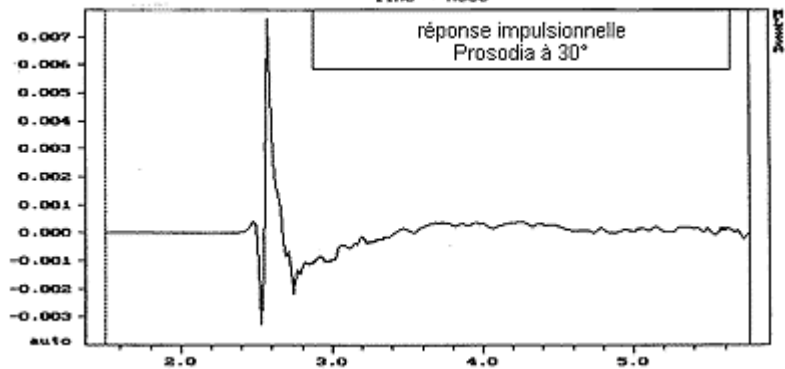
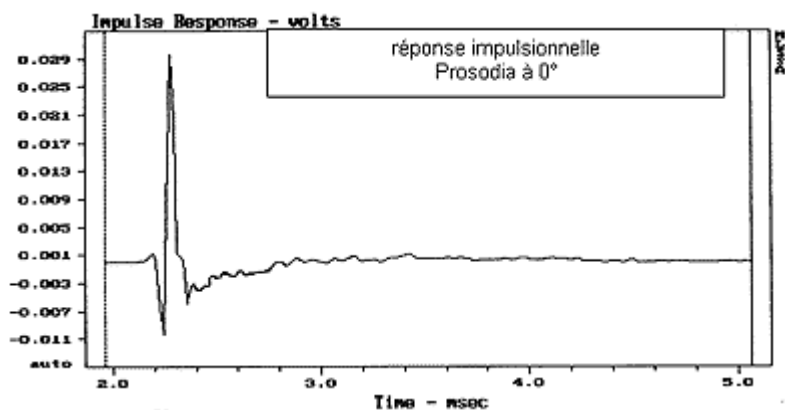
CAS CONCRETS ET ILLUSTRATIONS PRATIQUES

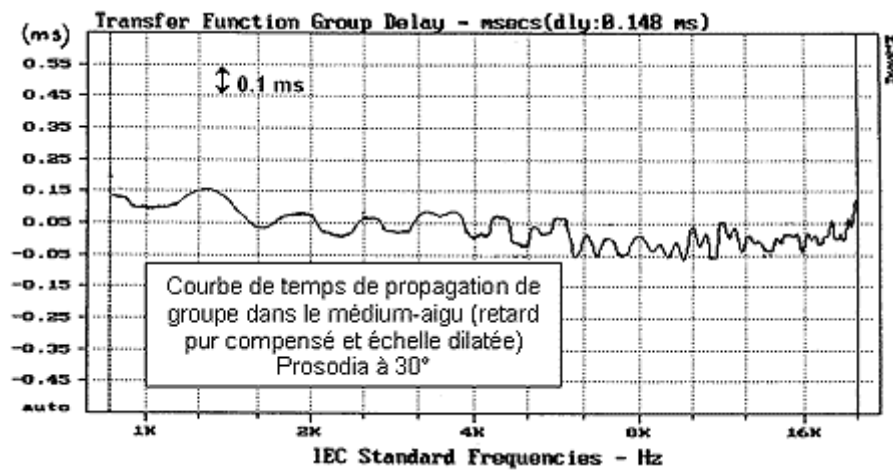
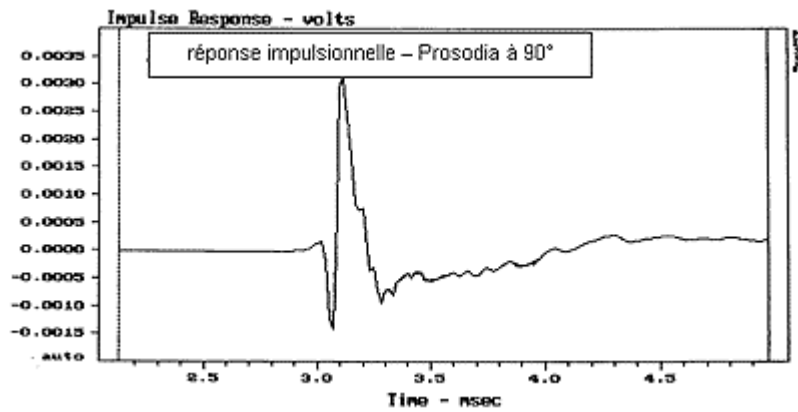
Les courbes de réponse en niveau et en phase d'une enceinte A2t Prosodia qui répond pleinement à ces exigences.





Réponse impulsionnelle à 0°, 30°, 60° et 90° et temps de propagation de groupe de la Prosodia:

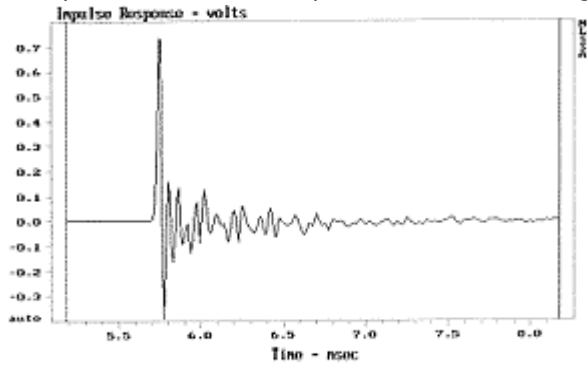




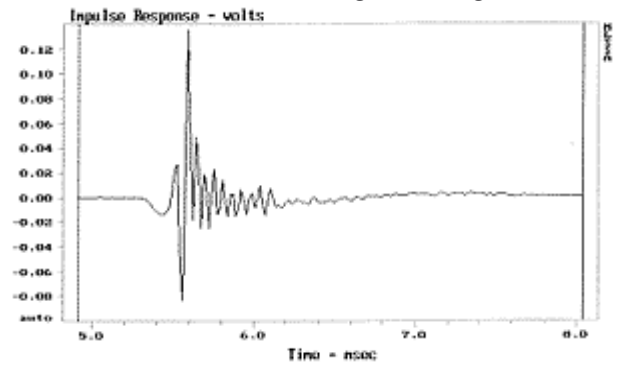
L'écoulement d'énergie très rapide autour des haut-parleurs ainsi que la régularité des courbes de réponse et de temps de propagation de groupe sont les conditions indispensables à la mise en œuvre optimale du processus d'égalisation et de réglage décrit ci avant.

Mesures d'enceintes réputées de qualité, mais ne répondant pas à ces critères :

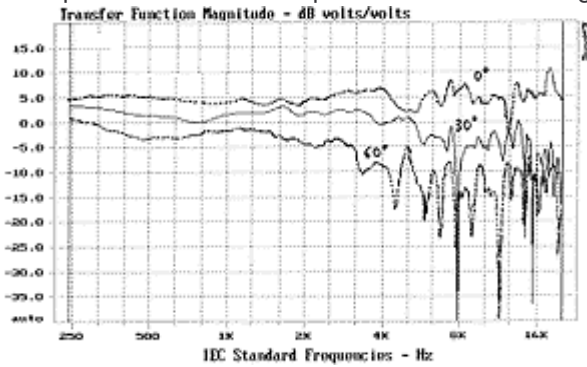
Réponse impulsionnelle
d'un panneau électrostatique utilisé en monitoring



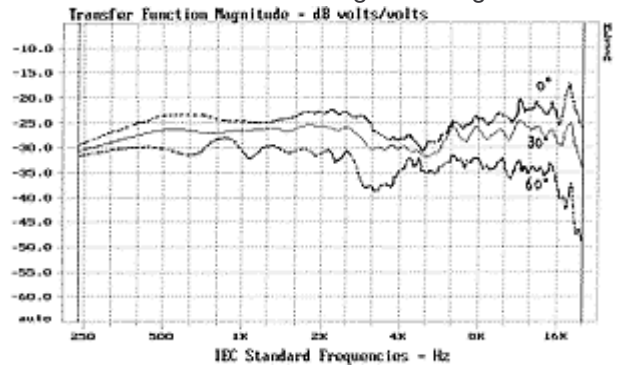
Réponse impulsionnelle
d'une enceinte monitoring haut de gamme



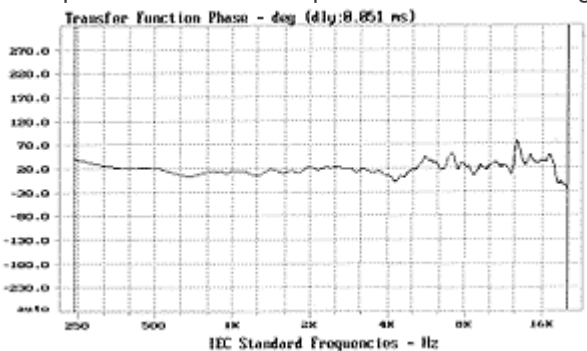
Courbes de réponse à 0°, 30° et 60°
d'un panneau électrostatique utilisé en monitoring



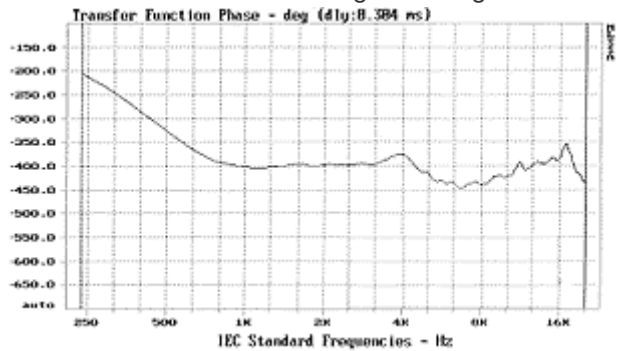
Courbes de réponse à 0°, 30° et 60°
d'une enceinte monitoring haut de gamme



Temps de propagation de groupe à 0°
d'un panneau électrostatique utilisé en monitoring

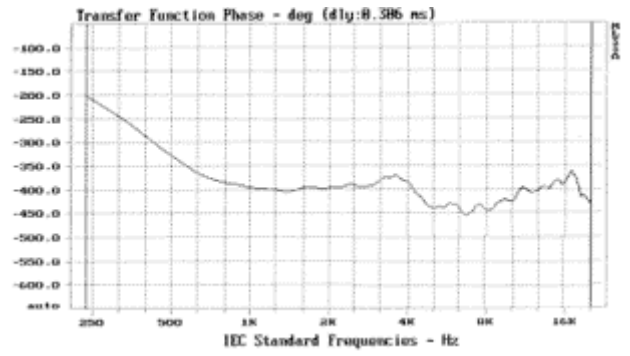
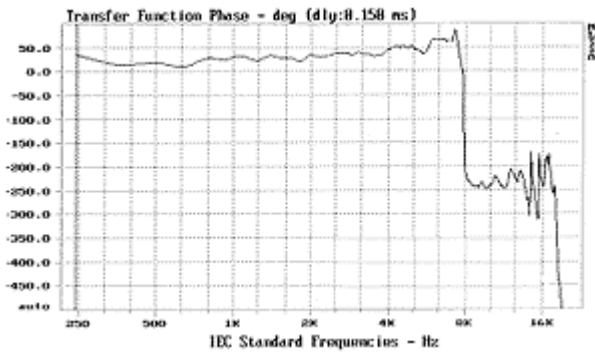


Temps de propagation de groupe à 0°
d'une enceinte monitoring haut de gamme

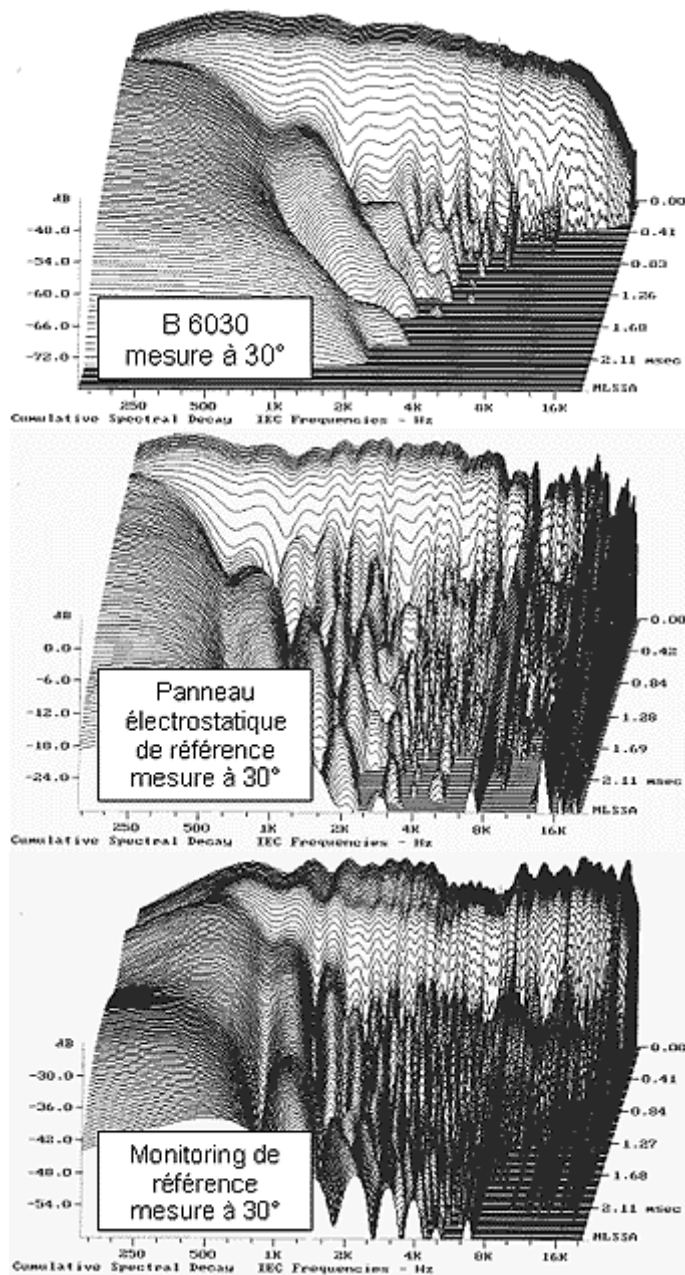


Courbe de phase à 30°
d'un panneau électrostatique utilisé en monitoring

Courbe de phase à 30°
d'une enceinte monitoring haut de gamme



Comparaison de la rapidité d'écoulement des énergies de différentes enceintes :



Le "traînage" commun à toutes les mesures dans les basses fréquences résulte d'une fenêtre de mesure se réduisant de plus en plus par rapport à la longueur d'onde et n'est par conséquent, pas significatif.

Ce phénomène de "dilution" dans le temps de l'énergie des fréquences élevées est perçu par l'oreille comme une augmentation agréable de la précision en localisation et en timbre de la source sonore enregistrée. Celui-ci peut être identique à l'effet produit par des réflecteurs acoustiques situés tout près des sources réelles enregistrées ou par des instruments de musique acoustiques enregistrés d'une manière naturelle (à l'aide d'un couple stéréophonique, par exemple). Si on prend le cas d'une voix enregistrée en monophonie et en proximité, celle-ci apparaîtra plus précise et plus définie bien que cet effet de précision soit dû à une imprécision du système d'écoute.

Cette imprécision, bien qu'agréable, produit un effet de masque empêchant l'analyse objective de la prise de son et du mixage.

La contrepartie à utiliser des enceintes dont l'écoulement d'énergie est très rapide, se traduit par une **interdépendance d'autant plus grande avec le local d'écoute** dont l'effet est de ce fait, moins masqué.

Le positionnement géométrique du système, le traitement acoustique et l'égalisation doivent être d'autant plus soignés et pris en compte, suivant l'analyse décrite précédemment.