

esmuc

Treball Fi de Grau

Tecnologies de so immersiu:

Una aplicació a la creació d'un viatge personal en tren

Estudiant: Arnau Abadal Casino

Especialitat/
Àmbit/Modalitat: Sonologia

Director/a: Perfecto Herrera

Curs: 2018-2019

Vistiplau
del director/a
del Treball

Extracte trilingüe

Català

Les múltiples tecnologies de so immersiu que tenim a dia d'avui permeten la creació d'espais sonors imaginaris que s'allunyen de la realitat. Degut a les seves característiques tècniques no són emprades rutinàriament, per la qual cosa estem més acostumats a l'escolta estèreo. Per tal de fer palès el seu potencial, en aquest treball es fa una contribució artística de so immersiu amb la creació de dues peces que representen un viatge personal en tren. Aquest objectiu requereix una recerca prèvia molt transversal que barreja diverses disciplines com són: Tècniques d'Enregistrament i Postproducció, Fonaments d'Acústica, Escolta Crítica, Estètica, i Composició Electroacústica, entre d'altres.

La recerca d'obres musicals que juguen amb el so dels trens és clau per tenir una base estètica de la qual partir. Crear una metodologia pròpia a l'hora de realitzar els enregistraments dels sons ferroviaris és imprescindible, com també l'estudi dels diferents sistemes de so espacial que es troben a l'actualitat (estèreo, binaural, surround, àudio basat en objectes, sound field i Wave Field Synthesis). Tot això conforma la base sobre la qual les composicions "Dia i Nit" i "Essència", incloses al treball, han estat realitzades i ideades per a reproduir-se en un sistema d'àudio multicanal 5.1.

Español

Las múltiples tecnologías de sonido inmersivo que tenemos a día de hoy permiten la creación de espacios sonoros imaginarios que se alejan de la realidad. Debido a que sus características técnicas no son empleadas rutinariamente, estamos más acostumbrados a la escucha estéreo. Con el fin de hacer patente su potencial, en este trabajo se hace una contribución artística de sonido inmersivo con la creación de dos piezas que representan un viaje personal en tren. Este objetivo requiere una búsqueda previa muy transversal que mezcla diversas disciplinas tales como: Técnicas de Grabación y Postproducción, Fundamentos de Acústica, Escucha Crítica, Estética, y Composición Electroacústica, entre otras.

La búsqueda de obras musicales que juegan con el sonido de los trenes será clave para tener una base estética de la cual partir. Crear una metodología propia a la hora de realizar las grabaciones de los sonidos ferroviarios será imprescindible, como también el estudio de los diferentes sistemas

de sonido espacial que encontramos actualmente (estéreo, binaural, surround, audio basado en objetos, sound field, y Wave Field Synthesis). Todo esto conformará la base sobre la cual las composiciones “Dia i Nit” y “Essència”, incluidas en el trabajo, han estado realizadas y ideadas para reproducirse en un sistema de audio multicanal 5.1.

English

The multiple immersive sound technologies that we enjoy nowadays allow us to create imaginary sound spaces that move away from reality. Because of their characteristics are not routinely employed, we are more used to the stereo listening. In order to show its potential, in this project an artistic contribution of immersive sound is made with the creation of two pieces that represent a personal trip by train. This objective requires a very transversal research that mixes several disciplines such as: Recording and Post-Production Techniques, Acoustic Foundations, Critical Listening, Aesthetics, and Electroacoustic Composition, among others.

The research of musical pieces involving the sound of trains will be the key for having the aesthetic base from which to start. We will need the creation of our own methodology when making recordings of railway sounds, as well as the study of the different spatial audio systems that we currently have (stereo, binaural, surround, object based audio, sound field, and Wave Field Synthesis). All this will form the basis on which the compositions “Dia i Nit” and “Essència”, included in this project, have been realized and devised for reproducing in a 5.1 multi-channel audio system.

*Al Perfecto i al Pau per la seva dedicació i recolzament,
al Gerard, al Fèlix, a l'Adrián, al Sixto, a l'Àlex, a l'Enric i al Ferran pel seu suport tècnic i creatiu,
i al Cels, al Josep, a la Mar i al Martí per acompanyar-me durant aquesta etapa.*

Índex de continguts

| | |
|---|-----------|
| Introducció..... | 3 |
| Motivació..... | 3 |
| Objectius..... | 4 |
| Metodologia..... | 4 |
| Mapa conceptual del treball..... | 5 |
| 1. Trens i música..... | 7 |
| 2. Enregistraments de camp..... | 10 |
| 2.1. Implicacions de l'enregistrament de camp..... | 11 |
| 3. La línia R1..... | 13 |
| 3.1. Història..... | 13 |
| 3.2. Característiques de cada sèrie..... | 14 |
| 3.2.1. Sèrie 447 de Renfe..... | 14 |
| 3.2.2. Sèrie 463-464-465 de Renfe..... | 15 |
| 3.2.3. Sonoritat de les sèries 447 i 465..... | 16 |
| 4. So immersiu..... | 18 |
| 4.1. Localització del so..... | 18 |
| 4.1.1. Localització en el pla horitzontal..... | 19 |
| 4.1.2. Percepció de l'elevació..... | 20 |
| 4.1.3. Factors que afecten la localització..... | 21 |
| 4.1.4. Percepció de la distància..... | 22 |
| 4.1.5. Funció de Transferència Relacionada amb el Cap (HRTF)..... | 23 |
| 4.2. Sistemes de so espacial..... | 24 |
| 4.2.1. Estèreo..... | 25 |

| | |
|---|-----------|
| 4.2.2. Binaural..... | 29 |
| 4.2.3. So Surround..... | 34 |
| 4.2.4. Àudio basat en objectes..... | 40 |
| 4.2.5. Sound Field..... | 43 |
| 4.2.6. Wave Field Synthesis (WFS)..... | 46 |
| 4.3. Productes per treballar amb so immersiu..... | 48 |
| 5. Creació del viatge sonor en tren..... | 49 |
| 5.1. Tria del sistema de reproducció..... | 49 |
| 5.1.1. Eines de postproducció de so immersiu..... | 50 |
| 5.2. Metodologia dels enregistraments..... | 50 |
| 5.2.1. Tècniques de microfonia..... | 51 |
| 5.2.2. Realització dels enregistraments..... | 52 |
| 5.3. El viatge immersiu..... | 53 |
| 5.3.1. La seva estètica..... | 53 |
| 5.3.2. Composició de les peces..... | 54 |
| 6. Resultats..... | 55 |
| 6.1. Les peces..... | 55 |
| 6.1.1. Dia i Nit..... | 55 |
| 6.1.2. Essència..... | 56 |
| 6.2. Mescla immersiva..... | 57 |
| 7. Conclusions..... | 59 |
| Bibliografia..... | 61 |
| Annex I..... | 65 |
| Productes hardware de so immersiu..... | 65 |
| Productes software de so immersiu..... | 67 |
| Annex II - Material adjunt..... | 72 |

Introducció

Motivació

Per desplaçar-me cada dia de casa a l'ESMUC utilitzava la línia de tren R1 de Rodalies Renfe. La seva multiculturalitat i immensa quantitat de passatgers que l'usen generen unes textures i línies melòdiques que s'entrellacen amb el so mecànic, industrial i de gran rang dinàmic dels trens. Tots aquests sons junts conformen un paisatge sonor que sempre m'ha semblat molt divers i tímbricament ric. És per aquest motiu que vaig decidir que volia immortalitzar l'essència dels meus viatges diaris.

A l'hora de reproduir els viatges tenia clar que volia anar més enllà del sistema estèreo estàndard al qual estava acostumat. També calia tenir en compte que el sistema que triés hauria de ser fàcil d'instal·lar, ja que els espais en general no estan preparats per anar més enllà de l'estèreo. Com bé diu Christian Marclay quan parla de l'art sonor exposat als museus:

“We still think of museum galleries as nineteenth-century galleries, like “How do we hang this on the wall, how do we light it?” But nobody knows anything about sound—how you hang a speaker, how you EQ it to the room... More and more museums have a lounge-type listening room, but there are still a lot of changes that need to happen before the art world is ready to present sound as art” (Licht, 2007).

Aquest fet em va generar diverses preguntes: Com podia presentar el viatge sonor? Quin sistema podria utilitzar per submergir el públic als viatges? Intentant trobar resposta a aquestes preguntes vaig topar amb el so immersiu. Aquest em permetia augmentar el ventall de possibilitats creatives a l'hora de crear espais sonors imaginaris i impossibles. Al mateix temps, però, havia de ser un sistema pràctic per poder-lo presentar.

Cal aclarir que el pes d'aquest treball no està en la composició del viatge, sinó en la cerca i exploració de tècniques de so immersiu i aproximacions estètiques al context del so ferroviari. Fins llavors només n'havia vist de passada algunes. És justament per aquest motiu que m'he endinsat en les diferents tècniques per posteriorment poder-les aplicar a les peces.

Objectius

L'objectiu principal del treball és la creació subjectiva d'un viatge personal en tren mitjançant l'estudi de les tècniques de so immersiu que tenim a dia d'avui. Tanmateix, per aconseguir-lo també caldrà fer recerca en altres àmbits que es detallen al següent apartat.

Metodologia

- Breu recerca d'obres musicals que treballin amb els sons dels trens per tal de conèixer les diverses estètiques i possibilitats compositives.
- Tria del mètode de reproducció de so immersiu. Suposarà:
 - L'anàlisi del paisatge sonor que cal enregistrar.
 - L'estudi de les tècniques de so immersiu.
- Enregistrament dels sons de la línia R1. Suposarà:
 - La tria de la tècnica d'enregistrament i material a utilitzar.
 - Dissenyar un guió a seguir en el moment d'enregistrar sons.
- Composició de les peces per a un sistema de reproducció d'àudio immersiu.
- Avaluació i anàlisi de les peces immersives i els sistemes emprats.

Mapa conceptual del treball

Aquest treball està format per 7 capítols. El primer de tots parla de l'interès musical que tenen els trens com també la seva aparició en diverses obres. Això ens portarà a parlar al capítol 2 dels enregistraments de camp i què cal tenir en compte a l'hora de realitzar-los. Al capítol 3 es farà un breu pas per la història de la línia R1, i posteriorment, una descripció més tècnica dels trens que hi circulen per poder entendre la seva sonoritat. Al capítol 4 es desenvoluparà el so immersiu, començant pels mecanismes de localització del so i posteriorment definint cadascuna de les tècniques de so immersiu. Al capítol 5 es parlarà del mètode seguit a l'hora de crear el viatge, passant per la tria del sistema de reproducció, la tècnica de microfonia i les decisions preses a la composició del viatge. Al capítol 6 es descriurà amb detall el resultat final que es presenta. Al capítol 7 es parlarà de les conclusions extretes del treball. Per acabar, es trobarà la bibliografia i seguidament els annexos.

1. Trens i música

Un dels grans canvis que ha sofert el paisatge sonor mundial el podem atribuir a la revolució industrial. El so que emeten les màquines produeix un efecte que R. Murray Schafer anomena *flat line* (línia plana), generat per l'impuls repetitiu i de ritme continu i elevat, com per exemple el d'un motor. Aquest efecte artificial al qual hi estem tan acostumats, i que sovint no en som conscients que és present al nostre dia a dia, rarament el trobem a la natura, amb l'excepció del so que generen alguns insectes com les cigales (a més, la intensitat sonora és molt menor) (Murray Schafer, 1977).

Aquesta nova sonoritat va despertar l'interès d'alguns artistes futuristes com Luigi Russolo, que va escriure el manifest *The art of Noise* (L'Art del Soroll). També va construir instruments musicals que generaven soroll, batejats amb el nom d'*Intonarumori*, fets que el van portar a ser considerat com un dels primers compositors de *noise music* (Murray Schafer, 1977).

La màquina que va exportar el so de la indústria moderna de les grans ciutats cap al camp va ser la locomotora de tren (Murray Schafer, 1977). No són pocs els compositors que al llarg de la història s'han interessat pels seus sons rítmics, continuats i tímbricament rics, que trenquen amb el paisatge sonor rural.

Arthur Honegger, l'any 1923 va compondre el poema simfònic *Pacific 231*¹. Aquesta peça fa referència a una locomotora de vapor, la Pacific. Apassionat de les locomotores, Honegger va declarar que aquesta obra no era la simple imitació dels sons de la locomotora, sinó que volia traduir en la textura orquestral l'experiència visual d'un tren d'alta velocitat, així com els seus sons (Epstein, 2018).

El 1948, Pierre Schaeffer va compondre una col·lecció anomenada *Cinq études de bruits* (Cinc estudis de sorolls). De les cinc peces que la conformen cal destacar la primera, anomenada *Étude aux chemins de fer*² (Estudi a les vies fèrries). És una peça de música concreta³ que es nodreix

1 Disponible a: <https://www.youtube.com/watch?v=wS6XAJd-9h8>

2 Disponible a: <https://www.youtube.com/watch?v=aL77mHnCrNs>

3 Gènere de música experimental que usa com a material primari sons enregistrats. Va ser desenvolupat l'any 1948 per Pierre Schaeffer.

d'enregistraments de sons de tren fets a l'estació de Batignolles, on Schaeffer juga amb la seva descontextualització provocant una escolta acusmàtica. Pretén presentar els sons com a pur material, desvinculant-lo de la seva procedència, per tal de poder emfasitzar els seus valors musicals com el ritme, la textura i l'alçada (Patrick, 2016).

El grup alemany de música electrònica Kraftwerk, l'any 1977 va publicar l'àlbum *Trans-Europe Express*⁴. A les bases d'algunes de les peces podem escoltar un ritme metàl·lic, repetitiu i poderós creat amb sintetitzadors, que recorda al d'una locomotora de tren. Karl Bartos, percussionista electrònic del grup, va explicar a la revista digital *The Quietus* que pel fet d'utilitzar el tren com a motiu principal estaven seguint el camí d'alguns com Pierre Schaeffer, qui va fer la primera peça de música concreta només usant el so dels trens (Murphy, 2017).

Onze anys més tard, al 1988, Steve Reich va escriure una peça de tres moviments per a quartet de cordes i cinta magnètica anomenada *Different Trains*⁵. En aquesta, els sons dels instruments de corda fregada es fusionen amb els de trens, diàlegs i altres sons mostrejats. Amb aquesta música, al llarg dels tres moviments Reich ens transporta a dos tipus de viatge en tren ben diferents l'un de l'altre. Un és el que feia de petit de Nova York a Los Angeles quan anava a visitar els seus pares separats. L'altre és el que feien els trens de l'Holocaust, que també podria haver fet ell de petit pel fet de ser jueu si hagués estat vivint durant aquells anys a Europa (Vázquez, 2002).

Per últim, l'àlbum més recent dels esmentats és *El tren fantasma*⁶, publicat el 2011 per Chris Watson. L'autor agafa els enregistraments del paisatge sonor que apareix al quart episodi de la sèrie documental *Great Railway Journeys of the World* (que ell mateix havia gravat). A diferència del documental, on el paisatge sonor queda sepultat darrere de la veu i la música, al disc el porta al primer pla donant-li protagonisme. D'aquesta manera posa en manifest i dramatitza el paisatge sonor constantment canviant entre el so industrial i el natural (Smith, 2016).

Els diferents compositors mencionats tracten el so dels trens de maneres diferents. A Schaeffer li interessa el so descontextualitzat, en canvi Kraftwerk imita sintèticament el so constant i repetitiu de la locomotora per contextualitzar-nos en una era industrialitzada. De la mateixa manera

4 Disponible a: https://www.youtube.com/watch?v=i_85fUIRImU

5 Disponible a: https://www.youtube.com/watch?v=1E4Bjt_zVJc

6 Disponible a: <https://chriswatsonreleases.bandcamp.com/album/el-tren-fantasma>

Honegger ho fa amb l'orquestra i Steve Reich amb el quartet, recreant aquest so mecànic. I per últim Chris Watson, que reivindica l'escolta activa d'un paisatge sonor molt contrastant, sense voler presentar-lo intacte de la mà humana. Com diu Frere-Jones "*Listen to your world. [...] It may be more interesting than all the things you buy to escape from it*" (Smith, 2016).

2. Enregistraments de camp

Algunes de les composicions esmentades anteriorment utilitzen enregistraments del so del tren com a material per compondre. Aquests són enregistraments de camp, o en anglès “field recordings”. Podríem definir aquest terme com a enregistraments d'àudio que es realitzen fora d'un medi controlat com seria un estudi de gravació. Aquesta definició tan general recull moltes disciplines: des de l'enregistrament documental que pretén ser el més objectiu i transparent possible per tal de representar auditivament una realitat, fins al gènere musical anomenat música concreta. No obstant, el que uneix a totes aquestes són les diferents maneres d'escoltar el nostre món que implicaran diferents consideracions estètiques, polítiques i socials. L'article *Sound Outside: The art of field recording* (Ableton, 2016) de què es parlarà al llarg d'aquest apartat ens aportarà diversos punts de vista d'aquest terme.

La separació entre objectiu i subjectiu a les diferents disciplines és molt difícil de delimitar. La simple tria del micròfon que utilitzarem i de la seva posició a l'espai ja està aportant una connotació subjectiva a l'enregistrament. Les paraules de Chris Watson ajuden a entendre-ho:

“[My approach] did expand and change direction a bit. I didn't regard the first pieces as pure documentation; I thought they stood as pieces themselves, and really there was nothing I could do to improve on that; I saw them as really quite sort of perfect "compositions", which came about through a process of documenting sounds. Then as I changed my recording techniques I began to hear individual aspects that I hadn't really heard or appreciated before, that really defied any sort of human-compositional techniques, so I became interested in just engaging with those and that's why the compositional technique for records like Outside the Circle of Fire was simply the microphone technique itself - that became sort of the main instrument, which was a change from Stepping into the Dark which was more atmospheric, more ambient, more recordings of space.” (Ableton, 2016).

Molts altres músics i artistes consideren que no existeixen els enregistraments purs o neutrals en els quals nosaltres no hi intervinguem. Com per exemple l'artista xinès Yan Jun, que opina que no hi ha divisió entre documentar i crear, ja que el simple fet de triar l'equip, la posició i prémer el

botó d'enregistrar són actes de composició. Tot i així sí que és cert que podem incidir més o menys sobre el resultat final.

També s'hauria de destacar un altre fet discutit per Lawrence English⁷. English diu que les condicions de l'era digital, les oportunitats de viatjar i l'accés abundant a pràcticament tot ha fet que els sons que ens envolten, fora de l'àmbit habitual de la majoria de les persones, es converteixin en redundants, i sovint, fins i tot culturalment "tòxics". Ell proposa que busquem noves perspectives que converteixin les experiències més habituals en situacions d'escolta profunda.

2.1. Implicacions de l'enregistrament de camp

Quan enregistrem intentem buscar les millors condicions acústiques, i aquestes les acostumem a trobar a l'estudi, però a dins no hi podem ficar un bosc o un tren. Sovint hem de sortir a l'exterior per copsar el so que busquem. En fer aquest canvi de medi ens podem trobar un seguit de problemes com el soroll del vent, els avions, els ocells i totes aquelles molèsties imprevisibles que no podem silenciar i que ens representaran un autèntic repte.

Degut que ens podem trobar amb tots aquest (i més) impediments, per poder fer un bon enregistrament de camp haurem d'haver pensat en un seguit de punts abans de començar a gravar (Marks, 2010):

- Definir l'objectiu de l'enregistrament: S'ha de saber si el so que volem capturar serà de gran importància dins del resultat final, o quedarà amagat entre altres sons. També és interessant saber en quin format serà reproduït el so.
- Redactar un llistat de tots els sons que necessitem: Ens permetrà saber quines tècniques d'enregistrament i equip haurem d'utilitzar, com també pensar en el lloc idoni on gravar.
- De quin pressupost disposem.
- Quin equip tenim disponible.

⁷ Nascut el 1976. És un compositor, artista i comissari, que actualment viu a Brisbane, Austràlia.

- Anticipar-nos a les condicions del lloc on gravarem i explorar la ubicació.

Un cop tinguem clars tots aquests punts serà el moment de decidir què necessitem realment per fer-ho possible.

En el cas d'enregistrar trens caldrà tenir en compte un altre seguit de punts més específic (Virostek, 2011):

- Predir el rang dinàmic del tren: Com bé sabem, un tren té un rang dinàmic molt gran, i sovint és molt difícil de preveure la seva evolució. Com que cadascun dels seus sons és d'interès cal decantar-se cap a una de les següents estratègies:
 - Augmentar el guany un cop ha passat el pic de màxima energia.
 - Establir un guany que permeti enregistrar tant els sons més discrets com els més forts.
 - Enregistrar amb dos micròfons, un amb el nivell atenuat pels sons amb nivells més alts i un altre pels que, per capturar-los, calgui un nivell de guany alt.
- Estar a una distància i angle adequats: No sempre ens interessa capturar els sons en un primer pla o en un cert angle per minimitzar el soroll de fons. A vegades també volem que prenguin espai ja que aquest també en forma part.
- No molestar i no ser molestat pels usuaris de la línia.
- Temps limitat a l'hora de muntar l'equip: El muntatge de l'equip requereix un cert temps, i el tren no espera. Cal ser-ne conscient.
- Aïllar els micròfons per tal de minimitzar sons que no desitgem: Per captar segons quins sons ens caldrà modificar el patró polar, col·locar algun tipus d'esponja o suspensió per evitar les vibracions del tren i fins i tot amagar-los per tal que l'entorn sonor actuï amb naturalitat.

3. La línia R1

3.1. Història

La línia R1 de Rodalies, antigament coneguda com a ferrocarril de Barcelona - Mataró, està catalogada com la més antiga de la Península Ibèrica. Es va inaugurar el 28 d'octubre de 1848 i estava delimitada per les ciutats de Barcelona i Mataró. La seva funció principal era comunicar ràpidament i eficaç Barcelona i els centres industrials que s'havien desenvolupat al llarg de la costa de l'actual Maresme. La flota estava formada per quatre locomotores batejades amb els noms de Mataró, Barcelona, Catalunya i Besòs.

L'empresari Miquel Biada i Bunyol i el financer Josep M. Roca van ser els impulsors d'aquesta línia, que transcorre per dins del primer túnel d'Espanya, el de Montgat (Siles, 2015).

Al cap de 9 anys, al 1857, es va prolongar fins Arenys de Mar. Dos anys més tard s'allarga fins a Tordera i finalment el 1861 fins a Maçanet-Massanes, actualment

la seva destinació final direcció Nord. No va ser fins l'any 1948, coincidint amb el centenari de la seva construcció, que es va electrificar el trajecte fins a Mataró.

Actualment l'R1, també coneguda com a línia del Maresme o línia de la Costa (Ferropedia, 2016), connecta Maçanet-Massanes amb Molins de Rei, comunicant la província de Girona amb la de Barcelona, des de la comarca de la Selva fins al Baix Llobregat (Carles Salmerón i Bosch; Museu del Ferrocarril de Vilanova i la Geltrú, 2010).



Il·lustració 1: Rèplica de la locomotora Mataró (Leris, 2009).



Il·lustració 2: El Túnel de Montgat el 1950 (Anònim, 1950).



Il·lustració 3: Recorregut de la R1 (Borràs, 2013).

L'any 2008 l'R1 va ser freqüentada per 102.214⁸ passatgers al dia, és a dir, més de 39 milions al llarg de l'any. Per tant, podem afirmar que és una línia que ha transportat gran densitat de viatgers de la xarxa de Rodalies de Barcelona.

Segons dades de la web de la Generalitat de Catalunya, el 2010 hi van circular una mitjana de 216 trens cada dia laborable, de la sèrie 447 i del tipus Civia (sèrie 465) (Generalitat de Catalunya, 2010).

3.2. Característiques de cada sèrie

Considerem interessant veure les característiques dels trens que circulen per la línia R1 ja que en dependrà la seva sonoritat.

3.2.1. Sèrie 447 de Renfe

Va arribar a la xarxa de Rodalies l'any 1993. A diferència de models anteriors aquest està exclusivament dissenyat per a la xarxa de Rodalies, que compta amb un gran nombre de parades i un gran tràfec de viatges. Per cobrir aquestes necessitats es va dotar els trens amb una millor distribució interior i una gran capacitat d'acceleració, que va provocar la reducció de la velocitat màxima a 120km/h.



Il·lustració 4: Unitat 447 de Rodalies de Catalunya (Verdugo, 2011).

8 Dades oficials de l'Estudio de Afors Nov. 2008.

Cada unitat de tren està formada per tres cotxes. Els dos dels extrems compten amb una cabina de conducció, estan motoritzats i únicament és el del mig un remolc. Per facilitar l'entrada i la sortida dels viatgers cada cotxe compta amb tres parells de portes de doble fulla.

Per tal de no quedar obsolet, l'any 2013, aquest model va patir reformes al seu interior. En el cas dels 447 que s'utilitzen en recorreguts regionals (i de tant en tant també circulen per la R1) se'ls van substituir els seients per butaques i es van col·locar espais per a bicicletes. En el cas dels trens de rodalies es va substituir el cotxe intermedi per un de nou amb pis baix i alguns element que porten la sèrie 465 (Ferropedia, 2015).

3.2.2. Sèrie 463-464-465 de Renfe

Aquesta sèrie, també coneguda com a Renfe Civia, va començar a entrar en funcionament a la xarxa de Rodalies l'any 2006. És un model de tren que es va crear des de zero a partir de l'experiència adquirida amb sèries anteriors com el 447. La novetat en aquest tren és la seva versatilitat vers la demanda d'espai interior, ja que es pot adequar el número de cotxes de la unitat amb un mínim de 3 i un màxim de 5. Aquesta sèrie pot estar formada per quatre tipus de cotxes que es poden incloure o no depenent de les necessitats de la línia: dos cotxes extrem (A1 i A2) amb cabina de conducció i pis normal; l'intermedi (A3) amb W.C. i pis rebaixat; i l'intermedi (A4) amb pis normal. El més habitual a l'R1 és el 465, que tal com indica l'última xifra del número esta formada per 5 cotxes en aquest ordre: A1, A4, A3, A4 i A2.



Il·lustració 5: Unitat 465 de Rodalies de Catalunya (Curses, 2011).

A diferència del 447 no hi ha cap cotxe remolcat. El primer i l'últim recolzen l'extrem lliure en un bogi⁹ remolc, i l'altre sobre un bogi motor compartit (il·lustració 6) amb el següent cotxe. D'aquesta manera els cotxes intermedis recolzen ambdós extrems sobre bogis de motors compartits. També incorporen cercols insonoritzadors amb els quals s'aconsegueix reduir l'emissió sonora produïda per les rodes, especialment a les corbes.

⁹ Carro giratori de dos o més eixos que suporta la caixa d'un vehicle.



Il·lustració 6: Bogi motor compartit de la sèrie 465 (Curses, 2011).



Il·lustració 7: Bogi de la sèrie 447 (Curses, 2011).

El motor és exactament el mateix que el de la sèrie 447 degut als seus bons resultats, i permet arribar a la mateixa velocitat que la sèrie anterior (120km/h).

Estructuralment, les caixes estan construïdes amb perfils d'alumini de grans dimensions, amb dos parells de portes de doble fulla. El seu interior compta amb altres materials lleugers com panells de resines fenòliques i materials estratificats, convertint aquesta sèrie en un tren molt més lleuger i silenciós. També cal destacar que tot el tren està connectat interiorment mitjançant un passadís diàfan sense portes, i la unió entre els cotxes es fa mitjançant ròtules (Ferropedia, 2016).



Il·lustració 8: Interior del sèrie 465 (Ricote, 2009).

3.2.3. Sonoritat de les sèries 447 i 465

En ser models exclusivament pensats per a un servei de Rodalies fa que les dues sèries tinguin moltes similituds quan parlem de la seva sonoritat. Es podria dir que els dos són "instruments" similars i que per tant produeixen sons similars. Per exemple, si comparem l'arxiu *1_portes447.wav*¹⁰ i *2_portes465.wav*, podem escoltar l'obertura i tancament de les portes de doble fulla de les dues sèries, com també l'avís de tancament de les portes. Es pot observar que els dos sons segueixen una mateixa forma però el timbre, en canvi, és diferent.

¹⁰ Els arxius es poden trobar a l'apartat "Material adjunt" de l'annex II d'aquest treball.

Els materials més lleugers i moderns del 465, o la insonorització de les rodes comporten una reducció de soroll. La diferència es pot apreciar comparant els arxius *3_rodas447.wav* i *4_rodas465.wav*, enregistrats en el mateix punt del tram situat entre les estacions de Sant Adrià de Besos i El Clot-Aragó.

Els motors de les dues sèries generen sons tonals que sobretot podem apreciar a l'arrancada. En el cas del 447, podem escoltar un D5 inicial que salta a un F#4, on comença un escombrat freqüencial fins al A5 (arxiu *5_cotxe1_447.wav*). A la sèrie 465 en canvi, primer sentim un G#4 que fa un salt de 5ª cap al D#5, on hi comença l'escombrat freqüencial fins a G#5 (*9_cotxe2_465.wav*). La disposició del motor (i per tant on el sentirem) també és diferent. Aquest fet el podem apreciar si escoltem l'arrencada del tren des de diferents cotxes. Si escoltem els arxius *5_cotxe1_447.wav* i *7_cotxe3_447.wav* enregistrats al centre dels cotxes 1 i 3 (els dels extrems), podem escoltar el motor. En canvi, si escoltem l'arxiu *6_cotxe2_447.wav* (el central), no escoltarem el motor, només el brunzit del climatitzador, ja que els seus bogis són remolc. En el cas de la sèrie 465, el motor pràcticament el sentirem a tots els cotxes indiferentment del seu número, tal i com podem apreciar als arxius *8_cotxe1_465.wav* i *9_cotxe2_465.wav* enregistrats al centre dels cotxes 1 i 2. També és cert que al cap del tren (també a la cua, ja que el tren és simètric) tal i com podem escoltar a l'arxiu *10_cap del cotxe1_465.wav*, el motor se sent menys degut que el bogi que té just a sota és remolc. Una altra característica que fa sentir el motor al llarg de tot el 465 és que aquesta sèrie no té els cotxes separats per portes a diferència del 447.

Tot i no ser una característica sonora de la mecànica dels trens, també és interessant escoltar el contrast entre dos viatges de la mateixa sèrie en horaris diferents. A l'arxiu *11_contrast_447.wav* podem escoltar la diferència entre un viatjar un divendres a les 15:15h i a les 22:54h amb un 447. Els divendres i dissabtes per la nit els trens tendeixen a anar plens, transportant passatgers cap a les zones d'oci nocturn. A l'arxiu *12_contrast_465.wav* també podem escoltar la diferència entre viatjar a les 15:39h, hora de poca afluència, o a les 18:29h, hora punta.

L'afluència de musics i pidolaires, també és una característica important del paisatge sonor de la línia. Als arxius *13_musics.wav* i *14_pidolaires.wav* podem escoltar una breu mostra.

4. So immersiu

El so immersiu ens permet generar un gran ventall de possibilitats creatives. Ens permet recrear la realitat, reinterpretar-la, reconstruir-la, redissenyar-la, etc. Però, què és el so immersiu?

Quan parlem de so immersiu ens referim a la sensació psicològica d'estar rodejats per fonts de so específiques com també so ambient. Aquesta la podem experimentar auditivament en un entorn d'escolta natural (és a dir en un espai acústic com els que ocupem durant el nostre dia a dia), o en un entorn d'escolta virtual creat a partir d'altaveus o auriculars, dissenyats per reemplaçar o augmentar l'entorn natural.

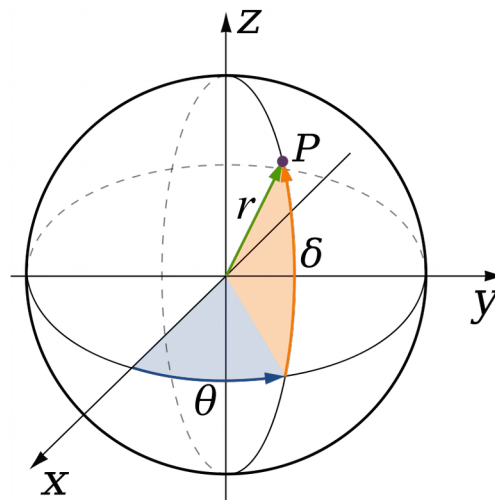
El so té l'habilitat de connectar l'oient amb una localització fixa mentre la resta d'informació sensorial canvia. Podem trobar un bon exemple al cinema, on s'utilitza molt aquest efecte per situar el públic a un espai concret mentre la perspectiva visual canvia constantment. El so ens proporciona tanta informació espacial que sense cap dels altres sentits, només amb l'oïda, ja podem traslladar-nos imaginàriament en un espai totalment diferent al qual està el nostre cos.

Per aconseguir una bona experiència, l'objectiu principal del so immersiu és recrear un entorn sonor que sigui el màxim proper al món real, o millor encara, un món que només pugui existir a l'espai virtual.

Gran part del contingut d'aquest capítol està adaptat del llibre *Immersive Sound* (Geluso & Roginska, 2018), un manual molt complet per a introduir-se al món del so immersiu.

4.1. Localització del so

La localització del so (P) en un espai de tres dimensions s'acostuma a representar amb un sistema de coordenades (il·lustració 9), on en el pla horitzontal hi tenim projectat l'azimut (angle θ), en el pla vertical l'elevació (angle δ) i la distància representada amb el radi (r).

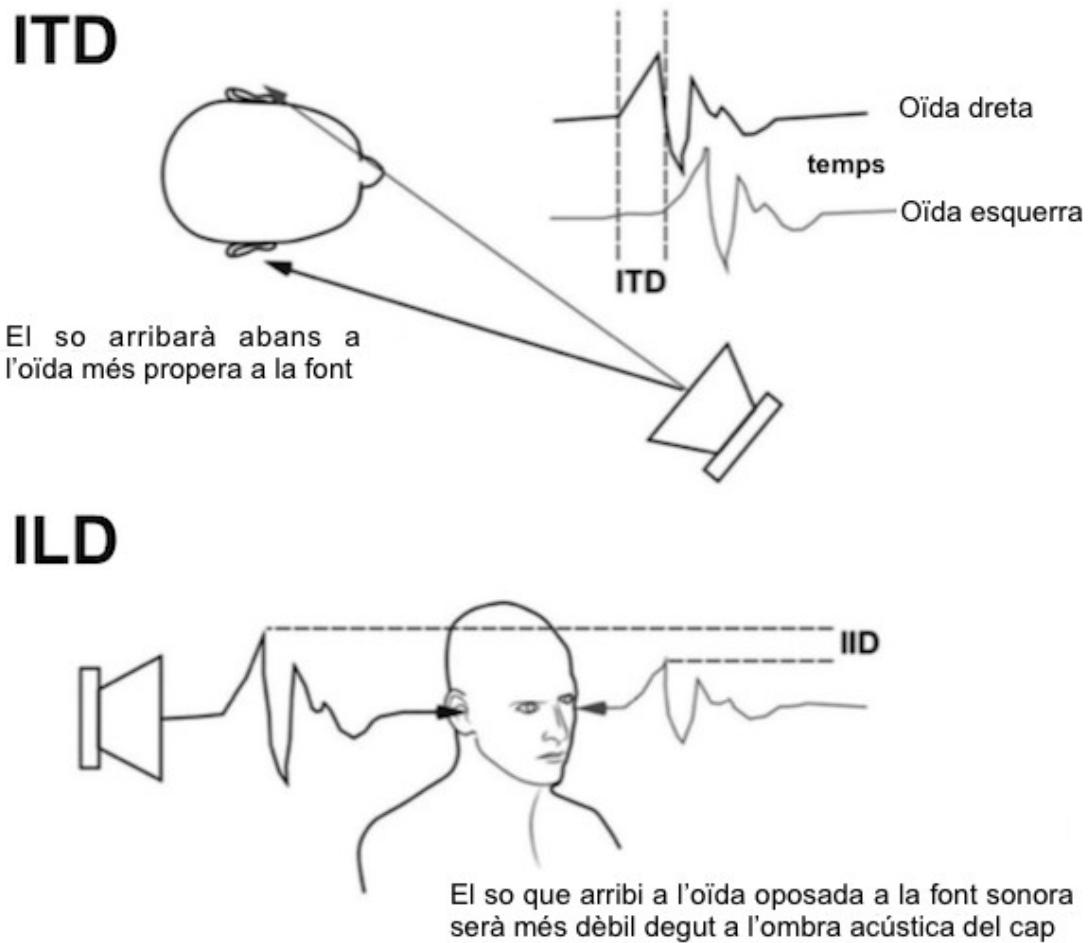


Il·lustració 9: Coordenades esfèriques (Wikimedia, 2008).

4.1.1. Localització en el pla horitzontal

El cervell, per poder localitzar les fonts sonores dins del pla horitzontal compara els senyals que rep a cada oïda. La teoria Duplex de Lord Rayleigh (1907) diu que aquest procés funciona gràcies a la diferència interaural de nivell (ILD: Interaural Level Differences) i la diferència interaural de temps (ITD: Interaural Time Differences). És a dir, un so que prové de l'esquerra arribarà amb més intensitat i amb més rapidesa al timpà esquerre que al dret, per tant el so arribarà amb una ILD i ITD respecte els dos timpans que ens permetran deduir la procedència d'aquest. Aquests mecanismes es complementen, d'aquí prové el terme “duplex”: quan un no és fiable, l'altre el substitueix.

Les ILD són menys efectives a les freqüències baixes degut que la mida mitjana del cap humà és més petita que la seva gran longitud d'ona, de manera que el cap no fa ombra acústica a les freqüències inferiors a 1.5 kHz. En canvi les ITD no seran efectives a les freqüències superiors a 1.5 kHz, degut que la mida de les longituds d'ona cau per sota del diàmetre del cap humà (López, 2015). Per aquest motiu, en aquestes condicions, el cervell no és capaç d'interpretar les diferències de fase (les ITD) entre el senyal de l'oïda esquerra i la dreta. Per tant, per a freqüències inferiors a 1.5 kHz el cervell utilitzarà les ITD, i per freqüències superiors, les ILD (Wenzel, Begault, Godfroy-Cooper, 2018).



Il·lustració 10: Les ITD i ILD (Geluso & Roginska, 2018).

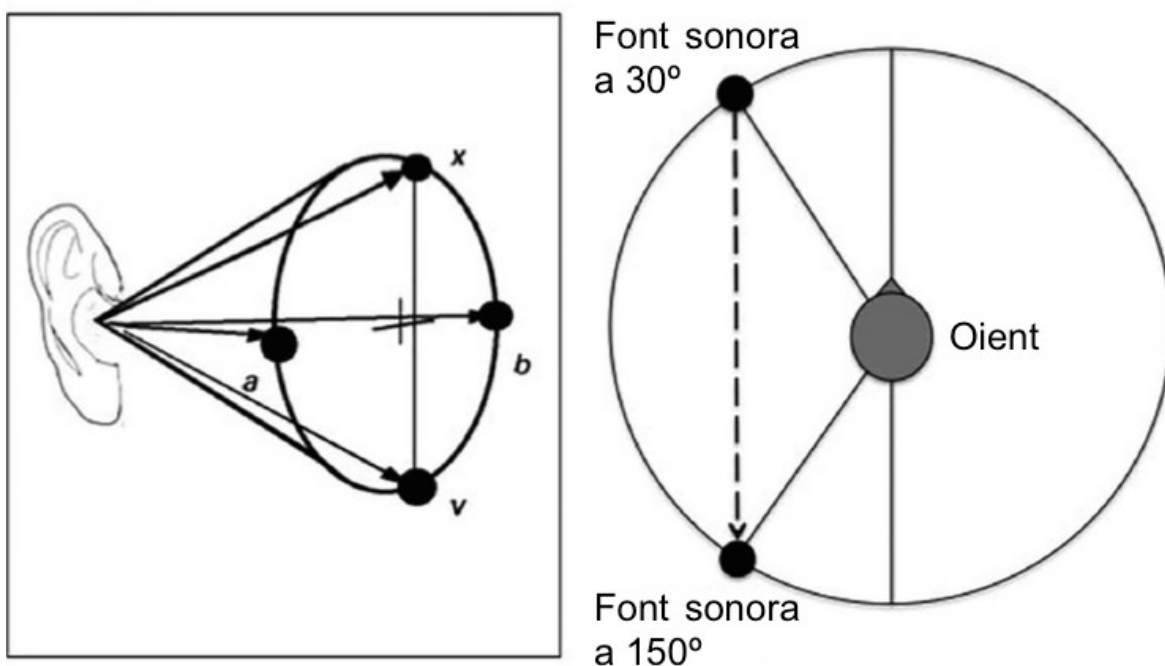
4.1.2. Percepció de l'elevació

En el pla vertical, en canvi, les diferències interaurals són molt petites i el cervell s'ha d'ajudar d'altres factors per poder localitzar el so, com per exemple el filtre que produeix la fisiologia humana (mida, densitat i forma del cap; el pavelló auditiu; el canal auditiu; la forma de les cavitats nasals i orals; i l'espatlla) en interactuar amb el so. El so que procedeix de dalt es trobarà un relleu diferent del nostre cos (que alhora produirà una reflexió, difracció i absorció diferent del so) que aquell que procedeix de baix (Wenzel, Begault, Godfroy-Cooper, 2018). Aquest filtratge s'anomena Funció de Transferència del Cap (HRTF: Head-Related Transfer Function).

4.1.3. Factors que afecten la localització

Tot i la gran quantitat d'informació que recull i interpreta el cervell, tenim un grau d'error que pot anar dels 5° als 20° en el moment de determinar la localització exacta d'una font sonora. La diferència angular més petita que podem arribar a percebre s'anomena angle mínim audible (MAA: Minimum Audible Angle). El MAA serà més petit quan el so sigui frontal (aproximadament 1° o 2°), i més gran quan sigui lateral (entre 90° i 270°). Per tant podem dir que localitzem amb més precisió els sons frontals (Wenzel, Begault, Godfroy-Cooper, 2018).

Un altre error que podem trobar és no saber discernir entre davant i darrere. Això és degut al con de confusió. Si imaginem un cap sense orelles com el de la il·lustració 11, el so que provingui de davant i el que provingui de darrere tindrà les mateixes ITD i IID, per tant no serà possible saber si el so ve de davant o de darrere. Uns quants estudis han mostrat que si permetem a l'oient moure el cap, pràcticament s'elimina aquest problema ja que les ITD i IID deixen de ser idèntiques.



Il·lustració 11: Con de confusió (Geluso & Roginska, 2018).

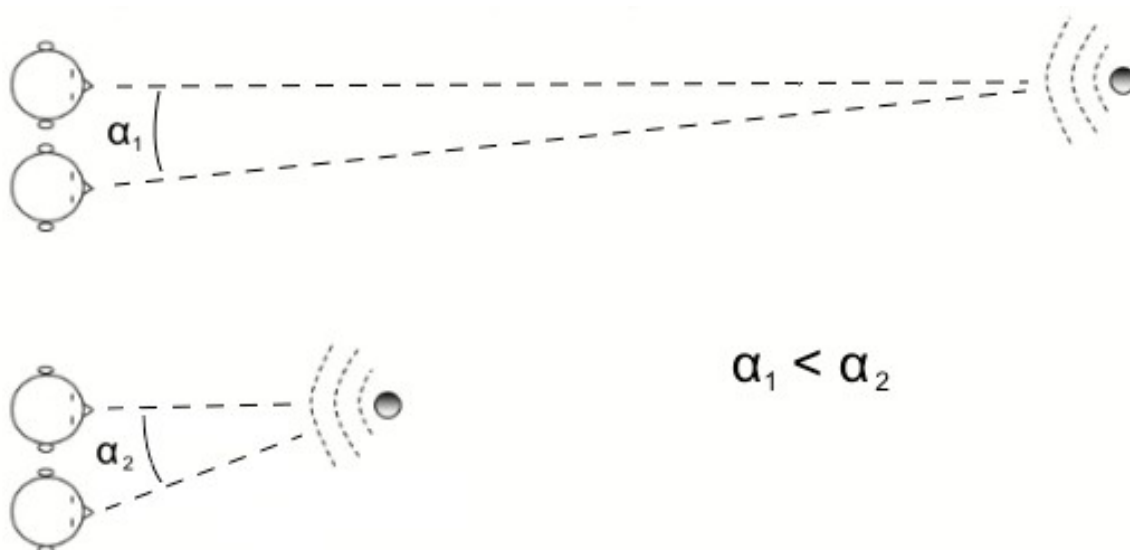
4.1.4. Percepció de la distància

Amb totes aquests mecanismes el cervell pot arribar a saber l'angle θ i δ de la font, però encara no hem parlat de r , és a dir, la distància que hi ha entre nosaltres i la font.

Per poder deduir la distància d'una font utilitzem diversos mètodes que es poden combinar entre ells. Un d'ells és la intensitat amb la qual ens arriba una font sonora. Si arriba amb poca intensitat interpretarem que està lluny. Aquest mètode té una manca, i és que per poder deduir la distància també hem de conèixer la potència amb la qual s'ha emès el so, ja que ens pot arribar igual de fort un mosquit a 2 cm de l'orella que un camió a 10 m.

Un altre mecanisme que tenim és el que es basa en l'atenuació de les altes freqüències. En la transmissió per aire, les altes freqüències pateixen una atenuació superior a la patida per les baixes. D'aquesta manera, només fixant-nos en el timbre de la font podrem deduir si està lluny o a prop. Quant més lluny estigui la font més atenuació a les altes freqüències patirà.

Per les altes freqüències tenim un altre mecanisme que es basa en la paral·laxi per moviment, és a dir, en canviar el punt d'escolta respecte l'objecte que volem localitzar. Aquest canvi de posició ens proporcionarà un canvi de balanç entre les oïdes. Si l'objecte és lluny de nosaltres, un moviment lateral produirà molt poc canvi en la diferència d'intensitat entre les oïdes. En canvi, si és a prop i ens movem d'esquerra a dreta, sí que notarem un canvi important d'intensitat degut que l'angle serà més gran (il·lustració 12).



Il·lustració 12: Paral·laxi per moviment (López, 2015).

Un excés de ILD també ens proporciona informació de distància. Per exemple, si tenim un insecte a l'orella esquerra el sentirem molt fort en aquesta i pràcticament gens a la dreta. La diferència d'ILD serà tan gran que deduirem que l'insecte està just al costat nostre.

Per últim, un dels mecanismes més importants és el que utilitza la relació de so directe i so reverberat. Quan la diferència d'intensitat entre ells sigui gran, voldrà dir que el so està aprop nostre, i si la diferència és petita voldrà dir que està lluny (López, 2015).

4.1.5. Funció de Transferència Relacionada amb el Cap (HRTF)

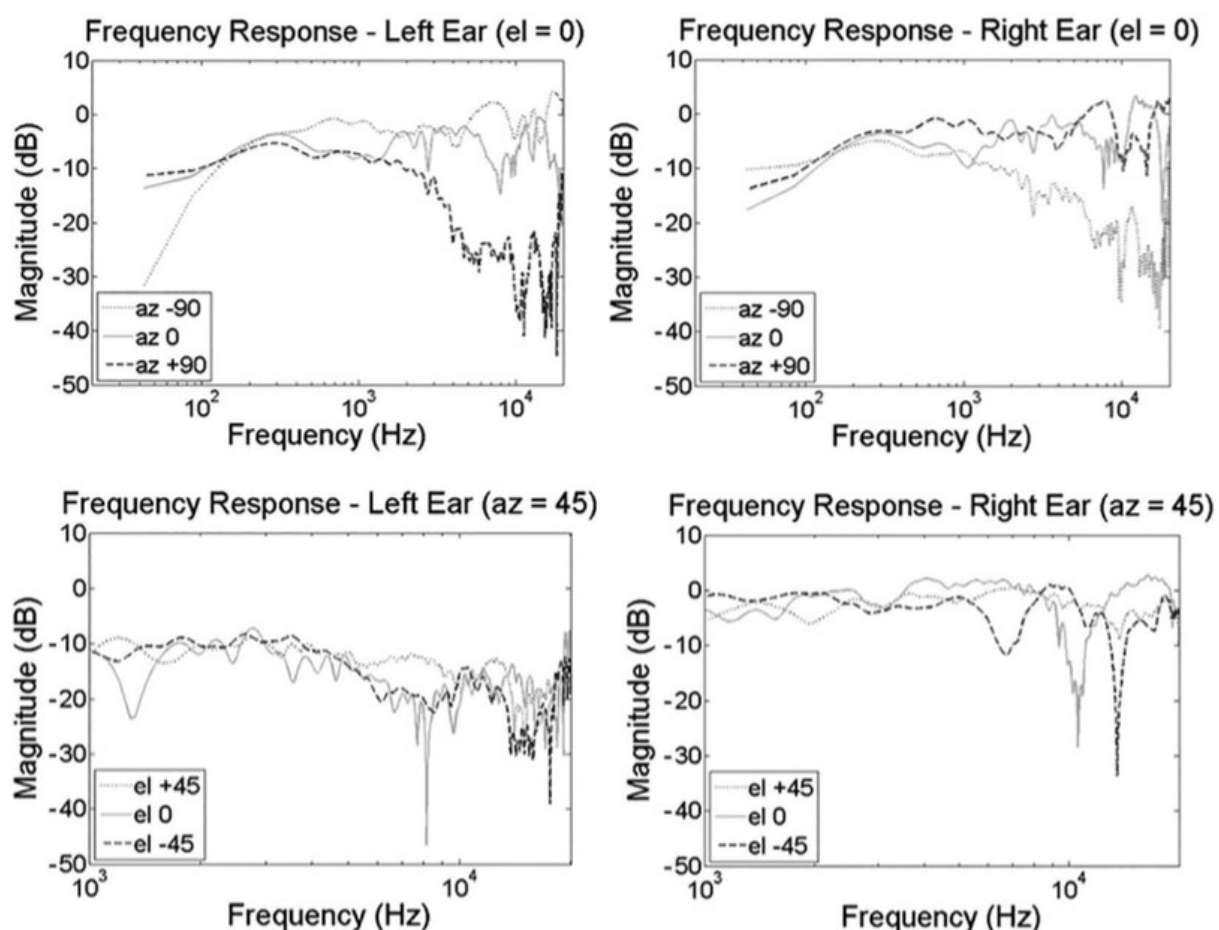
Tots els efectes que pateix un so des del seu origen fins que arriba al timpà es poden expressar en una única operació de filtratge. Com s'ha dit anteriorment, aquest filtre s'anomena HRTF. Aquesta funció engloba tots els aspectes direccionals de la percepció del so, és a dir les ITD, ILD i els efectes de reflexió, difracció i absorció produïts per la fisiologia humana.

Per mesurar l'HRTF situem una persona o maniquí en una sala anecoica a la qual se li insereixen a les orelles micròfons binaurals. Posteriorment es reproduïx un senyal de test (normalment un escombratge freqüencial) des d'un altaveu posicionat a un azimuth i elevació coneguts. Aquest senyal s'enregistra pels micròfons binaurals d'on s'extraurà la seva resposta impulsional, l'HRIR (Head Related Impulsional Response) (Roginska, 2018). Si fem la transformada de Fourier de l'HRIR obtindrem l'HRTF.



Il·lustració 13: Mesura de la HRIR en una sala anecoica (Geluso & Roginska, 2018).

Com podem veure a la il·lustració 14, l'HRTF depèn de l'angle d'azimut i d'elevació, com també de la freqüència de la font. Per tant s'haurà de repetir la mesura disposant l'altaveu en diferents punts (sempre mantenint el mateix radi), o com podem observar a la il·lustració 13 disposant de diversos altaveus a diferents angles d'azimut i elevació. Cal tenir en compte que les HRTFs seran diferents en cada oïda degut que fisiològicament no són exactament iguals.



Il·lustració 14: Respostes de magnitud de l'oïda esquerra i dreta derivades de les HRTFs prèviament mesurades (Geluso & Roginska, 2018).

4.2. Sistemes de so espacial

L'any 1881 l'enginyer francès Clément Ader va idear el *Teatrophone*, un sistema que permetia fer una retransmissió d'àudio espacial de l'Òpera de París a l'Exposició Internacional de l'Electricitat. Estava format per una sèrie de parells de transmissors situats a l'Òpera que enviaven el senyal d'àudio fins a dos receptors ubicats a l'exposició, on els oients se'ls col·locaven a les orelles i gràcies a les diferències interaurals transmeses per les dues línies podien escoltar música espacial.

Podríem dir que el *Teatrophone* va ser el primer sistema de reproducció d'àudio de dos canals que jugava amb so immersiu. No és fins a finals del segle XIX que comencen a aparèixer teories rigoroses sobre la localització del so 2D, que posteriorment s'han anat perfeccionant donant pas al 3D (Boren, 2018).

L'entorn d'escolta immersiva engloba un seguit de tècniques d'àudio que podem classificar de la següent manera (López, 2015):

- **Sistemes binaurals**, basats en la forma com escoltem.
 - Gravació binaural
 - Síntesi binaural
- **Sistemes basats en la panoramització**, que situen un senyal entre diversos altaveus mitjançant canvis d'amplitud.
 - Estèreo
 - Sistemes de so envoltant (*Surround*)
 - *Vector Base Amplitude Panning* (VBAP)
- **Sistemes de síntesi de camp acústic**, que sintetitzen el camp acústic basant-se en la física.
 - *Ambisonics*
 - *Wave Field Synthesis* (WFS)

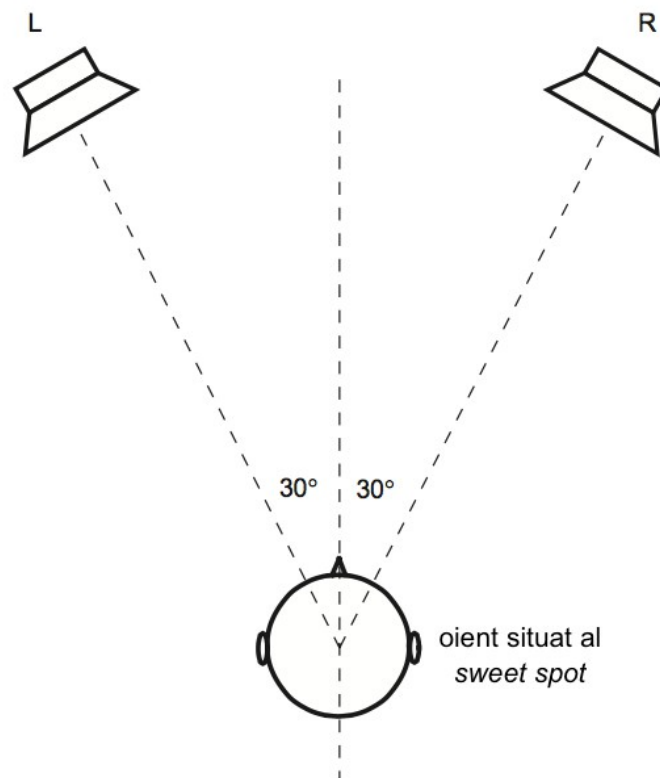
4.2.1. Estèreo

El 1933, l'enginyer anglès Alan Blumlein va establir les bases del so estereofònic a la seva patent, on hi deia que la informació sobre la fase i l'amplitud captada per dos micròfons direccionals es podia reconstruir amb només dos altaveus. D'aquesta manera, amb dos canals podia produir una imatge sonora estèreo.

El sistema

La configuració estèreo es caracteritza per una imatge horitzontal frontal representada per dos altaveus, que direccionen el so cap a un *sweet spot* (punt optim on s'ha de col·locar l'oient) ubicat a l'eix central perpendicular del pla format pels dos altaveus.

La ITU-R¹¹ estableix que l'angle format entre l'eix central i els altaveus ha de ser de 30° equidistants l'un de l'altre, de manera que si l'oient surt del *sweet spot* percebrà una imatge sonora distorsionada i inestable (Geluso, 2018).



Il·lustració 15: Esquema del sistema estèreo (Geluso & Roginska, 2018).

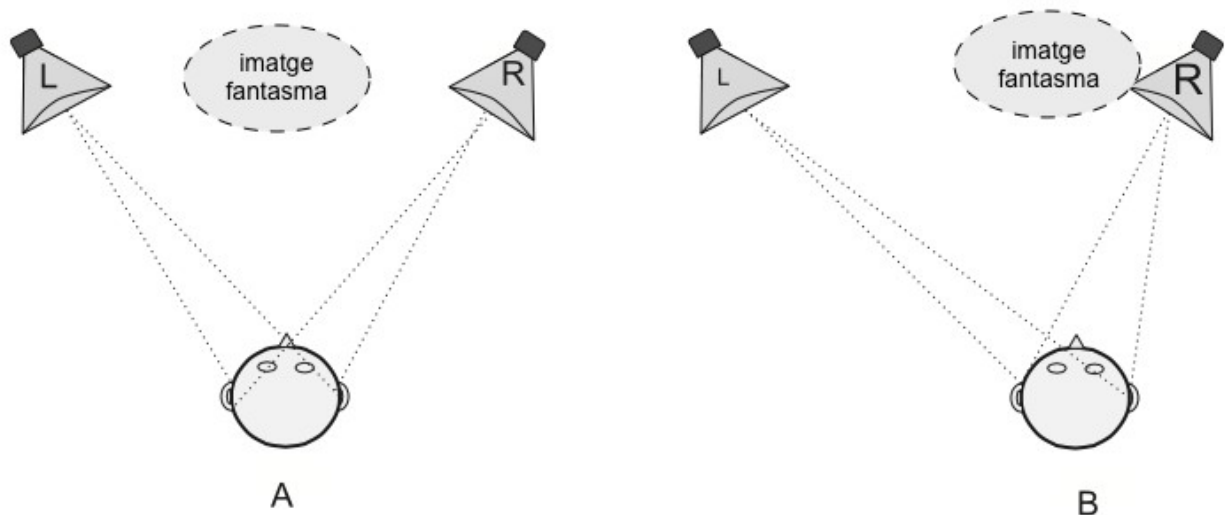
Per tal de millorar l'escolta estèreo es pot ajustar l'espaiat dels altaveus, la direcció, el seu patró polar i la seva resposta freqüencial. També afectaran l'escolta la mida, disseny i el tractament acústic de la sala d'escolta.

11 La ITU-R és el sector de la radiocomunicació de la *International Telecommunication Union*, la organització encarregada de coordinar i regular les telecomunicacions a nivell internacional.

La localització dels sons en el sistema estèreo funciona bàsicament gràcies a les ITD i ILD que hem vist anteriorment, degut que és una imatge frontal i només juga al pla horitzontal. Per tant, per a codificar el senyal en un programa estèreo, tindrem en compte les diferències de temps entre canals (ICTD: Inter-Channel Time Differences) que l'oïda captarà com a ITD, i les diferències de nivell entre canals (ICLD: Inter-Channel Level Differences) que l'oïda captarà com a ILD (Geluso, 2018).

Imatge fantasma

Amb el sistema estèreo, podem recrear imatges sonores entre els dos altaveus on realment no hi ha cap generador de so. Aquest efecte s'anomena imatge fantasma i l'aconseguim gràcies a les ILD i ITD. D'aquesta manera podem crear una finestra acústica entre dos altaveus. Aquest sistema però, té un defecte. Com s'ha comentat abans, només la podem percebre de forma òptima si l'oient està ubicat al *sweet spot*. En el cas que l'oient es mogui, les ITD i ILD prèviament codificades es rebran distorsionades (ja que la distància entre altaveus i oient ja no serà simètrica), i localitzarem la imatge fantasma en un punt en el qual no hauria de ser-hi (il·lustració 16).

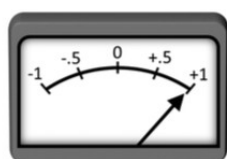
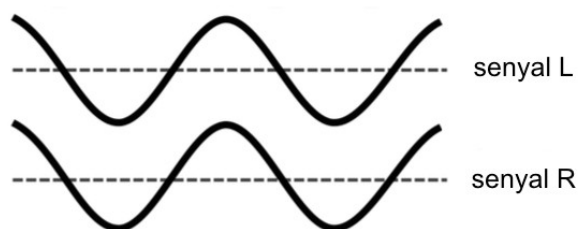


Il·lustració 16: Localització errònia de la imatge fantasma degut al desplaçament de l'oient (Geluso & Roginska, 2018).

Correlació de fase

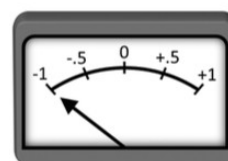
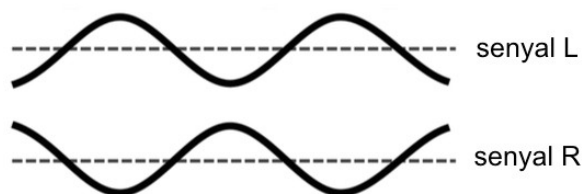
Quan treballem amb més d'un senyal d'àudio entra en joc la fase, ja que es combinen dues o més formes d'ona que poden sumar o restar. Si ho simplifiquem i ens imaginem dos senyals sinusoidals

idèntics, quan aquests estiguin en fase direm que la correlació és 1 (il·lustració 17). Quan estiguin una mica desfasats anirà cap al 0 i quan estigui completament a la contra l'una de l'altra tindran una correlació de -1 (il·lustració 18).



correlació de fase = +1

Il·lustració 17: Senyals idèntics fasats (Geluso; Roginska, 2018).



correlació de fase = -1

Il·lustració 18: Senyals idèntics contrafasats (Geluso; Roginska, 2018).

Aquest efecte el trobarem de la mateixa forma en els senyals estèreo. Com més correlació tinguin els dos canals voldrà dir que la imatge estarà més centrada i per tant tendirà a ser monofònica, i com més s'aproximi al 0 més estèreo serà. S'ha de tenir en compte que quan la correlació tendeix a 0 queda un buit sonor entre els dos altaveus. Si la correlació baixa del 0 serà indicador de que una porció de la codificació estèreo està invertida de fase en un dels canals. Això repercutirà quan vulguem convertir el senyal estèreo a mono, patint coloració espectral i/o pèrdua de guany (Geluso, 2018).

Creació de la imatge estèreo

Per aconseguir una bona imatge panoràmica cal repartir adequadament el so i col·locar-lo al lloc adient de tal manera que sigui intel·ligible i no quedin espais buits. Com ja s'ha parlat anteriorment, aquesta imatge l'aconsegurem amb ICTD, ICLD i equalització. Aquest procés el podem fer a l'hora d'enregistrar quan triem la tècnica de microfonia (XY, AB, MS, ORTF, parell estèreo...), o a l'hora de mesclar gràcies a programes que distribueixin amb diferents guanys (ILD) la font sonora, o la retardin en algun dels canals per dirigir-la cap a l'altre (ITD). Amb aquests mètodes, la imatge estèreo creada sempre estarà delimitada dins de l'angle que formen els dos altaveus.

4.2.2. Binaural

L'àudio binaural és segurament l'efecte espacial més fàcil de capturar, però el més difícil de realitzar a post-producció. És per aquest motiu que tot i ser la primera tècnica de so 3D, fins ara no ha començat a madurar (Boren, 2018).

Entenem per binaural un sistema de dos canals que distribueix el so a l'oïda esquerra i dreta de l'oient. El que el diferencia del sistema estèreo és que el so que entra a l'oïda ha estat tractat per una combinació de ITDs, ILDs i filtratge espectral per tal de mimetitzar l'escolta humana. Aquest procés el podem imposar naturalment als enregistraments binaurals, o artificialment a través del processament del senyal (Roginska, 2018).

El sistema

L'àudio binaural es pot reproduir a través d'altaveus o d'auriculars. Els auriculars tenen diversos avantatges, com el de proveir un entorn controlat d'escolta o d'aïllament del so ambient, o el de mantenir l'oient al *sweet spot*. Com qualsevol tecnologia també té uns certs desavantatges. Quan movem el cap, l'entorn acústic es mou amb nosaltres provocant una escolta no natural. Com a conseqüència, el con de confusió no es podrà eliminar.

Quan escoltem àudio a través dels auriculars perdem tota la informació que ens proporcionen les HRTFs, ja que l'àudio és distribuït directament a les orelles. Per tant, per localitzar el so només tindrem el suport de les ITDs i les ILDs.

Captura del so binaural

El principal repte del sistema binaural és reproduir les HRTFs que la pròpia fisiologia humana contribueix a generar. Per tal de recrear-les, com hem comentat abans, tenim el mètode natural. Aquest, el que pretenen és filtrar el so amb la pròpia estructura de la fisiologia humana. Això es pot aconseguir mitjançant el cap d'un maniquí que simuli el d'un humà (il·lustració 19).



Il·lustració 19: Micròfon Neumann KU-100 (Geluso & Roginska, 2018).

Al llarg de la història s'ha anat perfeccionant el material i la tecnologia fins al punt de tenir una rèplica estandarditzada pràcticament igual a la d'un cap real.

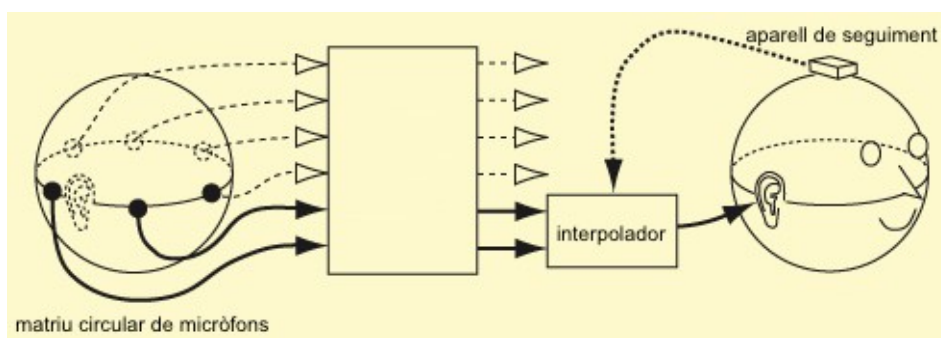
La reducció de la mida dels micròfons també ha permès poder crear-ne uns que es puguin introduir al canal auditiu per tal de captar el so filtrat tal i com ho fa el timpà (il·lustració 20) (Roginska, 2018).



Il·lustració 20: Micròfon binaural a l'interior d'una orella.

El problema principal de tots aquest mètodes és que la imatge 3D capturada té una perspectiva completament fixa. És a dir, un cop l'enregistrament està fet l'oient no pot rotar el cap i interactuar amb el sistema. Precisament, un requeriment imprescindible dels sistemes de realitat virtual (que cada vegada tenen més popularitat i aplicacions), és poder interactuar amb el so envoltant.

Per poder solucionar això, Algazi, Duda i Thompson (2004) van introduir el mètode MTB (*Motion-Tracked Binaural*). Aquest consisteix a gravar el so en 360° gràcies a uns micròfons disposats en una matriu circular del diàmetre del cap humà (il·lustració 21). Durant la reproducció de l'enregistrament a l'oient se li fa un seguiment dels moviments rotatius del cap (*head tracking*). D'aquesta manera podem determinar quin dels micròfons de la matriu s'aproxima més a la posició de la seva orella (il·lustració 21). En el cas que la posició de l'orella coincideixi entre la de dos micròfons es fa una interpolació del senyal capturat per aquests (Roginska, 2018). D'aquesta manera s'aconsegueix poder interactuar amb el so però només en el pla horitzontal.

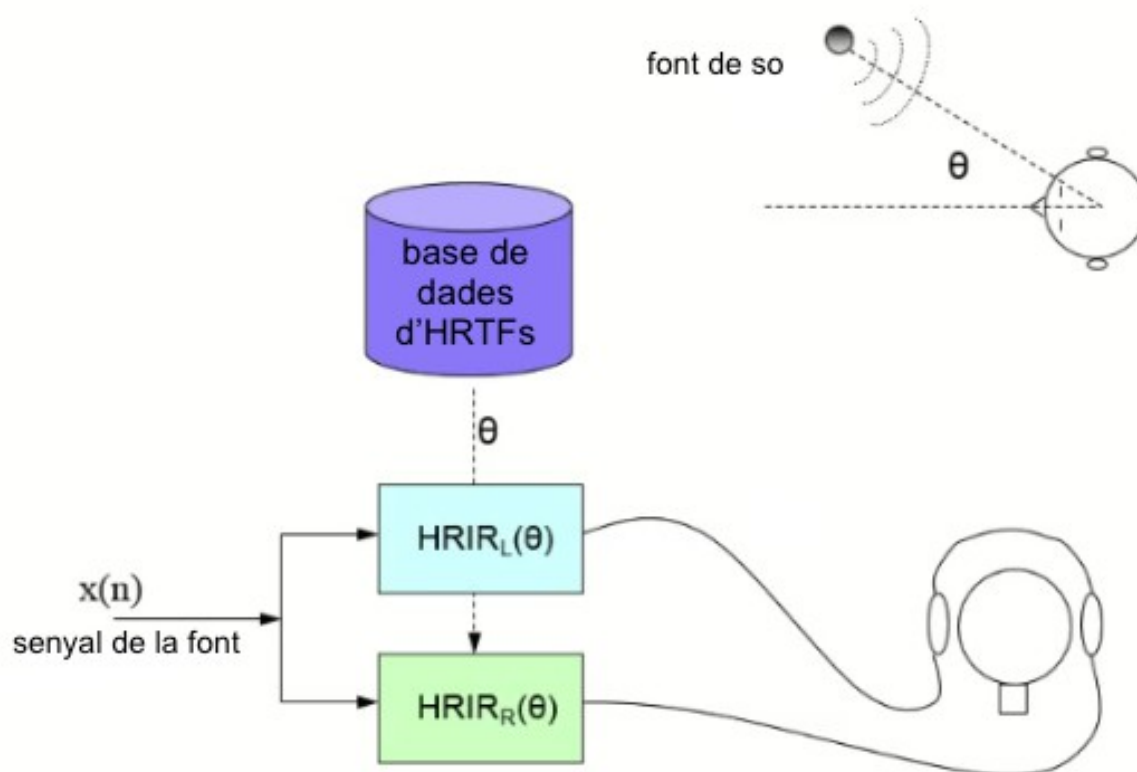


Il·lustració 21: Mètode *Motion-Tracked Binaural* (Algazi, 2004).

Síntesi binaural

Aquest mètode es basa en la síntesi de l'escolta binaural, és a dir, el filtratge artificial del so amb les HRTFs que anteriorment ja s'ha comentat com aconseguir. Treballa amb una base de dades que conté HRTFs de totes les direccions de l'espai, és a dir, de tota l'esfera que envolta l'oient.

La síntesi binaural, a diferència de les tècniques d'enregistrament binaural, sí que permet la rotació del cap en tots els eixos com també el canvi de posició. Això s'aconsegueix d'una forma similar al mètode MTB. Mitjançant el seguiment angular del cap de l'oient (*head tracking*) podem saber els seus angles d'azimut (θ) i d'elevació (δ). Amb aquests, podem buscar a la base de dades les HRTFs que corresponen a cadascuna de les oïdes per aquella direcció en concret. A partir de la seva resposta impulsional filtrarem el senyal de la font amb l'HRIR del canal L i del canal R, aconseguint l'espacialització de la font originalment monofònica (il·lustració 22) (López, 2015).



Il·lustració 22: Funcionament de la síntesi binaural per al pla horitzontal (López, 2015).

Un altre avantatge de la síntesi binaural és que podem tractar i canviar de posició el so, ja que les HRTFs les podem afegir posteriorment.

Tècniques HRTF

Les HRTFs juguen un paper molt important dins de l'àudio binaural. Cada individu té unes HRTFs diferents degut que tothom té una fisiologia diferent. Com a conseqüència, les HRTFs d'una persona segurament no funcionaran bé en una altra.

Per aconseguir una imatge auditiva acurada cal fer una mesura individual, però això és poc viable ja que requereix molt temps i eines (Roginska, 2018). L'altra via que podem trobar és personalitzar les HRTFs mitjançant:

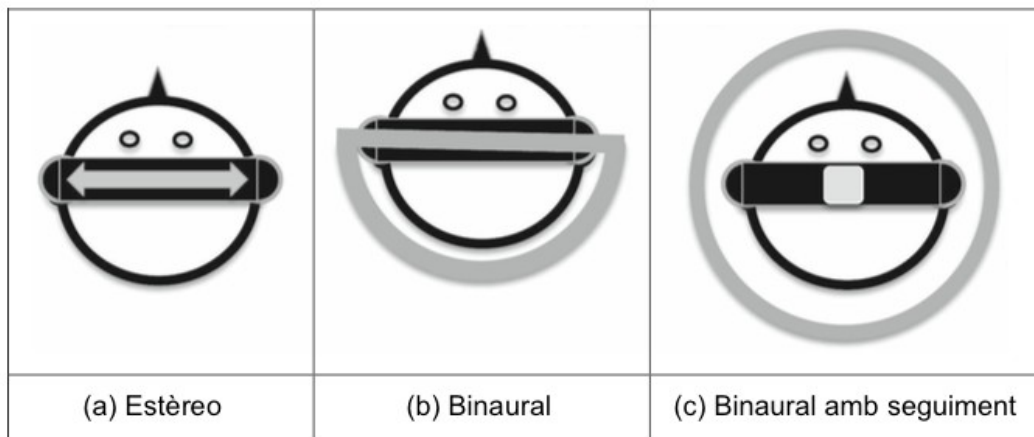
- La reconstrucció o modelatge de les HRTFs a partir de la mesura detallada de la geometria del cap i tors de l'oient.
- La descomposició de les HRTFs en filtres que representen diferents característiques de la fisiologia del cap i tors de l'oient.
- La tria auditiva entre un seguit d'HRTFs per trobar la que millor s'adapti.

L'última d'elles és la que més sovint es troba als programes que treballen amb àudio binaural, ja que és la més ràpida i econòmica.

IHL

L'acrònim IHL prové d'*Inside-the-Head Locatedness*, que es pot traduir com l'efecte o la sensació que les fonts sonores s'originen dins del nostre cap. Apareix sovint quan l'escolta no és natural, sobretot quan s'utilitzen auriculars, per tant, és un efecte que pot aparèixer a la reproducció binaural.

La millor forma de combatre l'IHL és generant un entorn d'escolta el més natural possible. Això es pot aconseguir amb una reverberació realista, HRTFs individualitzades i un entorn interactiu. Per tant en una escolta estèreo a través d'auriculars l'efecte IHL serà molt més present. Sentirem el so dins del nostre cap, repartit en una línia imaginària que va de l'orella esquerra a la dreta. En una escolta binaural estàtica aquest efecte serà menys present, i només passarà amb el so que localitzem al davant. Per últim en una escolta binaural amb *head tracking*, quant més realista sigui, més probabilitat hi haurà de sentir tot el so fora del nostre cap (Roginska, 2018).



Il·lustració 23: Localització del so en els diferents sistemes (Geluso & Roginska, 2018).

Mètodes de reproducció binaural

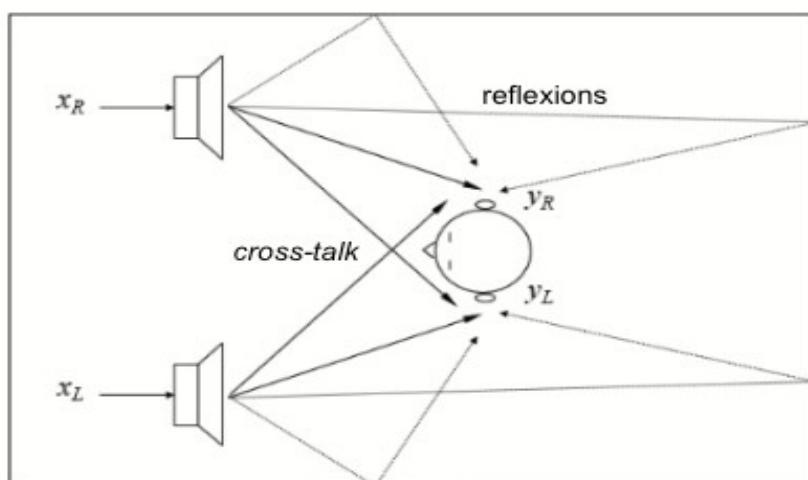
Auriculars

Podem reproduir àudio binaural de diferents maneres com per exemple amb auriculars tancats, oberts, *in-ear*, *earbud*, *multi-driver*, o *ear speakers*. Tots ells tenen característiques diferents que caldrà tenir en compte segon el tipus d'escolta que vulguem. Tanmateix caldrà equalitzar-los i calibrar-los si es vol una percepció espacial més bona.

Altaveus

Quan reproduïm un senyal binaural per altaveus, part del senyal L també arriba a l'oïda dreta i part del senyal R també arriba a l'oïda esquerra. Aquest fet s'anomena *cross-talk* (diafonia en català). Per culpa d'aquest, tot el balanç de ITD i ILD que recullen els enregistraments binaurals a través de les HRTFs es perd. També interfereixen a l'escolta les reflexions que produeix la pròpia sala (il·lustració 24).

Per solucionar aquest problema existeixen filtres que cancel·len el *cross-talk*. L'inconvenient d'aquest mètode és que la zona d'escolta encara es redueix més que en el sistema estèreo, de manera que l'oient ha d'estar molt centrat i quiet (López, 2015).



Il·lustració 24: Efecte *Cross-talk*, senyal directe i reflexions (López, 2015).

4.2.3. So Surround

El terme *so surround* o so envoltant l'acostumem a utilitzar per parlar de tot tipus de formats que requereixen més de dos altaveus. En aquest cas, quan parlem de *so surround*, ens referirem als principis estereofònics convencionals que van des de tres canals cap amunt i que només treballen en en pla horitzontal.

El cinema ha estat un dels grans impulsors del *so surround*. Steinberg i Snow (1934) van experimentar que la suma d'un tercer canal central (C) al sistema LR tenia l'efecte d'estabilitzar la gran imatge de les pantalles, ja que sense ell, al centre es percebia un buit sonor. *Fantasia* (1940) va ser primera pel·lícula comercial multicanal i va necessitar un sistema nou expressament creat per ella. El van batejar amb el nom de *Fantasound*, però era un sistema tant costós que no va durar més enllà dels primers anys de la Segona Guerra Mundial, tot i així va marcar un precedent. No va ser fins als anys 50 que el cinema va incorporar definitivament la reproducció estèreo. El canal C estava dedicat als diàlegs, i els canals LR a la música estèreo i els efectes sonors (Rumsey, 2018). Des de llavors han anat apareixent nous formats que han incrementat el nombre de canals i han portat als espectadors un so envoltant més acurat.

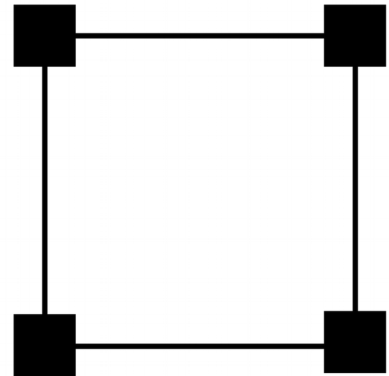
Per descriure la configuració d'aquests sistemes utilitzarem la nomenclatura internacional de l'ITU-R "*n-m estèreo*". On *n* és el nombre de canals frontals i *m* el dels laterals o posteriors, per tant els canals d'efectes no es tindran en compte aquí.

Sistemes de so

Sistema Quadrifònic (2-2)

Com bé indica el seu nom, aquest sistema està format per quatre altaveus formant entre ells un quadrat. Aquest sistema, es va llançar al públic la dècada dels 70 per aportar sensacions noves als consumidors, però no va tenir gaire èxit.

Els dos canals frontals, en estar en angle de 45° no eren compatibles amb el sistema estèreo de 30° . A més, en estar tant separats encara presentaven un major forat acústic al centre, i com a conseqüència una imatge frontal més pobre (Rumsey, 2018).



Il·lustració 25: Esquema del sistema quadrifònic (Lateiner, 2006).

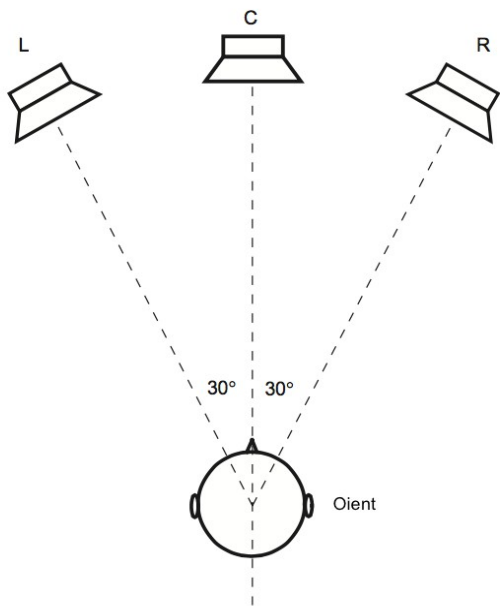
Estèreo de tres canals (3-0)

Com hem comentat abans, durant els anys 30, Steinberg i Snow van desenvolupar els precedents d'aquest sistema format per tres canals equidistant (il·lustració 26), l'esquerre (L), el central (C) i el dret (R). L'estàndard de la ITU-R diu que els altaveus L i R han d'estar a 30° , tot i així l'espai que omple el canal central ens permet col·locar-los a 45° (tenint en compte que perdrem la compatibilitat estèreo) (Rumsey, 2018).

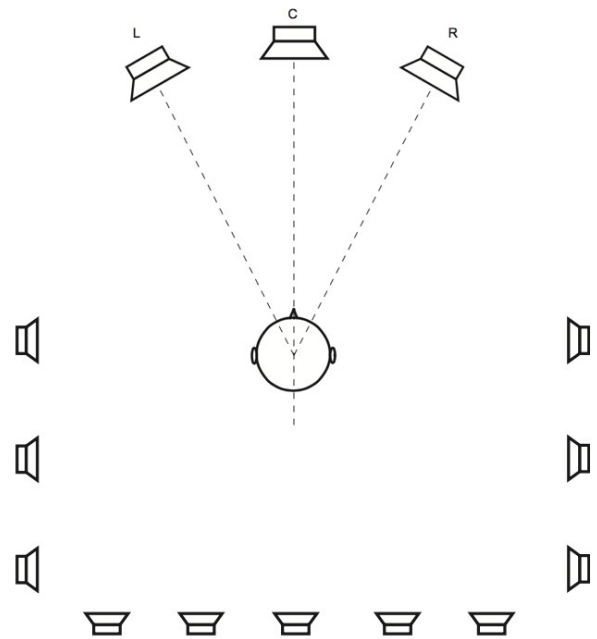
Surround de quatre canals (3-1 estèreo)

Aquest sistema (il·lustració 27) té la mateixa configuració frontal que el 3-0, però afegeix un canal *surround* (S) situat darrere de l'oient. Aquest sistema no té la intenció de generar imatges 360° , sinó de crear un efecte envoltant.

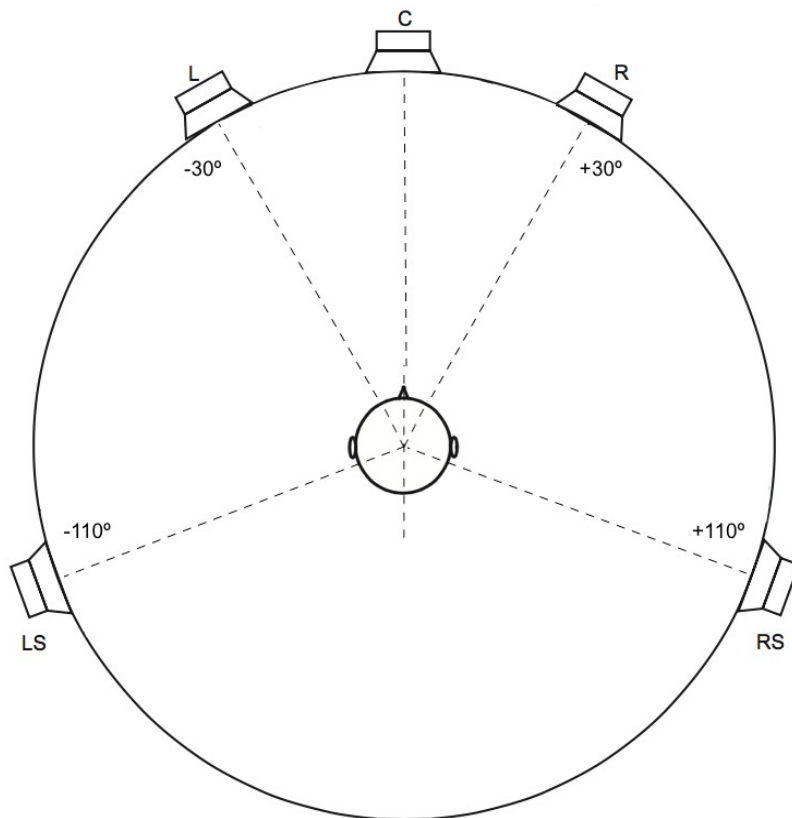
El canal S acostuma a alimentar una matriu d'altaveus, que pot ser més gran o més petita depenent de les dimensions de la sala. Els altaveus *surround* sovint es descorren electrònicament per augmentar el grau d'amplitud o difusió dels efectes *surround* (Rumsey, 2018).



Il·lustració 26: Esquema del sistema 3-0 segons la configuració estàndard de l'ITU-R (Geluso & Roginska, 2018).



Il·lustració 27: Esquema del sistema 3-1 (Geluso & Roginska, 2018).



Il·lustració 28: Esquema del sistema 3-2 segons l'ITU-R (Geluso & Roginska, 2018).

Sistema surround 5.1 (3-2 estèreo)

El sistema 3-2 o 5.1 (il·lustració 28) incorpora 2 canals més respecte el 3-1. L'efecte *surround* passa a ser estèreo, i a tot el sistema se li afegeix un canal d'efectes de baixa freqüència (LFE). Per tant tenim els canals L, C i R amb la configuració que ja coneixem, els canals *surround* LS i RS a -110° i $+110^\circ$ respectivament, i el LFE.

Tot i haver incrementat un canal *surround*, segueix quedant molt espai sense cobrir. És per aquest motiu que de la mateixa manera que al 3-1, l'estàndard ITU-R permet afegir altaveus addicionals als canals LS i RS per cobrir les regions buides en un angle de $\pm 60^\circ$ i $\pm 150^\circ$.

Tots els altaveus haurien de ser del mateix tipus (excepte el del canal LFE) per tal que el so sigui uniforme. Dit això, hi ha bons arguments per utilitzar altaveus dipolars per als canals *surround*, ja que aquests radien so en forma de vuit, de manera que el so que envien endarrere reflecteix a les parets i l'oient percep una major sensació d'espai. Per contra és més difícil crear una imatge sonora definida (Rumsey, 2018).

El canal LFE no està pensat per cobrir les baixes freqüències dels canals principals, si no que més aviat té la funció de reproduir efectes especials de baixa freqüència (per exemple explosions en el cas del cinema). Està limitat a un màxim de 120Hz per tal de no reproduir material que pertocaria als canals principals, i també ajuda a que no sigui fàcilment localitzable.

La configuració 5.1 d'altaveus millora el so envoltant i és compatible amb la imatge estèreo tradicional, però segueix no sent un bon disseny per l'àudio 360° (Rumsey, 2018).

Altres sistemes surround

Tots els sistemes *surround* que s'han descrit fins ara tenen una manca a l'hora de cobrir els 360° . És per aquest motiu que han anat apareixent altres combinacions de canals i altaveus per tal de millorar l'escolta *surround*.

Un d'aquests sistemes és el 7.1 o 5-2, que afegeix a la part frontal dos canals, un entre L i C (CL), i l'altre entre C i R (CR). D'aquesta manera s'aconsegueix una imatge frontal més sòlida i es pot incrementar l'angle dels canals LS i RS.

També hi ha altres combinacions del 7.1 com és el cas del 3-4, on hi tenim el convencional sistema de canals L, C, R, LS i RS, amb l'afegit de dos canals situats darrere l'oient anomenats *Back Surround* a més dels dos *surround* que ja incorpora el sistema 3-2. Amb aquests, el salt entre els altaveus frontals i els *surround* no és tan agressiu.

Un altre dels sistemes envoltants que amplien el nombre de canals és el 10.2 desenvolupat per THX¹². Aquest, afegeix al format 5.1 dos altaveus frontals alçats a 45°, un central al darrere (BS), i un LFE de més, per tal d'aconseguir més espacialització a les baixes freqüències (Rumsey, 2018).

Monitorització del so *surround*

Indiferentment del sistema, les sales on s'hi ha d'escoltar so envoltant han d'estar tractades acústicament amb materials absorbents i difusors. Això és així perquè la sala ha de ser neutra, ja que és el propi sistema el que recrea l'espai.

Alguns sistemes poden ser difícils d'instal·lar degut al l'espai físic que necessiten, sobretot en el cas dels canals LS i RS. Com també pot suposar un problema l'altaveu frontal (sovint coincideix amb la posició de la pantalla). Tots els altaveus han d'estar situats a la mateixa distància del *sweet spot*. En el cas que algun d'ells no respecti la distància i aquesta sigui menor a la resta, se li haurà d'aplicar un retard.

Com hem comentat abans, els altaveus dipolars són una opció per tal de crear un so *surround* difús. Aquests acostumen a anar bé quan només podem instal·lar dos altaveus per generar el so envoltant, o quan volem recrear en una sala petita l'espai d'una sala gran. Tot i així no seran una bona opció si el que ens interessa és una bona localització sonora envoltant.

Ja hem vist en altres sistemes que el canal LFE serveix pels efectes de baixa freqüència. Aquest altaveu interactua molt amb els modes de la sala, per tant serà clau la seva posició. Alguns d'aquests altaveus estan dissenyats per ser col·locats en un lloc concret de la sala, però d'altres no, de manera que haurem de buscar el lloc on millor s'ajusti (acostumen a ser cantonades).

12 THX és una empresa americana de sistemes de so fundada per George Lucas.

Sovint, els sistemes comercials per us domèstic no incorporen el canal LFE, o s'utilitza per les freqüències greus dels canals principals. Com a conseqüència es perd sensació estèreo a les freqüències baixes (Rumsey, 2018).

Creació de la imatge *surround*

A l'hora de produir so envoltant s'acostuma a utilitzar material monofònic panoramitzat i efectes artificials per tal de generar sensació d'espai. Per exemple, si es vol que una trompeta soni a un angle de 20°, caldrà treballar amb els dos canals més propers a aquest punt. En el cas del sistema 3-2 seran el canal C i R, i s'haurà de repartir el so entre aquests dos.

Tot i això, també hi ha un seguit de tècniques microfòniques que permeten captar el so natural en format *surround*. Les tècniques (sobre les quals no s'entrarà en detall) es poden separar en dos grups: aquelles basades en una única matriu de micròfons prou pròxims entre ells, i aquelles que tracten per separat els canals frontals i els posteriors.

També es pot treballar combinant àudio panoramitzat a la postproducció i àudio enregistrat amb tècniques microfòniques *surround*.

Height channels

Tal i com s'ha vist al sistema 10.2, els sistemes *surround* han anat ampliant canals cap al pla vertical fins a sistemes com el 22.2 de la NHK¹³, que disposa de tres capes d'altaveus a diferents alçades.

Quan es treballa en el pla vertical les ILDs i ITDs contribueixen molt poc a l'hora de localitzar fonts sonores. En aquest pla són més útils les modificacions espectrals que produeixen les espatlles, el cap i l'orella.

De la mateixa manera que jugant amb les ILDs i ITDs en el pla horitzontal s'aconsegueixen imatges fantasma, si es juga amb les HRTFs es poden aconseguir imatges fantasma al pla vertical (Kim, 2018).

¹³ La *Nippon Hōsō Kyōkai* (NHK), en català Corporació Emissora del Japó, és una emissora pública japonesa.

4.2.4. Àudio basat en objectes

El concepte d'àudio basat en objectes no és un mètode de reproducció de so espacial. És una forma d'estructuració de la producció de so espacial que ens pot ajudar a l'hora de treballar amb sistemes avançats com per exemple la WFS, el VBAP o l'*Ambisonics* (López, 2015).

El producte final d'àudio (també anomenat màster), quan està pensat per a reproduir-se als sistemes estèreo i *surround* acostuma a estar basat en canals, és a dir, àudio pensat per a reproduir-se en una configuració d'altaveus concreta. En aquest cas el màster està format per pistes d'àudio assignades a cada canal del sistema. D'aquesta manera l'àudio ja queda limitat i marcat a les prestacions del sistema.

La digitalització de l'àudio va permetre posar en pràctica l'àudio basat en objectes. Aquest es centra en els objectes sonors en comptes dels canals. Quan es treballa amb àudio basat en objectes, el màster final que s'obté està format per tantes pistes d'àudio amb metadades associades (posició de la font, direccionalitat, etc.) com objectes sonors hi ha a l'espai que es vol recrear.

Els objectes sonors acostumen a ser pistes d'àudio monofònic que han estat prèviament enregistrades o sintetitzades. Els enginyers d'àudio treballen amb aquests i els posicionen en espais 2D o 3D imaginari, és a dir, els hi atribueixen les metadades. D'aquesta manera creen espais sonors virtuals que posteriorment es podran descodificar i presentar al consumidor.

L'avantatge principal de l'àudio basat en objectes és que la mescla final de canals no es fa a la producció, sinó que es fa a la reproducció. Per tant, el reproductor d'àudio, basant-se en les posicions dels objectes i la dels altaveus decidirà quina serà la mescla més òptima. D'aquesta manera s'aconsegueix oferir experiències d'àudio més riques i immersives dins de cada entorn, ja que no té un nombre de canals preestablert, sinó que s'adapta al sistema i a les preferències dels consumidors (Tsingos, 2018).

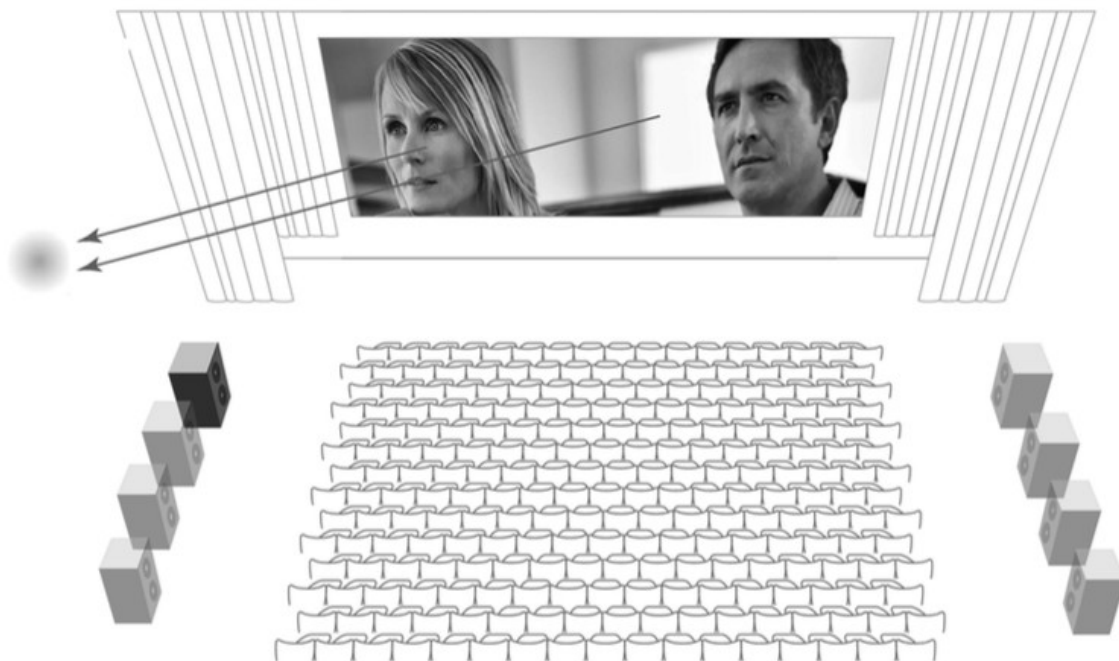
Representació dels objectes d'àudio

Per representar objectes en un espai cal una referència. Aquesta pot ser egocèntrica, en la qual es codifica la posició de l'objecte respecte a la de l'oient (el *sweet spot* és una referència egocèntrica).

També pot ser al·locèntrica, de manera que es pren com a referència un punt de l'espai sobre el qual s'està treballant.

Quan es treballa amb àudio basat en objectes, l'interès està en la relació que hi ha entre els objectes i l'entorn, de manera que serà més útil treballar amb una referència al·locèntrica, representada amb coordenades cartesianes. Posteriorment, si fes falta, es pot convertir a una escolta egocèntrica, per exemple en el cas de reproduir l'àudio a través d'auriculars.

Els altaveus de les sales modernes de cinema tenen un mapatge de la seva posició respecte una referència al·locèntrica. De manera que si els objectes sonors i els canals d'àudio utilitzen el mateix punt de referència, aconseguirem preservar la relació espacial entre els objectes i els canals en diferents sales. D'aquesta manera podem aconseguir que des de cadascuna de les butaques del cinema es localitzi un objecte sonor en un mateix punt, i que aquest també coincideixi amb la perspectiva dels personatges de la pel·lícula (il·lustració 29).



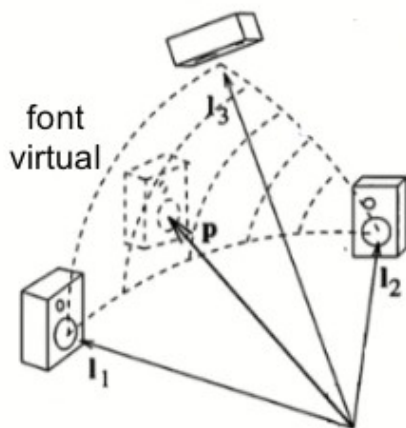
Il·lustració 29: Localització d'una font sonora dins de la sala de cinema (Geluso & Roginska, 2018).

Vector-Based Amplitude Panning (VBAP)

Per tal de codificar tota aquesta informació s'utilitzen algorismes de representació d'àudio, que assignaran un senyal monofònic individual a cadascun dels altaveus del sistema per tal de generar la percepció d'un esdeveniment auditiu en un punt de l'espai (Tsingos, 2018).

Alguns d'aquest algorismes, dels quals es parlarà més endavant, intenten recrear el camp sonor basant-se en els aspectes físics del so (WFS, Ambisonics). N'hi ha un altre de més simple que un cop més es basa en el *panning* (la panoramització del so.) Aquest s'anomena *Vector-Based Amplitude Panning* (VBAP).

El VBAP formalitza matemàticament el *panning* al domini vectorial, és a dir, la posició d'un objecte es representarà mitjançant un vector. El VBAP estén el *panning* del 2D al 3D i en comptes de repartir un so entre dos altaveus ho fa entre tres. Per calcular el guany de cada altaveu cal recórrer a les equacions que podem trobar a la il·lustració 30 (López, 2015).



$$p = g_1 l_1 + g_2 l_2 + g_3 l_3$$

$$p^T = g L_{123}$$

Obtenció dels guanys de cada altaveu

$$g = p^T L_{123}^{-1} = [p_1 \ p_2 \ p_3] \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{bmatrix}^{-1}$$

Il·lustració 30: Base matemàtica del VBAP on: p és el vector de la font imaginària; g el guany dels altaveus; l el vector de l'altaveu; L_{123} la matriu dels vectors dels altaveus; i p^T la matriu transposada del vector p (López, 2015).

Àudio basat en objectes vs. basat en canals

A l'hora de triar si es vol treballar amb àudio basat en objectes o basat en canals en dependrà l'àudio del qual es disposi i en quins sistemes es voldrà reproduir.

Com s'ha dit anteriorment, l'àudio basat en objectes treballa amb senyals monofònics (haurien de ser sense informació espacial, ja que aquesta se li afegeix posteriorment). En canvi quan es treballa amb àudio basat en canals, els enregistraments estèreo i *surround* ajuden a millorar la sensació de so immersiu. Per tant la tècnica d'enregistrament vindrà condicionada pel tipus de sistema on es vulgui reproduir l'àudio.

Per altra banda, l'àudio basat en objectes permet la reproducció en un gran nombre d'altaveus discrets, proporcionant una tímbrica més bona, una localització més precisa, i una millor relació àudio/vídeo i coherència espacial per a una gran sala d'escolta (Tsingos, 2018).

Per últim, cal recordar que l'àudio basat en canals està pensat per a reproduir-se en un sistema en concret. Per tant quan es vulgui passar d'un sistema a un altre caldrà mesclar-lo de nou. A diferència d'aquest, l'àudio basat en objectes no té predefinit cap sistema de reproducció, sinó que s'adapta a ell.

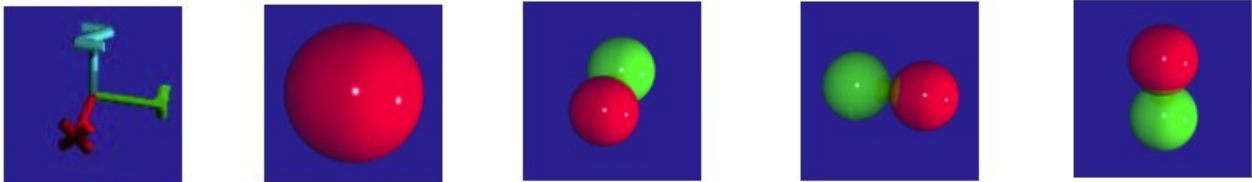
4.2.5. Sound Field

Rozenn Nicol defineix el terme *sound field* (camp sonor en anglès) com a la captura, reproducció i descripció de les ones sonores. A diferència de les tècniques binaural, estèreo i *surround*, que treballen amb les propietats de la percepció (ILDs, ITDs i HRTFs), el *sound field* treballa amb les propietats físiques de les ones sonores. Una altra característica del *sound field* és que a diferència de les tècniques tradicionals no té un focus frontal. Totes les direccions tenen la mateixa consideració.

Les bases del *sound field*, de la mateixa manera que molts dels sistemes anteriors, estan en les tècniques d'enregistrament estereofònic X/Y i M/S de Blumlein. El parell M/S, es pot entendre com un primer nivell d'enregistrament *sound field* restringit al pla horitzontal (Nicol, 2018). Per ampliar-ho al pla vertical caldrà esperar a l'aparició de la tecnologia *Ambisonics*.

Ambisonics

Michael Gerzon als anys 70 va presentar un sistema anomenat *Ambisonics*, on totes les direccions passaven a tenir la mateixa importància. El sistema, en la seva forma més bàsica (el primer ordre), codifica en quatre canals la informació direccional d'un camp de so tridimensional donat. Aquests s'anomenen: W, que informa sobre el gradient de pressió sonora de la font; X, sobre el gradient davant-darrere; Y, sobre el gradient esquerra-dreta; i Z, sobre el gradient dalt-baix (Hollerweger, 2008). Els senyals que condueixen els canals requereixen un processat previ abans de poder-se reproduir en un sistema d'altaveus. A diferència de la resta de formats *surround* multicanal, els seus canals no condueixen el senyal específic per cada altaveu sinó un codi anomenat *B-format*.



Il·lustració 31: Funcions de codificació Ambisonics de primer ordre. D'esquerra a dreta: Eix de coordenades, W, X, Y i Z (Hollerweger, 2008).

L'avantatge d'aquest format és que en estar codificat, el so *surround* es pot reproduir i adaptar a qualsevol sistema d'altaveus (binaural, sistemes *surround*, etc.) que tingui un descodificador de B-format. Per tant, un descodificador Ambisonics sempre estarà dissenyat per a un sistema d'altaveus específic, i un espai sonor codificat amb *B-format* es podrà reproduir en qualsevol sistema de descodificació Ambisonics. Aquesta tecnologia no va lligada a cap sistema d'altaveus. L'únic límit que posa és que el mínim nombre d'altaveus (L) necessaris ha de ser igual o superior al nombre de canals (N) Ambisonics: $L \geq N$

Ordres superiors

Per augmentar la "resolució del so 3D" es pot pujar d'ordre. L'inconvenient és que quan més es puja, més canals es necessiten (pel segon ordre seran necessaris 5 nous canals, i pel tercer 7). Per tant el nombre mínim d'altaveus també augmentarà:

$$N = (M+1)^2 \text{ per a la reproducció en 3D, i } N = 2M+1 \text{ per a la reproducció en 2D}$$

on N és el número de canals Ambisonics, i M l'ordre del sistema.

Cada augment d'ordre es combinarà amb els canals dels ordres anteriors. Per tant el segon ordre no només estarà format pels canals del 2n, sinó que també pels del 1r i el 0.



Il·lustració 32: Funcions de codificació *Ambisonics* de segon ordre. D'esquerra a dreta: R, S, T, U i V (Hollerweger, 2008).



Il·lustració 33: Funcions de codificació *Ambisonics* de tercer ordre. D'esquerra a dreta: K, L, M, N, O, P i Q (Hollerweger, 2008).

L'increment de la "resolució del so 3D" implicarà una localització de les fonts sonores molt més acurada (Hollerweger, 2008).

Sistema d'ordres mesclats

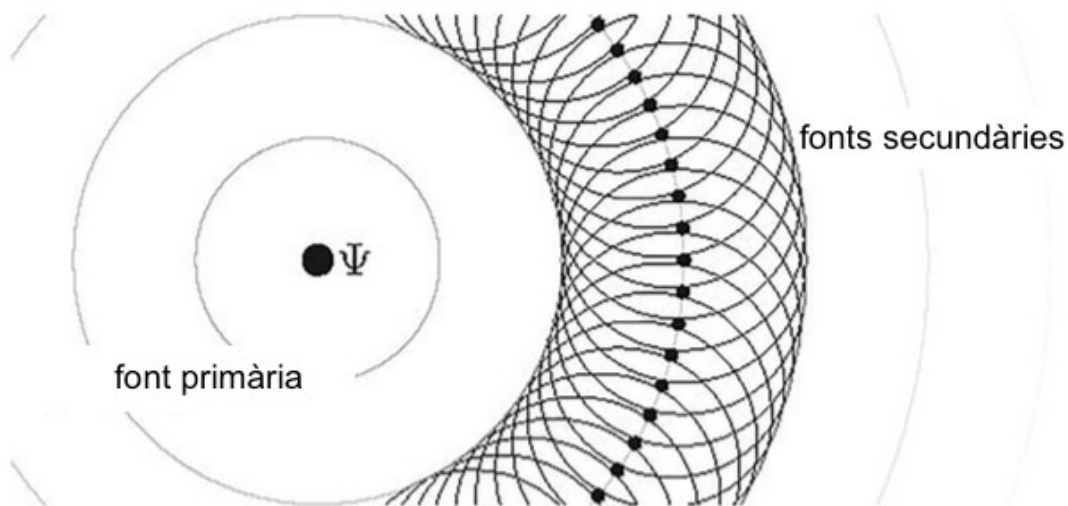
En alguns casos pot interessar reduir el nombre de canals, i com a conseqüència la resolució espacial. Per exemple, l'oïda dels humans està molt ben desenvolupada en el pla horitzontal, en canvi al vertical no tant. Per tant, no és necessària una gran resolució en el pla vertical, de manera que es poden reduir canals en aquest pla. Això ho podem aconseguir combinant diferents ordres. En el pla que vulguem tenir més definició utilitzarem ordres superiors, i al que no, inferiors. De manera que quedaran així el nombre de canals:

$$N = N_H + N_V = [2M_H + 1] + [(M_V + 1)^2 - (2M_V + 1)]$$

On N és el número de canals, M_H l'ordre horitzontal, i M_V l'ordre vertical del sistema. De manera que un sistema amb $M_H = 3$ i $M_V = 1$ necessitarà un mínim de 8 canals, a diferència d'un sistema de 2ⁿ ordre que necessitarà 9 canals i tindrà menys resolució horitzontal (Hollerweger, 2008).

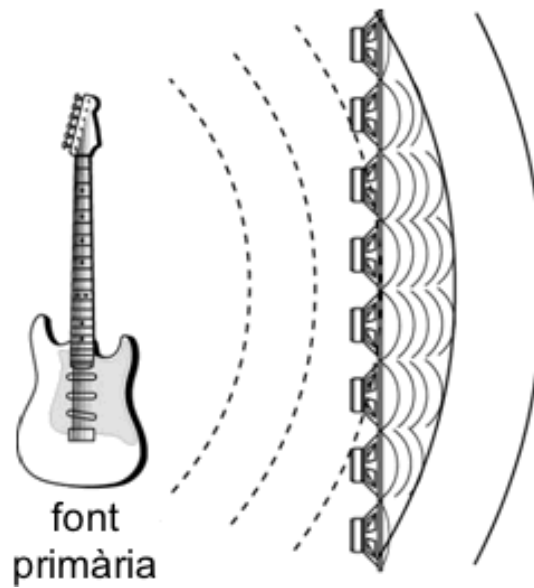
4.2.6. Wave Field Synthesis (WFS)

Els sistemes de so *surround* com el clàssic 5.1 tenen limitada l'àrea d'escolta (*sweet spot*), i en el cas de no ser-hi dins es percep una imatge acústica distorsionada. La WFS, és un dels mètodes de síntesi del camp acústic més precisos que existeixen, i permet solucionar aquest problema. Per fer-ho s'aprofita del principi de Huygens (Sporer; Brandenburg; Brix; Sladeczek, 2018). Aquest principi diu que un front d'ona pot ser substituït per fonts secundàries situades en el mateix front. D'aquesta manera s'aconsegueix produir el mateix camp acústic que el generat per la font original (il·lustració 34) (López, 2015).



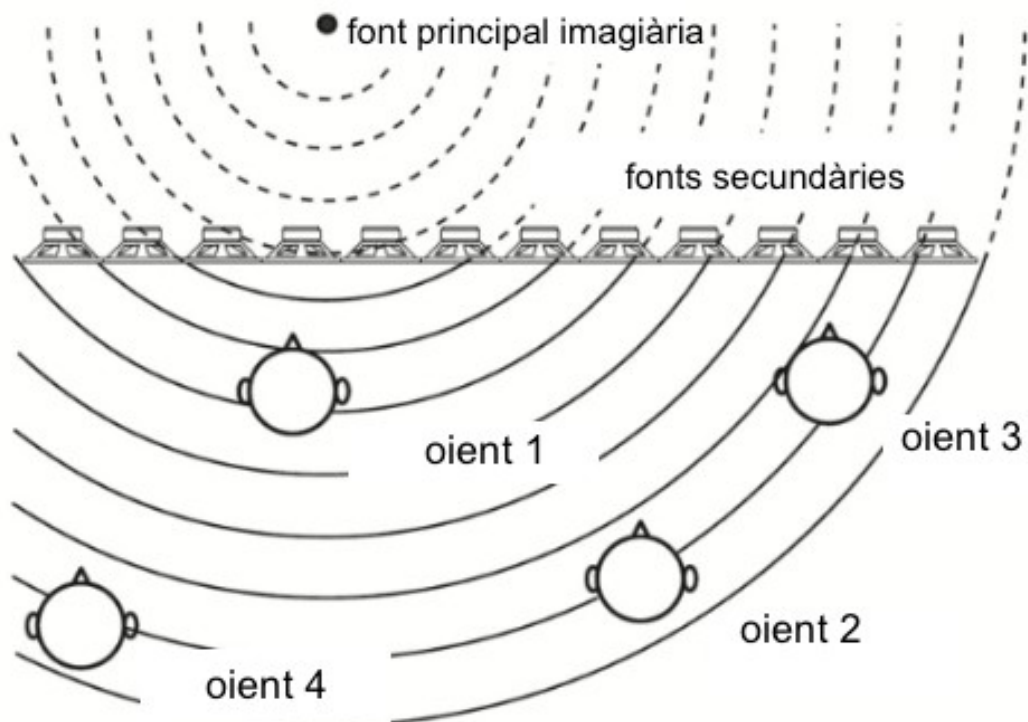
Il·lustració 34: Principi de Huygens (Geluso & Roginska, 2018).

Aquest mateix fenomen, es pot recrear amb una matriu d'altaveus col·locats en línia recta emetent el so de la font original. El senyal de cadascun d'aquests altaveus tindrà una modificació de retard i amplitud en funció de la seva distància amb la font original (il·lustració 35). És a dir, quan més lluny estigui de la font sonora més tard i amb menys amplitud reproduirà el seu so.



Il·lustració 35: Reconstrucció del front d'ona a partir de fonts secundàries (López 2015).

D'aquesta manera aconseguim que tots els oients de la sala puguin percebre l'origen del so en un mateix punt tal i com veiem a la il·lustració 36.



Il·lustració 36: Amplia àrea d'escolta de la WFS (López, 2015).

La WFS permet el lliure posicionament de les fonts, de manera que es pot classificar com un mètode d'àudio basat en objectes. Gràcies a això, a diferència dels sistemes de so estèreo i *surround*, tots els oients de la sala perceben amb realisme i fidelitat la posició de les fonts.

El desavantatge principal de la WFS és que requereix un gran nombre d'altaveus. Com més resolució sonora vulguem més altaveus necessitarem, i per tant més canals, ja que cadascun necessita el seu senyal individual. Aquest fet fa que sigui molt car i difícil d'instal·lar (Sporer; Brandenburg; Brix; Sladeczek, 2018).

4.3. Productes per treballar amb so immersiu

A l'annex I situat al final d'aquest treball, s'hi pot trobar un llistat dels diferents productes hardware i software que empren algunes de les tecnologies de so immersiu que s'han estudiat al llarg d'aquest capítol.

5. Creació del viatge sonor en tren

En aquest apartat es prendran les decisions tècniques i artístiques en base al coneixement adquirit en capítols anteriors.

5.1. Tria del sistema de reproducció

Després d'estudiar tots els sistemes de so immersiu es decideix apostar pel sistema 5.1 *surround* i el binaural per auriculars.

Motius pels quals s'ha escollit el 5.1:

- És el sistema d'altaveus més pràctic que permet jugar amb el so *surround* estèreo.
- Disposem del material necessari per treballar-hi.
- És un sistema que només treballa amb ITD i ILD, per tant ens estalviem problemes amb l'HRTF o l'IHL, que són complicats de resoldre.

Motius pels quals s'ha escollit el binaural:

- És el sistema de so immersiu que està més a l'abast de tothom. Només cal disposar d'auriculars i un aparell que reproduïxi àudio estèreo.
- Permet posar en pràctica l'enregistrament binaural i tècniques com l'*Ambisonics* amb els reptes que comporta.

Cada sistema té els seus avantatges i desavantatges. El punt fort del 5.1 és que és un sistema que s'ha utilitzat molt a l'hora de fer produccions de so envoltant, de manera que els programes que permeten treballar-hi estan molt desenvolupats. En canvi, el binaural és molt més difícil de manipular. Com ja s'ha vist anteriorment si enregistrem amb micròfons binaurals, posteriorment no podrem moure la imatge espacial. Només ho podrem fer amb sistemes d'enregistrament 3D més avançats com l'MTB.

Treballar amb aquests dos sistemes permetrà posar en pràctica dues tècniques diferents. El repte principal serà la compatibilitat d'un amb l'altre. Caldrà fer una mescla individualitzada per a cada sistema de reproducció, o els dos podran treballar amb la mateixa?

5.1.1. Eines de postproducció de so immersiu

A l'hora de treballar amb el material d'àudio s'utilitzarà el programa ProTools, que permet panoramitzar en 5.1. Com es pot veure al llistat que hi ha a l'annex I d'aquest treball, hi ha un gran nombre d'eines que permeten treballar amb so immersiu binaural. S'ha optat per provar les gratuïtes i escollir-ne una per fer la mescla binaural.

5.2. Metodologia dels enregistraments

Tal i com s'ha vist a l'apartat *Implicacions de l'enregistrament de camp*, cal tenir ben preparada la metodologia dels enregistraments. L'objectiu principal és generar matèria prima pel viatge sonor en tren, tenint en compte que la peça es reproduirà en un sistema de so immersiu.

En no voler ser una còpia realista de l'espai sonor de la R1, sinó una experimentació creativa de so immersiu, serà més interessant captar individualment els diversos sons. D'aquesta manera es podrà jugar molt més amb la seva espacialització. Tot i així, també serà interessant enregistrar àudio binaural per posar a prova les seves possibilitats a la mescla per auriculars.

Un cop coneixem l'objectiu dels enregistraments cal tenir en compte un seguit de problemes que presenten els sons de la línia R1:

- No es poden enregistrar dins de l'estudi.
- En general no es pot localitzar amb exactitud el seu punt d'origen degut que molts d'ells són mecànics i provenen de sota el tren.
- La gran majoria no és possible aïllar-los l'un de l'altre, ja que el propi so ambient els enllaça o sovint es barregen l'un amb l'altre.
- En general els sons que ens trobarem a cada viatge són poc previsibles, ja que depenen molt del tren i dels seus passatgers.

Una solució a aquests inconvenients podria ser l'ús de llibreries sonores específiques de trens, que ajudarien a tenir elements nets. Tot i així, s'ha decidit que tots els sons que apareixeran a les peces formaran part dels enregistraments realitzats i així experimentar el repte que suposa.

5.2.1. Tècniques de microfonia

Per enregistrar aquests sons cal recórrer a aparells portàtils. En no disposar de pressupost, es recórrerà al material que es té a l'abast. En aquest cas es treballarà amb una gravadora *ZOOM H6*. Com ja s'ha dit, els sons són difícils de localitzar, per resoldre aquest problema s'utilitzarà el parell XY de micròfons cardioïdes de condensador de la pròpia *ZOOM H6* (il·lustració 37), disposant-los a 120° l'un de l'altre. D'aquesta manera es podrà cobrir un gran angle. També s'utilitzarà el micròfon de condensador cardioïde *Audio-Technica Pro35* (il·lustració 38), que amb l'ajut d'una perxa permetrà enregistrar amb una certa proximitat alguns sons de difícil accés. Per últim, per enregistrar el so en format binaural es posaran a prova dos parells de micròfons que es col·loquen a l'orella. Un serà el micròfon electret omnidireccional *OKM II classic* de *Soundman*, i l'altre un parell d'electrets omnidireccional implementats per l'Àlex Barrachina¹⁴.



Il·lustració 37: Gravadora de ma *ZOOM H6* amb el parell XY acoblat (Zoom North America, 2013).



Il·lustració 38: Micròfon *Audio-Technica Pro35* (Audio-Technica, 2006).

14 Professor i cap de departament de Sonologia de l'Escola Superior de Música de Catalunya (2019).



Il·lustració 40: Micròfons *Soundman OKM II Classic* (Soundman, 2011).



Il·lustració 39: Micròfon interaural implementat per l'Àlex Barrachina.

Com que els sons són poc previsibles, l'estratègia que s'utilitzarà és gravar el viatge des del seu inici fins al seu final per tal de captar tot allò que vagi succeint. Aquest procés s'haurà de repetir nombroses vegades, enregistrant el viatge des de diferents punts i amb diferents micròfons per poder captar els diversos sons. A partir d'aquestes pistes d'àudio es mostrejaran els diversos sons que hi apareguin.

5.2.2. Realització dels enregistraments

El període de dates en el que es fan els enregistraments va del 17 de setembre del 2018 fins al 27 de febrer del 2019, realitzant-se en dies laborables entre les 08.00h i les 23:30h. S'escolleixen els horaris en funció del que es vol gravar, per exemple: per l'enregistrament de la megafonia, es tria un horari de baixa afluència per tal que els passatgers no hi interfereixin.

A l'hora de realitzar els enregistraments, s'ha de tenir en compte l'horari del tren per tal de tenir l'equip muntat i preparat per poder començar a gravar uns minuts abans que arribi. Com també que s'ha de disposar de l'espai i autonomia suficient per poder realitzar la gravació de tot el trajecte.

El material generat a partir dels enregistraments cal classificar-lo per tal que després sigui fàcil de manipular. Cada arxiu d'àudio té un número de referència i una etiqueta que aportarà informació

ràpida del que conté. Amb el número de referència, podrem buscar informació més detallada de cada arxiu a la taula adjuntada a l'apartat *Material adjunt* de l'annex II. Un exemple podria ser "102_Anunci_parada_447PB_T_001". On: "102" és el número de referència de l'arxiu; "Anunci_parada" el tipus de so que conté; "447PB" el tipus de tren (sèrie 447 planta baixa); "T" l'horari (tarda); i "001" el número d'arxiu del seu tipus.

5.3. El viatge immersiu

5.3.1. La seva estètica

El viatge vol donar protagonisme i portar al primer pla els sons de la línia R1 sovint titllats de tòxics. Molta gent els sent cada dia però no hi para atenció. Com diu Chris Watson *"We hear everything but we hardly ever listen"* (Jeffries, 2016). Aquests sons, que són l'empremta de la R1, ens ajudaran a endinsar-nos i capturar l'essència dels múltiples viatges en tren que ha fet l'autor en aquesta línia.

"Recording these sound effects well will bring out characteristics of the type of train itself and evoke the country and spirit of where the train was recorded" (Virostek, 2011).

De la mateixa manera que l'obra *El tren fantasma* de Chris Watson, aquest viatge intenta recrear des del punt de vista de l'autor la sensació i l'emoció de ser allà. Com bé diu Watson:

"That's the great thing about sound; it doesn't need any great artistic justification. People get sound very directly. It strikes into our hearts and imaginations in a very unique way."

Una altra obra de Chris Watson que recull molt bé el que es vol presentar al viatge de la R1 és *The Town Moor – A Portrait In Sound*. Una obra pensada per a reproduir-se en un sistema d'àudio 3D basat en la tecnologia *Ambisonics*. La peça és una pel·lícula sense imatge, que usa la narrativa del so per crear una imatge acústica envoltant del parc de Town Moor (Watson, 2016).

Totes aquestes idees estètiques acabaran consolidant-se en forma de dues peces. La primera mostrarà el contrast entre viatjar de dia i de nit, i la segona serà un viatge en el qual es buscarà reflectir-hi l'essència de la R1.

5.3.2. Composició de les peces

El primer pas a seguir a l'hora de crear les peces és escoltar el material enregistrat d'on es mostrejaran i classificaran els diversos sons.

Un cop fet el primer pas s'inicia la part més compositiva seguint les idees esmentades anteriorment. Es treballarà amb el programa *Ableton Live*, ja que és una eina pensada per l'edició d'àudio des d'una vessant més artística i és l'eina a la qual s'està més habituat.

La mescla de les peces al format 5.1 i binaural es farà amb el programa *ProTools*, una eina professional de mescla i postproducció d'àudio, i també la que hi ha instal·lada al hardware del sistema 5.1 del qual es disposa.

6. Resultats

6.1. Les peces

Aquest treball culmina en una part pràctica formada per dues peces de so immersiu batejades amb el nom de “Dia i nit” i “Essència”. Les dues són un viatge a la realitat i la ficció dels trajectes en tren que fa l'autor. Estan formades exclusivament per sons de la línia R1 seleccionats entre més de 19h d'àudio enregistrat.

6.1.1. Dia i Nit

Aquesta peça representa el contrast entre viatjar de dia i de nit. Al matí estem més adormits. El viatge és tranquil i el so repetitiu del tren es converteixi en un mantra, que fins i tot pot produir somnolència. Però sovint, els avisos de parada presenten volums massa forts, o fins i tot distorsionats que trenquen amb la monotonia silenciosa.

Un viatge en tren ben entrada la nit és rauxa, tot el contrari del matí. Transporta cap a les zones d'oci nocturn una gran quantitat de passatgers que acaben generant un tumult del qual no se'n pot escapar. Sovint hi ha incidències a la línia, un detonant que fa incrementar encara més la densitat de passatgers. El paisatge sonor és molt dens i atapeït. Es pot intentar obviar però al final se'ns fica al cap. El xivarri emmascara el so de la maquinaria del tren, que podem arribar a escoltar si tanquem els ulls i ens imaginem el tren buit. Per trobar la calma caldrà canviar de cotxe: el contigu potser és més buit.

El 5.1 ens dona la possibilitat de generar amb facilitat diferents plans sonors. A l'inici de “Dia i Nit” podem escoltar un pla estèreo darrere de l'oient on hi circulen automòbils, i un pla frontal estèreo on hi circula el tren. També ens permet situar sons a posicions on no acostumen a estar. És el cas del “propera parada” que va apareixent i acumulant-se en diferents punts que sorprenen a l'oient, creant un clímax que es deixa caure a la següent part de la peça (del minut 2:08 al 4:40).

Cap al final, al minut 6:21, s'hi poden escoltar diversos talls on desapareix la barrera sonora emmascaradora que generen els passatgers, deixant pas al so de tren que proporciona uns segons

de calma i desperta l'interès del seu timbre. Per sota, una reverberació digital genera una onada creixent molesta, que representa la densitat sonora de la qual no aconseguim desempallegar-nos fins que canviem de vagó.

A la peça s'utilitzen tres plug-ins. Un és un *delay* inserit al canal dels “propera parada” per tal de simular el retard que es produeix entre els múltiples altaveus de la megafonia del tren. També hi tenim un plug-in de reverberació amb una cua molt llarga, que no pretén recrear sinó crear planxes i textures sonores. Per últim també hi tenim un plug-in d'equalització que juga dos papers. És creatiu a la primera part de la peça, on per tal de donar força a l'aparició del tren s'exageren els subgreus. Al final de la peça, en canvi, és correctiu. Al tenir molta densitat sonora i diverses capes sumades, interessa equalitzar les pistes individualment per tal que no xoquin l'una amb altra.

6.1.2. Essència

Cada línia de tren té una empremta diferent. Transporta amunt i avall passatgers que acaben representant l'essència d'aquell territori. En aquesta peça farem un viatge amb alguns d'aquests personatges, submergint-nos en les seves veus que s'entrellacen i formen una polifonia amb el so constant del tren. Un so sovint titllat de tòxic, que sentim cada dia però no escoltem. Estarem disposats aquest cop, però, a parar amb atenció l'orella?

A “Essència”, es busca el joc entre les diverses melodies que generen els elements sonors dels viatge. Podem trobar moments on s'utilitza la ralentització dels sons, provocant l'aparició de timbres i ritmes interessants que a velocitat normal no es poden percebre. Es pot trobar un exemple al minut 2:20 on s'alenteix el clàxon del tren i posteriorment l'avís de tancament de portes.

Els ritmes repetitius de la maquinaria del tren donen peu a jugar amb els músics del tren. A partir del minut 6:40, podem percebre com el so mecànic del tren es va sincronitzant amb el tempo de la cançó, prenent-li el relleu i el protagonisme.

A la peça hi podem trobar un plug-in de reverberació que en alguns moments ens ajudarà a crear espais ficticis o a crear textures harmòniques. El to que produeix el motor al minut 1:30 queda suspès i s'entrellaça amb la resta de sons que van apareixent posteriorment. Un altre plug-in que

podem trobar és un equalitzador. En aquest cas s'utilitza sobretot de forma creativa per omplir i buidar diferents zones espectrals al llarg de la peça, com també per filtrar sons i situar-los en diferents espais. Al minut 5:40 quan estem dins del bany del tren, podem sentir filtrat amb un passa-baixes un caixó musical, generant la sensació que els músics estan tocant fora. Per últim, al final hi tenim una distorsió que ajuda a fer més agressiu el punt clímax que tanca la peça.

6.2. Mescla immersiva

Com s'ha comentat al capítol anterior, s'havia decidit fer una mescla 5.1 i una altra binaural. El principal problema a l'hora de treballar amb dos sistemes diferents és la seva compatibilitat.

El primer pas va ser buscar eines que permetessin passar d'una mescla a l'altra. Una bona opció era treballar amb *Ambisonics*, ja que permet adaptar la mescla als canals que necessita cada sistema. El principal problema és que quan treballem amb *Ambisonics*, és millor utilitzar objectes sonors nets i monofònics. Com que enregistrar els sons del tren amb aquestes característiques és pràcticament impossible, es va triar l'estratègia del *surround* virtual esmentada pel professor José Javier López Monfort en un dels seus vídeos (López, 2015). López explica la possibilitat de processar els canals del sistema 5.1 a través d'una HRTF genèrica com si els 6 altaveus fossin 6 fonts imaginàries situades a les direccions estàndard del 5.1. L'inconvenient d'aquest mètode és que utilitza una mescla pensada per un sistema 5.1 per generar la mescla binaural.

Per crear la imatge envoltant de les peces es va jugar amb els enregistraments estèreo i binaurals disposats als canals L/R i LS/RS, creant plans que es superposen l'un amb l'altre. Un clar exemple, com s'ha comentat abans, el podem trobar a l'inici de "Dia i Nit". Les peces també compten amb enregistraments monofònics repartits al llarg dels 360° donant més dinamisme i sorpresa, com per exemple els "propera parada" de "Dia i Nit". Aquesta mescla presenta un problema, i és que al treballar amb enregistraments estèreo, el canal central queda buit i només s'omple amb sons i efectes puntuals. Generalment, el canal central del 5.1 s'utilitza pels diàlegs, però com que en aquest cas estem parlant d'una peça musical, queda justificada l'absència de so en aquest durant la major part de les peces.

Finalment, l'àudio enregistrat en binaural s'ha utilitzat com si fos un estèreo AB. Els micròfons *Soundman* emprats només capten amb precisió les ITDs i ILDs entre les dues oïdes. Al ser unes càpsules que sobresurten molt del pavelló auditiu no arriben a captar del tot bé el filtre HRTF produït per la pròpia fisiologia de l'oïda.

Un cop fetes les mescles, es van processar amb els plug-ins¹⁵ *AMBEO Orbit*, *Anaglyph*, *IEM Plug-in* i *MyBino*. que permeten monitoritzar binauralment el sistema 5.1 creant el *surround* virtual. Cadascun d'ells aplica el seu filtre d'HRTF que transforma força la equalització i la mescla. De tots ells, el que més conserva l'original és *MyBino*. Un altre defecte que no soluciona cap dels tres plug-ins és l'efecte IHL, és a dir, en tots els casos dona la sensació que el so es genera dins del cap en comptes de generar-se al voltant. Degut que el processat del *surround* virtual es carrega gran part de la mescla immersiva, s'ha decidit que finalment només es presentaran les peces¹⁶ mesclades per a 5.1. Tot i així, també s'adjunten en binaural per tal de documentar el resultat obtingut amb *MyBino*. Es recomana que primer s'escolti la peça immersiva amb un sistema de so 5.1, ja que la binaural no és representativa del resultat final.

15 Podeu trobar més informació dels plug-ins a l'annex I del treball.

16 Podeu trobar les peces a l'annex II del treball.

7. Conclusions

Al llarg de l'estudi de les diverses tecnologies de so immersiu, s'ha pogut veure com algunes d'elles s'utilitzen des de fa dècades. Tot i així, estan en constant evolució i han avançat molt des de llavors. De la mateixa manera que el cinema ho va fer al seu moment, ara és la realitat virtual qui està accelerant el seu desenvolupament. No obstant, s'ha pogut veure com les tecnologies més modernes com l'*Ambisonics* encara són limitades en alguns aspectes. És el cas de la poca adaptabilitat que tenen amb les característiques físiques (per tant les HRTFs) de cada oient.

L'estudi de la metodologia dels enregistraments de camp ha estat clau per poder fer una bona preproducció. En tractar-se d'enregistraments per a peces immersives i la inexperiència en aquest terreny ha fet que s'infravalores la seva dificultat i apareguessin diversos entrebancs que han allargat el procés. En general ha faltat una planificació més bona i un equip més específic.

Els trens, han estat la font d'inspiració de molts autors, i així ho han fet un cop més. Ha estat tot un repte la creació dels dos viatges immersius tant des del punt de vista compositiu com tècnic. Al mateix temps, aquest treball també ha servit per aprendre i posar en pràctica noves formes d'entendre la música i sonoritzar-la, com també veure les dificultats de treballar en aquest terreny poc habitual.

També s'ha pogut veure com el sistema 5.1 està força desenvolupat però la seva difusió és difícil. En tractar-se d'un sistema que necessita un cert nombre d'altaveus i en una disposició concreta fa descartar la seva instal·lació a molts espais i és difícil trobar sales on es puguin reproduir les peces.

La creació del viatge personal en tren ha passat per diverses fases ben diferents l'una de l'altra. Des de la recerca de les tecnologies de so immersiu, fins a la recerca de les obres musicals que juguen amb el so dels trens. La combinació de totes elles ha estat necessària per complir satisfactòriament l'objectiu d'aquest treball.

Bibliografia

- Ableton. *Sound Outside: The art of field recording* [en línia]. Artists, 20 de juny, 2016 [Consulta: agost 2018]. Disponible a: <<https://www.ableton.com/en/blog/art-of-field-recording/>>.
- Algazi. *Mètode Motion-Tracked Binaural* [en línia]. University of California, 2004. [Consulta: març 2019]. Disponible a: <<https://www.ece.ucdavis.edu/cipic/spatial-sound/mtb/>>
- Audio-Technica. *Micròfon Audio-Technica Pro35* [en línia]. Audio-Technica, 2006. [Consulta: març 2019]. Disponible a: <https://www.audio-technica.com/flags/pdfs/AT06_EUR_ESP.pdf>
- Boren, Braxton (2018). *History of D Sound. Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 40–62). Nova York; Londres: Routledge. ISBN 97811389000.
- Borràs, Bernat. *Recorregut de la R1* [en línia]. Ferropedia, 2013. [Consulta: novembre 2018]. Disponible a: <http://www.ferropedia.es/mediawiki/index.php/Archivo:Trenscat_termometre-R1.gif>
- Curses. *Bogi de la Sèrie 447* [en línia]. Wikimedia Commons, 2011. [Consulta: novembre 2018]. Disponible a: <https://commons.wikimedia.org/wiki/File:Castellbisbal_080_2011-05-02.jpg>
- Curses. *Bogi motor compartit de la Sèrie 465* [en línia]. Ferropedia, 2011. [Consulta: novembre 2018]. Disponible a: <http://www.ferropedia.es/mediawiki/index.php/Archivo:Bogie_-_Civia.jpg>
- Curses. *Unitat 465 de Rodalies de Catalunya* [en línia]. Ferropedia, 2011 [Consulta: novembre 2018]. Disponible a: <http://www.ferropedia.es/mediawiki/index.php/Archivo:Castellbisbal_059_2011-04-25.jpg>
- Epstein, Louis. *Honegger's "Pacific 231" - review of the premiere* [en línia]. Musical Geography. [Consulta: març 2019]. Disponible a: <<https://musicalgeography.org/honeggers-pacific-231-review-of-the-premiere/>>
- Ferropedia. *Renfe Civia* [en línia]. Ferropedia, 27 de març 2016 [Consulta: setembre 2018]. Disponible a: <http://www.ferropedia.es/mediawiki/index.php/Renfe_Serie_465>.
- Ferropedia. *Renfe Serie 447* [en línia]. Ferropedia, 26 de març 2015 [Consulta: setembre 2018]. Disponible a: <http://www.ferropedia.es/mediawiki/index.php/Renfe_Serie_447>.
- Geluso, Paul (2018). *Stereo. Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 63–87). Nova York; Londres: Routledge. ISBN 97811389000.

Geluso, Paul; Roginska, Agnieszka (2018). *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*. Nova York; Londres: Routledge. ISBN 97811389000.

Generalitat de Catalunya. *Línia R1: Dades rellevants* [en línia]. Gencat, 25 de novembre 2010 [Consulta: Octubre 2018]. Disponible a: <http://rodalies.gencat.cat/ca/sobre-rodalies/linies-i-estacions/servei_rodalia_barcelona/r1/index.html>.

Hollerweger, Florian. *An introduction to Higher Order Ambisonic* [en línia]. Florian Hollerweger's website, 2008. [Consulta: Octubre 2018]. Disponible a: <<http://flo.mur.at/writings/HOA-intro.pdf/view>>

Inductiveload. *Coordenades esfèriques* [en línia]. Wikimedia Commons, 2008. [Consulta: gener 2019]. Disponible a: <[https://commons.wikimedia.org/wiki/File:Spherical_Coordinates_\(Latitude,_Longitude\).svg](https://commons.wikimedia.org/wiki/File:Spherical_Coordinates_(Latitude,_Longitude).svg)>

Jeffries, Stuart. *Newcastle's most buzzing night spot? A cowpat on the Town Moor*. *Art and design* [en línia]. The Guardian, 2016. [Consulta: març 2019]. Disponible a: <<https://www.theguardian.com/artanddesign/2016/jun/21/newcastle-town-moor-chris-watson-cabaret-voltaire>>

Kim, Sungyoung (2018). Height Channels. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 221–243). Nova York; Londres: Routledge. ISBN 97811389000.

Lateiner. *Esquema del sistema quadrifònic* [en línia]. Wikimedia Commons, 2006. [Consulta: gener 2019]. Disponible a: <[https://commons.wikimedia.org/wiki/File:4_0_channels_\(quadrophonic\)\(quadrophonie\)_label.svg](https://commons.wikimedia.org/wiki/File:4_0_channels_(quadrophonic)(quadrophonie)_label.svg)>

Leris, Bonaventura. *Rèplica de la locomotora Matarò* [en línia]. Wikimedia Commons, 2009 [Consulta: novembre 2019]. Disponible a: <https://commons.wikimedia.org/wiki/File:CHBcn-Matar%C3%B3_built_Jones%26Potts-1848_Replica_MTM-1948.jpg>

Licht, Alan (2007). *Sound Art: Beyond Music, Between Categories*. Nova York: Rizzoli. ISBN 9780847829699.

López Monfort, José Javier. *Sonido Espacial* [en línia]. Universitat Politècnica de València: 2015 [Consulta: gener 2019]. Disponible a: <<https://www.youtube.com/watch?v=8ap3KzqiHhQ>>
<<https://www.youtube.com/watch?v=qn1wqz9z7uk>>
<<https://www.youtube.com/watch?v=PNtn4VqXVNM>>
<https://www.youtube.com/watch?v=3cfyuah_7lk>
<https://www.youtube.com/watch?v=PfOUsDL_89o>

<<https://www.youtube.com/watch?v=MuoUfL-X-HU>>

<<https://www.youtube.com/watch?v=yJMdxjVaNV4>>

Marks, Aaron. *Aaron Marks Special: A Practical Guide to Field Recording (Part 1)* [en línia].

Designing Sound: 22 d'octubre 2010 [Consulta: setembre 2018]. Disponible a:

<<http://designingsound.org/2010/10/22/aaron-marks-special-a-practical-guide-to-field-recording-part-1/>>.

Murphy, Collen. *The Story of Kraftwerk "Trans-Europe Express"* [en línia]. Classic Album Sundays,

2017. [Consulta: Gener 2019]. Disponible a: <<https://classicalbumsundays.com/album-of-the-month-kraftwerk-trans-europe-express/>>

Murray Schafer, R. (1977). *The Soundscape; Our Sonic Environment and the Tuning of the World* (Destiny Bo). Rochester, Vermont.

Nicol, Rozenn (2018). Sound Field. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 276–310). Nova York; Londres: Routledge. ISBN 97811389000.

Patrick, Jonathan. *A guide to Pierre Schaeffer, the godfather of sampling* [en línia]. The Vinyl

Factory, 2016. [Consulta: Gener 2019]. Disponible a:

<<https://thevinylfactory.com/features/introduction-to-pierre-schaeffer/>>

Ricote, Ricardo. *Interior del Sèrie 465* [en línia]. Flickr, 2009. [Consulta: novembre 2018]. Disponible a: <<https://www.flickr.com/photos/ricote/4519265125>>

Roginska, Agnieszka (2018). Binaural Audio Through Headphones. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 88–123). Nova York; Londres: Routledge. ISBN 97811389000.

Rumsey, Francis (2018). Surround Sound. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 180–220). Nova York; Londres: Routledge. ISBN 97811389000.

Salmerón i Bosch, Carles; Museu del Ferrocarril de Vilanova i la Geltrú. *Línia R1: Història* [en línia].

Gencat, 25 de novembre 2010 [Consulta: octubre 2018]. Disponible a:

<http://rodalies.gencat.cat/ca/sobre-rodalies/linies-i-estacions/servei_rodalia_barcelona/r1/index.html>.

Siles, Gregor. *El primer ferrocarril. Barcelona-Mataró, 1848* [en línia]. Món Empresarial, 2015.

[Consulta: Novembre 2019]. Disponible a:

<<http://www.monempresarial.com/2015/10/30/primer-ferrocarril-barcelona-mataro-1848/>>

- Smith, Jacob. (2016). The chance meeting of a goose and a plover on a turntable: Chris Watson's wildlife sound recordings. *Sound Studies*, 151–164.
- Soundman. *Micròfons Soundman OKM II Classic* [en línia]. Soundman, 2011. [Consulta: març 2019]. Disponible a: <<http://www.soundman.de/en/products/>>
- Sporen; Brandenburg; Brix; Sladeczek (2018). Wave Field Synthesis. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 311–332). Nova York; Londres: Routledge. ISBN 97811389000.
- Tsignos, Nicolas (2018). Object-Based Audio. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 244–275). Nova York; Londres: Routledge. ISBN 97811389000.
- Vázquez, Pablo. *Steve Reich - Different Trains* [en línia]. Filomusica, 2002. [Consulta: gener 2019]. Disponible a: <<http://filomusica.com/filo29/grana.html>>
- Verdugo, Jordi. *Unitat 477 de Rodalies de Catalunya* [en línia]. Ferropedia, 2011. [Consulta: novembre 2019]. Disponible a: <http://www.ferropedia.es/mediawiki/index.php/Archivo:447_Renfe_-_Figueres_-_2011-08-08_-_Jordi_Verdugo.jpg>
- Virostek, Paul. *Field Report: Diesel Train Sound Effects with Microphone Comparison* [en línia]. Creative Field Recording, 26 de juliol 2011 [Consulta: agost 2018]. Disponible a: <<https://www.creativefieldrecording.com/2011/07/26/field-report-diesel-train-sound-effects-with-microphone-comparison/>>.
- Watson, Chris. *The Town Moor – A Portrait In Sound* [en línia]. Chris Watson, 2016. [Consulta: març 2019]. Disponible a: <<https://chriswatson.net/2016/06/24/the-town-moor-a-portrait-in-sound/>>
- Wenzel; Begault; Godfroy-Cooper (2018). Perception of Spatial Sound. *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio* (p. 5–39). Nova York; Londres: Routledge. ISBN 97811389000.
- Wikimedia Commons. *El Túnel de Montgat el 1950* [en línia]. Wikimedia Commons, 1950. [Consulta: novembre 2019]. Disponible a: <https://commons.wikimedia.org/wiki/File:T%C3%BAnel_de_Mongat.jpg>
- Zoom North America. *Gravadora de ma ZOOM H6* [en línia]. Zoom North America, 2013. [Consulta: març 2019]. Disponible a: <<https://www.zoom-na.com/products/field-video-recording/field-recording/h6-handly-recorder>>

Annex I

Productes hardware de so immersiu

3Dio Omni Binaural Microphone

L'Omni Binaural de 3Dio, és un micròfon professional format per 4 parelles de càpsules disposades de tal manera que permeten enregistrar àudio binaural en 360°. Gràcies al mètode *Motion-Tracked Binaural* (MTB) podem aconseguir una reproducció interactiva amb l'espai sonor.



Tècnica d'enregistrament: Binaural.

DPA 5100



El DPA 5100 és un micròfon professional compacte pensat per enregistrar àudio en format 5.1. La seva versatilitat permet muntar-lo en un peu de micròfon o suspès del sostre, com també pot suportar condicions meteorològiques intenses.

Tècnica d'enregistrament: Surround 5.1.

Neumann KU100

El Neumann KU100 és un micròfon en forma de cap humà. És el micròfon binaural referent del mercat i una excel·lent eina per l'enregistrament binaural d'estudi, però el seu preu és força elevat.

Tècnica d'enregistrament: Binaural.



Rode NT-SF1

El micròfon professional Rode NT-SF1 està format per 4 càpsules de condensador disposades en una matriu tetraèdrica. Aquestes permeten crear enregistraments en 3D gràcies a la tecnologia *Ambisonics*. Posteriorment es pot convertir l'arxiu d'àudio en altres formats com el binaural, l'estèreo o el 5.1 entre d'altres.

Tècnica d'enregistrament: Ambisonics.



Soundman OKM II Classic



Els OKM II Classic de Soundman són uns micròfons binaurals prou assequibles i de bona qualitat. L'inconvenient principal és que sobresurten força de l'orella, per tant no capta del tot bé el filtre que produeix el pavelló.

Tècnica d'enregistrament: Binaural.

Zoom H3-VR

La gravadora de ma ZOOM H3-VR també compta amb 4 càpsules disposades en una matriu tetraèdrica. És un model molt més econòmic i portàtil que el de Rode.

Tècnica d'enregistrament: Ambisonics.



Productes software de so immersiu

AMBEO Orbit

AMBEO Orbit és un plug-in gratuït que permet mesclar àudio binaural gràcies a la tecnologia *Ambisonics*. És una bona eina per començar ja que és molt fàcil de manipular. Permet treballar amb àudio mono i/o estèreo, i també recrear diferents espais segons el tipus de reflexions. El desavantatge principal de l'eina és que és molt limitada. No permet allunyar les fonts sonores de l'oient, ni rotar el cap, ni canviar les HRTFs.

Sistema de reproducció d'àudio: Binaural.

Entitat responsable: Sennheiser.

Enllaç web: <https://en-us.sennheiser.com/ambeo-blueprints-downloads>



Ambi Head

Ambi Head és un plugin professional per convertir senyal *Ambisonics* (fins al 3^r ordre) en àudio binaural 3D, i que permet la manipulació d'escenes 3D (rotació). Incorpora filtres HRTF (AMBEO Neumann KU 100 HRTF, i Youtube 360 HRTF) específicament dissenyats per l'escolta immersiva, com també permet importar-ne.

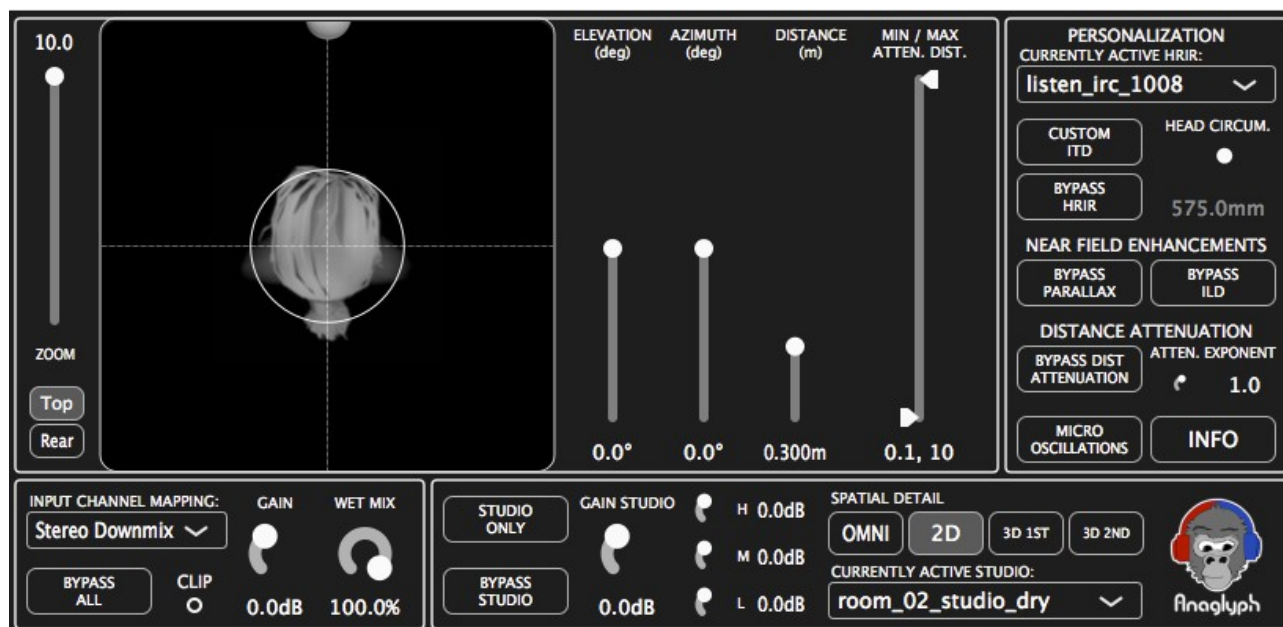
Sistema de reproducció d'àudio: Binaural.

Entitat responsable: Noisemakers.

Enllaç web: <https://www.noisemakers.fr/ambi-head-hd/>

Anaglyph

L'Anaglyph s'auto defineix com a eina d'espacialització binaural d'alta definició. És un plug-in gratuït que permet modificar bastants paràmetres, com per exemple l'HRTF, o personalitzar les ITDs segons la dimensió del cap de l'oient. Té en compte l'atenuació per distància i també intenta minimitzar l'efecte IHL. El principal desavantatge és que no permet treballar amb àudio estèreo i el converteix en mono. L'altre inconvenient és que no permet *head-tracking*.



Sistema de reproducció d'àudio: Binaural.

Entitat responsable: Anaglyph va ser desenvolupat al llarg de nombrosos projectes de recerca amb el *Centre National de la Recherche Scientifique* i la Universitat de Sorbonne.

Enllaç web: <http://anaglyph.dalembert.upmc.fr/index.html>

DearVR pro

DearVR pro és un plug-in d'àudio 3D immersiu. El seu avantatge principal és que permet manipular bastants paràmetres de l'espai acústic en el qual estem localitzant les fonts sonores. Per contra, la part perceptiva (HRTF) està força tancada.



Sistema de reproducció d'àudio: Binaural, Ambisonics.

Entitat responsable: Dear Reality.

Enllaç web: <https://www.dearvr.com/products/dearvr-pro>

Facebook360 spatial workstaion

Facebook360 spatial workstation és un plug-in gratuït de disseny d'àudio per vídeos 360° o de realitat virtual. És una eina fàcil d'utilitzar que contràriament no permet modificar cap paràmetre dels filtres HRTF. Està pensada per fer una especialització d'àudio ràpida i senzilla.

Sistema de reproducció d'àudio: Binaural, Ambisonics.

Entitat responsable: Facebook360.

Enllaç web: <https://facebook360.fb.com/spatial-workstation/>

IEM Plug-in Suite

IEM Plug-in Suite són una sèrie de plug-ins gratuïts de codi obert que permeten treballar amb la tecnologia *Ambisonics* fins al setè ordre. Inclou un compressor multibanda, codificadors d'àudio al domini *Ambisonics*, descodificadors d'*Ambisonics* a binaural o altres sistemes d'altaveus, i un mesurador de l'energia de la imatge espacial.

Com molts altres plug-ins, aquest tampoc permet modificar paràmetres dels filtres HRTF, tot i que sí que permet equalitzar l'àudio en funció dels auriculars que s'utilitzaran. Una altra característica que el diferencia d'altres, és que et permet escollir l'ordre d'*Ambisonics* al qual vols treballar.

Sistema de reproducció d'àudio: Binaural, Ambisonics.

Entitat responsable: Institute of Electronic Music and Acoustics.

Enllaç web: <https://plugins.iem.at/>

Ircam Hear v3

L'Ircam Hear v3 és un plug-in professional que ens permet escoltar binauralment una mescla estèreo o *surround* fins a un màxim de 10 canals, compatible amb Dolby Atmos. Com molts altres, es basa en la mesura de les l'HRTFs feta en el cap i tors d'un maniquí, en aquest cas del KEMAR (Knowles Electronics Maniki For Acoustic Research), de manera que no podem personalitzar-les.

Sistema de reproducció d'àudio: Binaural.

Entitat responsable: Flux.

Enllaç web: <https://www.flux.audio/project/ircam-hear-v3/>

MNTN Spatial Sound System

MNTN Spatial Sound System és un software professional que permet treballar amb el so immersiu des d'una basant més artística. Pensat per portar el so immersiu a cinemes 360°, instal·lacions multimèdia, i actuacions en directe. La característica principal d'aquesta eina és que permet treballar amb una gran quantitat de canals d'entrada i sortida, fins a un màxim de 128.

Sistema de reproducció d'àudio: Ambisonics, multicanal, binaural.

Entitat responsable: MNTN.

Enllaç web: <https://mntn.rocks/>

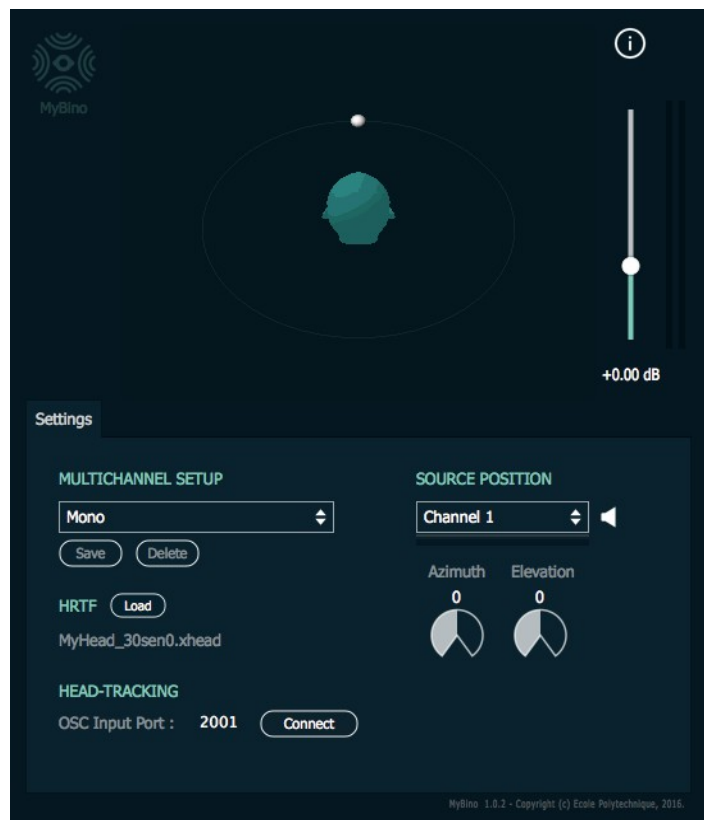
MyBino

MyBino és un plug-in gratuït que permet monitoritzar binauralment diversos sistemes d'altaveus multicanals (estereo, 5.1, 7.1, etc.). També permet treballar amb àudio basat en objectes per tal de posicionar a l'espai 3D una font sonora. És una eina senzilla però que ofereix prestacions interessants com poder carregar diversos filtres HRTF i fer *head-tracking*.

Sistema de reproducció d'àudio: Binaural.

Entitat responsable: Centre de Mathématiques Appliquées de l'Ecole Polytechnique.

Enllaç web: <http://www.cmap.polytechnique.fr/xaudio/mybino/>



Annex II - Material adjunt

Aquest enllaç conté els exemples d'àudio, les peces, la fitxa tècnica dels enregistraments i aquest treball en format digital. Per consultar el material podeu escanejar el codi QR o fer clic sobre l'enllaç.



<https://drive.google.com/open?id=1CYnqYNOI5UDI-gLxRITY3yOX4fhX3pTF>