b<>com *Spatial Audio Toolbox* [Render Spk2Bin]

Bernard Lagnel
Décembre 2019

https://b-com.com/fr/bcom-spatial-audio-toolbox-render-plugins

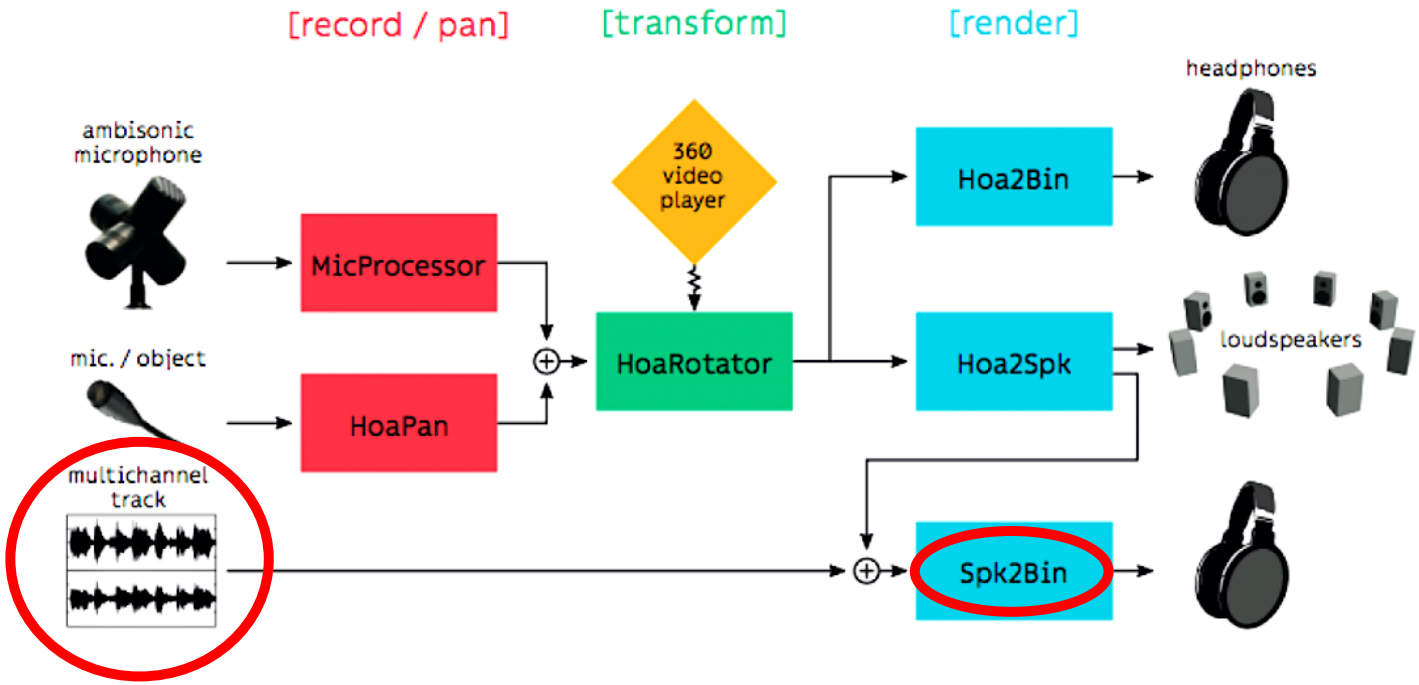
b<>com *Spatial Audio Toolbox*

Partager https://b-com.com/fr/bcom-spatial-audio-toolbox-render-plugins

Plugins pour le rendu de contenu audio immersifs

Le vif intérêt du grand public pour les expériences immersives est une des principales tendances de ces dernières années dans l'industrie du divertissement. **Dans ce contexte, l'audio a un rôle majeur à jouer** et de nouvelles techniques telles que l'ambisonique, l'audio orienté-objet, le head-tracking et la binauralisation contribuent de manière significative à son potentiel d'immersion.

Applications



b<>com présente un ensemble de plug-ins audio permettant la création de contenus et d'expériences véritablement immersifs. Les plug-ins **b<>com *Spatial Audio Toolbox*** s'appuient sur la technologie HOA (Higher Order Ambisonics), pour fournir un flux de travail complet, de la capture au rendu dans un casque ou sur haut-parleurs. Le HOA, devenu la référence audio pour la réalité virtuelle et augmentée, facilite la lecture de scènes sonores binaurales complexes avec head-tracking.

Fonctionnalités clés

- Haut-parleurs virtuels binauraux
- Standard VST ou plug-in AAX
- Base de données HRTFs
- HRTFs en conditions anéchoïques et réverbérés
- Jusqu'à 32 canaux d'entrée
- Cliquez pour isoler (Click to Solo)
- Interface graphique intuitive et conviviale

Bénéfices

- Écoutez votre mixage audio 3D avec un casque audio standard
- Travaillez avec votre station audionumérique préférée
- Choisissez les filtres qui vous conviennent le mieux
- Choisissez votre environnement virtuel
- Travaillez avec tous les principaux formats audio 2D et 3D
- Monitorez chaque canal de haut-parleur
- Commencez à créer immédiatement

Le pack [Render plugins] se compose de 3 différents plug-ins :

b<>com *Spatial Audio Toolbox* [Render Spk2Bin]

Le plug-in [Render Spk2Bin] restitue l'audio 3D basé sur les canaux dans un casque audio, en utilisant une approche de haut-parleur virtuel. Il comprend une base de données de fonctions de transfert relatives à la tête (HRTF) permettant de convertir les principaux formats audio 2D et 3D (22.2 etc.) en audio binaural. Chaque canal peut être sélectionné individuellement pour une surveillance rapide de configurations complexes. De plus, le plug-in inclut des HRTFs réverbérés, qui ajoutent un effet de pièce (room effect) réaliste et améliorent l'immersion.

b<>com *Spatial Audio Toolbox* [Render Hoa2Spk]

Le plug-in [Render Hoa2Spk] peut être inséré dans une station audionumérique en tant qu'effet spatial avec 4, 9, 16 ou 25 canaux d'entrée et 32 canaux de sortie. L'entrée est un signal HOA (Higher Order Ambisonics) qui peut atteindre l'ordre 4, et utilise la convention ACN/SN3D (ambiX). La sortie est un signal multi-canal adapté à la disposition des haut-parleurs sélectionnés. Plusieurs configurations de haut-parleurs en sortie sont supportées, de stéréo à 30,2.

b<>com *Spatial Audio Toolbox* [Render Hoa2Bin]

Le plug-in [Render Hoa2Bin] peut être inséré dans une station audionumérique en tant qu'effet spatial avec 4, 9, 16 ou 25 canaux d'entrée et 2 canaux de sortie. L'entrée est un signal Higher Order Ambisonics qui peut atteindre l'ordre 4, et utilise la convention ACN/SN3D (ambiX). La sortie est un signal binaural pour casque audio. Plusieurs ensembles d'HRTFs peuvent être sélectionnés, avec ou sans réverbération.

Merging et b<>com développent une solution complète pour améliorer l'immersion par le son

B-COM Render Plugin Pack w/ Spk2Bin, Hoa2Spk, & Hoa2Bin (Electronic Download)

SKU: MTPSPBCRNDER MFR: PSP-BC-RENDER

Download Version

Le pack :

Our Price

$768.00

As low as $128.00/mo for 6 months++

Add to Cart

+ wishlist

05 sep 2016

2016

Merging Technologies présente actuellement ses nouveaux produits à l'IBC, le salon international de référence pour l'industrie audiovisuelle à Amsterdam. Au programme : la toute dernière version de sa station digitale audio Pyramix qui intègre les plugins de spatialisation sonore de b<>com.

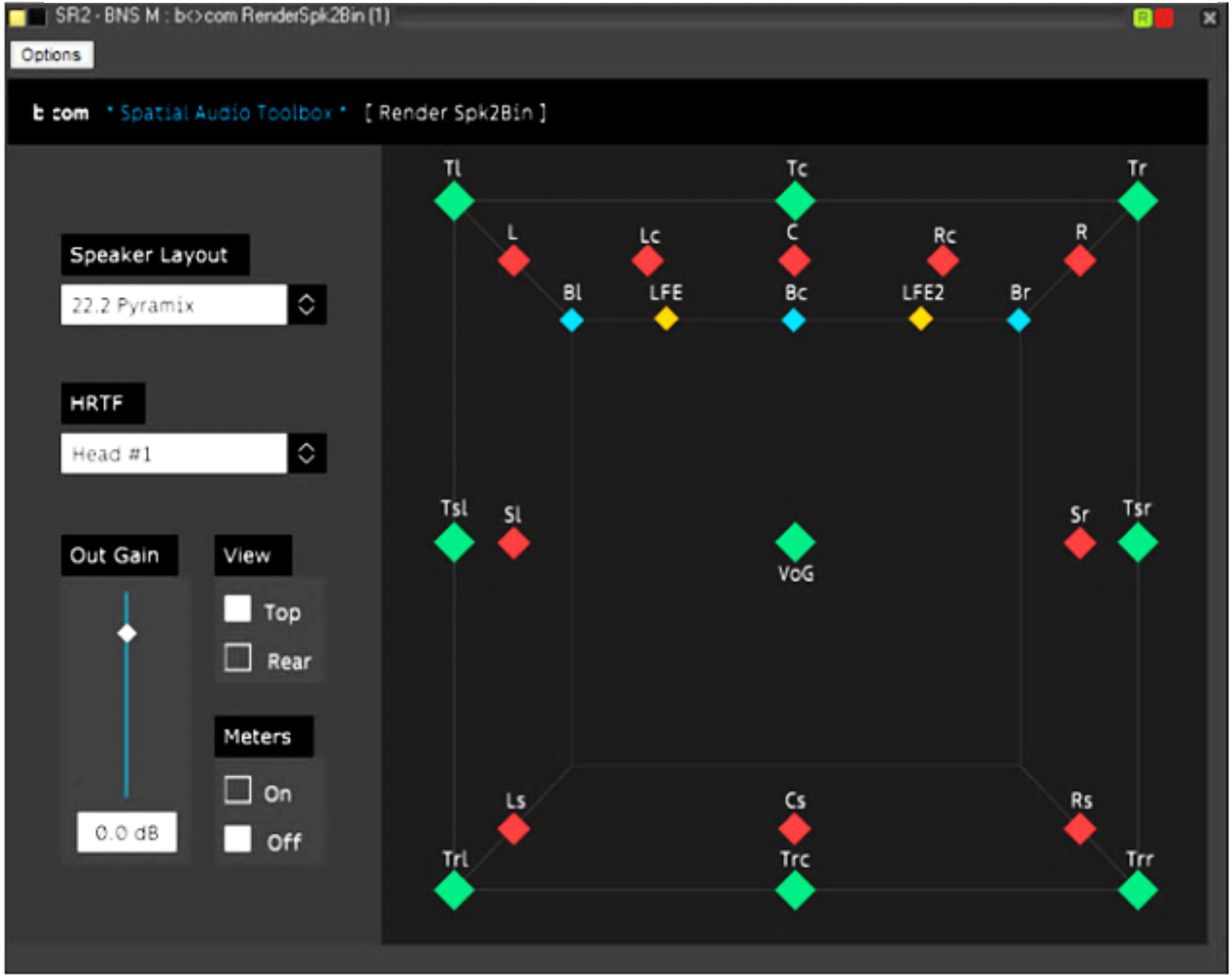
Depuis plus de 25 ans, Merging Technologies développe des solutions de mastering, production et post-production audio et vidéo pour les professionnels des médias et du divertissement. Fournisseur de stations audio et de convertisseurs audio sur réseau IP Horus et Hapi, **Merging s'est rapproché de b<>com, premier acteur sur le marché à offrir une solution professionnelle adaptée à la manipulation du format High Order Ambisonics**.

Afin de faciliter l'adoption de la technologie ambisonique par les professionnels, les équipes de b<>com ont développé une suite logicielle intégrable dans des stations de post-production. Elle permet aux créateurs de contenu de positionner des sources dans un espace sonore et d'y appliquer des traitements spatiaux (rotation, focus, masquage). **Ces plugins audio assurent l'essentiel des traitements depuis la captation jusqu'au rendu sur haut-parleurs ou sur casque (binaural)**.

« Cette suite logicielle s'inscrit parfaitement dans les outils que nous voulons mettre à disposition pour la production de bande son immersive. Elle a, par ailleurs, été développée conjointement avec des utilisateurs français de la station Pyramix et en particulier le panner 3D. Nos clients les studios apprécient de pouvoir manipuler facilement ce tout nouveau format audio grâce à des outils professionnels », explique Maurice Engler, Project Manager & Regional Manager chez Merging Technologies

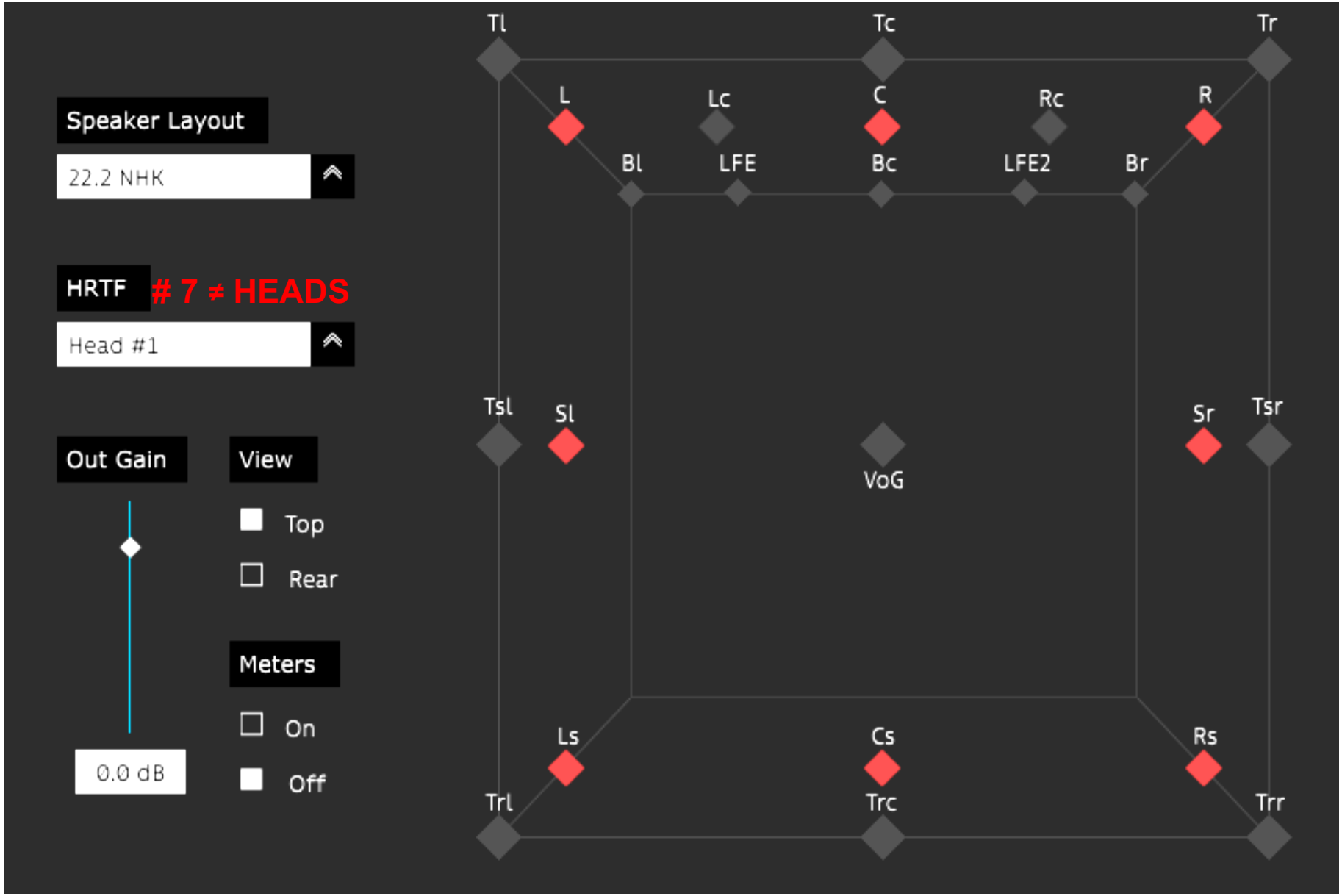
« La production de contenus audio au format ambisonique est un enjeu fort pour accompagner le développement de contenus immersifs, notamment en réalité virtuelle et vidéo 360. Avec cette suite logicielle, nous proposons une solution « professional grade » efficace et totalement intégrable à la plupart des stations de post-production audio du marché », complète Jean-Yves Aubié, Responsable du laboratoire Nouveaux Formats Media chez b<>com.

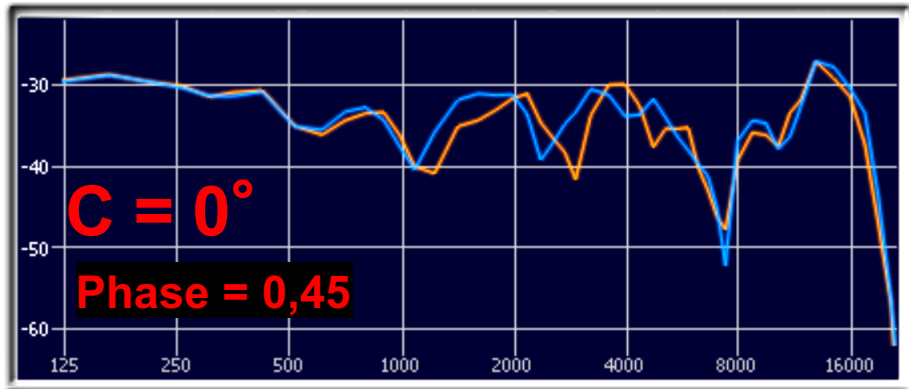
b<>com *Spatial Audio Toolbox* [Render Spk2Bin]

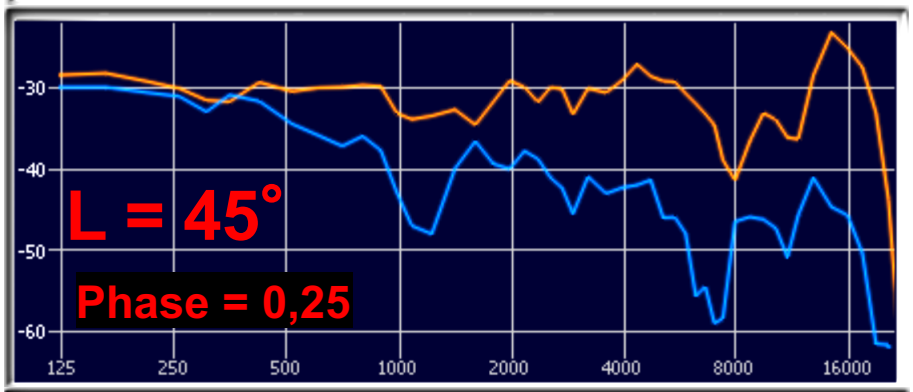


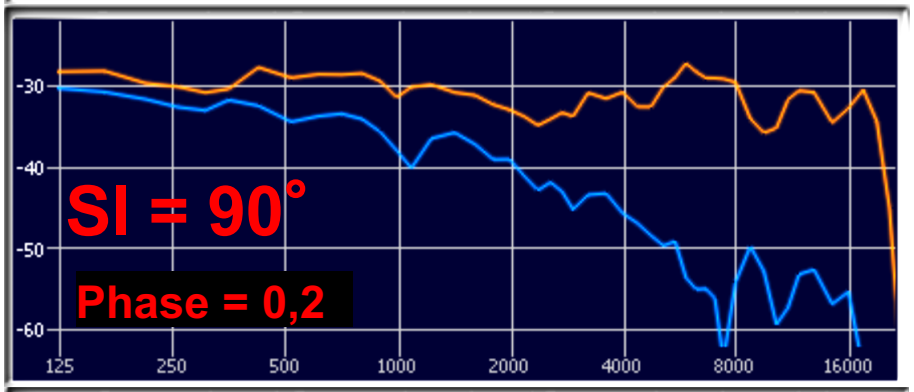
b<>com *Spatial Audio Toolbox* [Render Spk2Bin]

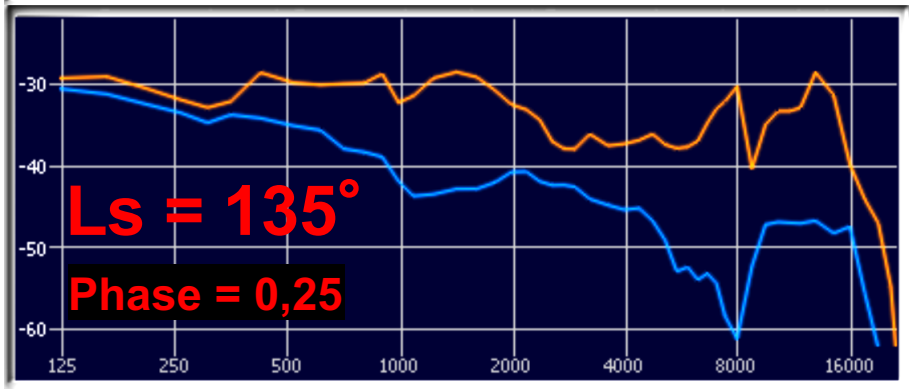
Mise en service d’un 8.0 à partir d’un 22.2 NHK :



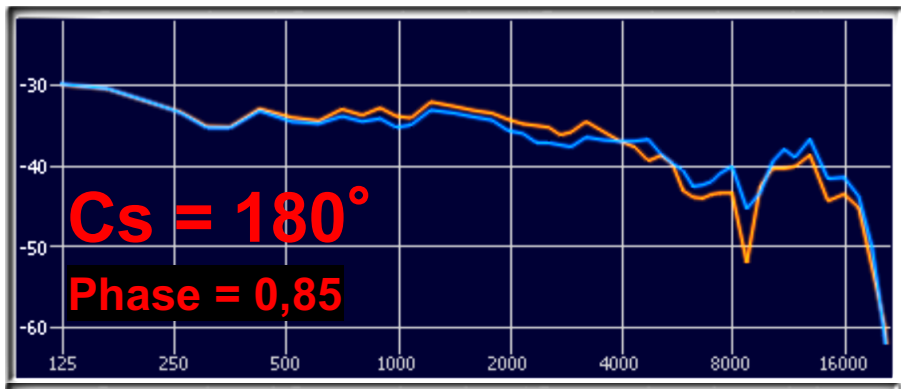

C = 0°
Phase = 0,45

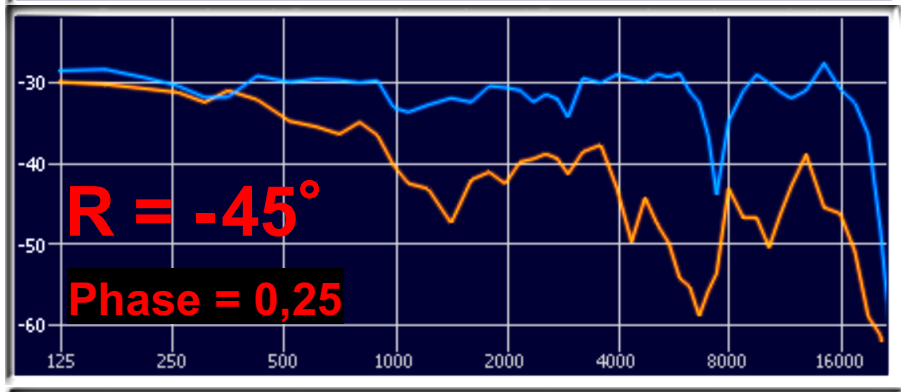

L = 45°
Phase = 0,25

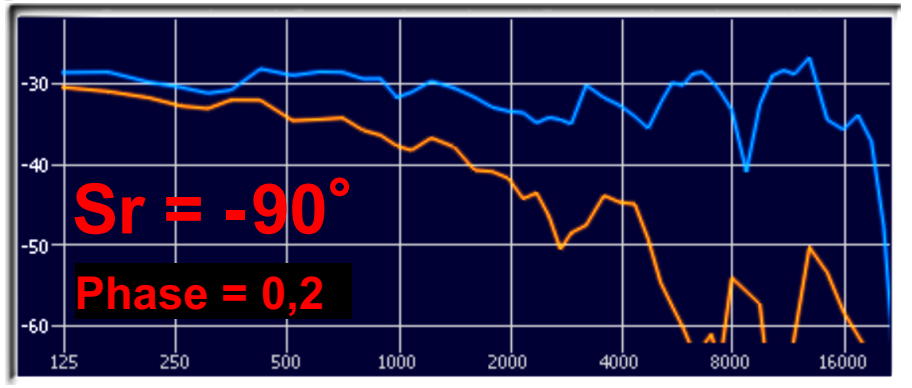

Sl = 90°
Phase = 0,2

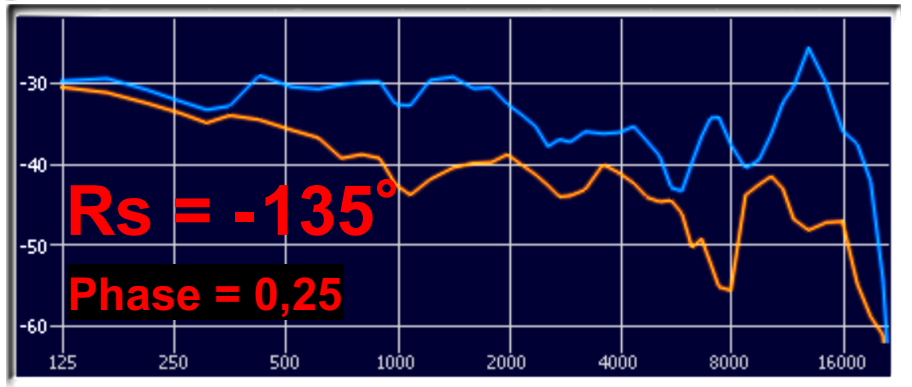

Ls = 135°
Phase = 0,25

HRTF HEAD #1

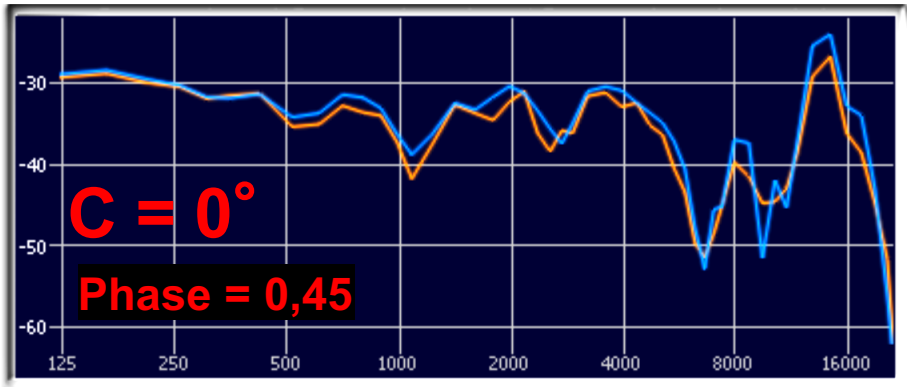

Cs = 180°
Phase = 0,85

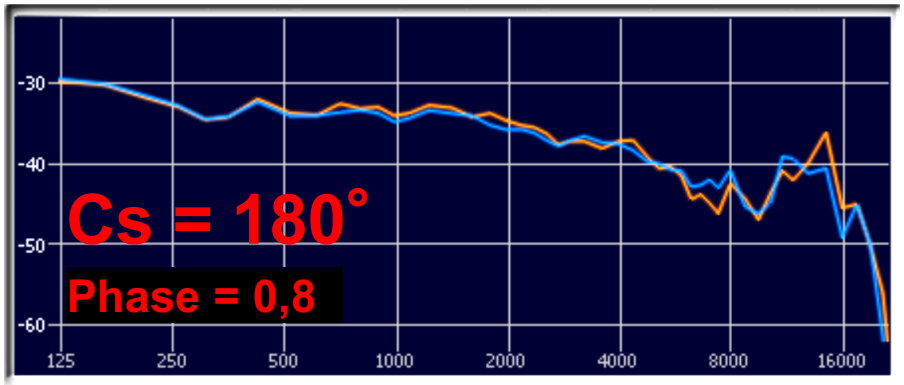

R = -45°
Phase = 0,25

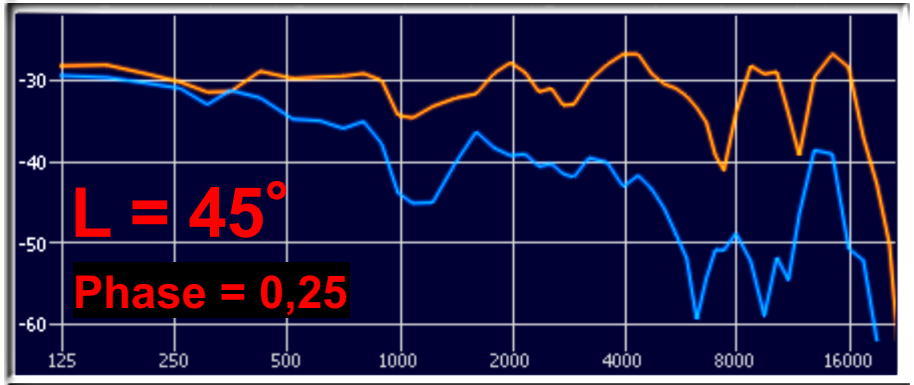

Sr = -90°
Phase = 0,2

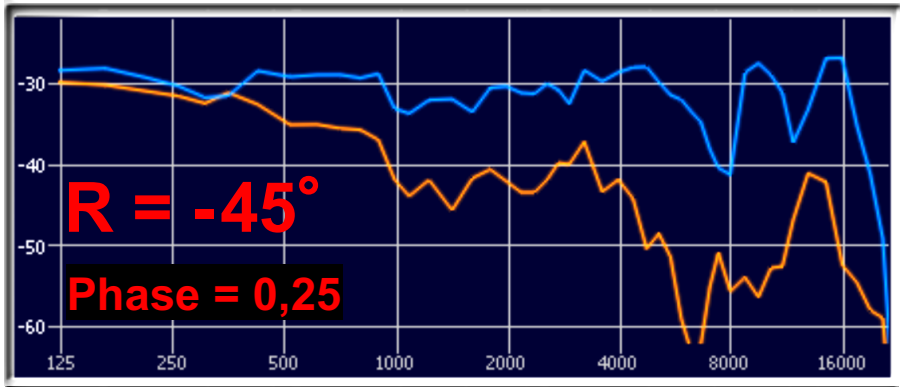

Rs = -135°
Phase = 0,25

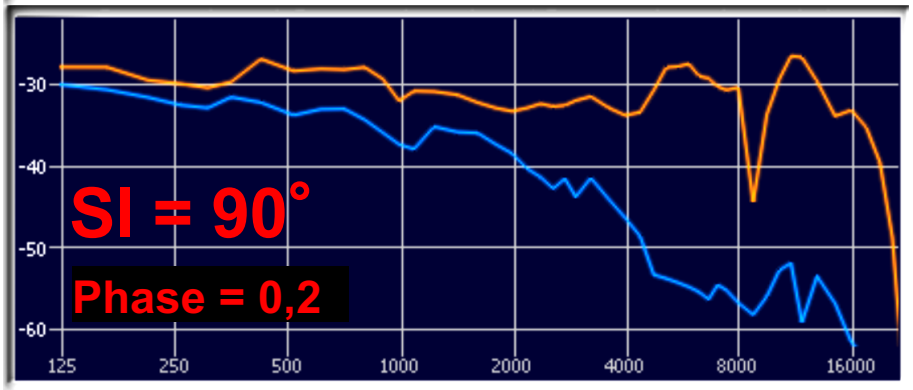
HRTF HEAD #2

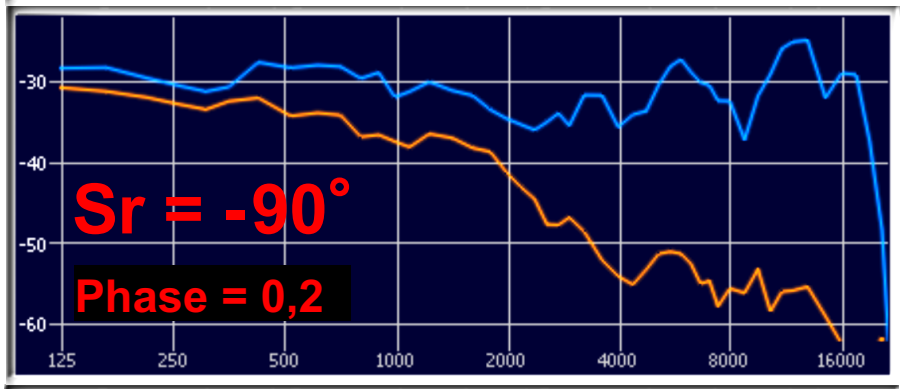

C = 0°
Phase = 0,45
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

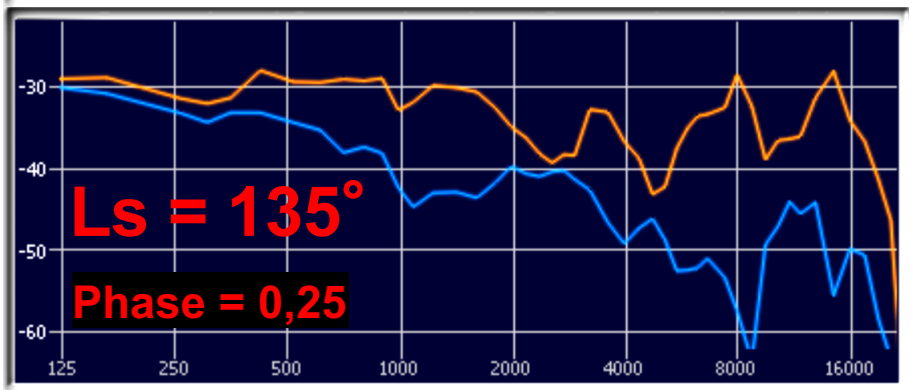

Cs = 180°
Phase = 0,8
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

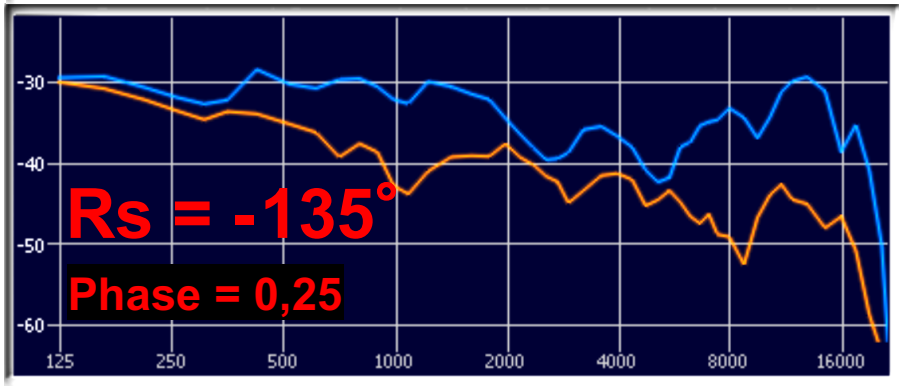

L = 45°
Phase = 0,25
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000


R = -45°
Phase = 0,25
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

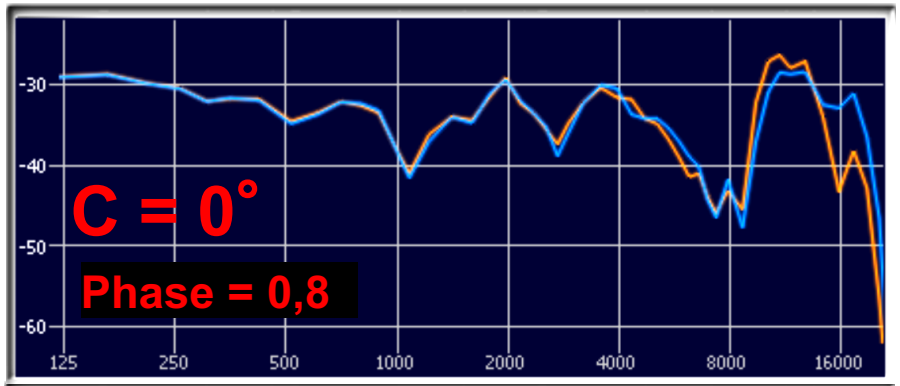

Sl = 90°
Phase = 0,2
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

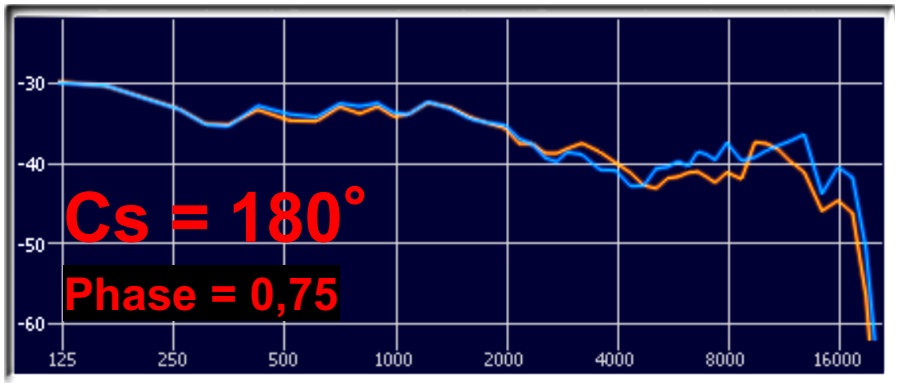

Sr = -90°
Phase = 0,2
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

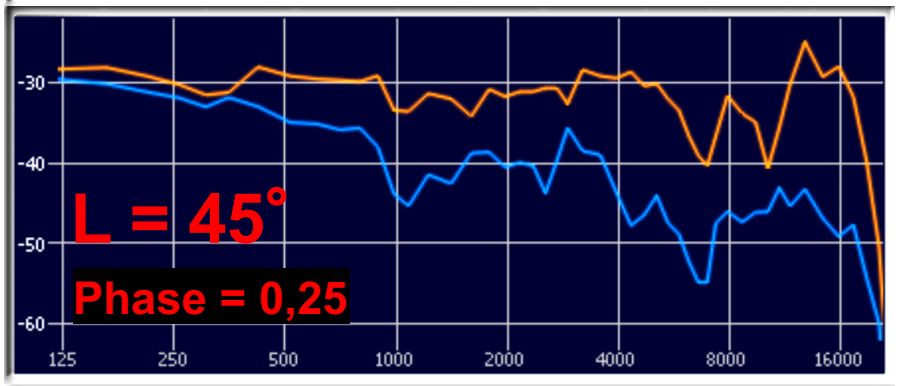

Ls = 135°
Phase = 0,25
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

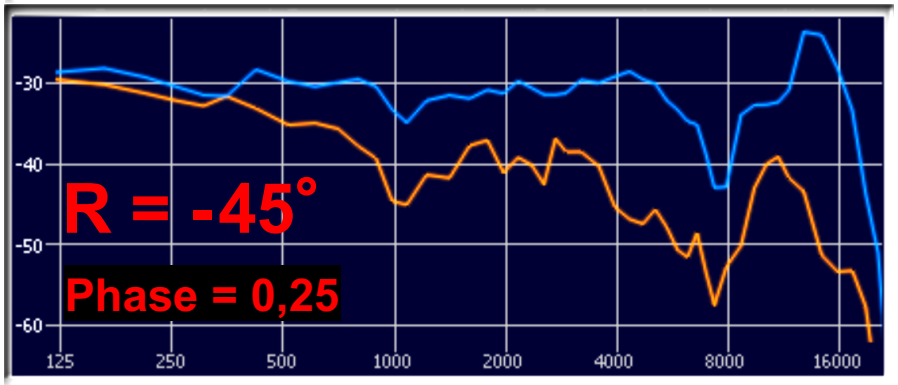

Rs = -135°
Phase = 0,25
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

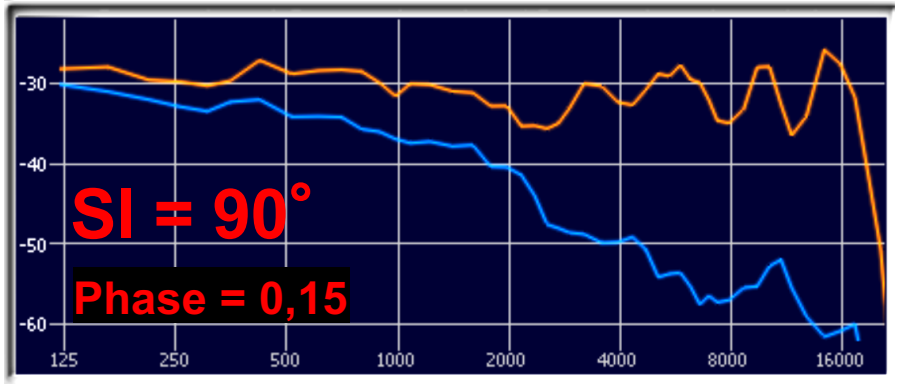
HRTF HEAD #3

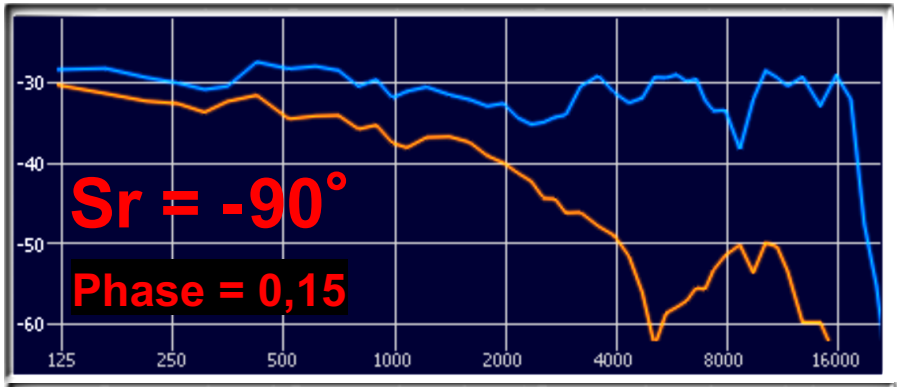

C = 0°
Phase = 0,8

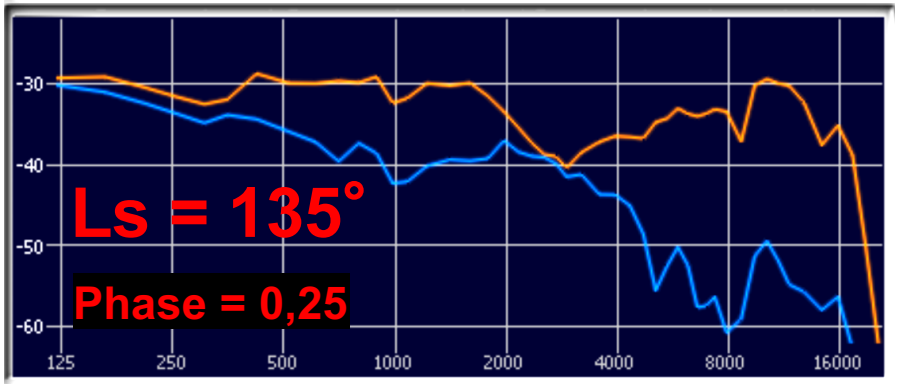

Cs = 180°
Phase = 0,75

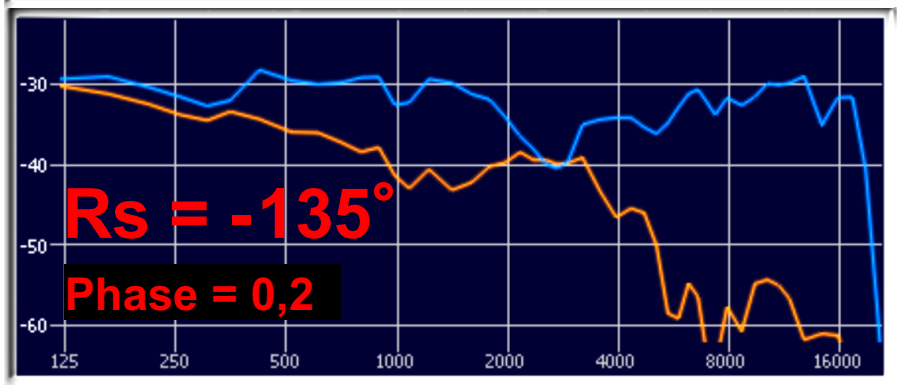

L = 45°
Phase = 0,25


R = -45°
Phase = 0,25

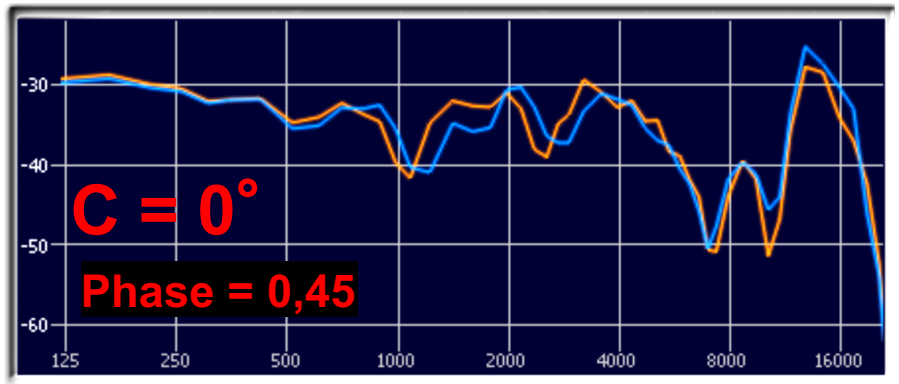

Sl = 90°
Phase = 0,15

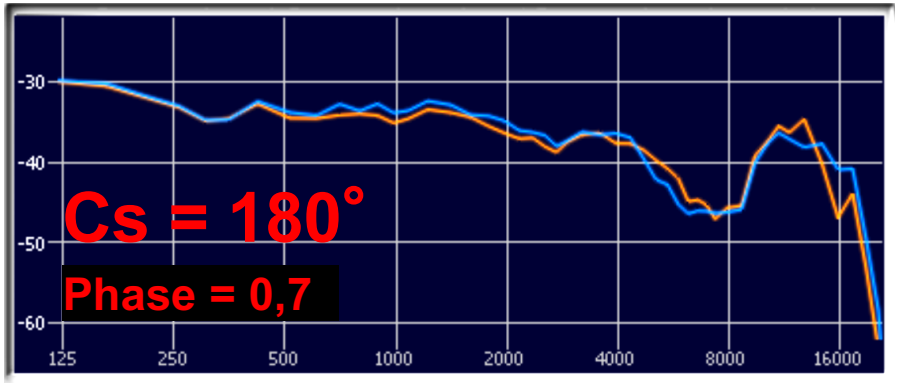

Sr = -90°
Phase = 0,15

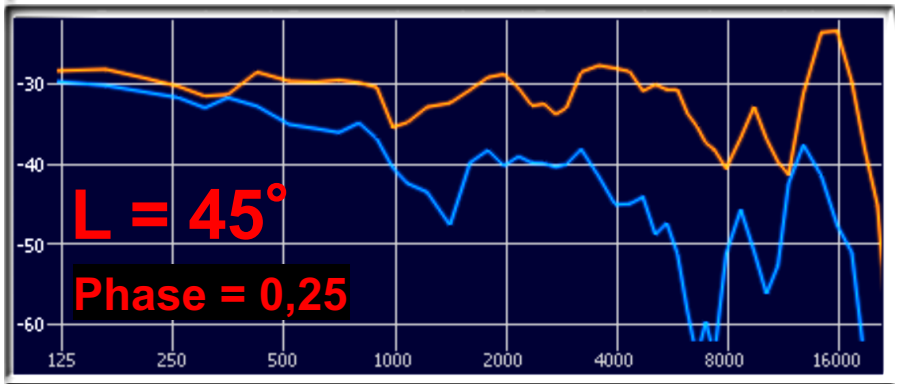

Ls = 135°
Phase = 0,25

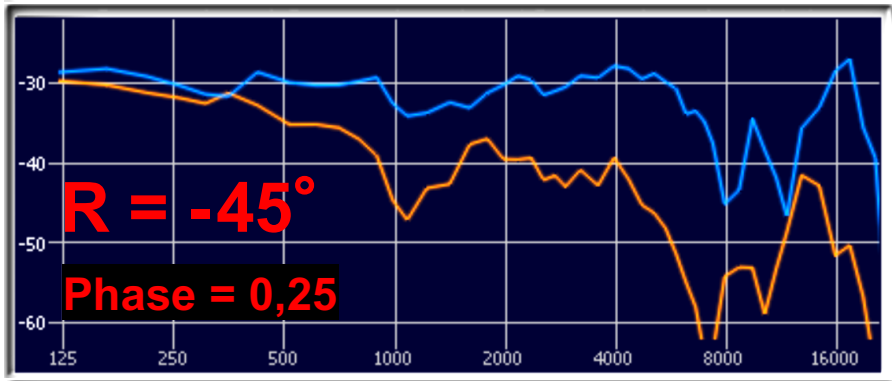

Rs = -135°
Phase = 0,2

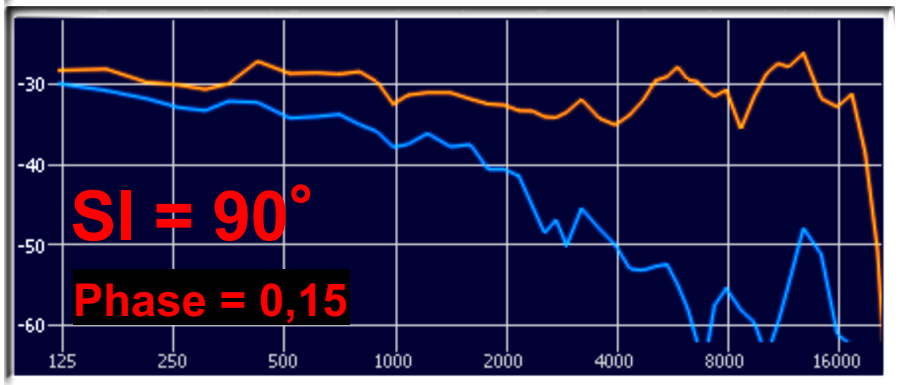
HRTF HEAD #4

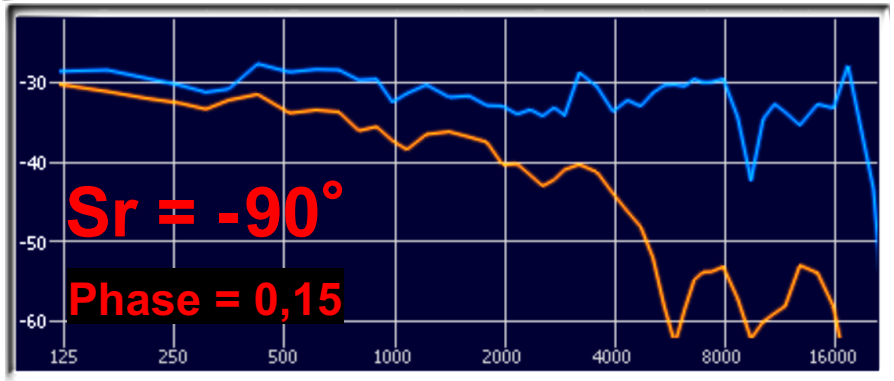

C = 0°
Phase = 0,45

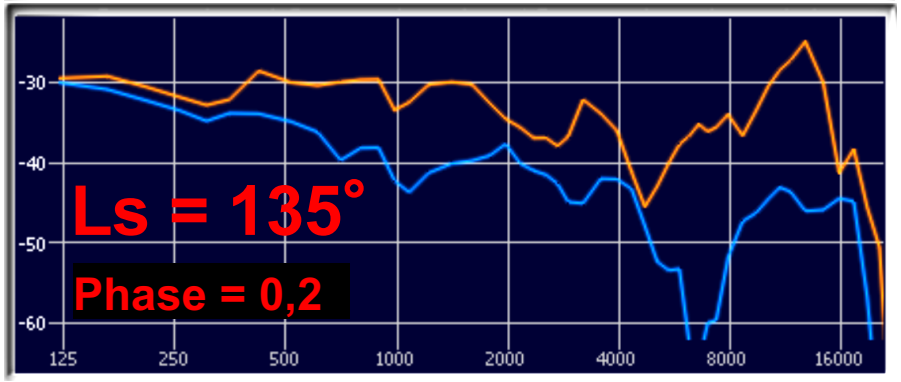

Cs = 180°
Phase = 0,7

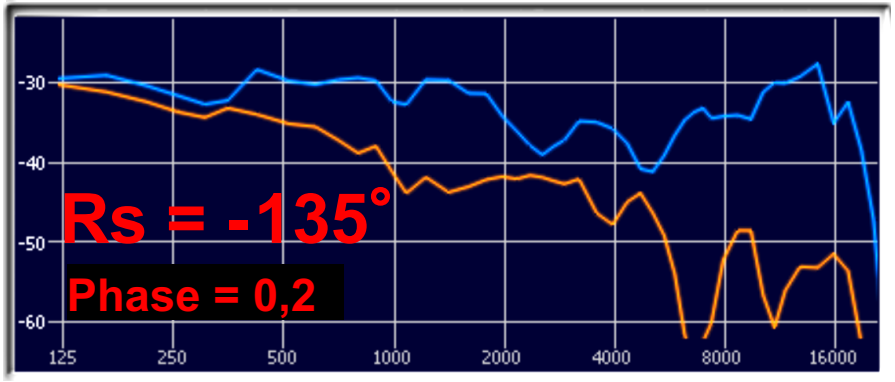

L = 45°
Phase = 0,25


R = -45°
Phase = 0,25

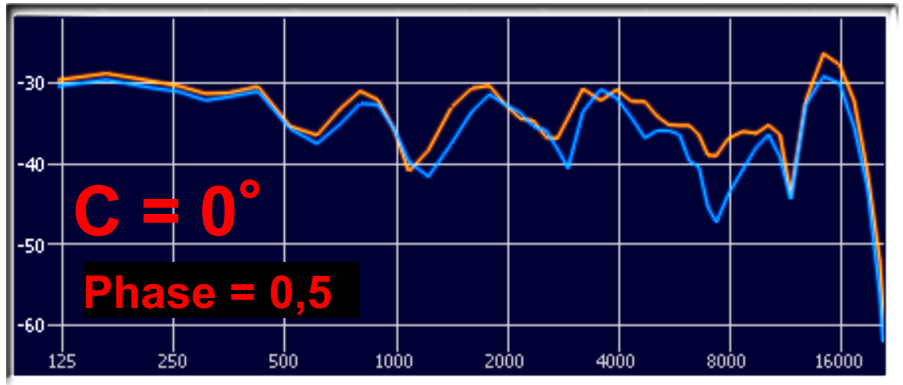

Sl = 90°
Phase = 0,15

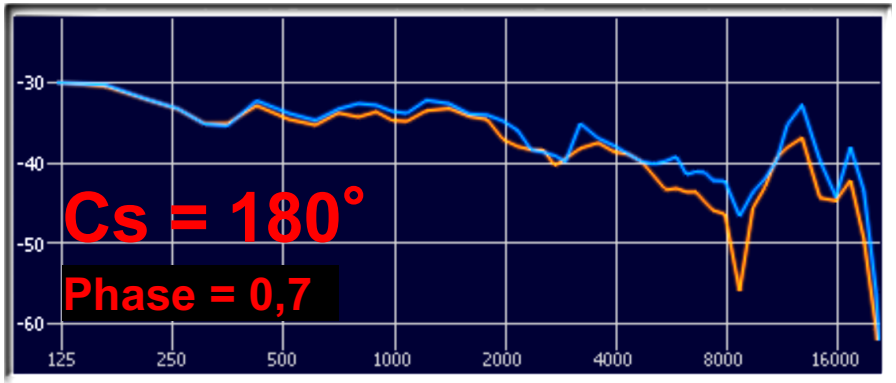

Sr = -90°
Phase = 0,15

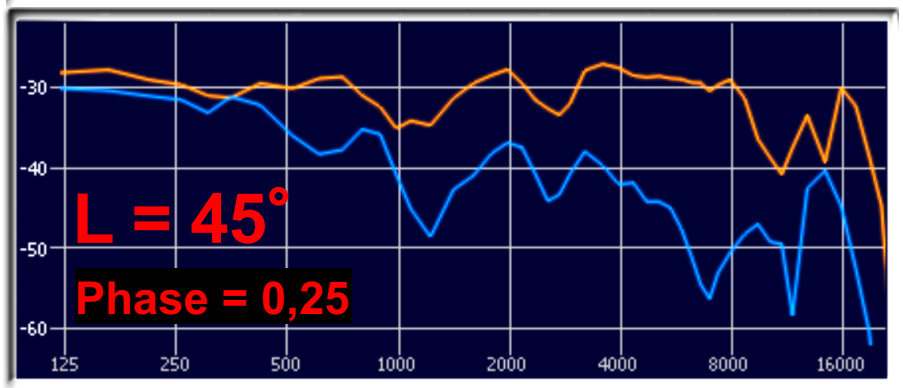

Ls = 135°
Phase = 0,2

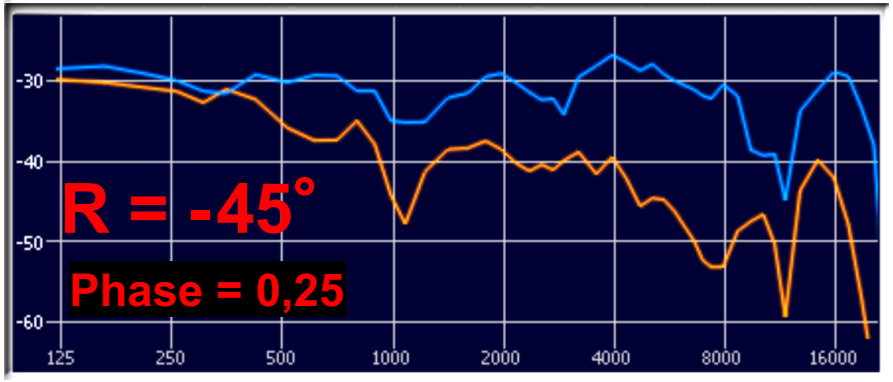

Rs = -135°
Phase = 0,2

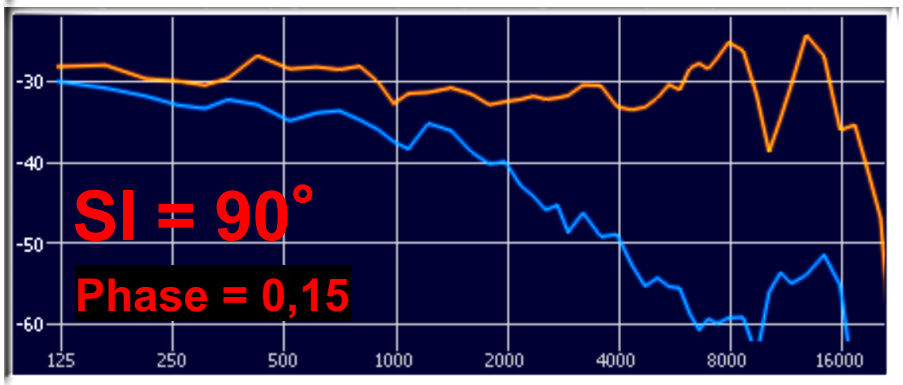
HRTF HEAD #5

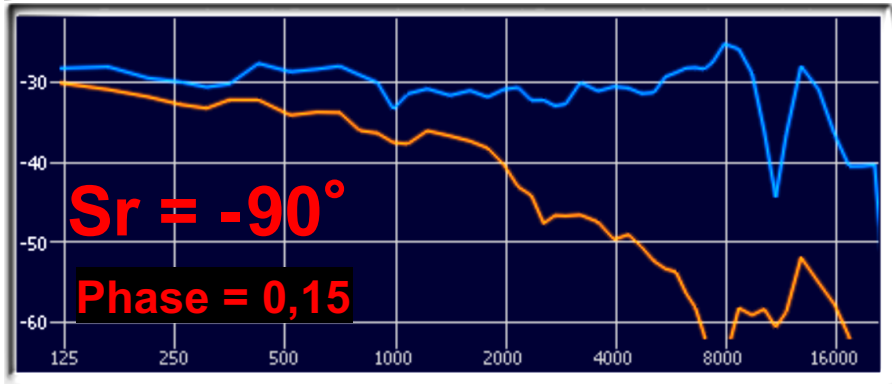

C = 0°
Phase = 0,5

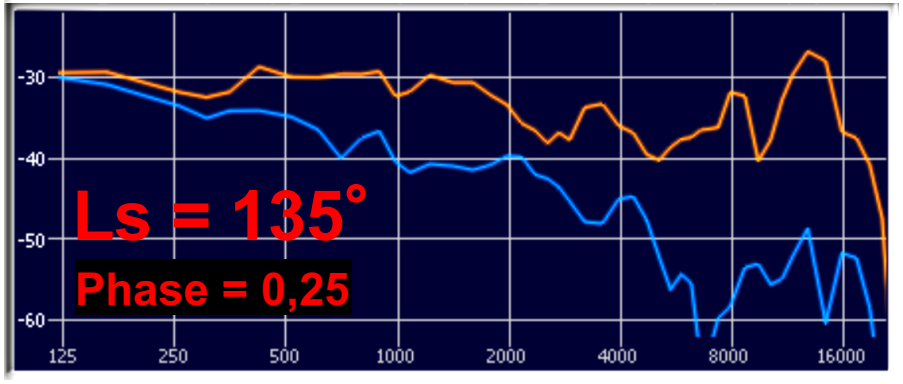

Cs = 180°
Phase = 0,7

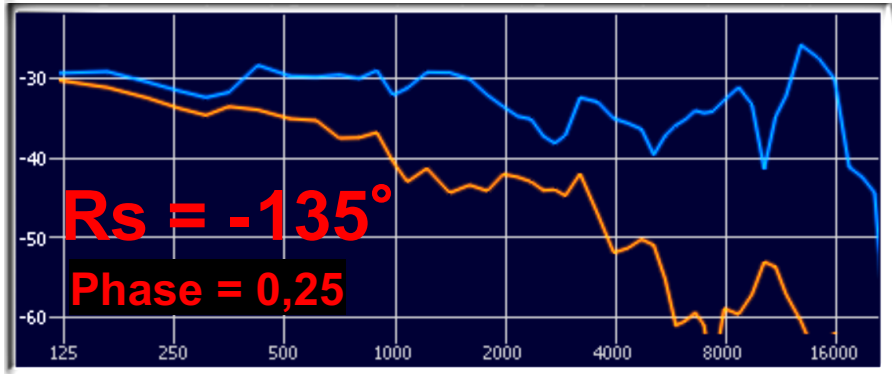

L = 45°
Phase = 0,25


R = -45°
Phase = 0,25

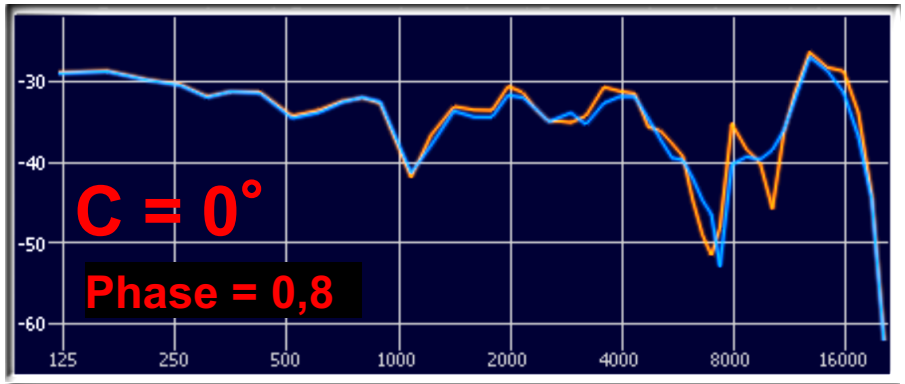

Sl = 90°
Phase = 0,15

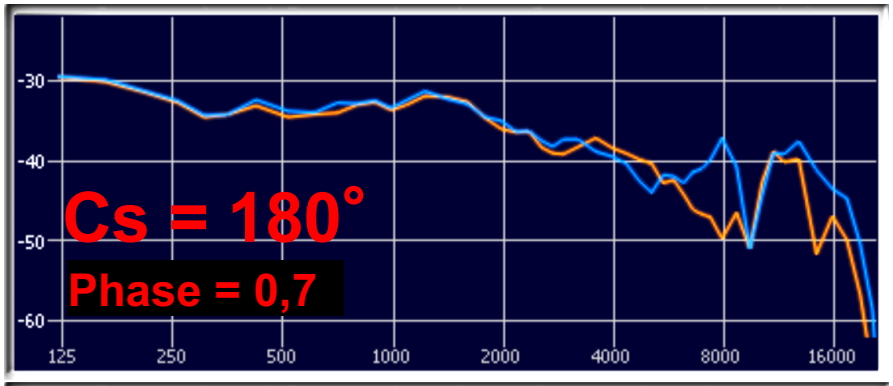

Sr = -90°
Phase = 0,15

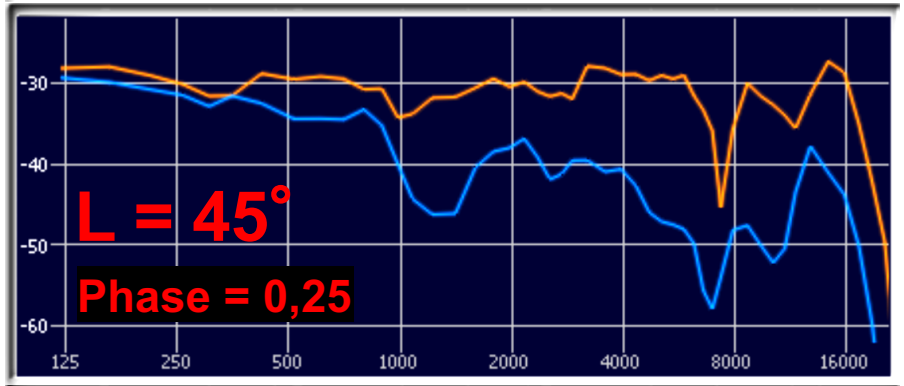

Ls = 135°
Phase = 0,25

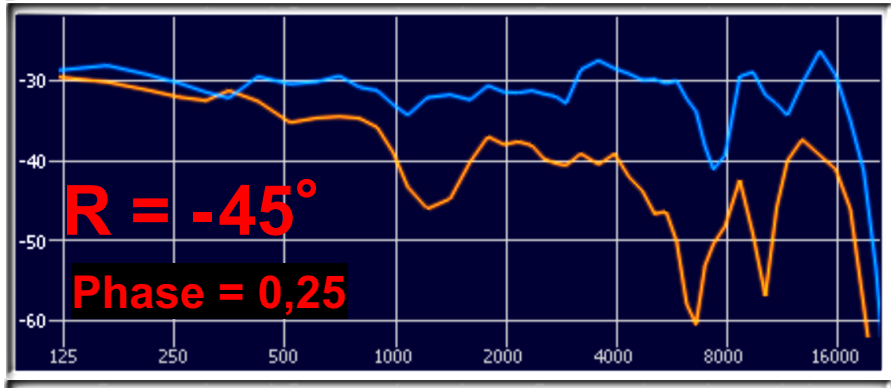

Rs = -135°
Phase = 0,25

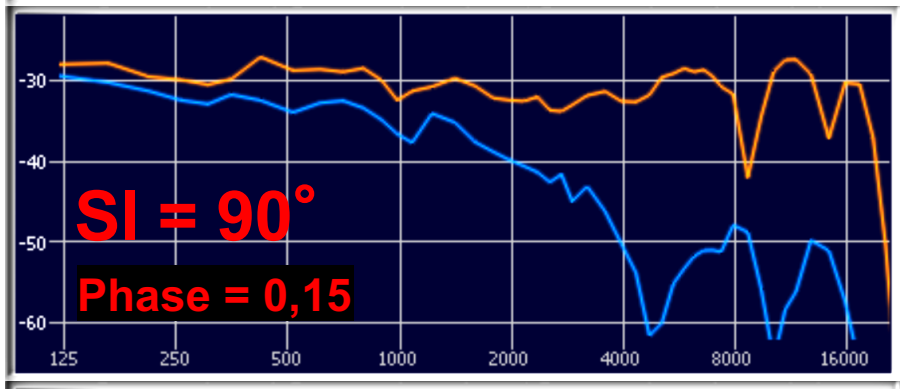
HRTF HEAD #6

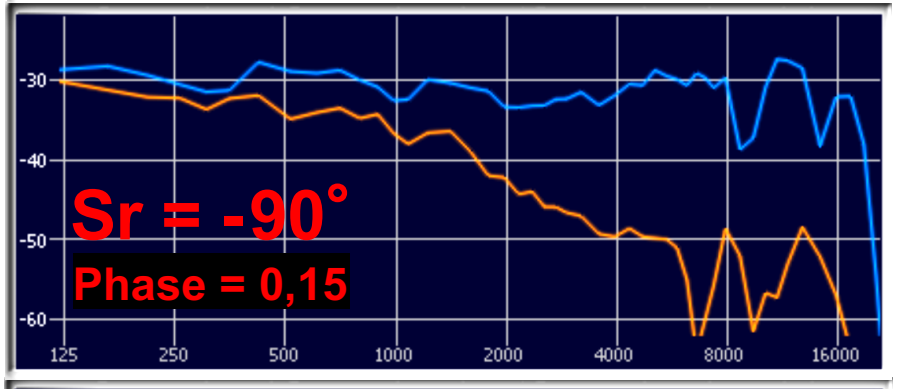

C = 0°
Phase = 0,8
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

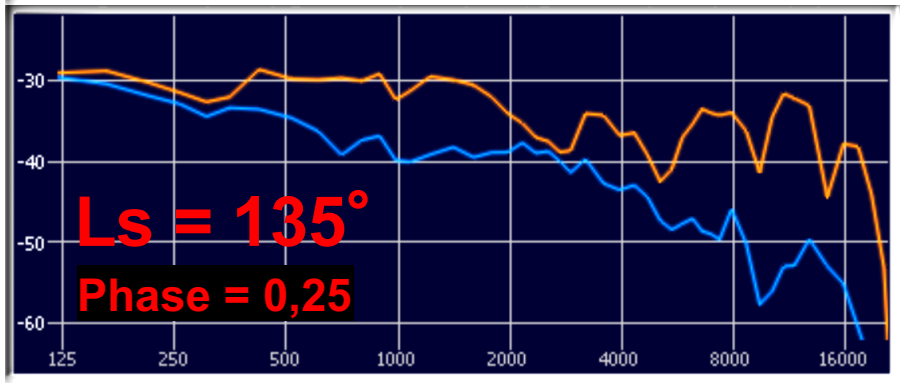

Cs = 180°
Phase = 0,7
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

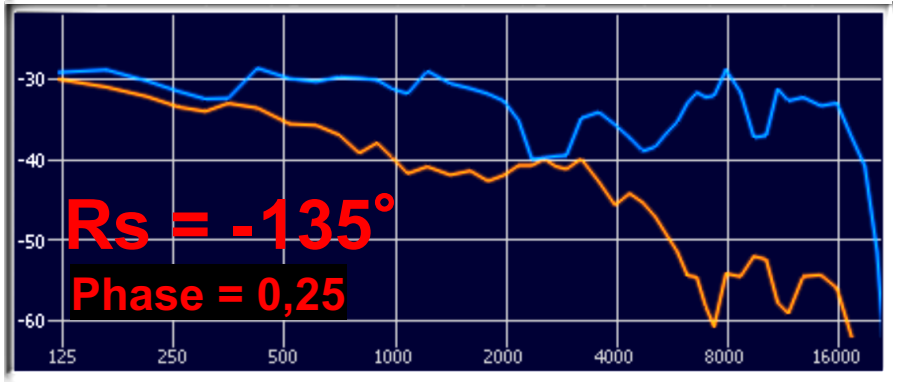

L = 45°
Phase = 0,25
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000


R = -45°
Phase = 0,25
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

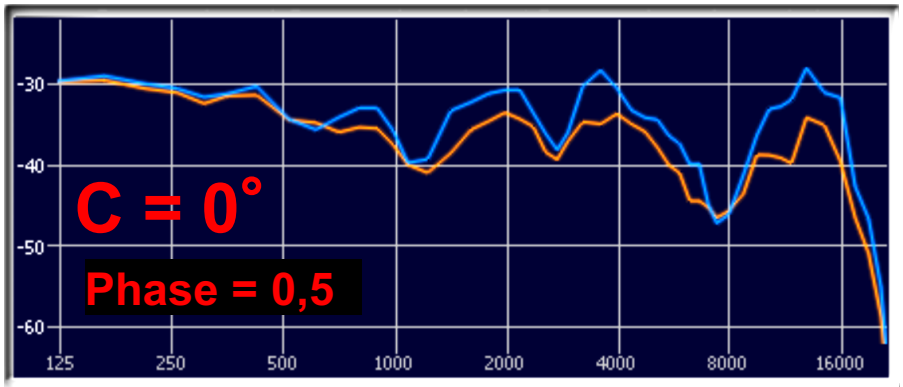

Sl = 90°
Phase = 0,15
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

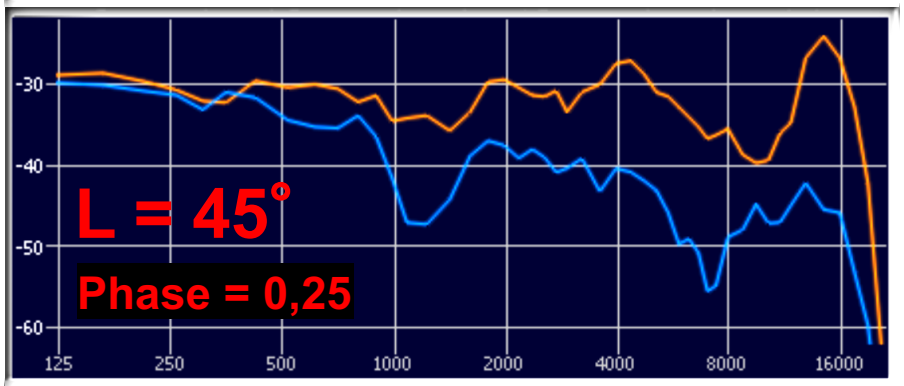

Sr = -90°
Phase = 0,15
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

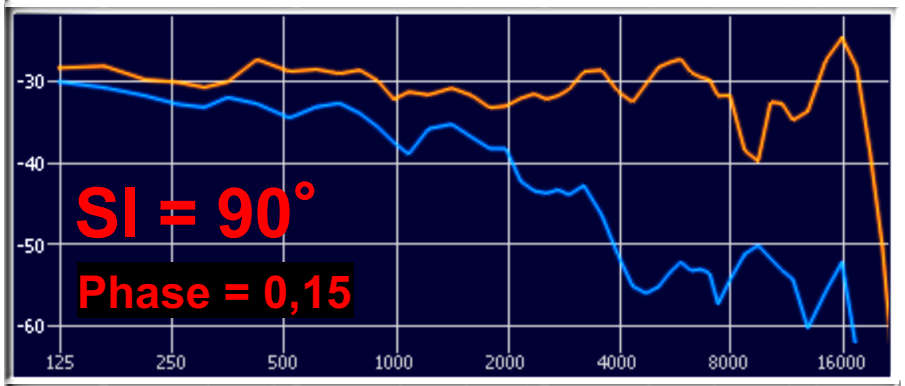

Ls = 135°
Phase = 0,25
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

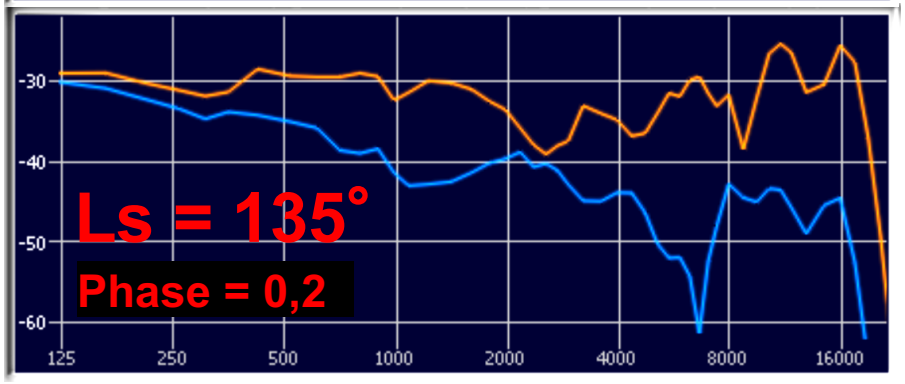

Rs = -135°
Phase = 0,25
-30
-40
-50
-60
125
250
500
1000
2000
4000
8000
16000

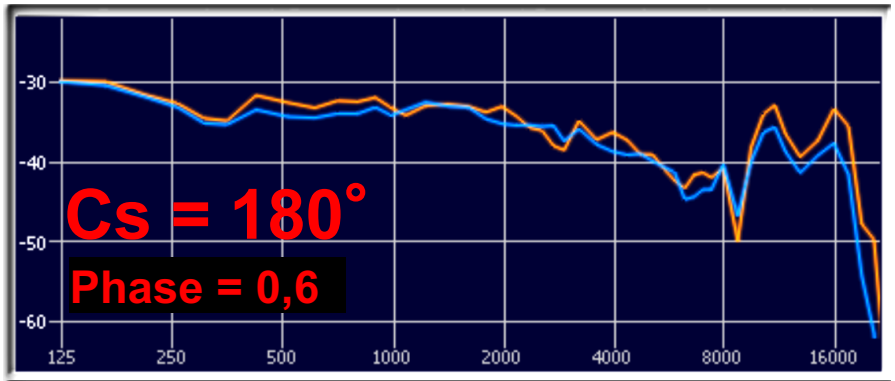
HRTF HEAD #7

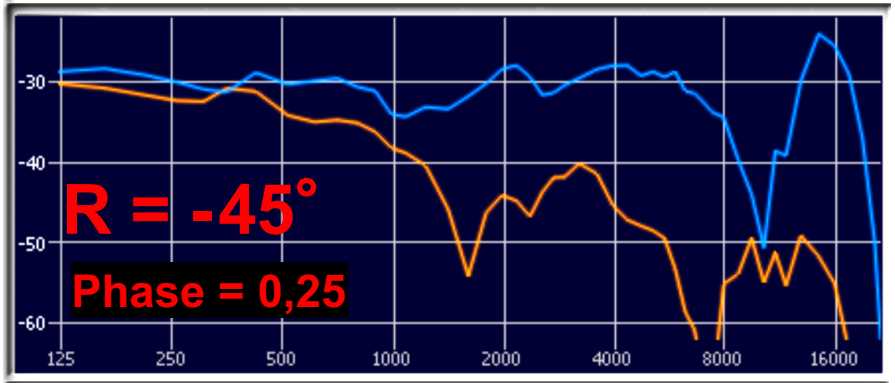

C = 0°
Phase = 0,5

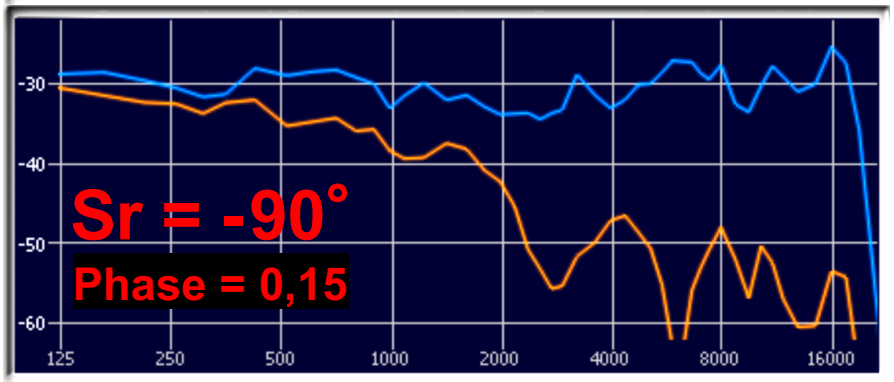

L = 45°
Phase = 0,25

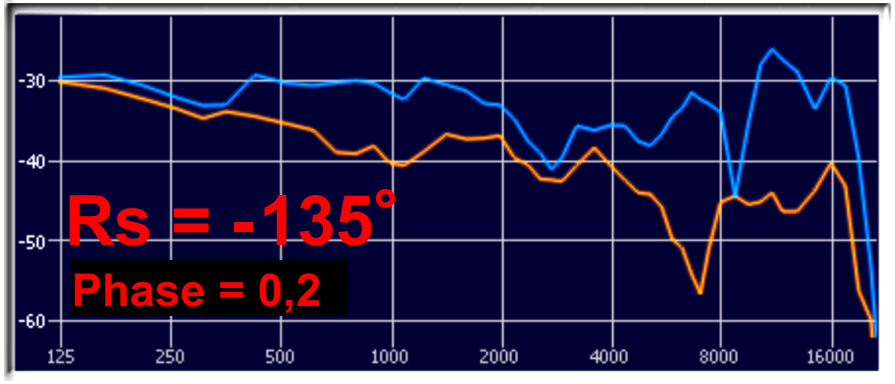

Sl = 90°
Phase = 0,15


Ls = 135°
Phase = 0,2


Cs = 180°
Phase = 0,6

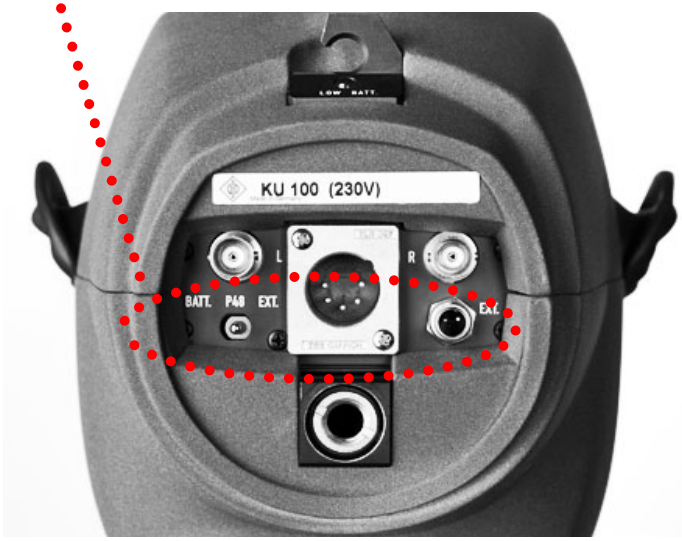

R = -45°
Phase = 0,25


Sr = -90°
Phase = 0,15


Rs = -135°
Phase = 0,2

En 1992 : NEUMANN KU 100 (≈ 7400 €)

BATT. pile 9v -- P48. fantôme 48v -- EXT. secteur 220v.



(Sorties : 1 XLR 5 + 2 BNC)

Les Oreilles sont anthropométriques (G et D identiques pour les mesures).

2 micros du système KM 100 (circuit de sortie sans transfo)

- **Tête égalisée pour une réponse linéaire en champ diffus.**

  Écoute compatible avec des haut-parleurs

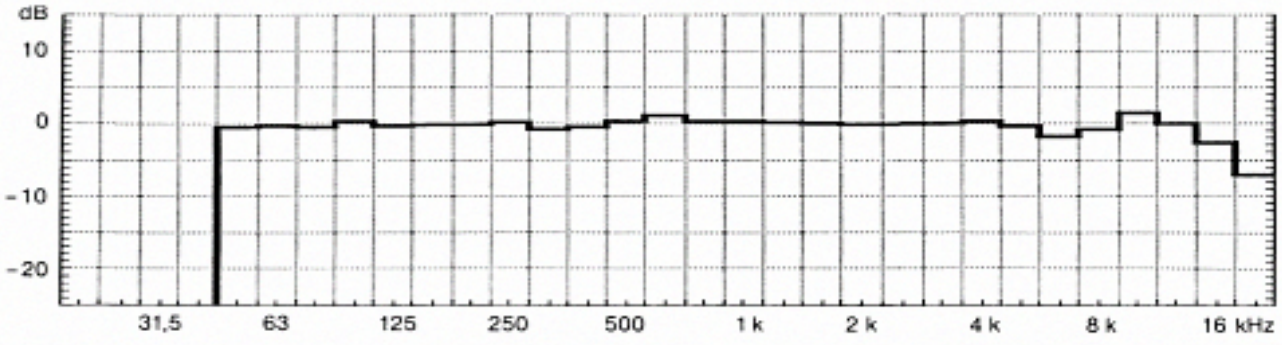
- **Filtre coupe bas à 40 Hz ou 150 Hz et atténuation de 10 dB.**

Photos : www.neumann.com et www.madooma.com

NEUMANN KU100

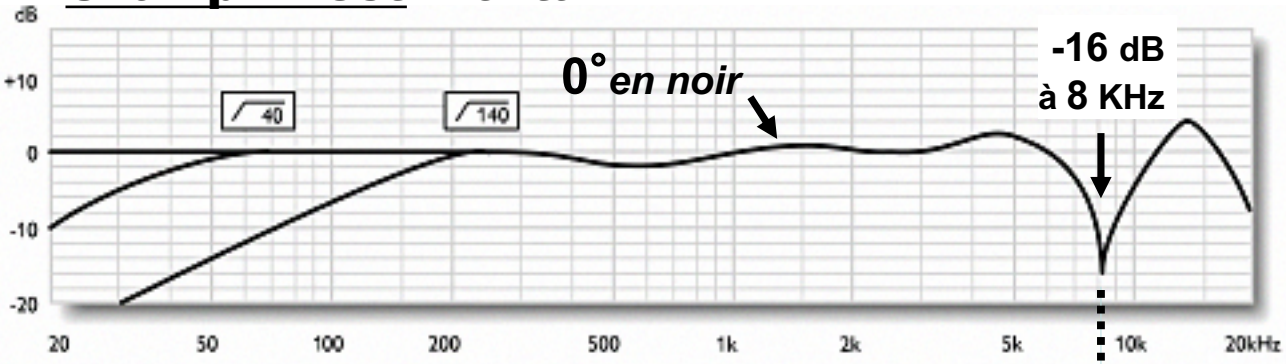
Documents Neumann

Champ Diffus :



± 2 dB

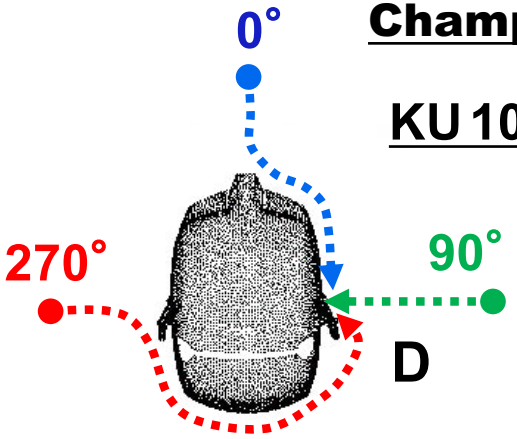
Champ Direct *Frontal :*



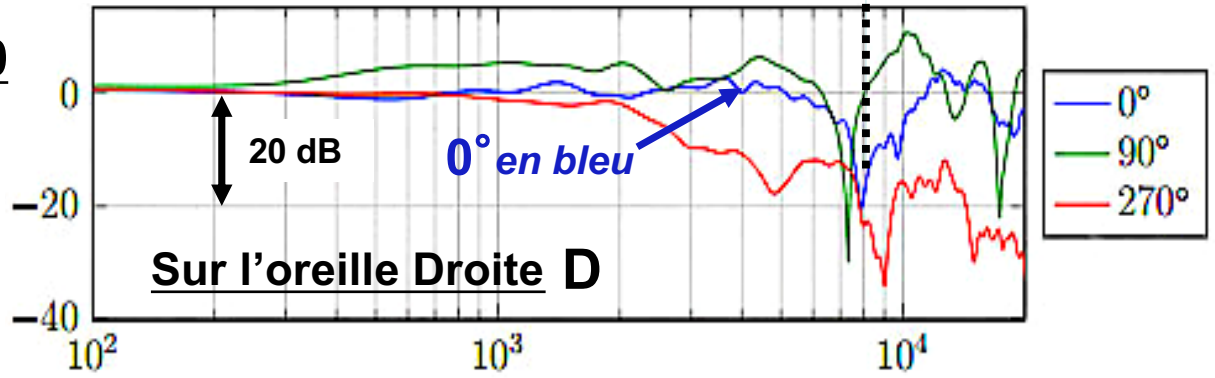
0° *en noir*

-16 dB
à 8 KHz

Champ Direct *Frontal à 0° :*



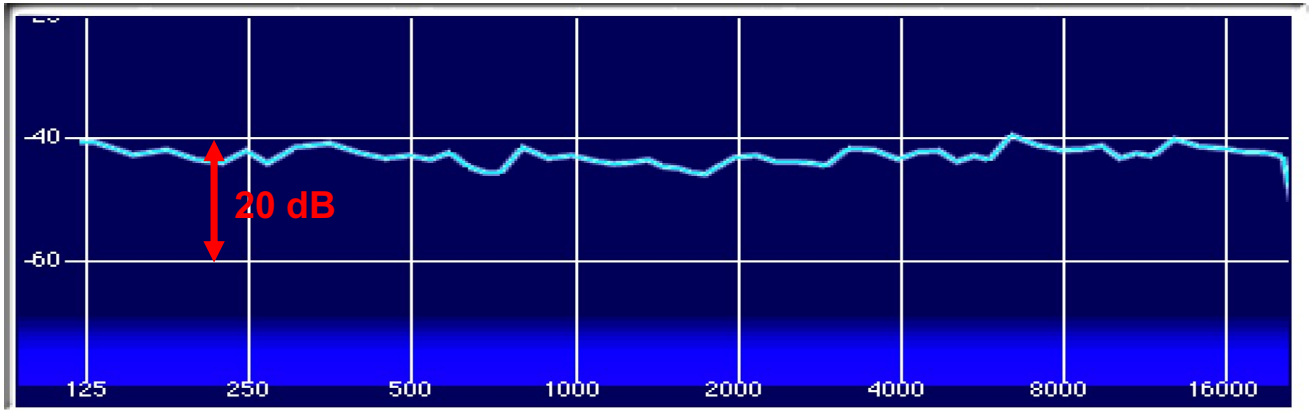
KU100



Jakob Vennerød
Trondheim, June 2014

IN

Bruit Rose

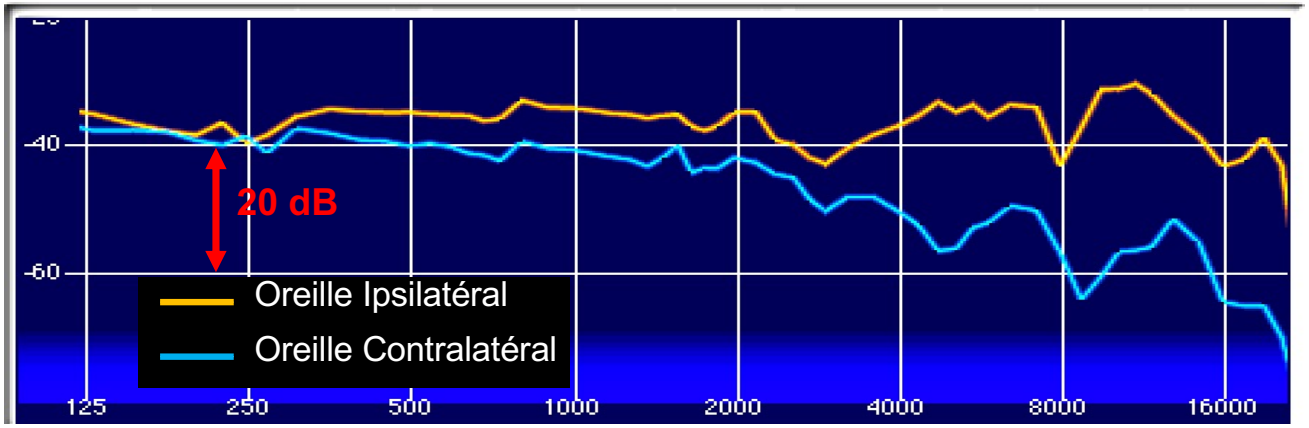


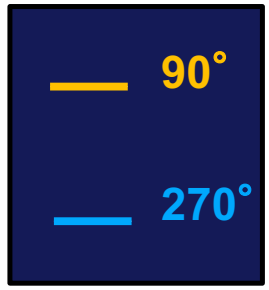
SENNHEISER

OUT

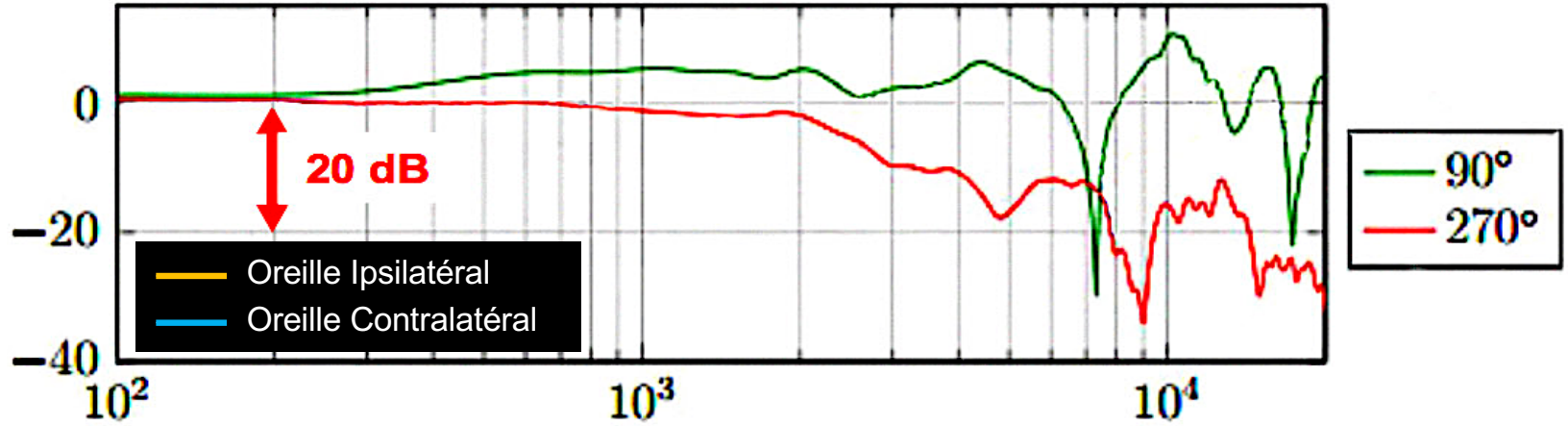
ORBIT

Clarity = 0%





Jakob Vennerød
Trondheim, June 2014



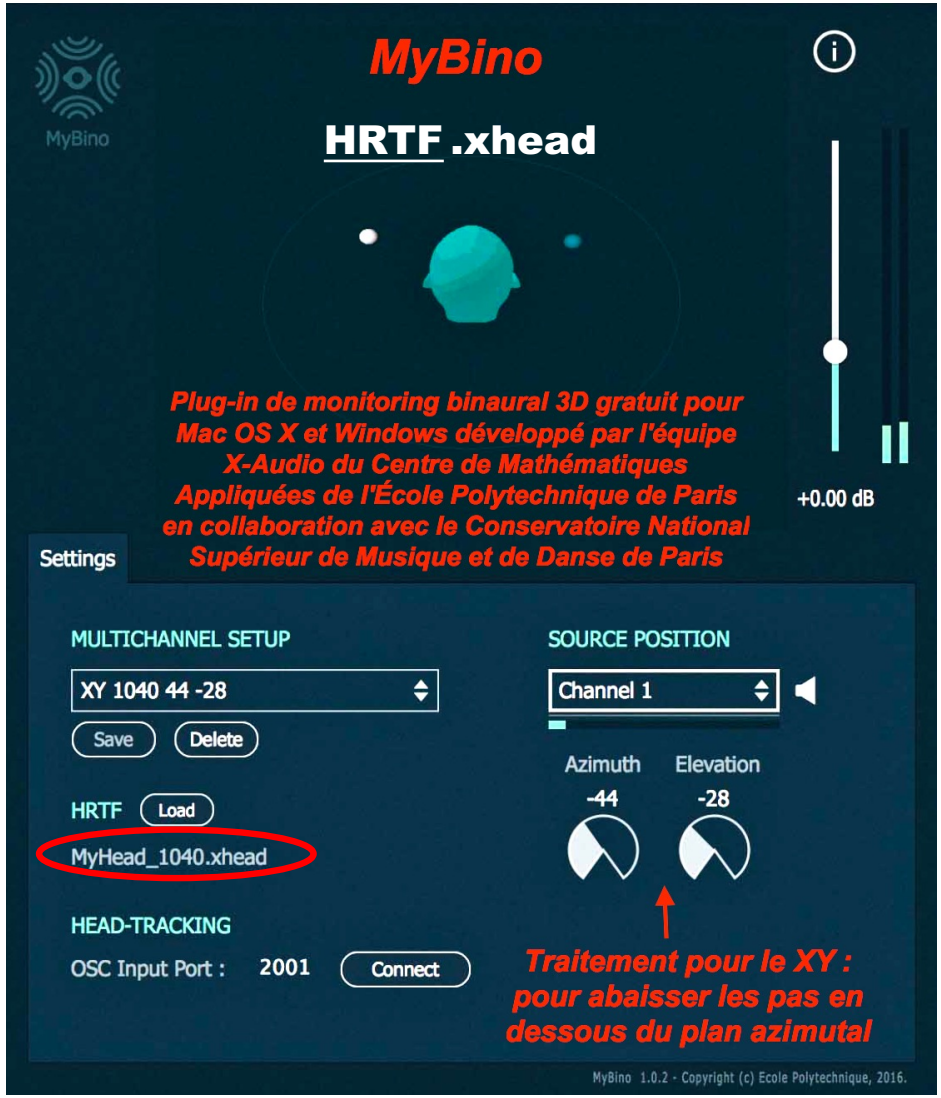
https://www.researchgate.net/profile/Jakob_Vennerod/publication/278404811_Binaural_Reproduction_of_Higher_Order_Ambisonics_A_Real-Time_Implementation_and_Perceptual_Improvements/links/5580323408ae87edac4c9091/Binaural-Reproduction-of-Higher-Order-Ambisonics-A-Real-Time-Implementation-and-Perceptual-Improvements.pdf

Binaural de synthèse

2 Plug-ins de Monitoring *Binaural* 3D …… *Gratuits*

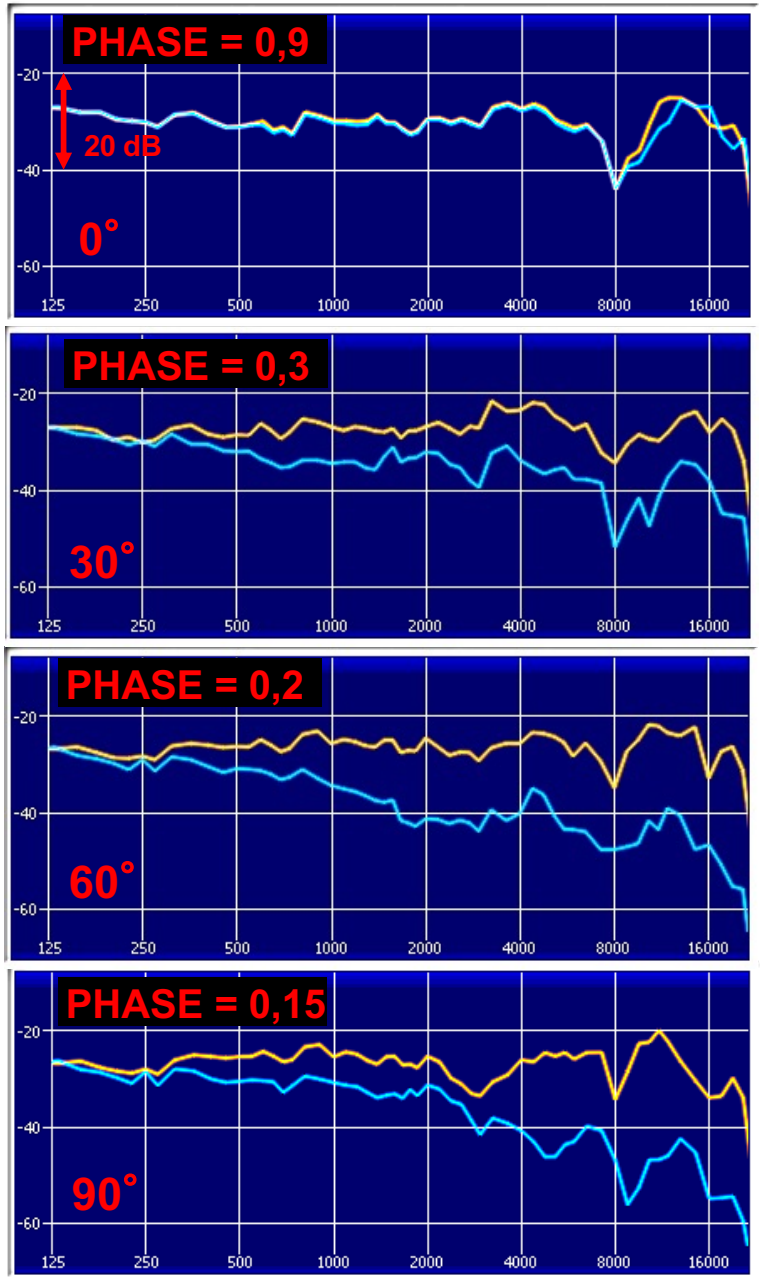
https://fr-fr.sennheiser.com/ambeo-blueprints-downloads

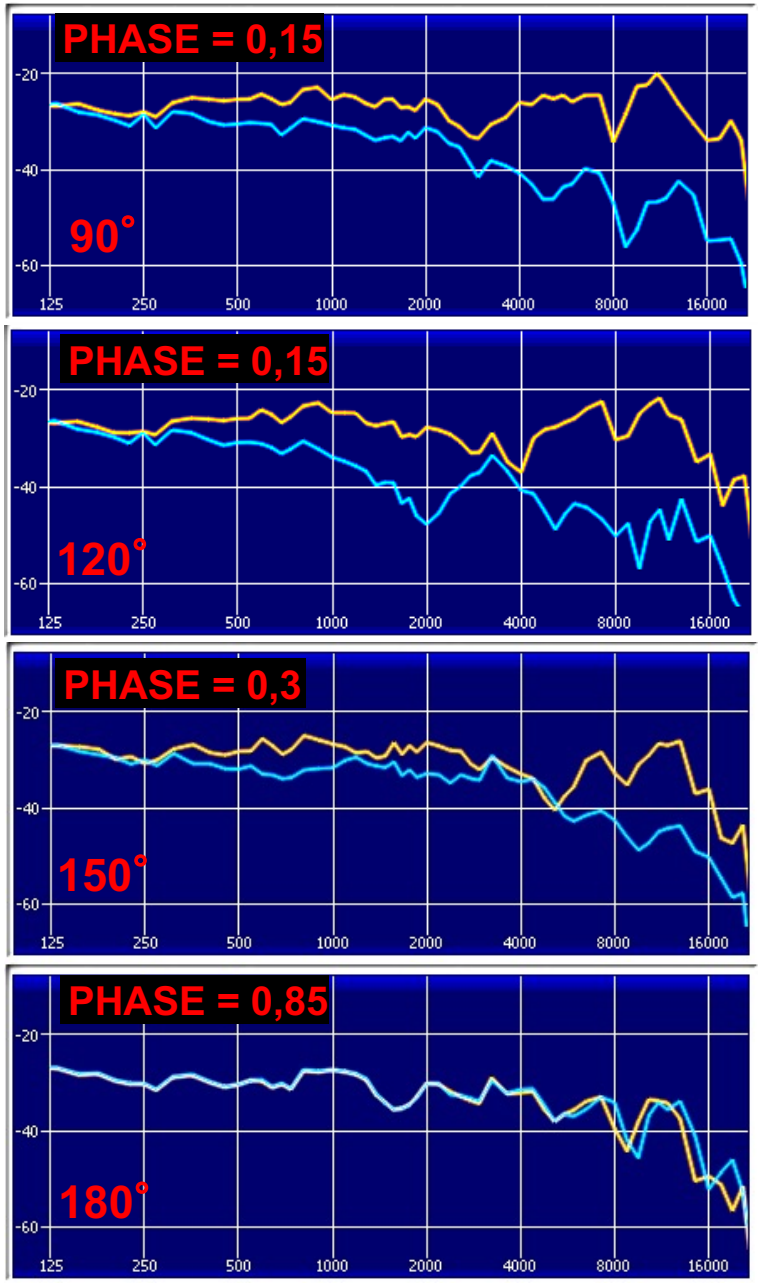
http://www.cmap.polytechnique.fr/xaudio/mybino/



ORBIT : HRTF KU100 pour une source Mono…….. CLARITY 0 %

SENNHEISER



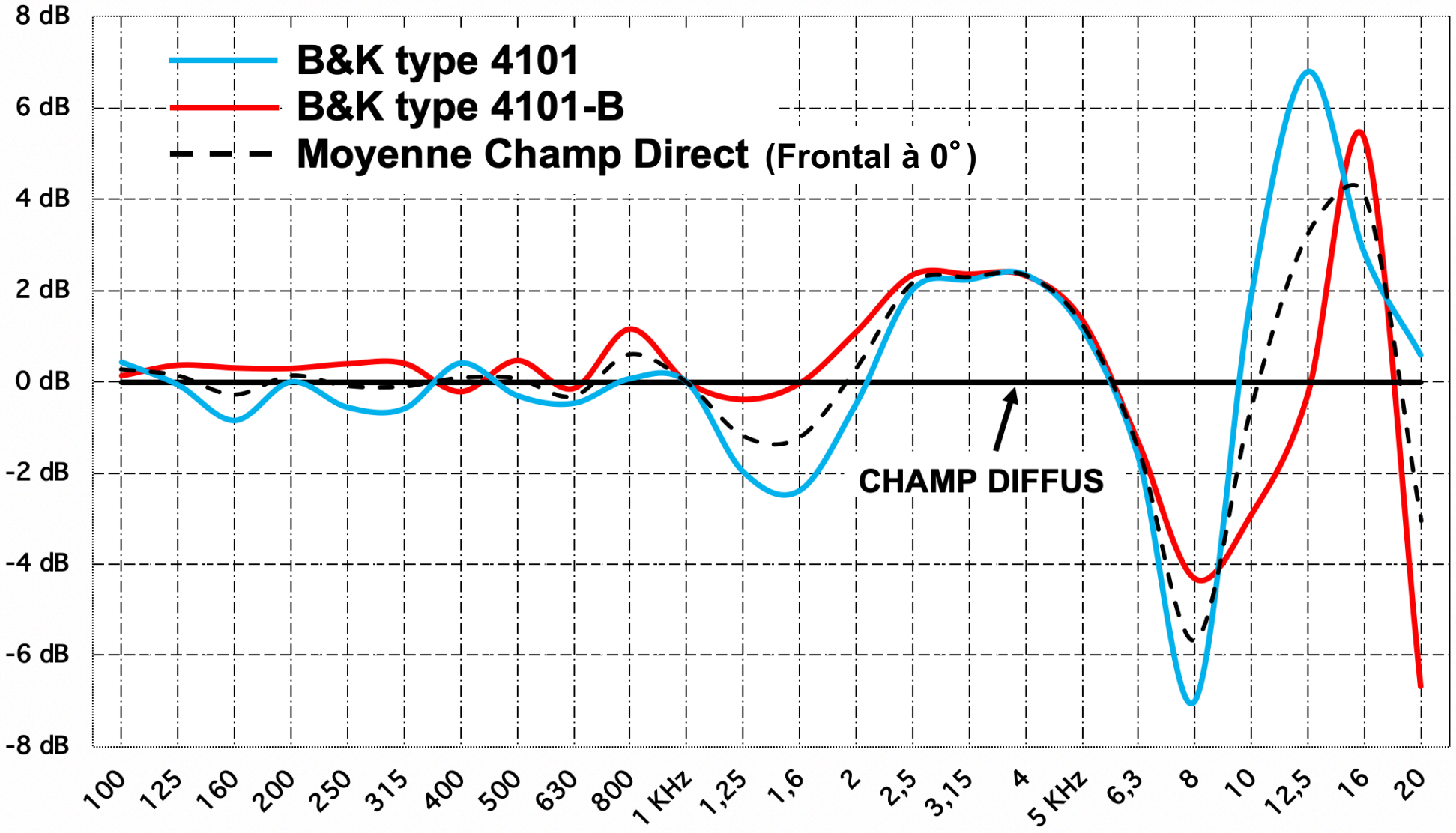


Oreille Ipsilatéral
Oreille Contralatéral

Oreille Ipsilatéral
Oreille Contralatéral

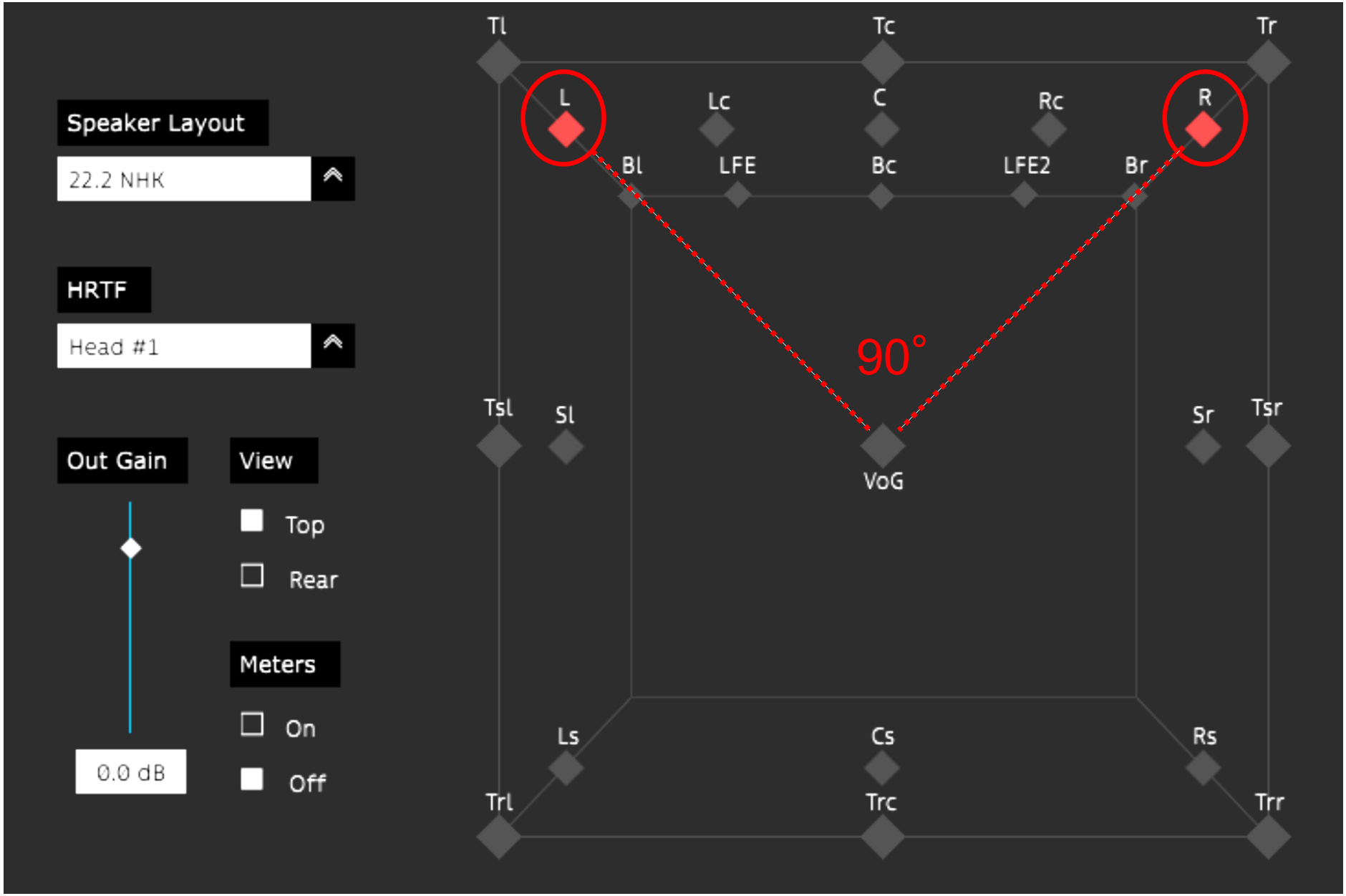
Réponse en Champ Direct des B&K pour une correction Champ Diffus linéaire :

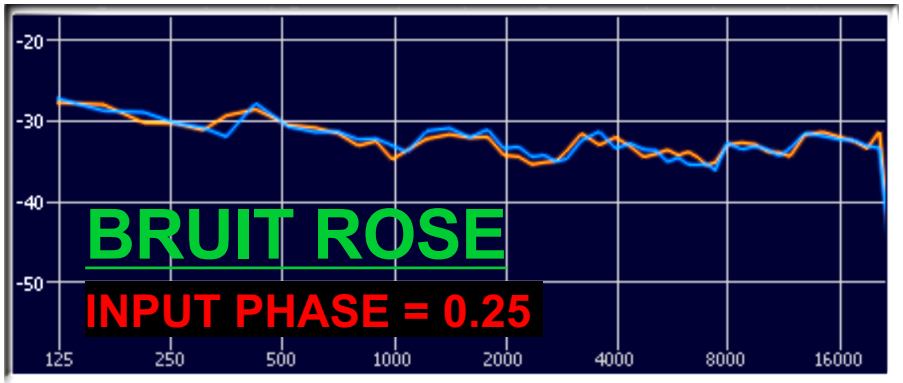
https://www.lesonbinaural.fr/EDIT/PDF/DPA_4560.PDF

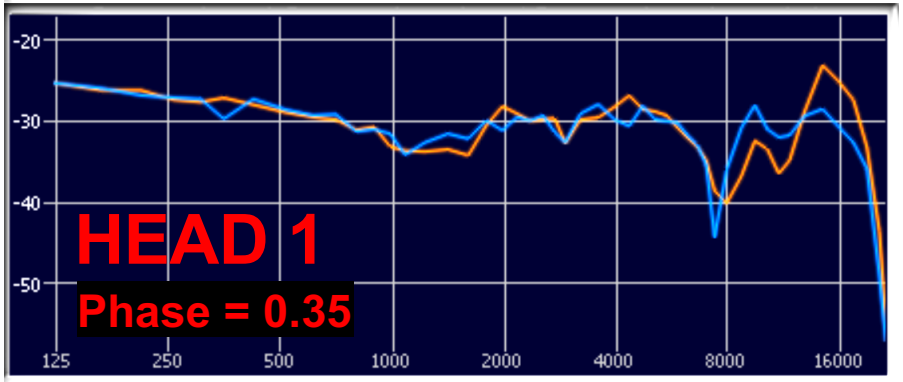


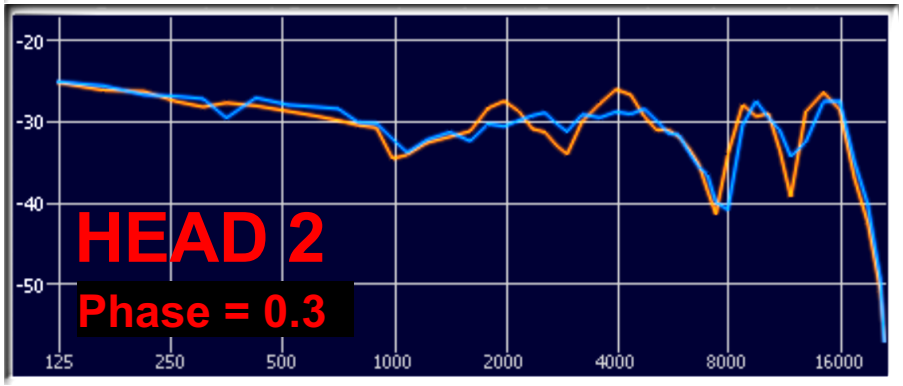
b<>com *Spatial Audio Toolbox* [Render Spk2Bin]

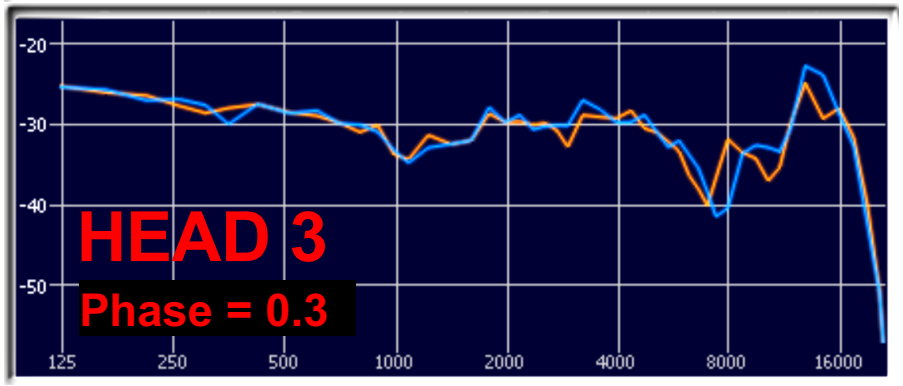
Mise en service des canaux **L** et **R** :



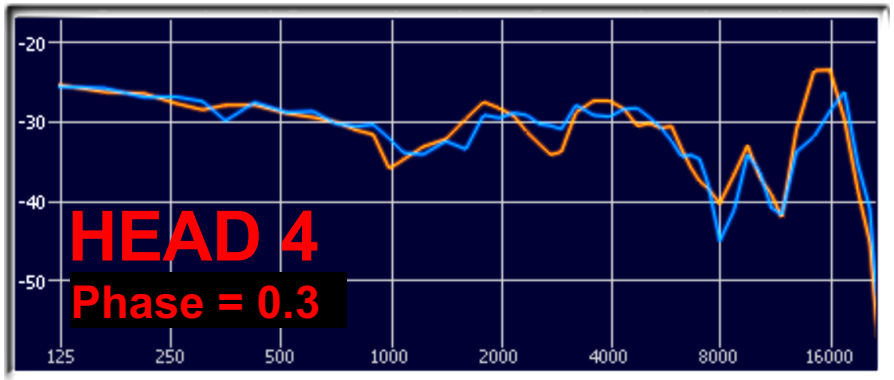

BRUIT ROSE
INPUT PHASE = 0.25

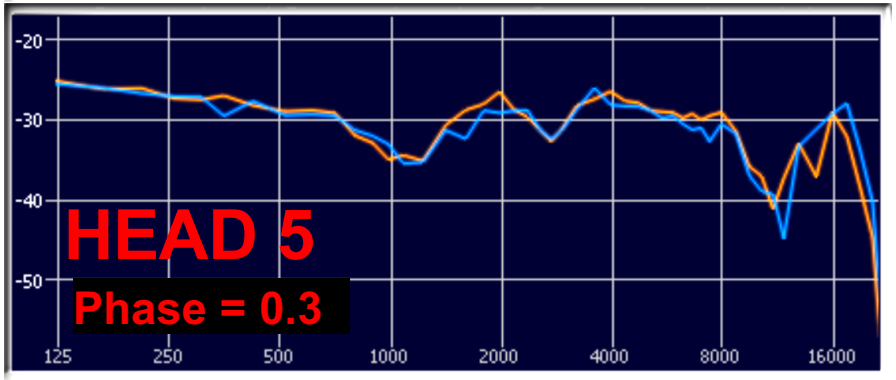

HEAD 1
Phase = 0.35

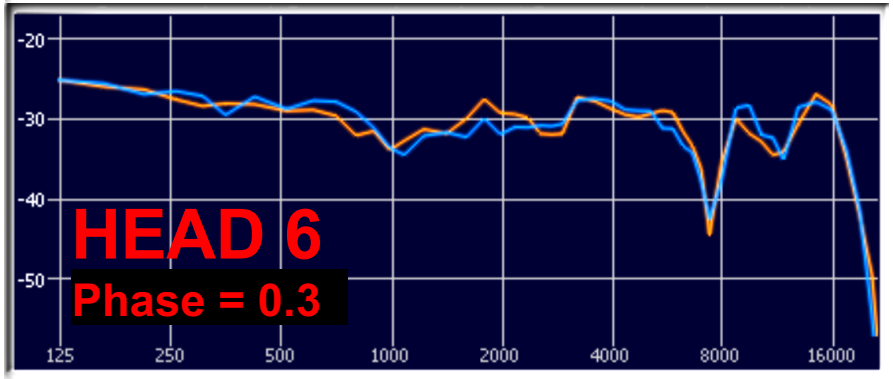

HEAD 2
Phase = 0.3

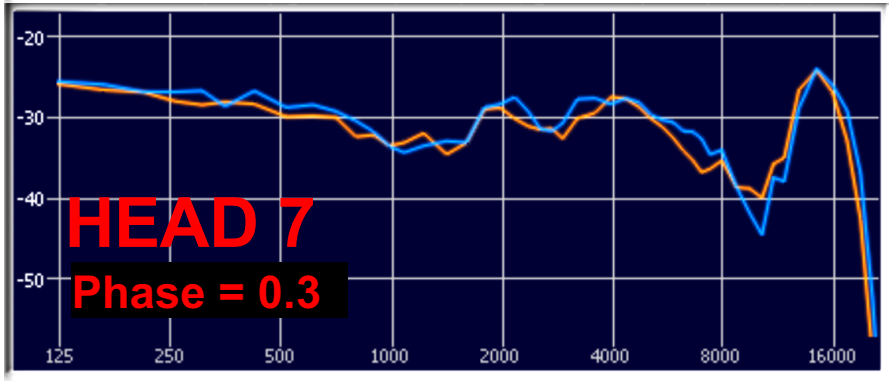

HEAD 3
Phase = 0.3

HRTF L + R à 45°


HEAD 4
Phase = 0.3


HEAD 5
Phase = 0.3


HEAD 6
Phase = 0.3


HEAD 7
Phase = 0.3

HYPER RADIO

SON 3D

Expertise

HRTF SÉLECTIONNÉES PAR RADIO FRANCE

https://www.lesonbinaural.fr/EDIT/PDF/hrtf_rf_21_22_23_24_26_27.PDF

BERNARD LAGNEL
Ingénieur du son Radio France

juin 2018

Octobre 2015

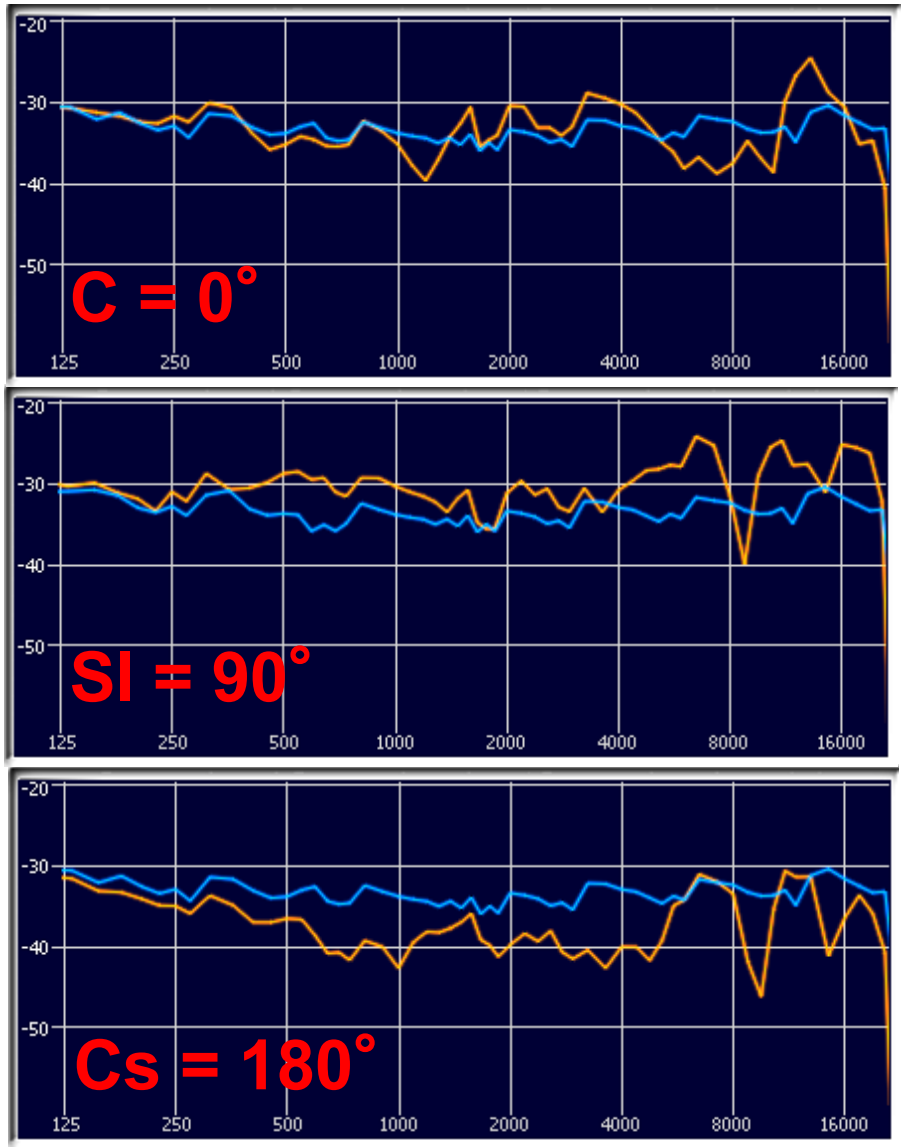
HYPER RADIO

SON 3D

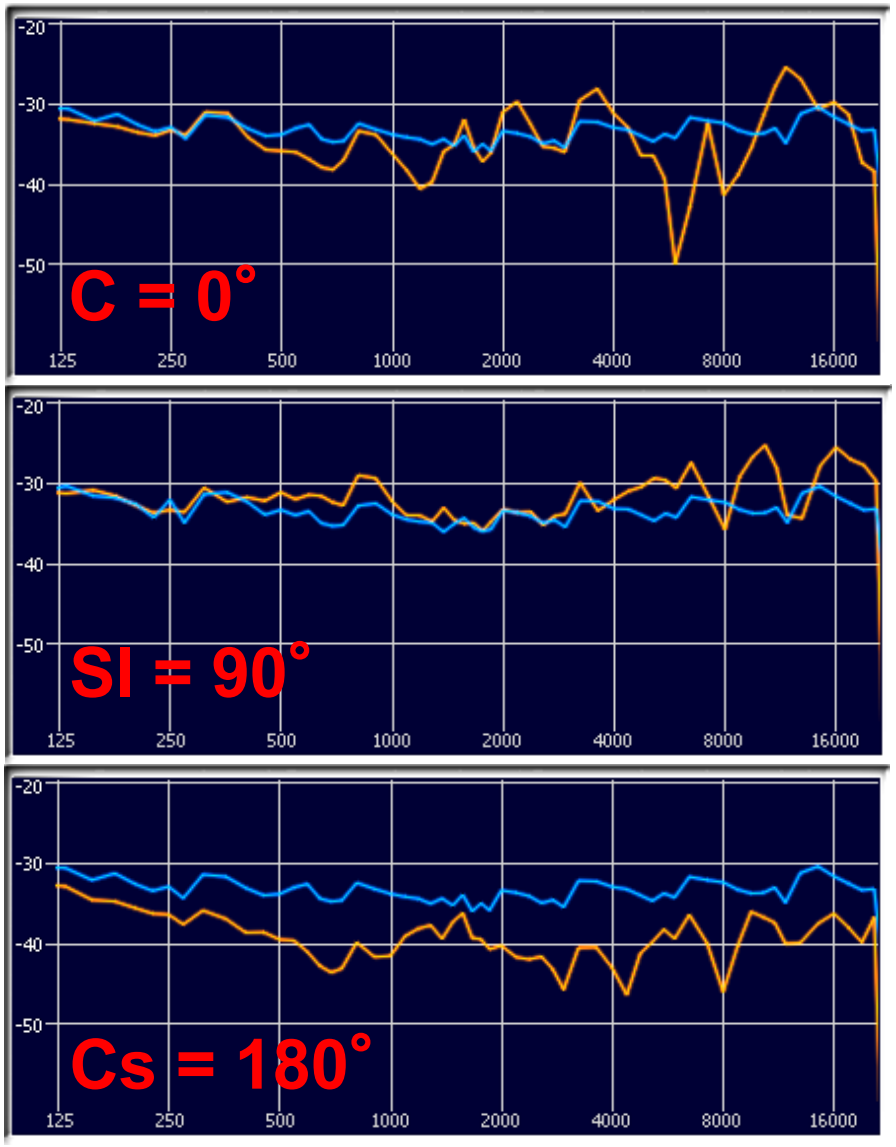
Oreille Ipsilatéral L

BRUIT ROSE IN

HRTF RF 21 A



HRTF RF 22 B



https://www.lesonbinaural.fr/EDIT/PDF/hrtf_rf_21_22_23_24_26_27.PDF

Octobre 2015

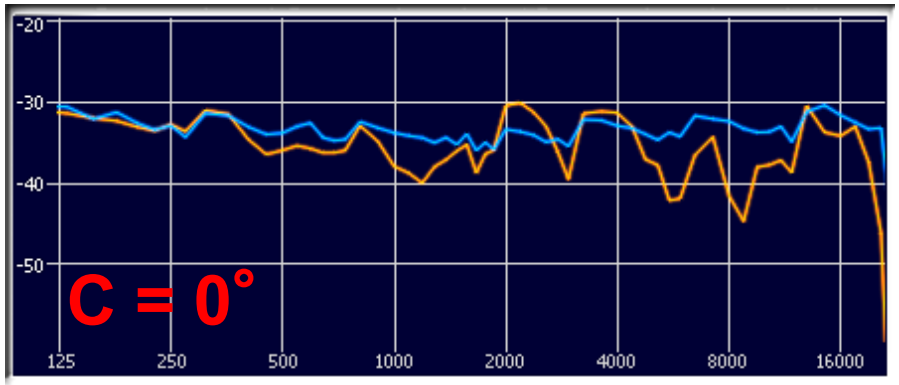
HYPER RADIO

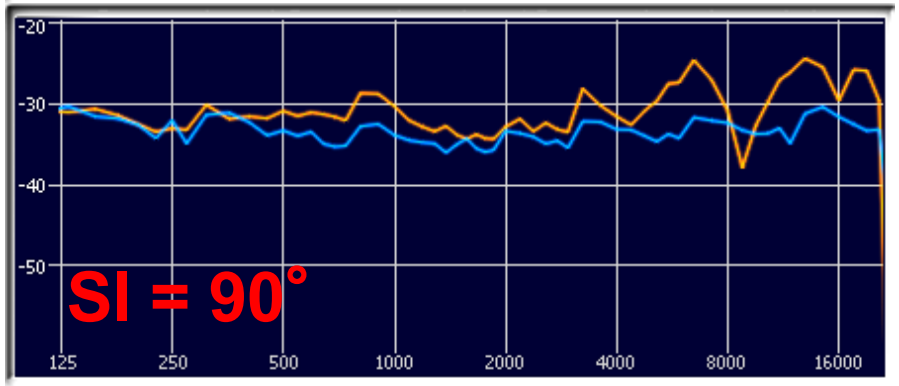
SON 3D

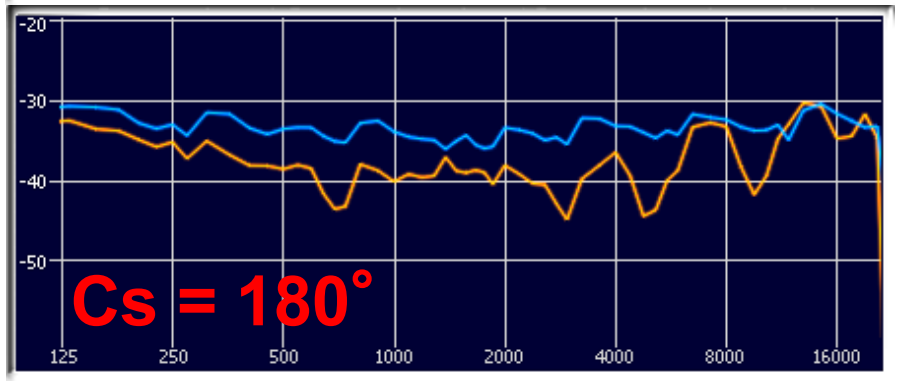
Oreille Ipsilatéral L

BRUIT ROSE IN

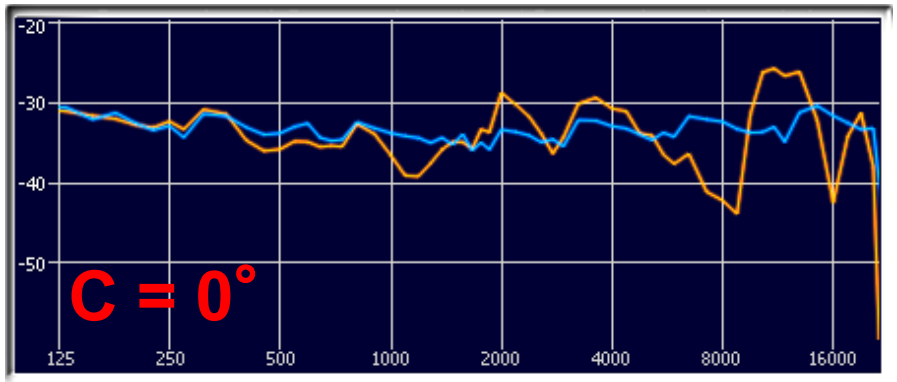
HRTF RF 23 C

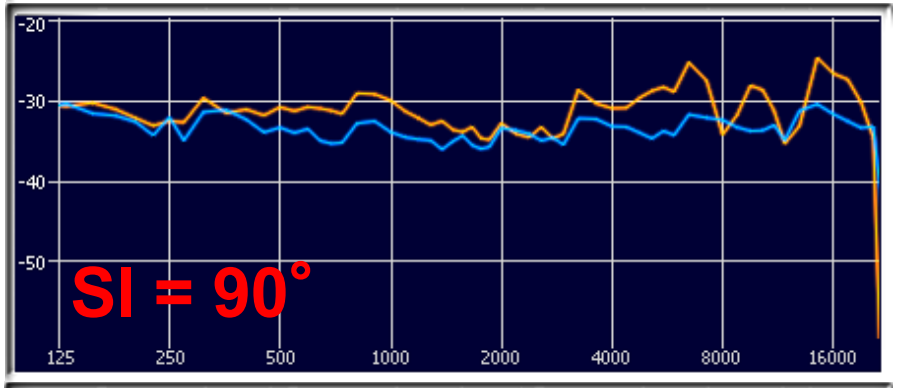


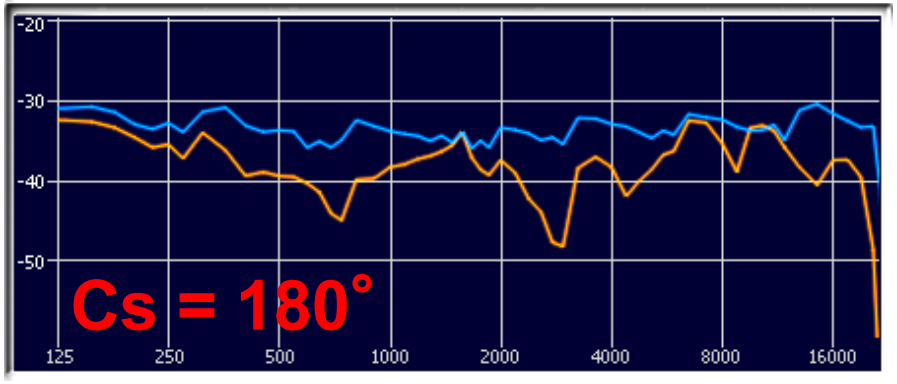




HRTF RF 24 D







https://www.lesonbinaural.fr/EDIT/PDF/hrtf_rf_21_22_23_24_26_27.PDF

Octobre 2015

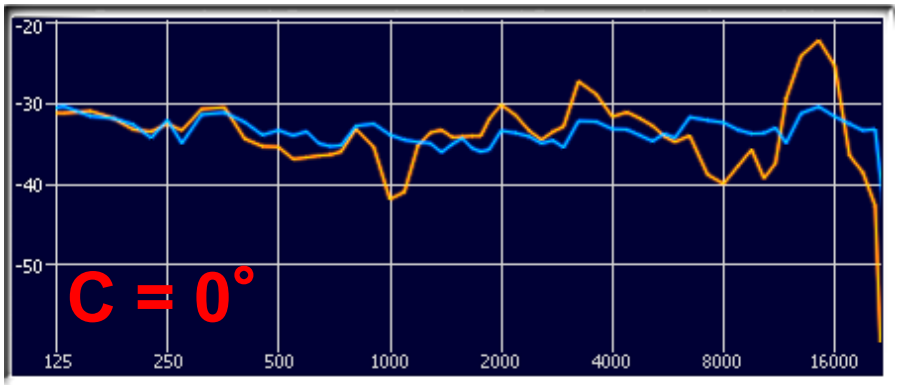
HYPER RADIO

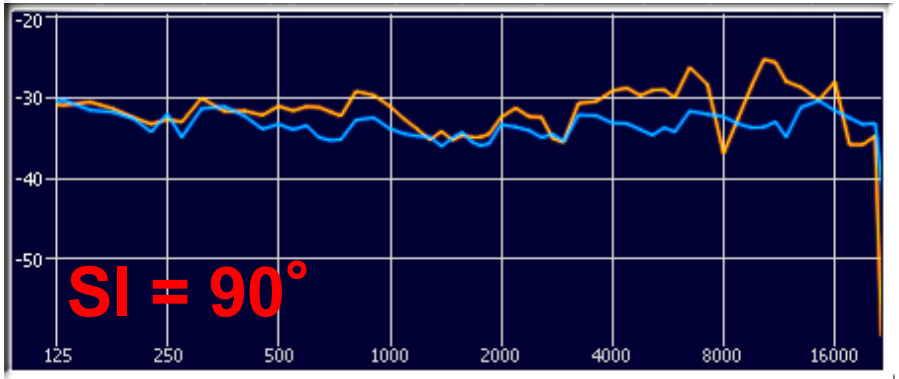
SON 3D

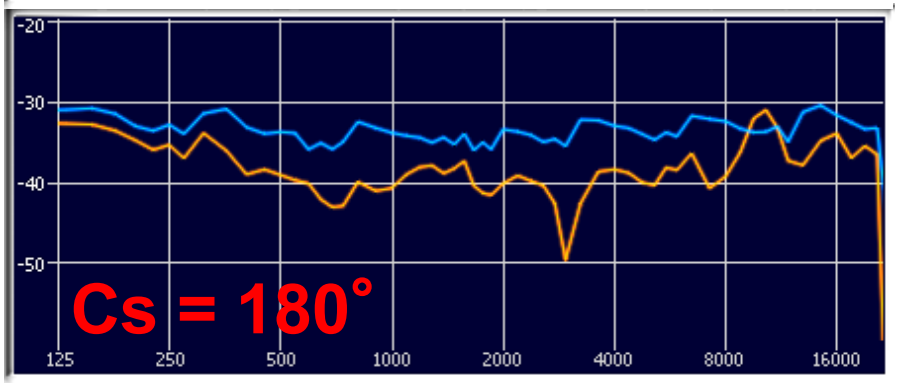
Oreille Ipsilatéral L

BRUIT ROSE IN

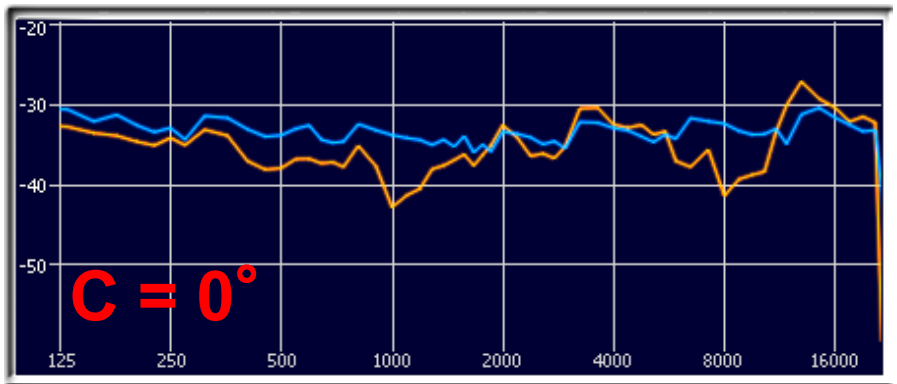
HRTF RF 26 E

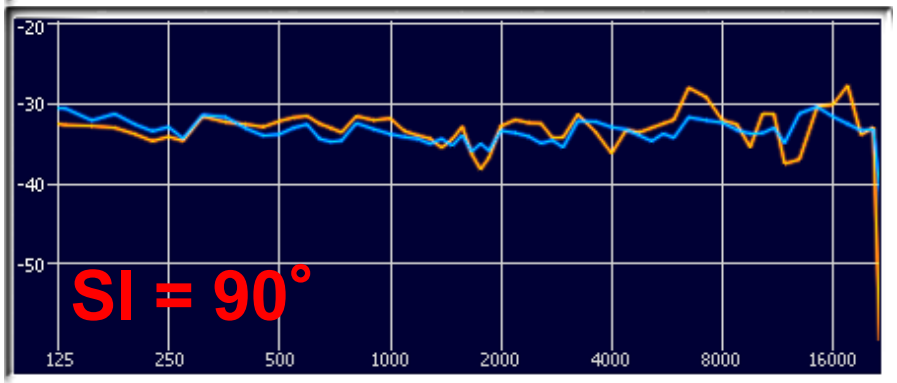


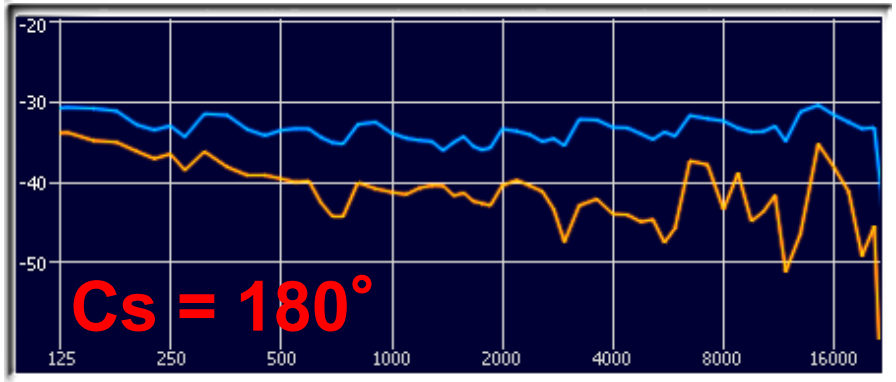




HRTF RF 27 F





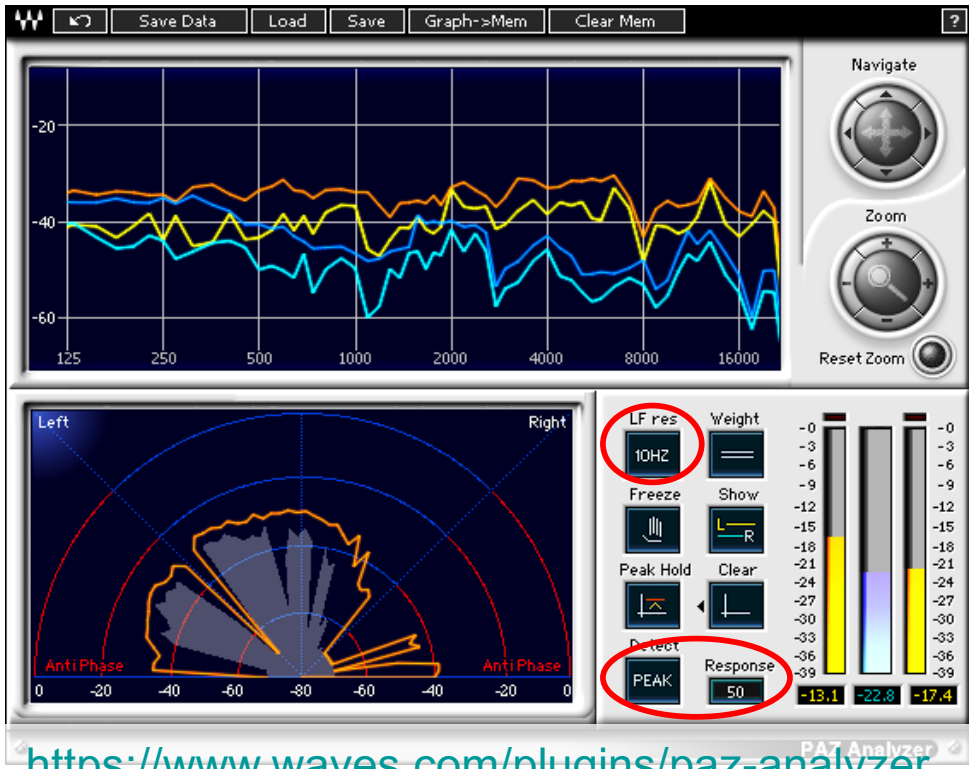


https://www.lesonbinaural.fr/EDIT/PDF/hrtf_rf_21_22_23_24_26_27.PDF

BINAURALIZER de **NOISEMAKERS**

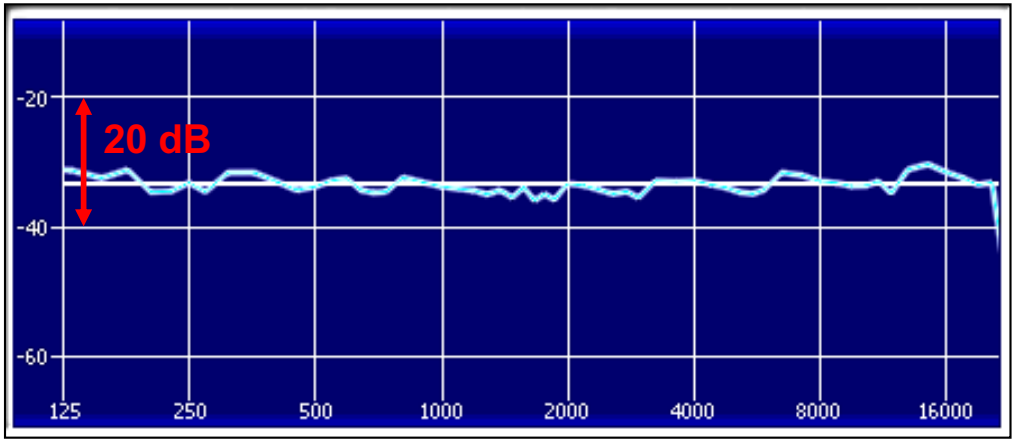
https://www.noisemakers.fr/product/binauralizer/

OUTPUT : ***PAZ*** de **WAVES** pour l'**ILD**

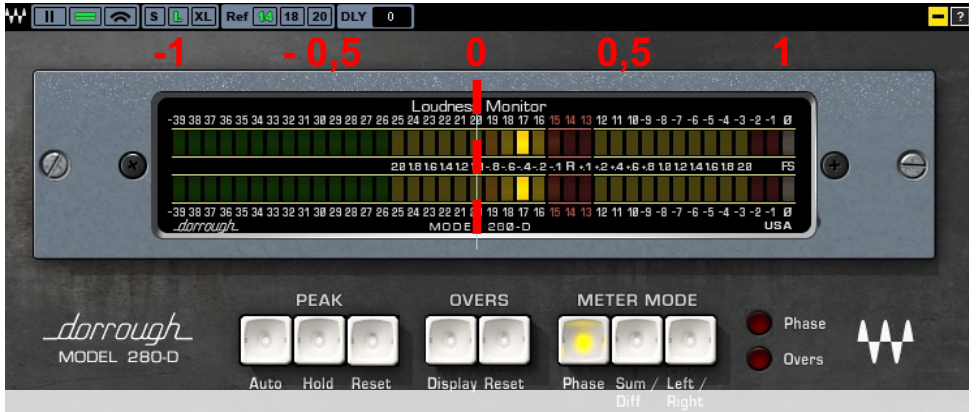


https://www.waves.com/plugins/paz-analyzer

INPUT : ***BRUIT ROSE***



OUTPUT : ***Dorrough*** de **WAVES** pour la **PHASE**



https://www.waves.com/plugins/dorrough-stereo

b com

L'AUDIO: L'AUTRE DIMENSION DE LA REALITÉ VIRTUELLE

SOMMAIRE

/ **Les auteurs** /

/ **L'audio : l'autre dimension de la réalité virtuelle** /

/ **Lexique** /

/ **Aller plus loin** /

/ **A propos** /

/ Les auteurs /

Ce livre blanc a été rédigé par l'équipe Hypermédia de b<>com et particulièrement par Clément Cerles et ses collègues du laboratoire Nouveaux Formats Médias. **Clément Cerles**, est un jeune ingénieur du son diplômé de l'Ecole Nationale Supérieure Louis Lumière à Saint Denis. Chez b<>com, il est spécialiste en son spatialisé (HOA) et mène également une activité de sound designer.

Le laboratoire **Nouveaux Formats Médias** étudie les nouvelles caractéristiques et dimensions de l'image (2D et 3D) et du son (spatialisé) afin d'augmenter le réalisme et l'impact des contenus. Il développe des outils et solutions visant à faciliter l'adoption et la production de ces futurs formats par les professionnels. Les équipes développent en ce moment une suite logicielle d'audio spatialisée basée sur le format Ambisonics. Celle-ci permettra aux créateurs de contenu de positionner des sources dans un espace sonore et d'y appliquer des traitements spatiaux (rotation, focus, masquage).

INTRODUCTION

Selon le cabinet Digi Capital, le marché de la réalité virtuelle devrait atteindre les 30 milliards de dollars d'ici 2020.

La RV est un sujet très à la mode, ce qui contribue à créer cette sensation d'omniprésence. Les derniers salons tels que le Mobile Word Congress ou le CES, l'ont prouvé en regroupant une multitude d'acteurs du domaine. Aujourd'hui, les équipementiers promettent une réelle expérience immersive mais comme le souligne très bien **Steven Spielberg** : *« Une expérience audiovisuelle ne sera jamais immersive tant que l'on regardera un rectangle, qu'il s'agisse d'un écran de cinéma, de TV, de tablette, smartphone ou d'ordinateur. Nous devons nous débarrasser de ça et devons mettre le spectateur au cœur de l'expérience, afin que peu importe où il regarde, il soit intégré dans une expérience tridimensionnelle. C'est ça le futur ».* Quel que soit l'usage, l'expérience de réalité virtuelle est finalement une promesse de téléportation. Elle est aujourd'hui en partie restreinte pour tout ce qui est contenus dits « 360 degrés » car le spectateur est limité dans ses mouvements et n'a pas la possibilité de se déplacer à sa guise dans le contenu. Les acteurs du jeu vidéo l'ont bien compris et offrent depuis longtemps aux gamers d'évoluer librement dans le jeu et ce dans un univers au photoréalisme de plus en plus hallucinant.
Les paramètres à prendre en compte et les critères qui permettront à la technologie d'être totalement acceptée sont très nombreux. La question du son est primordiale. Selon le studio **Felix & Paul**, spécialisé dans la RV, *« La spatialisation et le sound design sont des composantes essentielles pour assurer une expérience pleinement immersive aux utilisateurs ».* L'audio représentera en effet un facteur essentiel pour assurer une expérience convaincante capable de retenir l'attention de l'utilisateur, et de lui redonner l'impression d'une écoute naturelle à 360°.

«L'audio : l'autre dimension de la réalité virtuelle» est extraite du livre blanc «Réaliser son premier film en Réalité Virtuelle et Vidéo 360» qui sera édité cet automne, à l'occasion de Screen4all, le forum des technologies disruptives et des nouveaux usages pour le cinéma et la télévision. François Klein producteur attaché aux nouvelles technologies vidéo, s'est allié à b<>com pour son expertise dans les nouveaux formats audio. Le ton est pratique, destiné aux producteurs, réalisateurs de contenus en réalité virtuelle.

Dans un premier temps, est proposé un état des lieux des différentes technologies associées au son 3D : formats de représentation et de restitution. Puis sont exposés des défis et contraintes imposés par ces nouveaux formats : prise de son, post-production, rendu, compression.

1 /LES FORMATS DE REPRÉSENTATION ET RESTITUTION

Les trois formats capables de représenter une scène sonore 3D

Le format Multicanal (ou 'channel-based', voir Lexique) est le format usuel en audio (musique, cinéma, radio, etc) aujourd'hui. Dans un fichier audio multicanal, chaque canal correspond à une enceinte fixe (les canaux gauche et droite en stéréo par exemple) ou à un groupe d'enceintes recevant toutes le même signal (les canaux surround au cinéma). Ainsi l'extension au « son 3D » est venue par l'ajout de canaux supplémentaires pour la hauteur, correspondant à des enceintes fixées au plafond. Auro 3D, soutenu par Barco, propose ainsi son nouveau format multicanal pour l'exploitation cinéma : l'Auro 11.1.

Le format Objet (ou 'object-based', voir Lexique) mis en avant par les nouveaux standards au cinéma (Dolby Atmos, DTS :X) repose sur la description du mouvement des sources dans le temps. Ainsi, contrairement à une chaîne de production classique où la bande son est produite sur un nombre de canaux donnés (comme le 5.1, la stéréo etc), le format objet décline chaque source en objet : le fichier audio et un fichier de métadata décrivant ses déplacements au cours du temps. Un calculateur de rendu (renderer) présent sur le lieu de diffusion se charge de spatialiser chaque objet en fonction de la configuration des haut-parleurs du lieu (dolby CP850 pour le cinéma p.ex).

Le format Ambisonics est un format 'scene-based' (voir Lexique), c'est-à-dire qu'il décrit non plus des sources individuelles (comme les formats 'channel-based' ou format objet), mais le champ sonore dans son intégralité (en 2D ou 3D). La précision de l'environnement sonore repose sur le nombre de composantes (correspondant à des canaux) utilisées pour sa description (pour leur calcul, voir lexique). Résultat : l'Ambisonics apporte une flexibilité dans sa manipulation car elle est « indépendante » du dispositif de restitution. Ce format impose un décodage pour pouvoir être écouté (en binaural ou sur n'importe quelle configuration de haut-parleurs), il est également associé à un type de microphone spécifique, nous reviendrons plus tard sur ce point.

Chacun de ces formats impose un dispositif de restitution

Certains avec des positions d'enceintes fixes pour le format multicanal (configuration 5.1, 7.1, etc), des positions plus flexibles pour les formats objet et Ambisonics. D'autres solutions de rendu « transverse » existent, associées à ces différents formats.

L'écoute binaurale est une technique de restitution recréant une écoute en trois dimensions au casque ou via une paire d'oreillettes (elle ne fonctionne pas directement sur haut-parleurs). Elle repose sur la recréation de l'écoute naturelle en modélisant

l'influence de la forme de nos oreilles et de notre corps sur les sons qui nous entourent : ce sont les filtres HRTF (Head Related Transfer Function) propres à chaque individu. L'effet fonctionne de façon optimale si l'on dispose de filtres HRTF personnalisés. Si ce n'est pas le cas on peut obtenir un effet convaincant avec les différents profils HRTF disponibles. Des profils génériques existent, comme celui du mannequin KEMAR HATS qui simule une oreille « standard ». Chaque paire de filtres HRTF correspond à une position d'écoute, aussi un profil d'écoute correspond à un grand nombre de paires de filtres HRTF (pour les différentes positions d'écoute à 360°). Une écoute binaurale d'un contenu multicanal est possible en positionnant un haut-parleur virtuel à chaque position d'écoute donnée (-30° et +30° pour de la stéréo p.ex) et en décodant le tout. Il en va de même pour le format objet (Dolby et DTS proposent des décodages en binaural dans leur format propriétaire, respectivement Atmos for VR et X :Headphones) et pour le format Ambisonics.

Le mannequin KEMAR HATS (kemar.us) simule une oreille «standard»

La WFS (Wave Field Synthesis) est une technique de restitution dérivée du format Objet, où chaque source est représentée par un fichier audio et des métadata de localisation. Néanmoins, le décodage repose sur une reconstruction du champ sonore (comme le format Ambisonics) mais nécessite un très grand nombre de haut-parleurs. Ce dispositif permet notamment de simuler une source focalisée, c'est-à-dire placée à l'intérieur du dispositif, comme si elle était à côté de vous.

D'autres techniques de rendu 3D existent mais sont moins répandues et certaines d'entre elles restent expérimentales, aussi nous ne nous attarderons pas dessus. Chaque format propose à sa manière un « son 3D », il convient donc de bien avoir en tête pour chacun, le format audio des fichiers, les traitements possibles et le système de rendu souhaité : haut-parleurs ou casque. C'est ce qui permettra de choisir le type de spatialisation sonore le plus approprié à l'expérience en Réalité Virtuelle souhaitée. Chacun de ces formats fait également appel à des compétences particulières importantes dans une production de son spatialisé pour de la vidéo 360.

2 / LES NOUVEAUX DÉFIS DE TRAITEMENT DU FORMAT AUDIO

Enregistrement / Prise de son

Un tournage en 360° VR impose de nouvelles contraintes sur la prise de son, dont il est important d'avoir conscience pour tirer le meilleur de chaque prise.

En effet lors d’une prise de vue à 360° (ou d’une prise de vue fragmentée pouvant couvrir un angle de vision allant de 90° jusqu’à 210°), le perchman perd énormément en latitude en comparaison à un tournage classique. Il doit trouver des astuces pour récupérer le son en direct : utilisation de micro HF ou de microphones reliés à l’action (voir Strangers avec Patrick Wilson par exemple). Plus généralement, l’angle occupé sous le dispositif permet de placer un ou plusieurs microphones en fonction de l’encombrement et des possibilités de masquage en post-production. De même, une utilisation détournée du décor peut permettre plus facilement d’y cacher des micros « pot de fleur » ou même un perchman. L’ingénieur son de tournage devra adapter les solutions existantes avec les nouvelles contraintes de la prise de vue en 360VR.

Avec l’ouverture du champ de vision, il faut adapter la captation audio. Les dispositifs de captation multicanal « classiques» (croix IRT, Double MS, ABXY, ...) ou des systèmes de captation binaurale peuvent être utiles pour la captation de sons seuls, mais peuvent également s’avérer encombrants au pied de la caméra. Par exemple, un microphone Ambisonics suffisamment compact pour être placé au pied de la caméra 360, hors du champ de vision ou prenant peu de place pour être enlevé facilement en post-production, peut aussi être envisagé car il est capable d’enregistrer à 360°, avec l’information de hauteur. De plus, grâce à l’encodage en format Ambisonics, les signaux enregistrés par ce micro pourront être modifiés à posteriori (on pourra faire tourner l’image sonore par exemple). Néanmoins, son utilisation doit être faite avec précaution car, en l’absence de directivité, tout bruit parasite interfèrera et entrainera un éloignement des voix comparées à un micro canon. Un bon exemple : la scène de fin de I philip, d’Okio Studio, dans laquelle une personne parle au bord de la mer. L’ambiance est très enveloppante, l’actrice, pourtant située à quelques mètres, nous parle avec une voix lointaine. Les bruits ont été contrôlés et donnent un rendu plus subjectif.

Dans I Philip, cet effet « comme dans la vraie vie » nous étonne dans nos habitudes cinématographiques, où les voix sont généralement bien définies vis-à-vis de l’ambiance.

Le monitoring d’un tel micro sur le plateau peut se faire via un décodage B-Format vers stéréo, intégré dans la plupart des enregistreurs du commerce (Cantar X, Nagra VI, 788T, ...).Un décodage binaural pour le monitoring en tournage est encore expérimental et doit passer par une installation dédiée, généralement via un décodage logiciel. A noter aussi, la plupart des équipements de captation VR360 ne disposent pas de système de synchronisation (WC ou Timecode), il convient donc de revenir au clap pour l’enregistrement en attendant de nouveaux équipements. La synchronisation entre les micros d’appoint et le microphone « principal » (Ambisonics p.ex) reste nécessaire. La gestion des silences raccord est la même qu’en tournage classique : à enregistrer également par le micro Ambisonics si on l’utilise.

En résumé, en amont du tournage, il est important de bien cerner le format audio associé à la vidéo 360 en partenariat avec le responsable de la post-production son, car celui-ci va impacter toute la chaîne de production du tournage jusqu’au rendu final.

Processing / Post-production

L’introduction du « son 3D » dans une chaîne de production audio n’est pas nouvelle, le cinéma et le jeu vidéo s’intéressent à ces procédés depuis leur création, mais se basent vis-à-vis de l’écran 2D. La radio fournit déjà des contenus à écouter en binaural mais sans support visuel. Ainsi, l’immersion amenée par la technologie 360 VR est inédite et doit aller de pair avec une écriture audio adaptée pour produire les meilleurs contenus. Beaucoup de productions réalisées en 360 VR négligent l’audio, dégradant ainsi considérablement l’expérience vécue.
Ainsi, on ne peut pas concentrer l’attention sur l’ensemble du champ de vision sans fatiguer le spectateur ; c’est vrai aussi pour l’audio. Il est essentiel de bien distinguer les plans sonores (action en cours, ambiance de fond, etc.) si l’on veut concentrer l’attention de l’auditeur; ou les flouter complètement si on veut laisser le spectateur explorer l’environnement. Des sources sonores (une voix off ou de la musique) peuvent également être internalisées, c’est-à-dire entendues à l’intérieur de la tête, en comparaison à l’environnement sonore entendu comme « à l’extérieur », créant une différence de plan très intéressante pour la narration.

Micro Eigenmike, tournage de Refuge par b<>com

L’utilisation de la spatialisation doit être en accord avec le contenu

Toutes les sources ne sont pas intéressantes à spatialiser en 3D ! De plus, attention à la précision de la technologie utilisée pour spatialiser ces sources. Toutes les techniques ne se valent pas. Certaines peuvent même créer un effet indésirable de confusion entre des sources audio trop proches. La vision peut néanmoins améliorer ce phénomène si on associe chaque source à un personnage/objet identifiable (effet ventriloque). L’impression de mouvement lors d’un travelling pour un auditeur peut également être favorisée si l’on place des sources ponctuelles fixes sur le parcours, donnant ainsi un point de repère pour le mouvement.

Dans la vie courante, très peu d’éléments émettent du son autour de nous, aussi il est préférable de ne pas surcharger l’auditeur d’informations. De même pour les sources en mouvement : l’effet de l’hélicoptère passant

au-dessus de la tête marchera une fois, mais fatiguera s'il est trop répété. Il faut également prendre en compte la perception psycho acoustique, biaisée par nature, si l'on veut un rendu réaliste. Nous percevons mieux les sources sonores placées devant nous ; les sons situés à l'arrière pouvant être entendus comme placés à l'avant. La confusion avant-arrière est naturellement présente dans notre écoute. Cette perception peut être améliorée si le son a déjà été entendu auparavant. La perception latérale est également imprécise. Quant à la perception en hauteur, elle est inférieure à celle sur le plan horizontal, et dépendante de la nature du signal. Les mouvements de tête apportent aussi des informations complémentaires sur la localisation des sources. Il convient donc de placer les sons en accord avec cette perception. Par exemple, un son ponctuel placé à l'arrière fera naturellement tourner la tête pour mieux le percevoir. Cela peut être un effet de narration intéressant si on veut focaliser l'attention de l'utilisateur, ou un effet de gêne si cette focalisation n'est pas désirée. L'utilisation de la zone arrière/zénithale en audio est toutefois plus intéressante pour l'enveloppement que pour la spatialisation.

Le workflow de postproduction audio variera en fonction du format audio choisi. Il existe aujourd'hui des solutions de plugins s'intégrant dans les stations de travail audionumériques standards et amenant différentes possibilités. Nous n'entrerons pas trop dans les détails techniques, car il importe au responsable de la post-production de choisir la solution adaptée aux besoins du film et à l'intégration avec sa chaîne de production. Signalons quand même les possibilités offertes par les différents formats.

Micro contrôlé par logiciel TSL SoundField SPS200-K3

Différents formats : différentes solutions en post-production

La production pour écoute binaurale peut se faire de différentes manières. Il est possible de repartir d'une chaîne de travail standard 'channel-based' (un montage en 5.1 p.ex), puis d'effectuer une synthèse binaurale à la dernière étape, juste avant l'export. Il s'agit d'une simulation au casque d'une écoute sur haut-parleurs virtuels. Une chaîne de travail entièrement binaurale peut également être envisagée : on inclut un panner binaural par piste à spatialiser. Des solutions hybrides entre les deux sont également possibles : si l'on veut combiner un montage 5.1 (ou plus), des pistes stéréo ou mono non spatialisées (une voix off ou de la musique p.ex) et des prises de son en binaural natif (i.e réalisées avec des micros dans les oreilles). L'utilisation d'un effet de salle sur le mix final peut aider à externaliser l'écoute pour apporter davantage de réalisme par rapport à une écoute sur haut-parleurs. Quelle que soit la solution choisie et une fois l'export binaural réalisé, le fichier présentera un point de vue fixe à l'auditeur, qu'il sera impossible de manipuler (pour le faire pivoter avec un dispositif de headtracking p.ex). Un export binaural dynamique (qui suit les mouvements de l'auditeur) peut alors être possible. Nous aborderons ce sujet dans la partie «Rendu».

Le format Ambisonics offre un large éventail d'outils avancés pour l'enregistrement et la diffusion de contenus audio 3D. Cette technologie repose sur un format audio 3D générique capable de représenter le champ sonore, et ce indépendamment de la configuration de sortie. Ainsi, utiliser ce format peut permettre, en partant du même contenu Ambisonics, de réaliser un export à la fois vers une configuration de haut-parleurs (5.1, 7.1, etc) ou vers une écoute binaurale, à condition de disposer du décodeur associé. Cette notion de décodage est fondamentale, car le format Ambisonics HOA est un descriptif de la scène sonore

qui ne s'écoute pas par nature : il doit être décodé. A noter que la précision du rendu final est liée à la précision du format Ambisonics. Si le contenu est encodé à l'ordre 1 (4 canaux), le rendu risque d'être assez peu précis. En augmentant l'ordre, on augmente la précision mais également le nombre de canaux à transmettre. Un compromis doit donc être choisi entre la précision désirée et le poids de données voulues. Ce format permet aussi d'effectuer des transformations spatiales dynamiques comme une rotation, un zoom sur la scène sonore (si on veut concentrer l'attention sur une portion de l'espace en particulier), masquer une zone de l'espace, etc. Il est particulièrement adapté pour un rendu spatial dynamique, que ce soit sur haut-parleurs ou pour une écoute binaurale. Il est également possible d'associer ce format à d'autres, si l'on souhaite par exemple effectuer une rotation sur une seule partie du mixage, tandis que d'autres sources restent fixes : la musique par exemple.

Choisir une solution de monitoring audio cohérente avec le format visé est essentiel : il est tout à fait possible de réaliser un montage son en 5.1 avec une écoute sur enceintes puis d'effectuer un export pour une écoute binaurale. Il faudra cependant prendre un soin particulier à choisir le profil HRTF utilisé pour l'export et contrôler à l'écoute si la compatibilité est respectée. On peut également laisser l'auditeur choisir son profil HRTF en passant par un player dédié (c'est le cas dans le lecteur Nouvoson de Radio France, créé par Orange). Un outil permettant de contrôler le mixage avec la vidéo 360 VR est également nécessaire via un casque de VR ou une vue interactive. On pourra utiliser un lecteur externe ou celui de la station de travail elle-même si celle-ci le permet.

Casque Sennheiser HD 650

Rendu

Le rendu associé à l'audio spatialisé pour la VR360 est intrinsèquement lié au support choisi et dépend du degré d'interactivité voulu par le producteur. Souhaite-il un rendu statique, ou une écoute interactive, liée par exemple aux mouvements de l'auditeur ? De même souhaite-il un rendu uniquement au casque, à l'écoute sur haut-parleur ou compatible sur les deux dispositifs ?

Les grands enjeux de ces questions sont le dispositif (casque, haut-parleurs) et les solutions de restitution (codec, renderer) dont l'utilisateur dispose. Pour l'instant, aucune standardisation n'ayant émergée, un grand nombre de lecteurs vidéo existent avec des possibilités différentes vis-à-vis de l'audio : par exemple la plateforme video 360 de Youtube permet la lecture de fichiers ambisonics pour les équipements fonctionnant avec Android. Il n'est donc pas exclu qu'à l'avenir les plateformes de diffusion s'ouvrent vers des formats plus évolués (5.1, binaural) et guident les nouveaux standards audio. En conséquence, pour l'instant la garantie de la restitution spatialisée de l'audio passe par le développement d'une application dédiée à la lecture de contenu intégrant un renderer audio avec un format de fichier dédié (propriétaire ou non) ou par l'export sous format spécifiques aux différentes plateformes (youtube, milk VR, ...). On peut citer pour exemple le lecteur du GearVr, l'application VRSE, etc. Notons tout de même que la norme MPEG-H 3D Audio (qui intègre le format Ambisonics) devrait bientôt être déployée (en Corée du Sud notamment) et permettrait de répondre à cette problématique. On peut citer également les solutions de Dolby (Atmos for VR), du Fraunhofer (Cingo) ou de b<>com Reality Surround déjà existantes.

Coup d’oeil aux solutions de rendu statique. Si l’on considère un rendu binaural, les possibilités reviennent à choisir un profil de HRTF cohérent avec le mix réalisé afin d’en fournir un export. Cette étape peut être réalisée en amont (avec un fichier audio unique) ou au moment de la lecture (si le player intègre un décodeur binaural permettant la sélection de différentes paires d’HRTF - cf nouvoson). De même, si l’on souhaite un rendu sur haut-parleurs, il faut s’assurer que le player comprenne un renderer adapté au format (Ambisonics, Atmos, etc) ou penser à effectuer un export en amont correspondant à la configuration de haut-parleurs visée. Pour résumer, un export audio « statique » en vue d’un contenu 360VR impose de connaître les possibilités du lecteur utilisé pour le rendu. Il est indispensable de choisir et bien connaitre la solution de rendu dès le début du projet.

b<>com utilise le micro Eigenmike

Pour un rendu dynamique, il faut d’abord définir la latitude accordée à l’utilisateur. La rotation de la bande son avec les mouvements de tête est la première chose venant à l’esprit pour de la VR360. Mais au vu des possibilités offertes par les différents formats, d’autres applications et interfaçages sont possibles et évidemment ouverts aux producteurs de contenus. Il suffit de voir les différents accessoires prévus pour accompagner la VR pour entrevoir de nouvelles possibilités. Le format objet ouvre également de nombreuses possibilités puisque chaque objet est indépendant et manipulable en fonction des métadata transmises avec lui (position, volume, etc).
Comme pour le rendu statique, le rendu dynamique est contraint par la taille et le nombre de canaux gérés par l’application (avec le HOA, ce nombre peut grimper rapidement) ainsi que les capacités du renderer, qui doit à la fois interpréter les informations spatiales transmises par le casque de Réalité Virtuelle et manipuler la trame sonore audio en conséquence.

Compression / Débit

L’audio spatialisé nécessite en général un plus grand nombre de canaux, ce qui implique une taille de données plus importante. Cet enjeu est important pour les applications de VR360 généralement dédiés à des équipements ayant une taille mémoire réduite ou passant par du streaming avec les contraintes de débit associées. Ainsi dans cette partie, nous nous

intéresserons seulement aux codecs lossy (avec perte de données) qui répondent à cette problématique, car même si des standards existent déjà, les choses sont variables en fonction des formats de représentation (multicanal, objet, Ambisonics).

Le format **multicanal** existe depuis le début du cinéma dans les années 20 , aussi il existe un grand nombre de codecs associés. Pour la transmission, les principaux codecs sont : Dolby avec le format AC-3 et son extension EAC-3, Fraunhofer avec l'HE-AAC (extension du format AAC, qui permet d'aller jusqu'à 48 canaux). Ces deux codecs nécessitent l'achat d'une licence pour être décodés. Ce sont les principaux codecs présents sur les équipements aujourd'hui (TV, Internet, Console, Blu-ray ...). Citons également l'alternative open source Opus, intégrée dans la plupart des navigateurs web, pouvant aller jusqu'à 256 canaux.

Le format **Objet** est apparu plus récemment. Aussi même si des formats normalisés existent, comme l'AC-4 de Dolby ou sont en attente d'intégration comme le MPEG-H 3D, il existe pour l'instant peu d'équipements contenant les décodeurs appropriés. Par exemple l'arrivée de l'Atmos et du DTS-X pour les processeurs de diffusion home-cinéma est très récente.

Quant à l'**Ambisonics**, et malgré son ancienneté, il n'a pas fait l'objet d'une normalisation durable. Cependant, il est inclut dans le standard audio MPEG (MPEG-H 3D). Des solutions propriétaires existent, et imposent de disposer d'un renderer associé en bout de chaîne. Le format Opus pourrait être une solution d'intégration, mais nécessite d'être développé.

Face à cette profusion de formats, il est bon de mentionner que des conteneurs existent afin de pouvoir encapsuler plusieurs formats, c'est le cas du MDA (Multi Dimensionnal Audio) développé par DTS, qui permet de faire cohabiter plusieurs formats multicanaux et objets dans le même conteneur. Le format ADM (Audio Definition Model), développé dans la même optique, fait cohabiter plusieurs formats multicanal, objet et HOA dans le même fichier.

LEXIQUE

Ambisonics : Format de captation, synthèse, traitement et rendu d’environnement sonore. Il est caractérisé par un ordre de précision : 1 (appelé aussi B-Format) ou plus (HOA pour Higher Order Ambisonics). Pour chaque ordre M, on utilise un nombre de canaux spécifique pour décrire la scène sonore. Ce nombre est calculé comme ceci : pour un rendu 2D : 2M +1, pour un rendu 3D : $(M+1)^2$. Pour un ordre de précision de 4 (recréant les performances de l’oreille humaine), le format HOA contiendra 9 canaux pour un rendu 2D, et 25 canaux pour un rendu 3D.

B-Format : Format Ambisonics à l’ordre 1.

Binaural : Format d’écoute deux canaux « recréant » l’écoute naturelle (doit impérativement être écouté au casque). Une prise de son en binaural natif consiste à enregistrer avec des micros mis dans les conduits auditifs, « enregistrant » le son perçu par nos oreilles.

Channel-based : Format audio pour lequel chaque piste correspond à un canal de restitution fixe. Par exemple pour un fichier stéréo, deux canaux correspondent aux enceintes gauche et droite. Il en va de même pour les formats plus complexes (5.1, 7.1 Auro 3D 9.1, etc).

HOA : Format Ambisonics à l’ordre > 1.

HRTF (Head Related Transfer Function) : modélisation de l’impact du corps humain sur l’écoute.

Object-based : Format audio basé sur une représentation par « objet ». Un objet contient une source audio plus des métadonnées décrivant sa position dans l’espace au cours du temps. Il nécessite une étape de décodage pour être écouté (ex : Dolby ATMOS, DTS :X).

Scene-based : Format audio où la scène sonore est décrite entièrement et indépendamment du système de restitution (haut-parleurs, casque). Ce format nécessite une étape de décodage pour être écouté (ex : Ambisonics, WFS)

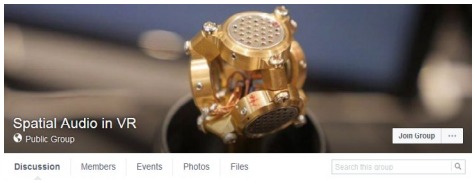
WFS (Wave Field Synthesis) : Technique de restitution d’environnement sonore nécessitant un grand nombre de haut-parleurs.

ALLER PLUS LOIN

La page A Sound Effect est une bibliothèque en ligne propsant des contenus Ambisonics entre autres
www.asoundeffect.com/ambisonics-primer/

Nouvoson, le premier site web innovant dédié à l'audio spatialisé
http://nouvoson.radiofrance.fr/

Le blog lesonmulticanal, les dernières nouveautés sur le son 3D
www.lesonmulticanal.com



Le groupe Facebook, Spatial Audio in VR :
www.facebook.com/groups/positional.audio.in.vr/

The Sursound list: pour les chercheurs et producteur de son multicanal qui veulent échanger et débattre
https://mail.music.vt.edu/mailman/listinfo/sursound

/ A propos /

b<>com

Grâce à ses innovations, l'Institut de Recherche Technologique (IRT) b<>com contribue à la transformation digitale européenne. Ses 230 chercheurs développent des outils, produits et services qui facilitent la vie quotidienne. Ils se concentrent sur deux domaines de recherche : l'hypermédia (images ultra haute définition, son 3D, contenus intelligents, réalités virtuelle et augmentée...) et les réseaux ultra haut débit plus agiles (cloud, cyber-sécurité, ultra haut débit mobile, résilience des réseaux, Internet des objets...). Parmi les multiples domaines d'application de ces technologies, la e-santé permet à b<>com de participer à la révolution numérique du système médical. Issu d'un partenariat public/privé, l'IRT réunit les meilleurs experts issus du monde industriel et académique sur son campus de Rennes, et ses sites de Lannion et Brest.
www.b-com.com

Livre blanc «Réaliser son premier film en Réalité Virtuelle et Vidéo 360»

L'essor de la Réalité Virtuelle dans la production de programmes a bouleversé la façon d'écrire et de produire en proposant un nouveau format d'image panoramique à 360°. Le spectateur, alors placé dans l'espace narratif en vue subjective, devient un acteur. Cette conception bouleverse toutes les structures d'écriture et valeurs de plans établies par plus de 120 ans de cinéma. Tout est à réinventer : de la prise vue à la distribution.
Quelle forme narrative utiliser ? Quel point de vue donner à l'utilisateur ? Doit-il être contemplatif ou peut-il interagir avec le contenu ? Comment filmer à 360° et quelles caméras privilégier ? Comment mettre en place un workflow technique ? Des questions auxquelles "Réaliser son premier film en Réalité Virtuelle et Vidéo 360" vous aidera à répondre. Publié par Génération Numérique, éditeur des magazines Mediakwest et Sonovision, ce livre blanc rassemble les principales connaissances actuelles pour la réalisation d'un projet VR.
Ecrit par François Klein, producteur, avec la participation d'experts de l'Institut de Recherche Technologique b<>com, le livre est enrichi d'interviews et d'avis de professionnels.

François Klein, auteur de «Réaliser son premier film en Réalité Virtuelle et Vidéo 3D»

François Klein est un producteur attaché aux nouvelles technologies vidéo. Diplômé de Gobelins l'École de l'image, il met les convergences artistique et techniques au cœur de ses activités. Rompu à la production de projets en vidéo, animation, stéréoscopie-3D et transmédia, François s'associe en 2015 à Consortium-C pour créer Digital Rise, un studio de développement dédié à l'accompagnement des créatifs dans l'élaboration et la mise en forme de fictions innovantes en Réalité Virtuelle et stéréoscopie-3D.

{innovation}

comes with b com

Merci de votre attention

Site : **https://www.lesonbinaural.fr**

Mail : **b.lagnel@gmail.com**