
Introducció a l'acústica fisiològica: fonació i percepció

**Producció de la veu i caracterització del
sistema auditiu humà**

PID_00221938

Pere Artís Gabarró

Temps mínim de dedicació recomanat: 2 hores



**Pere Artís Gabarró**

Enginyer de telecomunicació per la UPC, especialista en acústica, soroll i vibracions. Després d'haver treballat a Applus, Enginyeria i Arquitectura La Salle i en diferents laboratoris acústics, acredita una sòlida experiència en la consultoria i formació acústica.

Actualment, és gerent de Keacoustics, enginyeria especialitzada en projectes que inclouen assajos acústics, xarxes de monitorització acústica, gestió de la contaminació acústica en *smartcities*, control de soroll, acústica subaquàtica, simulacions numèriques, i formació. Apassionat de l'acústica, membre de telecos.cat i de l'associació catalana de consultors acústics, participa en la redacció de legislació.

Primera edició: febrer 2019

© Pere Artís Gabarró

Tots els drets reservats

© d'aquesta edició, FUOC, 2019

Av. Tibidabo, 39-43, 08035 Barcelona

Realització editorial: FUOC

Cap part d'aquesta publicació, incloent-hi el disseny general i la coberta, no pot ser copiada, reproduïda, emmagatzemada o transmesa de cap manera ni per cap mitjà, tant si és elèctric com químic, mecànic, òptic, de gravació, de fotocòpia o per altres mètodes, sense l'autorització prèvia per escrit dels titulars del copyright.

Índex

Introducció	5
Objectius	6
1. Parla	7
1.1. Producció de la veu	7
1.2. Propietats de la veu	9
1.2.1. Intel·ligibilitat	9
1.2.2. Espectre	9
1.2.3. Directivitat	10
2. Anatomia del sistema auditiu humà	12
2.1. L'oïda externa	13
2.2. L'oïda mitjana	14
2.3. L'oïda interna	15
3. Percepció acústica	18
3.1. Llindars auditiu i de dolor	18
3.2. Sonoritat i nivell de sonoritat	19
3.3. Bandes crítiques	22
3.4. Emmascarament	23
Resum	25
Exercicis d'autoavaluació	27
Solucionari	29
Bibliografia	30

Introducció

L'objectiu d'aquest mòdul és fer una introducció al coneixement de l'òrgan fonador de la veu humana i l'òrgan receptor dels estímuls acústics.

Amb relació a la producció de la veu, es proporcionen les bases per comprendre el funcionament de generació del so i les seves principals característiques: espectre, directivitat i freqüències, que porten informació sobre la intel·ligibilitat.

Tot seguit s'aborda l'anatomia del sistema auditiu humà, relacionant cadascuna de les parts amb els comportaments psicoacústics més rellevants: percepció de la freqüència i de la sonoritat.

Objectius

Els principals objectius d'aquest mòdul són els següents:

- 1.** Conèixer la generació de so del sistema auditiu humà.
- 2.** Conèixer els paràmetres de la veu humana: espectre, directivitat i informació sobre intel·ligibilitat de la paraula.
- 3.** Distingir les diferents parts del sistema auditiu humà i relacionar-les amb la percepció acústica.
- 4.** Identificar els paràmetres més rellevants de la percepció acústica: sonoritat i tonalitat.
- 5.** Entendre què és l'emascarament acústic i què són les bandes crítiques.

1. Parla

La parla és una propietat fonamental de les persones: el to de veu, el volum o el timbre ens permeten identificar estats anímics del nostre interlocutor, i també identificar amb quin subjecte estem parlant.

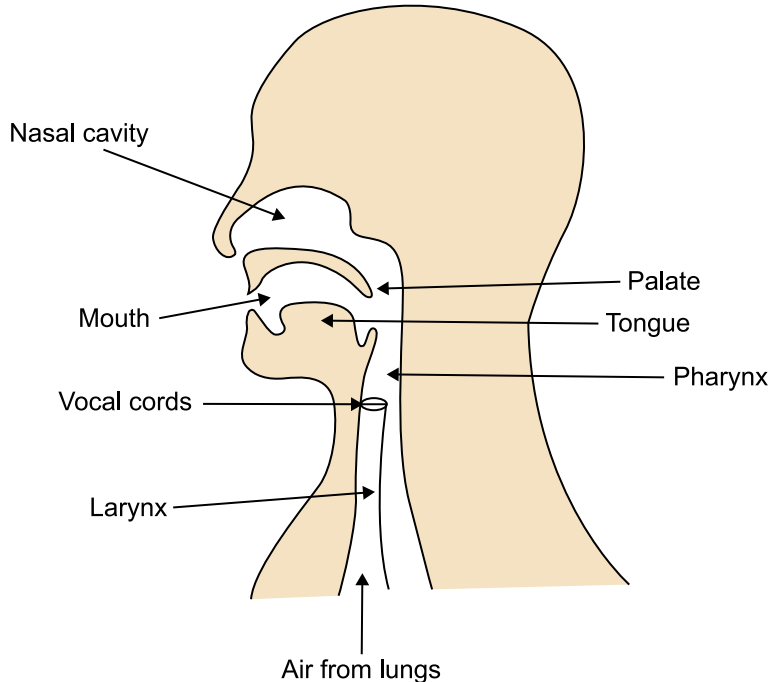
En aquest breu apartat s'explica el procés de generació de la veu humana, fent èmfasi en la separació entre sons sords i sonors, i explicant qualitativament què confereix el to i timbre de veu de cada persona.

Tot seguit, es descriuen les principals propietats de la veu humana, com són l'espectre emès i la directivitat del missatge bucal.

1.1. Producció de la veu

En la producció de la veu humana hi intervenen tots els elements que constitueixen el denominat **tracte vocal**. Tal i com es mostra a la figura 1, el tracte vocal està format per la **laringe**, la **faringe**, la **cavitat nasal** i la **cavitat bucal**.

Figura 1. Il·lustració del tracte vocal



Font: Kinsler, Frey, Coppens i Sanders (2000).

L'aire impulsat pels pulmons és conduït cap al tracte vocal, del qual, com a resultat dels mecanismes de modulació dels diferents elements que el componen, poden sorgir dos tipus de sons: els **sons sonors** i els **sons sords**.

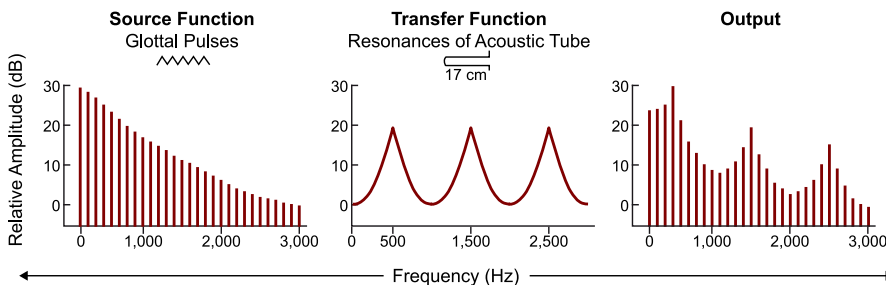
- Els sons sonors inclouen totes les vocals, les consonants sonores (/d/, /g/, /m/, etc.) i la veu cantada.
- Els sons sords són les consonants fricatives (/f/, /s/, etc.) i les oclusives (/p/, /t/, etc.).

En la generació dels sons sonors hi intervenen les **cordes vocals**, ubicades a la laringe: l'aire provinent dels pulmons fa vibrar les cordes vocals, provocant-ne una modulació. La durada d'un cicle complet, denominat **pols glotal**, és de 8 ms pel gènere masculí.

La figura 2 il·lustra el procés generatiu d'una vocal. El pols glotal presenta l'espectre de la primera gràfica, on es pot observar una freqüència fonamental i els corresponents harmònics. Per a la veu masculina, la freqüència fonamental és al voltant de 125 Hz; per a la veu femenina, als 250 Hz; i, pels infants, sobre els 300 Hz. Val a dir que tots aquests valors són molt aproximats i que per a diferents individus hi poden haver grans diferències. L'espectre originat per les cordes vocals és ric en harmònics i decau en freqüència segons la relació $1/f^2$.

A partir d'aquí, l'espectre és modificat per la faringe, la cavitat nasal i la cavitat bucal. El conjunt de cavitats es pot veure com un tub corbat de 15 cm de longitud i 70 cm³ de volum, que constitueix un filtre acústic capaç de modificar els harmònics. El conjunt de modulacions que apliquen aquests elements configura la veu característica de cada persona, el timbre de veu de cadascú.

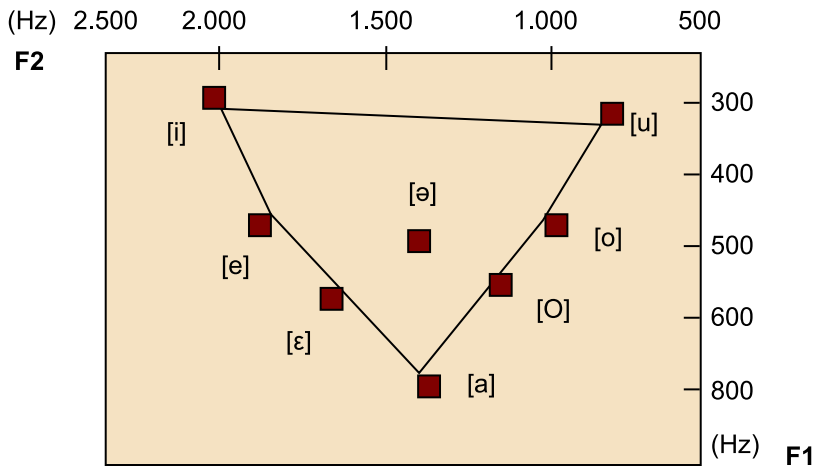
Figura 2. Procés de generació d'una vocal



Font: Jacobsen, Poulsen, Rindel, Gade i Ohlrich (2001).

Tal i com s'aprecia a la imatge central (*Transfer Function*) de la figura 2, hi ha tres harmònics fonamentals, denominats **formants**, corresponents a la faringe i les cavitats nasals i bucals. En el cas de la veu femenina, els formants estan situats, aproximadament, a: F1 = 500 Hz, F2 = 1,6 kHz i F3 = 3 kHz, essent un 20% superiors per a la veu masculina, i un 20% inferiors per a la veu infantil. En tot cas, la posició dels formants a l'espectre és característica de cada vocal i idioma. A la figura 3 es mostren els formants de les vocals catalanes (el conegut com a triangle vocàlic), que venen caracteritzades bàsicament pels formants F1 i F2:

Figura 3. Triangle vocàlic del català



Durant la generació dels sons sords, les cordes vocals resten obertes, de tal forma que el flux d'aire provinent dels pulmons no experimenta modulacions. Per a aquest tipus de sons també existeixen els tres formants principals.

1.2. Propietats de la veu

A continuació, es descriuen les propietats més rellevants de la veu humana: la intel·ligibilitat de la paraula, l'espectre i la directivitat.

1.2.1. Intel·ligibilitat

La informació freqüencial de les vocals està a les baixes freqüències, a diferència de les consonants, que tenen una distribució espectral més contínua i més informació a mitges i altes freqüències. **La intel·ligibilitat de la paraula està relacionada amb percebre correctament les altes freqüències.** Així doncs, la intel·ligibilitat de la paraula bàsicament vindrà donada per:

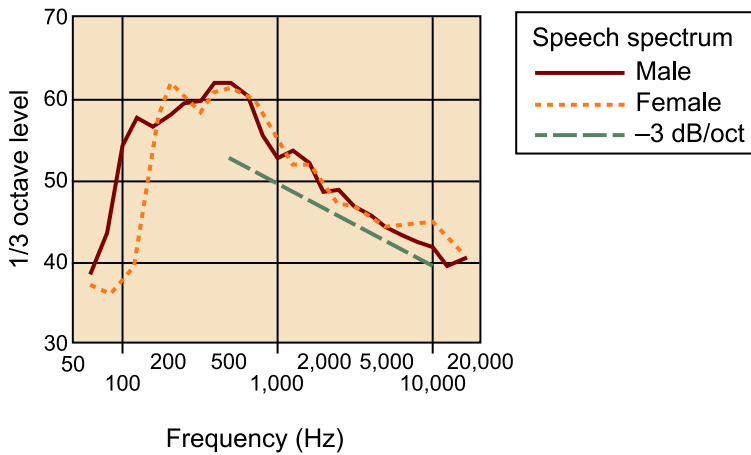
- Les consonants.
- Les bandes d'octava de 2 kHz i 4 kHz.

1.2.2. Espectre

La figura 4 mostra els espectres amitjanats de 18 mostres de 12 idiomes diferents. El valor global de la veu humana masculina, mesurada a 1 metre en front de la boca, es troba al voltant del 65 dB, essent 3 dB superior al de la veu femenina.

Com és d'esperar, **la veu masculina té més presència de baixes freqüències que no pas la femenina, precisament a causa de la durada del pols glotal.** Per a ambdós gèneres, **la informació màxima es troba al voltant de la banda de terç d'octava de 500 Hz, i a partir d'aquest màxim la relació és de -3 dB/octava** (és a dir, una pèrdua de 3 dB cada vegada que es dobla la freqüència).

Figura 4. Espectre de la veu humana



Font: Jacobsen, Poulsen, Rindel, Gade i Ohlrich (2001).

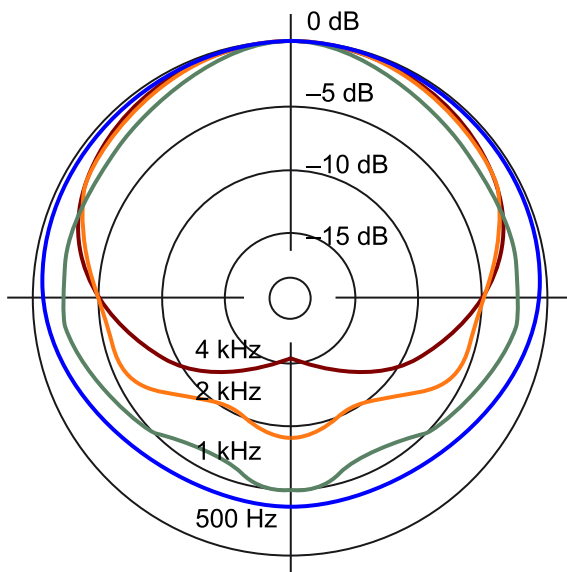
S'observa també que, a causa del funcionament del tracte vocal, la informació es distribueix en l'espectre independentment de l'idioma.

Finalment, es pot observar com les màximes components freqüencials són a les freqüències mitges, tot i que, com s'ha esmentat anteriorment, la intel·ligibilitat de la paraula s'associa a les bandes d'octava de 2 kHz i 4 kHz.

1.2.3. Directivitat

La directivitat de la veu humana ve donada tant pel tracte vocal com per la morfologia del cap. Per a altes freqüències, la veu humana és més directiva, i per la part frontal del cap sempre emetem més energia acústica que per la part posterior.

Figura 5. Diagrames de directivitat de la veu humana



Font: Carrión Isbert (1998).

Cal fixar-se, tal i com il·lustra la figura 5, que també radiem energia per la part posterior del cap: major a baixes freqüències i menor a altes freqüències. Mentre que a la banda de 500 Hz radiem aproximadament 5 dB menys per la part posterior que per la frontal, per a freqüències més elevades aquesta diferència és major: de fins a 15 dB menys a la banda d'octava de 4 kHz. També es pot apreciar com apareixen més lòbuls secundaris en el diagrama de directivitat a mesura que la freqüència creix.

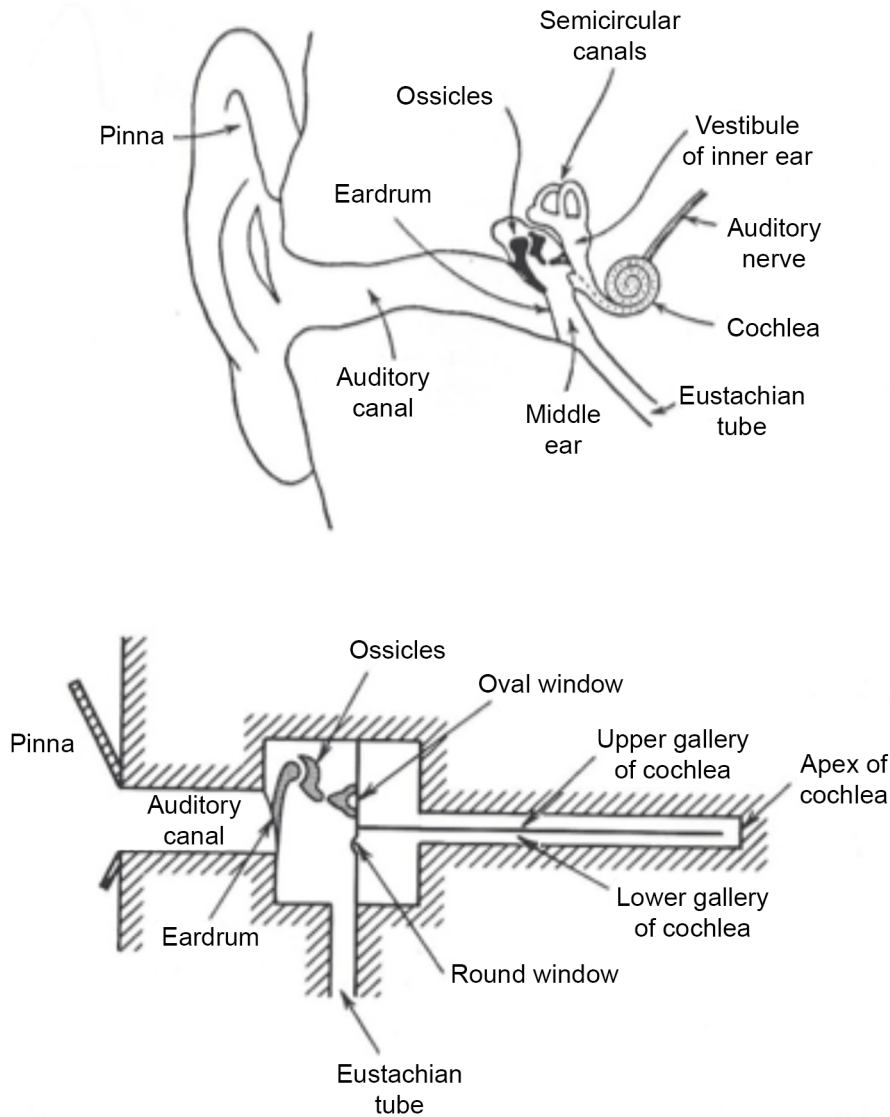
2. Anatomia del sistema auditiu humà

El sistema auditiu humà és l'òrgan receptor que actua com un transductor, tot traduint pressió sonora a impulsos nerviosos. En termes generals, el normal funcionament del sistema auditiu està dins dels límits següents:

- Rang freqüencial de 20 Hz a 20 kHz.
- Marge dinàmic de 5 dB a 140 dB (a 1 kHz).

Tal i com s'il·lustra a la figura 6, el sistema auditiu està format per tres parts principals: l'oïda externa, l'oïda mitjana i l'oïda interna. El so arriba al pavelló auditiu de l'oïda externa i és conduït fins al timpà a través del canal auditiu. El sistema d'adaptació d'impedàncies de l'oïda mitjana, format per una cadena d'ossets, condueix el senyal acústic fins a la finestra oval de l'oïda interna, que alberga un fluid. A la membrana basilar de l'oïda interna s'hi troben les cèl·lules ciliades, que transformen els desplaçaments del fluid en estímuls nerviosos que es transmeten al cervell mitjançant el nervi auditiu.

Figura 6. Representació del sistema auditiu humà



Font: Fastl i Zwicker (1999).

Als apartats següents, s'estudien més detalladament cadascuna d'aquestes tres parts principals del sistema auditiu humà.

2.1. L'oïda externa

L'oïda externa està formada pel pavelló i el canal auditiu. És en aquestes dues zones on es dona la màxima sensibilitat d'audició.

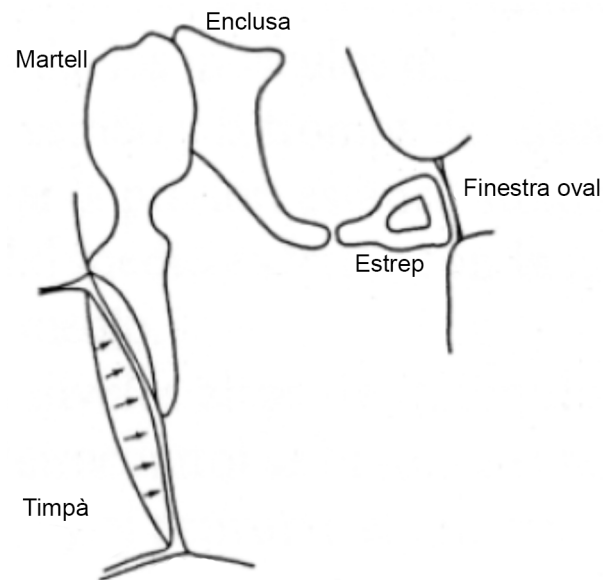
El **pavelló auditiu** actua com a protector del canal auditiu i exerceix un paper clau en la localització de fonts sonores. La geometria del pavelló afavoreix les reflexions i difraccions, de tal forma que el senyal que arriba a l'oïda ja conté informació de la direcció del so. El fet de tenir dues orelles facilita la identificació d'espais acústics i minimitza les interferències.

El **canal auditiu** assegura que el so no arriba directament al timpà. És un conducte d'aproximadament 28 mm de llargada i 8 mm de diàmetre, que ressona a $\lambda/4$ (i a les seves múltiples senars: $n\lambda/4$, $n = 3, 5, \dots$), donant lloc, per tant, a una freqüència de ressonància d'uns, aproximadament, 3,7 kHz (i els seus corresponents harmònics senars: 11,2 kHz, 18,5 kHz, etc., aproximadament).

2.2. L'oïda mitjana

Tal i com es mostra a la figura 7, l'oïda mitjana està formada per la membrana timpànica, una cadena de tres ossets (martell, enclusa i estrep), dos finestres (oval i rodona) i la trompa d'Eustaqui. Els quatre primers elements s'encarreguen d'adaptar les impedàncies de l'oïda mitjana i de l'oïda interna, que són diferents, mentre que la trompa d'Eustaqui és un element de protecció.

Figura 7. Representació de l'oïda mitjana



La vibració mecànica del timpà, una membrana d'1 mm de gruix que actua a mode de transductor de pressió sonora a energia mecànica, es transmet a la cadena d'ossos: el martell es connecta amb el timpà, i l'estrep amb la finestra oval que condueix a l'oïda interna.

Degut a que l'oïda mitjana alberga aire i l'oïda interna conté aigua, es produeix un desacoblament d'impedància acústica entre ambdues regions (provocat pel canvi de mitjà de propagació del so). La relació d'impedàncies d'ambdós mitjans resulta en unes pèrdues per desacoblament (P) d'aproximadament uns 33,5 dB:

$$Z_{om} = \rho_{aire} \cdot c_{aire} \cong 1,81 \text{ kg/m}^3 \cdot 345 \text{ m/s} = 624,45 \text{ rayl}$$

$$Z_{oi} = \rho_{aigua} \cdot c_{aigua} \cong 1000 \text{ kg/m}^3 \cdot 1400 \text{ m/s} = 1,4 \cdot 10^6 \text{ rayl} \quad (1)$$

$$P = 10 \log_{10} \left(\frac{Z_{oi}}{Z_{om}} \right) \cong 10 \log_{10} \left(\frac{1,4 \cdot 10^6}{624,45} \right) \cong 33,5 \text{ dB}$$

on Z_{om} i Z_{oi} són les impedàncies acústiques de l'oïda mitjana (és a dir, la impedància acústica característica de l'aire) i l'oïda interna (és a dir, la impedància acústica característica de l'aigua) expressades en **rayls**, respectivament; ρ_{aire} i ρ_{aigua} són les densitats de l'aire i l'aigua, respectivament; i c_{aire} i c_{aigua} són les velocitats de propagació del so a través de l'aire i l'aigua, respectivament.

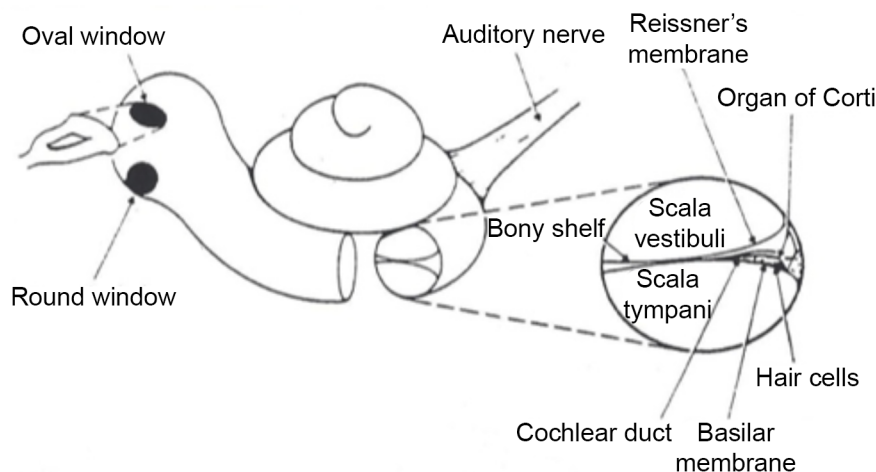
La desadaptació d'impedàncies la solucionen la cadena d'ossets i la diferència entre les àrees de la membrana timpànica i la finestra oval. La cadena d'ossets produeix un efecte palanca de factor 1,3. Aquest fet, unit a una relació de les àrees de la membrana timpànica i la finestra oval de 30:1, resulta en una reducció en les pèrdues per desadaptació d'impedàncies, passant dels 33,5 dB anteriors a únicament 5 dB.

Pel que fa la funció de la trompa d'Eustaqüi, aquesta comunica l'oïda mitjana amb la cavitat nasal i té la funció de mantenir constant la pressió atmosfèrica al voltant de la membrana timpànica. Hi ha un conjunt de músculs units a l'estrep i al martell que s'activen quan el cervell detecta un so excessivament elevat. Aquesta activació és un **reflex acústic** que fa augmentar la rigidesa dels músculs, tot disminuint la transmissió del so a l'oïda interna. Atès que el temps de resposta de l'oïda és de 20 a 40 ms, el reflex acústic no protegeix l'oïda de senyals transitoris.

2.3. L'oïda interna

Tal i com es mostra a la figura 8, l'oïda interna està constituïda pel laberint, el vestíbul i el cargol. El **laberint** és una cavitat òssia que conté un conjunt de canals semicirculars encarregats de regular el sistema d'equilibri. El **vestíbul** comunica aquests canals semicirculars amb el **cargol**, que alhora connecta l'oïda interna amb la mitjana a través de la finestra oval i la finestra rodona.

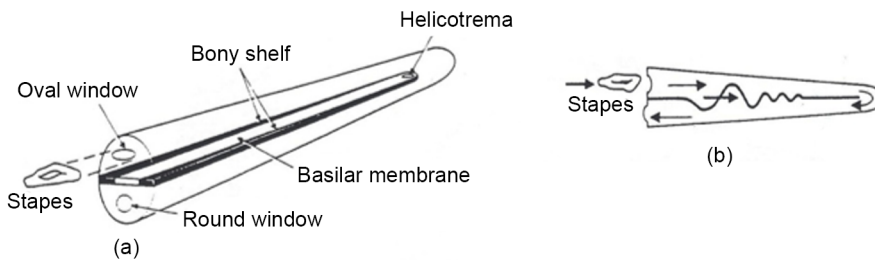
Figura 8. Representació de l'oïda interna



El cargol conté la **còclea**, un tub enrotllat en espiral que fa 2,75 voltes, amb una longitud total de 32 mm. Està dividit en tres seccions: **rampa timpànica**, **rampa vestibular** i **rampa coclear**. La rampa timpànica es comunica amb l'oïda mitjana mitjançant la finestra rodona, i la rampa vestibular ho fa per la finestra oval. Les dues rampes anteriors es comuniquen a l'helicotrema, que es troba a l'extrem del cargol. La rampa coclear conté la **membrana basilar**, on s'hi allotja l'**òrgan de Corti**. Aquest òrgan conté les anomenades **cèl·lules ciliades**, que responen amb impulsos elèctrics davant de les vibracions. El **nervi auditiu** recull tot el conjunt d'impulsos elèctrics amb les terminacions nervioses i ho envia cap al cervell.

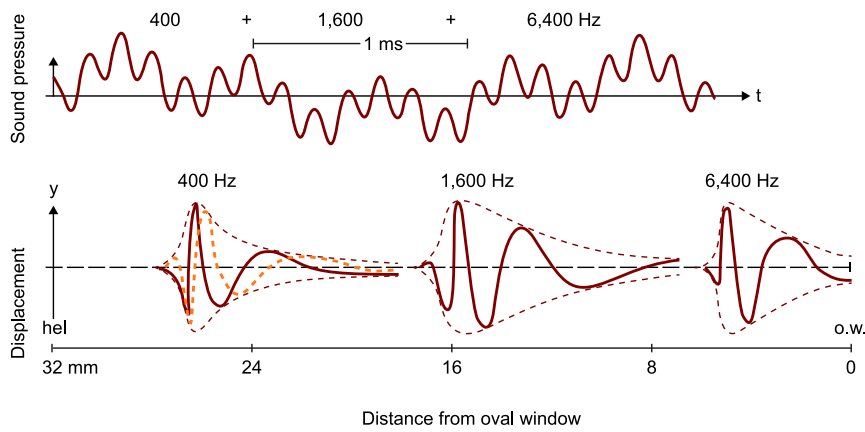
Quan l'estrep colpeja la finestra oval, s'origina una ona de pressió que es propaga a través del fluid coclear fins a la finestra rodona. L'ona de pressió causa l'excitació de les cèl·lules ciliades, tal com s'observa a la figura 9, que mostra un esquema de la còclea desenroscada.

Figura 9. Oïda interna: membrana basilar



En funció de la freqüència del so, s'excita una zona o una altra de la membrana basilar. **Les freqüències baixes exciten les cèl·lules ciliades properes a l'helicotrema, mentre que les freqüències altes la zona de la finestra oval.** La figura 10 mostra precisament aquest efecte. A la gràfica superior s'observa el senyal temporal amb presència de tres freqüències: 400 Hz, 1.600 Hz i 6.400 Hz. La gràfica inferior mostra l'esquema d'excitació de la membrana basilar, essent excitada una zona específica per cada freqüència. Cal fer notar que, tot i que es tracta de freqüències pures, s'exciten diferents cèl·lules al mateix temps. Alhora, per cadascuna de les freqüències es dona un pendent en el perfil de pujada i baixada marcadament asimètric, fet que, com es detalla en el subapartat 3.4, està estretament relacionat amb el fenomen de l'emascament.

Figura 10. Excitació de la membrana basilar en funció de la freqüència



Font: Jacobsen, Poulsen, Rindel, Gade i Ohlrich (2001).

El temps de resposta de les cèl·lules ciliades és d'1 ms, corresponent a 1 kHz. Aquesta limitació en la resposta freqüencial se soluciona a partir d'excitar-se diferents conjunts de cèl·lules que permeten la detecció de freqüències més elevades. Aquest comportament explica perquè la presbiacúcia (la pèrdua de capacitat auditiva a causa de l'envelliment) comença abans a altes freqüències que a baixes freqüències.

3. Percepció acústica

Si bé un senyal acústic es caracteritza, en termes objectius, per la seva amplitud i freqüència, la percepció acústica, de caire subjectiu, vindrà donada per la sonoritat i la tonalitat.

La **sonoritat** és una sensació d'amplitud, que no té per què mantenir una relació lineal amb l'amplitud. En efecte, un senyal de 70 dB a 2000 Hz no té per què donar la mateixa sensació de sonoritat que un soroll blanc de mateixa intensitat. A mesura que s'exciten més parts de la membrana basilar, la sensació de sonoritat augmenta. La sonoritat és directament proporcional al nombre de cèl·lules ciliades activades.

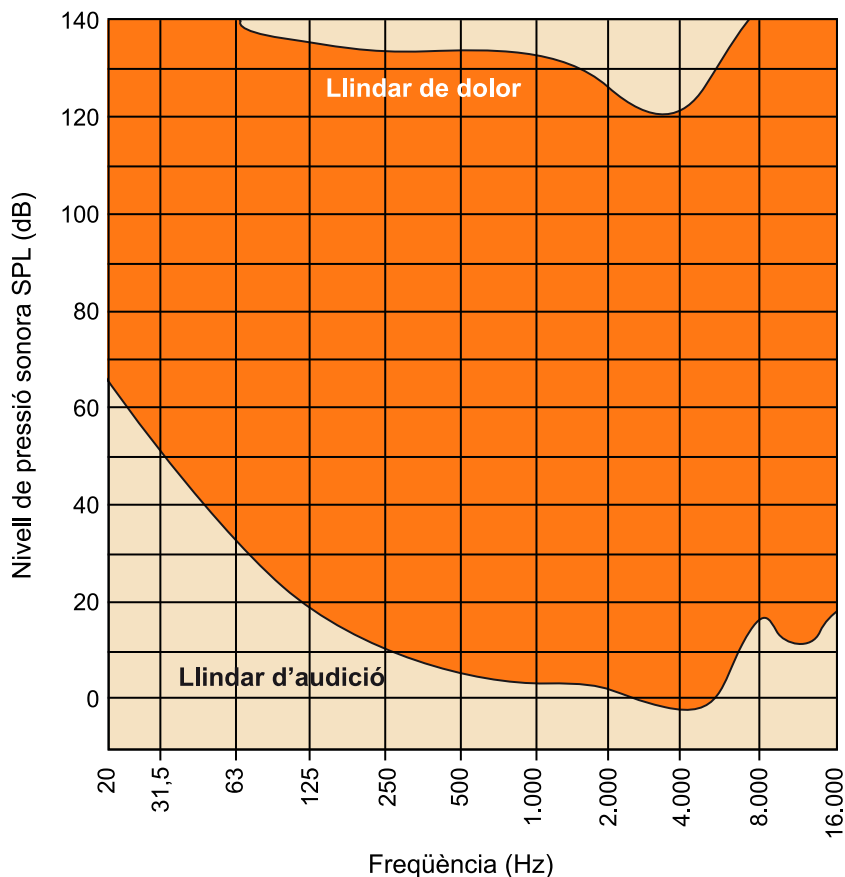
La **tonalitat** depèn de la posició del desplaçament de la membrana basilar: tonalitats associades a baixes freqüències es corresponen a desplaçaments propers a l'helicotrema, mentre que tonalitats associades a altes freqüències es corresponen a desplaçaments propers a la finestra oval. La tonalitat respon més linealment amb la freqüència del senyal acústic que no pas la sonoritat amb l'amplitud.

Tot i així, també cal considerar altres aspectes rellevants com l'emascarament, tractat en aquest apartat (veure subapartat 3.4), o l'escala de mels, disponible en la bibliografia rellevant (Fastl i Zwicker, 1999).

3.1. Llindars auditiu i de dolor

El **llindar auditiu** i el **llindar de dolor** del sistema auditiu humà són funció de la freqüència i determinen, respectivament, **el mínim nivell de pressió sonora necessari per a ser percebut** i **el màxim nivell de pressió sonora suportable per a l'oïda humana**.

Figura 11. Marges d'audició del sistema auditiu humà

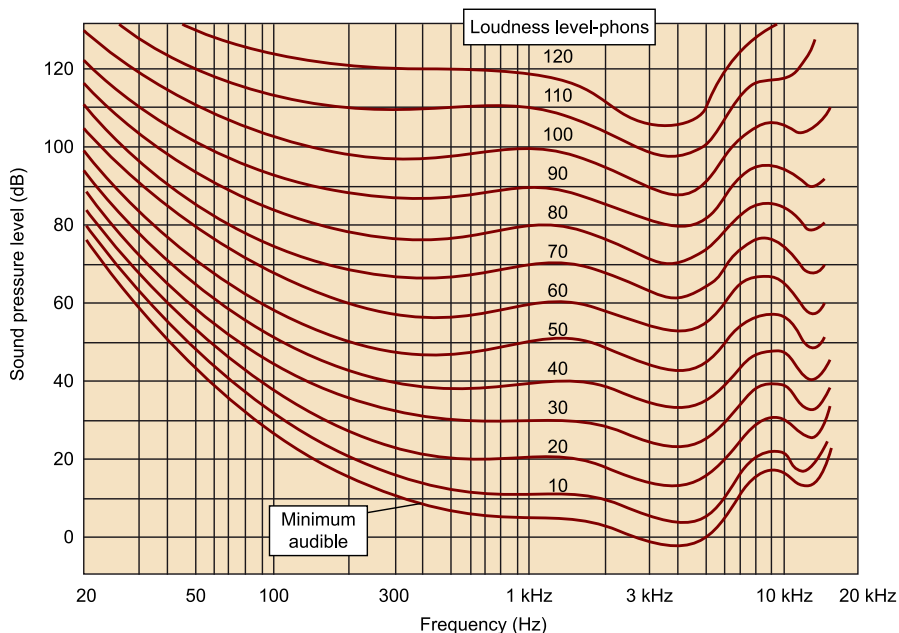


3.2. Sonoritat i nivell de sonoritat

El **nivell de sonoritat**, o *loudness level*, és un terme subjectiu que permet estimar la sonoritat d'un so. El **nivell de sonoritat d'un so determinat** és igual al nivell de pressió sonora d'un to pur a 1 kHz que té el mateix nivell de sonoritat que aquest so. La unitat del nivell de sonoritat és el **phon**.

La figura 12 mostra les **corbes d'igual sonoritat** segons Robinson i Dadson (1956) que han estat adoptades per la normativa internacional ISO 226:2003. **Aquestes corbes són vàlides per a tons purs i no haurien de ser utilitzades per a sons complexos, ja que no tenen en compte el fenomen de l'emascament ni de la integració temporal.** Per a sons complexos s'utilitzen mètodes analítics, com els de Stevens o Zwicker, que queden fora de l'abast d'aquest mòdul.

Figura 12. Corbes d'igual sonoritat conforme Robinson i Dadson



La corba d'igual sonoritat d'un valor de N phons és aquella a la qual li correspon un valor de pressió sonora de N dB a la freqüència de 1 kHz. Per tant, a tall d'exemple, a la corba d'igual sonoritat de 50 phons li corresponen 50 dB de pressió sonora ($50 \text{ dB}_{\text{SPL}}$) a 1 kHz.

El nivell de pressió sonora requerit per aconseguir un nivell de sonoritat de N phons a una determinada freqüència ve indicat pel valor de la corba d'igual sonoritat de N phons a aquella determinada freqüència. Així doncs, un nivell de pressió sonora de $20 \text{ dB}_{\text{SPL}}$ a 1 kHz representa un nivell de sonoritat de 20 phons, mentre que per aconseguir el mateix nivell de sonoritat a 50 Hz es necessita un nivell de pressió sonora de $52 \text{ dB}_{\text{SPL}}$ (aproximadament, uns 32 dB més). Aquest fet és equivalent a dir que l'oïda humana és menys sensible a baixes freqüències que a altes freqüències per a nivells de pressió sonora de 20 dB.

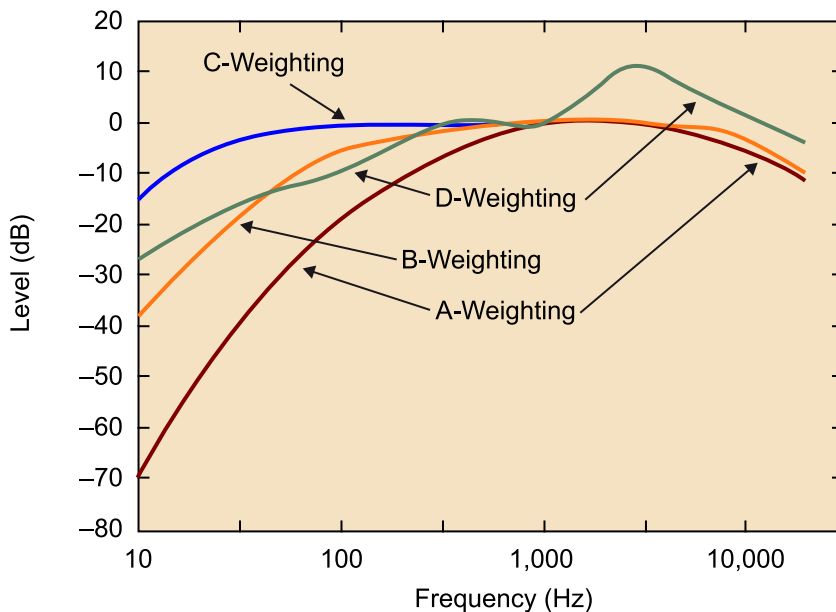
També s'observa que, a mesura que el nivell de pressió sonora augmenta, les diferències de sensibilitat són menors, de tal forma que per aconseguir un nivell de sonoritat de, per exemple, 100 phons a 50 Hz, cal un nivell de pressió sonora de $110 \text{ dB}_{\text{SPL}}$ (només 10 dB major, aproximadament, que a 1 kHz).

És també rellevant el fet que entre 3 kHz i 4 kHz es percep la mateixa sonoritat que a un to a 1 kHz amb un valor inferior de pressió sonora. Aquest comportament es deu a la freqüència de ressonància del conducte auditiu extern (veure subapartat 2.1), que ressona precisament en aquest marge freqüencial (a uns 3,7 kHz, aproximadament), i també als seus harmònics senars (12 kHz i 20 kHz, aproximadament), on també es dona aquest increment de sensibilitat.

Així mateix, cal notar que **les inverses de les corbes d'igual sonoritat es poden interpretar com les respostes en freqüència de l'oïda humana en funció del nivell de sonoritat**. De fet, és a partir d'aquestes corbes que es defineixen les ponderacions freqüencials més utilitzades (figura 13):

- La ponderació A, com la inversa de la corba de 40 phons, és usada àmpliament a les legislacions de molts països per mesurar la contaminació acústica.
- La ponderació C, com la inversa de la corba de 100 phons, és plana en la majoria del rang freqüencial i s'utilitza per avaluar nivells elevats i valors de pic.
- També s'han definit les ponderacions B (la inversa de 70 phons) i D (per a soroll d'avions), que avui en dia no s'utilitzen gaire.

Figura 13. Ponderacions freqüencials

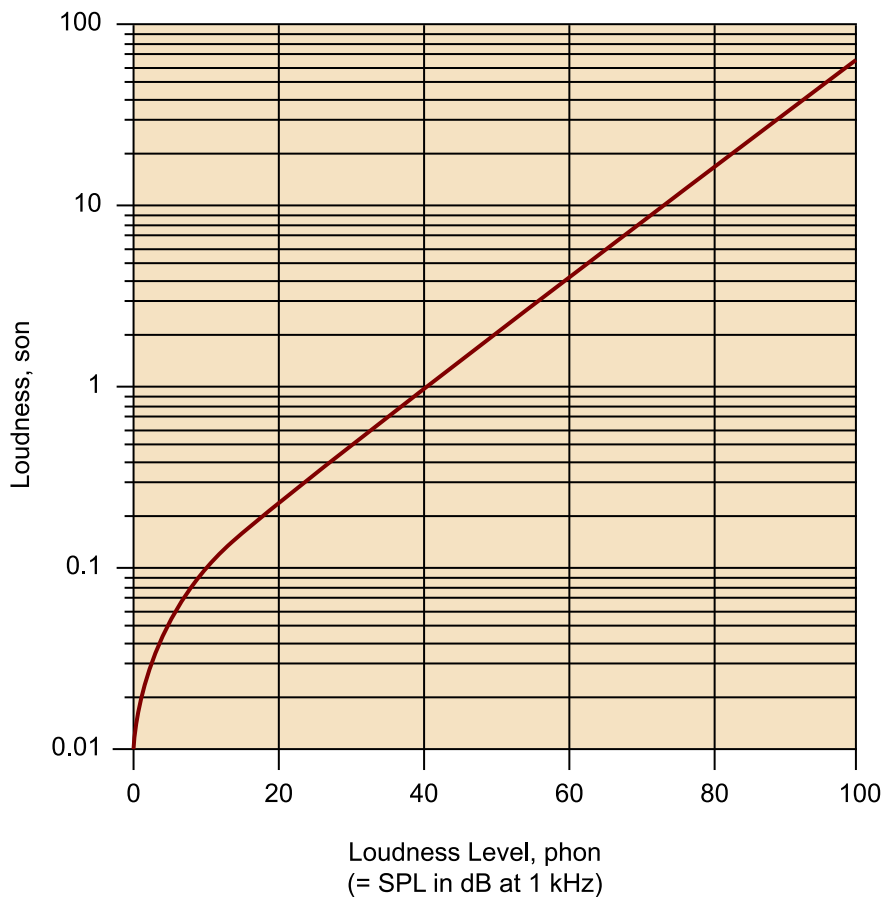


Finalment, cal fer notar que un to pur de 1 kHz de 90 dB_{SPL} no té el doble de sonoritat que un to pur de 45 dB_{SPL}. Per evitar aquesta confusió, s'estableix el son com a unitat de la sonoritat, que queda definida per la relació potencial de Stevens:

$$S = 2^{\left(\frac{NS-40}{10}\right)} \quad (2)$$

essent S la sonoritat, expressada en sons, i NS el nivell de sonoritat, expressat en phons. Aquesta relació es representa a la figura 14, on s'observa que, **per a una persona sana, s'incrementa en 10 phons el nivell de sonoritat en doblar la sonoritat, que és equivalent a incrementar en 10 dB de pressió sonora un to pur de 1 kHz**.

Figura 14. Relació entre la sonoritat i el nivell de sonoritat per a una persona sana



Aquesta relació s'acostuma a establir com a genèrica i és utilitzada àmpliament. De la mateixa manera, queda establert que, **perquè hi hagi d'haver un mínim canvi de percepció possible, ha de donar-se un increment d'1 dB**, i que un increment de 5 dB implica un clar canvi de nivell en la percepció.

La percepció de la sonoritat no és immediata, ja que el funcionament de l'oïda i el temps de reacció del cervell requereixen d'un cert temps per assolir la sensació de sonoritat. Aquest temps se situa entre 100 ms i 300 ms.

3.3. Bandes crítiques

Els moviments de la membrana basilar de l'oïda interna es poden entendre com un analitzador freqüencial, on el pic de l'envolupant s'hi posiciona en funció de la freqüència. Així, a la membrana basilar es poden definir 24 **bandes crítiques** (o **Barks**), d'amplada 1,3 mm cadascuna. Cada banda crítica es correspon a una zona en que s'exciten un conjunt de cèl·lules. L'ample de banda de cada banda crítica és del voltant de 100 Hz per a freqüències inferiors a 400 Hz, mentre que s'incrementa a partir dels 400 Hz.

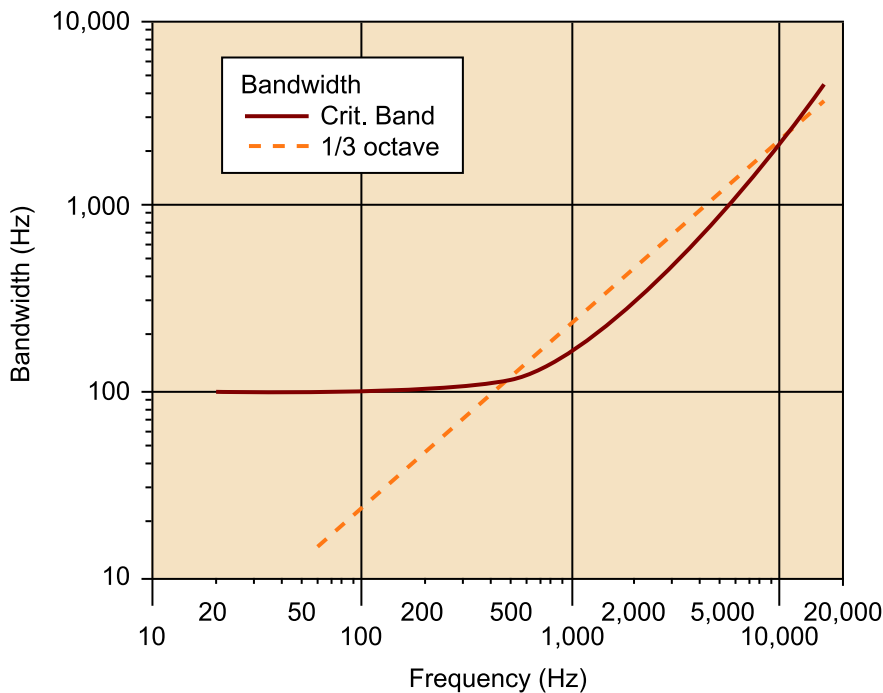
Les diferents bandes crítiques es poden obtenir a partir de la següent fórmula empírica:

$$CB = 25 + 75 (1 + 1,4f^2)^{0,69} \quad (3)$$

essent CB l'amplada de la banda crítica en Hz, i f la freqüència en kHz.

Les bandes de terç d'octava presents als analitzadors freqüencials són una aproximació de les bandes crítiques definides en l'òrgan de Corti, tal i com es pot veure a la figura 15.

Figura 15. Relació entre la freqüència central i l'ample de banda de les bandes crítiques i de les bandes de terç d'octava



3.4. Emmascarament

Tal i com ja s'ha explicat (veure apartat 2), el sistema auditiu humà presenta una determinada selectivitat freqüencial, però també és capaç de discriminar dues freqüències diferents.

El fenomen de l'**emascarament** es dona quan un **so esdevé menys perceptible a causa de la presència d'un so emmascarant**, fet que està estretament lligat a les bandes crítiques del sistema auditiu humà.

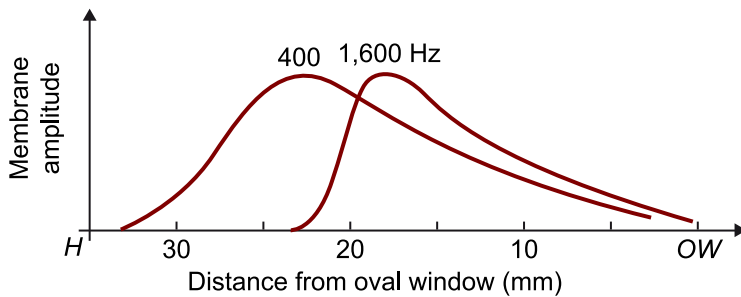
Es pot comprovar que, si s'emeten dos tons purs d'igual sonoritat però diferents freqüències **en bandes crítiques diferents**, la sonoritat serà el doble que si únicament es disposés d'un únic to. **Però si els dos tons estan situats dins la mateixa banda crítica, la sonoritat disminuirà a causa de l'emascarament.**

El fenomen de l'emascament està estretament lligat als moviments de la membrana basilar. Una explicació detallada del fenomen es pot trobar a la bibliografia (Fastl i Zwicker, 1999). En aquest apartat únicament es descriu succintament l'emascament des del punt de vista de dos tons purs.

Si es disposa de dos tons purs prou propers, es pot donar el cas que el de freqüència menor emmascari el de freqüència superior. A la Figura 16 s'observa l'amplitud de la membrana basilar per a dos tons purs de diferent freqüència (400 Hz i 1,6 kHz) en funció de la distància de la finestra oval: el to de 400 Hz està emmascarat parcialment el to de 1.600 Hz, i l'emascararia totalment si la seva amplitud fos major. A causa de l'asimetria del perfil d'amplitud de la membrana basilar, difícilment un to d'alta freqüència, doncs, podrà emmascarar un to de baixa freqüència.

Figura 16. Emmascament: resposta de la membrana basilar humana per a diferents freqüències

Human basilar membrane response for different frequencies



Els diferents mètodes de compressió d'àudio es basen en fenòmens com els descrits en els apartats de la present secció, denominats **psicoacústics**, de manera que únicament codifiquen allò que és percebut pel sistema auditiu humà.

Resum

En aquest mòdul s'han introduït els conceptes bàsics de generació de la veu humana i del sistema auditiu humà.

La veu humana es produeix al sistema fonador, on les cordes vocals i les cavitats buco-nasals hi exerceixen un paper fonamental. La generació dels sons sords es realitza fent ús de les cordes vocals, mentre que pels sons sords aquestes no hi intervenen. Les cavitats imprimeixen el timbre de cada veu humana, allò que ens permet identificar una veu coneguda enmig d'un batibull de persones.

L'anàlisi freqüencial del senyal auditiu d'una vocal condueix a l'aparició dels formants, que permeten diferenciar les vocals entre sí. S'ha fet esment que la informació del missatge va lligada a les consonants i a les bandes d'octava de 2 kHz i 4 kHz.

Pel que fa el sistema auditiu humà, es pot distingir entre l'oïda externa, mitjana i interna. L'oïda externa dirigeix les ones de pressió sonora cap al timpà, que alhora queda protegit pel conducte auditiu extern. Els dos pavellons auditius confereixen la identificació de la posició de la font en l'espai. Una vegada l'ona impacta sobre el timpà, la vibració es tradueix a la cadena d'ossets de l'oïda mitjana, encarregada d'adaptar les impedàncies de l'aire i del fluid que omple l'oïda interna. L'oïda interna alberga l'òrgan de Corti, que conté les cèl·lules ciliades susceptibles de ser excitades enfront un senyal acústic. Aquestes cèl·lules estan connectades al cervell mitjançant un seguit de neurones que condueixen l'impuls elèctric al còrtex cerebral.

L'estructura de l'oïda interna i el mode d'excitació de les cèl·lules ciliades condicionen la percepció sonora, que es pot caracteritzar en una primera aproximació per la tonalitat i la sonoritat.

Pel que fa la sonoritat, les corbes d'igual sonoritat indiquen el grau de sensibilitat de l'oïda humana per a cada banda freqüencial i per a cada rang de valors de pressió sonora. S'ha descrit la complexitat de l'oïda com a receptor i s'han presentat les principals propietats amb relació al llindar de dolor i d'audició.

A partir de les corbes de sonoritat es defineixen les conegudes corbes de ponderació freqüencial A i C, àmpliament utilitzades en infinitat d'aplicacions de l'enginyeria acústica.

Si bé la sonoritat té a veure amb quant de fort percebem un senyal, la tonalitat es relaciona amb quant de greu o d'agut és. S'han introduït les bandes crítiques de la membrana basilar i s'han relacionat amb les bandes d'octava dels analitzadors freqüencials emprats habitualment en aplicacions d'acústica i d'àudio.

Finalment, també s'ha fet èmfasi amb l'emascarament acústic, estretament lligat al comportament de l'oïda interna i utilitzat àmpliament pels algorismes de compressió d'àudio per eliminar informació no rellevant.

Exercicis d'autoavaluació

1. Amb relació al sistema auditiu humà, quina resposta és correcta?

- a) El marge dinàmic arriba fins als 140 dB, podent-se considerar que aquest valor màxim és el llindar del dolor.
- b) El sistema auditiu humà presenta una resposta en freqüència menys plana per a valors de pressió sonora elevats que no pas per a valors petits.
- c) La resposta (a) és certa i, a més, l'oïda és més sensible a baixes freqüències que a altes freqüències.
- d) Totes les respostes anteriors són correctes.

2. Amb relació a les corbes de sonoritat, quina resposta és correcta?

- a) Indiquen la sensibilitat del sistema auditiu humà en funció de la freqüència i en funció del valor de pressió sonora.
- b) Del seu estudi es pot extreure que l'oïda és més sensible a 1 kHz que a 100 Hz, de tal forma que, per aconseguir una percepció sonora de 20 phons a 100 Hz, cal un valor proper a 40 dB_{SPL}, gairebé 20 dB més que a 1 kHz.
- c) La corba de 70 phons correspon a la ponderació freqüencial C.
- d) Les respostes (a) i (c) són correctes.

3. Quina de les següents afirmacions és correcta?

- a) La funció de la cadena d'ossets de l'oïda mitjana és fer palanca i reduir la desadaptació d'impedàncies de l'aire i l'aigua.
- b) Hi ha una cèl·lula ciliada per cada banda freqüencial.
- c) L'ample de banda dels filtres de terç d'octava coincideix exactament amb l'ample de cada Bark de la membrana basilar.
- d) El timpà podria estar situat directament a l'exterior del pavelló auditiu.

4. Quina de les següents afirmacions és correcta?

- a) El decrement de sensibilitat del sistema auditiu a les bandes de 3-4 kHz té a veure amb la longitud del canal auditiu extern.
- b) Per a valors de pressió sonora elevats (> 80 dB), caldria aplicar la ponderació freqüencial A per tal de tenir en compte la resposta freqüencial de l'oïda.
- c) Per a freqüències molt elevades, no cal tenir en compte mai la ponderació freqüencial.
- d) Les ponderacions freqüencials A, B, C i D prenen el mateix valor per a 1 kHz: 0 dB.

5. Quines de les següents afirmacions són correctes?

- a) El timpà actua com un transductor acústic-mecànic.
- b) Si sentim un espetec molt fort sense estar avisats, no podem respondre immediatament a l'impacte i no podem evitar la transmissió de l'impacte a l'oïda interna. Aquest fet es deu a que el transitori de l'oïda no és nul.
- c) Sempre estarem protegits dels impactes, ja que l'oïda interna incorpora un mecanisme a base d'ossets que la protegeix enfront forts impactes.
- d) Cap de les afirmacions anteriors és correcta.

6. Quines de les següents afirmacions són correctes?

- a) L'emascament està estretament lligat a l'excitació de les cèl·lules ciliades de la membrana basilar.
- b) La resposta (a) no és correcta: només és cert que els tons greus poden arribar a emascarar completament els tons aguts.
- c) Les bandes crítiques són emulades pels analitzadors freqüencials d'octaves.
- d) En rebre l'excitació d'un to pur, s'activa una banda de la membrana basilar, no únicament una cèl·lula.

7. Quines de les següents afirmacions són correctes?

- a) Les cordes vocals intervenen en la generació dels sons sords.

- b) Les cordes vocals intervenen en la generació dels sons sonors.
- c) Les cavitats del sistema fonador humà configuren el timbre de veu de cada persona.
- d) La freqüència fonamental del pols glotal indica que les dones tenen una veu més aguda que els homes.

8. Quines de les següents afirmacions són correctes?

- a) A partir dels formants de les vocals és possible identificar la vocal emesa pel sistema fonador humà.
- b) La directivitat humana únicament depèn del tipus de tracte vocal de cada individu.
- c) La intel·ligibilitat de la veu humana es relaciona amb les consonants, sobre les bandes d'octava de 100 a 200 Hz.
- d) La veu humana en directe és més directiva per a altes freqüències que per a baixes freqüències.

Solucionari

Exercicis d'autoavaluació

1. a

2. a

3. a

4. d

5. a c

6. a c d

7. b c d

8. a d

Bibliografia

- Ballou, G. (2015). *Handbook for Sound Engineers*. Focal Press.
- Borden, G. J. i Harris, K. S. (1980). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech* (1a. ed.). Lippincott Williams & Wilkins.
- Carrión Isbert, A. (1998). *Diseño acústico de espacios arquitectónicos*. Edicions UPC.
- Davies, G. i Jones, R. (1990). *The Sound Reinforcement Handbook*. Yamaha Corporation.
- Fastl, H. i Zwicker, E. (1999). *Psychoacoustics: Facts and Models* (2a. ed.). Springer Series in Information Sciences.
- Jacobsen, F., Poulsen, T., Rindel, J. H., Gade, A. C., i Ohlrich, M. (2001). *Fundamentals of Acoustics and Noise Control*. Department of Electrical Engineering, Technical University of Denmark.
- Jung, W. (2005). *Op Amp Applications Handbook*. Elsevier.
- Kinsler, L. E., Frey, A. R., Coppers, A. B., i Sanders, J. V. (2000). *Fundamentals of Acoustics* (4a. ed.). John Wiley & Sons, Inc.
- Metzler, B. (1993). *Audio Measurement Handbook*. Audio Precision.
- Robinson, D. W. i Dadson, R. S. (1956). A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics*, 7(5), 166-181.
- Tramaine, H. M. (1969). *The Audio Cyclopedia* (2a. ed.). Howard W. Sams & Co.