

Disclaimer: il seguente documento contiene degli appunti rielaborati, pertanto questo non sostituisce né il testo consigliato dal docente né tantomeno il materiale didattico fornito.

LA PSICOACUSTICA

Il prof consiglia di andare a vedere i video della slide 18 del primo ppt sulla psicoacustica.

La psicoacustica è la scienza che studia come l'uomo percepisce il suono, l'acustica studiava invece il suono come fenomeno fisico generale.

Orecchio umano, curve isofoniche e mascheramento sono i 3 gruppi di argomenti che coprirà la psicoacustica.

Il sistema uditivo umano ha 2 modalità di funzionamento: fisiologico (vera e propria percezione di un suono) e cognitivo (emozioni/ stimoli/ sensazioni scaturiti dall'ascolto del suono).

Il funzionamento cognitivo è ancora misterioso in molti aspetti.

La comprensione sonora si compone di più fasi: riconoscimento del suono (se è una voce, una canzone, un clacson...), riconoscimento della sorgente sonora (capire l'origine del suono), comprensione del suono (le funzioni cognitive dell'apparato uditivo si attivano e il cervello comincia a fare delle ipotesi circa la zona nella quale si trova l'umano; per esempio: udire il motore di una macchina può indurre l'uomo a pensare di trovarsi nei pressi di una città; udire una certa sequenza di parole fa recepire al nostro cervello cosa il nostro interlocutore vuole comunicarci).

L'orecchio costruisce una sorta di mappa della zona circostante, quando deve comprendere la sorgente di un suono che ha recepito; in questo senso l'udito consente una visione periferica dell'ambiente in cui si trova l'ascoltatore.

Come funziona genericamente il sistema cognitivo:

1. Il suono viene udito e attraverso delle fibre raggiunge la corteccia celebrale
2. La corteccia celebrale seleziona le diverse componenti del suono (frequenze, toni...)
3. Le informazioni ricavate alla fase 2 vengono inoltrate a diverse aree del cervello (la maggiore parte di queste sono localizzate nella parte destra del cervello)
4. Sfruttando le aree di memoria del cervello si cerca di capire se il suono appena giunto all'orecchio è stato udito in passato; in questa fase si verificano le diverse reazioni/emozioni/ sensazioni che il suono suscita nell'ascoltatore

L'orecchio ha il compito convertire la pressione dell'aria causata dal suono sul timpano in impulsi nervosi (segnali elettro-chimici).

L'orecchio è diviso in: orecchio esterno, orecchio medio e orecchio interno.

Esterno

Comprende il padiglione auricolare (ossia la parte cartilaginea). Il padiglione auricolare è formato dal lobo (la parte inferiore) e dalla pinna (la parte superiore); la pinna è organizzata in "creste" (queste pieghe particolari che assume il padiglione auricolare) allo scopo di catturare le onde sonore così che vengano convogliate verso l'interno. Una volta che il suono è convogliato verso la fessura di ingresso, esso attraversa il condotto uditivo; la larghezza del condotto uditivo produce un effetto di risonanza a 2000Hz (ossia un fenomeno che amplifica i suoni di un certo valore di frequenza) sulle onde la cui frequenza appartiene al range delle frequenze medie (500-8000 Hz). Il mezzo di trasmissione sonora tipico dell'orecchio esterno è l'aria (a meno che l'ascoltatore non sia sott'acqua, ma nella maggiore parte dei casi sarà sulla terra ferma).

Medio

L'orecchio esterno finisce a partire dal timpano (è una membrana), raggiunto questo si inizia a parlare di orecchio medio. Dopo il timpano c'è una specie di camera riempita d'aria, al cui interno vi sono 3 piccole ossa ossia: martello, incudine e staffa (il martello è attaccato al timpano, l'incudine è attaccato al martello e la staffa all'incudine). L'aria che si trova nell'orecchio medio giunge ad esso mediante un condotto chiamato **tromba di Eustachio**, questo condotto prende l'aria che porta a questa parte di orecchio dalla faringe; la funzione svolta da questa parte dell'orecchio è di bilanciare la pressione all'interno della cavità con quella all'esterno dell'orecchio (alla quale è soggetta il timpano). L'aria all'interno dell'orecchio medio esercita una pressione che serve come risposta alla pressione esercitata dall'area esterna sul timpano nel momento in cui

viene captato un suono; questa differenza di pressione fa sì che il timpano vibri. Variazioni della pressione esterna fanno rendere conto all'uomo della presenza dell'aria all'interno di questa cavità.

Martello e Incudine fungono da leva per fare sì che la staffa colpisca la coclea, essi si comportano come una leva di primo genere perché il fulcro è al centro e le forze stanno ai due estremi della leva. Il meccanismo formato da Martello e Incudine viene innescato dalle vibrazioni del timpano causate dalla pressione dell'aria quando in essa si propaga un suono.

La staffa ha un'area superficie minore rispetto al timpano quindi se il timpano riceve una certa pressione data da una certa forza, ipotizzando che questa forza venga propagata così come giunge al timpano anche alla staffa, dato che questa ha una superficie minore rispetto al timpano, la pressione che essa eserciterà sulla coclea sarà maggiore di quella che il timpano ha ricevuto. (Questi fatti li si evincono dalla relazione $p=F/S \rightarrow$ superficie e pressione sono inversamente proporzionali!)

Detto in parole povere, queste 3 ossa hanno il compito di amplificare le vibrazioni che giungono al timpano.

Il muscolo stapedio è un muscolo posto fra martello e incudine che si irrigidisce quando arrivano suoni ad alta intensità all'orecchio; in questa maniera la staffa picchia meno forte sulla coclea contenendo la forte pressione che è invece arrivata al timpano e evitando che l'apparato uditivo si danneggi (questo muscolo è insomma l'unico meccanismo difensivo naturale dell'apparato uditivo che permette di attenuare l'intensità di ricezione dei suoni la cui intensità molto alta).

Il muscolo stapedio tuttavia ha dei limiti:

1. Se arriva un suono che eguaglia o supera la soglia del dolore, il muscolo stapedio non riesce a fare fronte a questa forte pressione e non riesce a mantenersi rigido, come conseguenza la coclea viene colpita con molta violenza causando appunto dolore.
2. Essendo lo stapedio un muscolo se questo è affaticato a causa di molti sforzi (esempio: l'ascoltatore è stato per molto tempo ad ascoltare musica ad alto volume) esso risponderà in maniera meno pronta qualora vengano recepiti suoni forti.

La staffa colpisce la componente che decreta la fine dell'orecchio medio e l'inizio dell'orecchio interno, ossia la **finestra ovale** (questa è una parte della coclea). Il mezzo fisico sul quale si propagherà il suono (in realtà non è un vero e proprio suono, quello è stato già udito in entrata all'orecchio esterno, quella impressa sulla finestra ovale è una pressione che è stata generata dalla ricezione di un suono e quindi la si intende come tale) arrivati a questo punto, non è più l'aria bensì un liquido oleoso e viscoso (questo è meno comprimibile dell'acqua \rightarrow comprimibilità minore coincide con velocità di propagazione del suono maggiore). La velocità di propagazione in questo secondo step sarà quindi più VELOCE perché nei liquidi il suono viaggia più velocemente; insomma, ciò che ha fatto fino a questo momento l'orecchio, è stato amplificare il suono e velocizzarne la trasmissione!

Da che il suono si stava propagando nell'aria, adesso di sta propagando in un liquido è perciò avvenuta una **rifrazione** all'interno dell'orecchio.

Dai discorsi fatti si è evinto che non è bene sforzare troppo il muscolo stapedio, infatti la legge impone che sul lavoro: se i suoni sono fra gli 80 e gli 85 dB è consigliabile prendere precauzioni circa la protezione dell'apparato uditivo; dagli 85 agli 87 dB le protezioni uditive diventano obbligatorie; dagli 87 dB in su le tonalità sonore sono illegali e si è legittimati a sospendere le attività lavorative.

DPI: Dispositivi di protezione individuale (sono quei cuffioni da cantare enormi).

DPC: Dispositivi di protezione collettiva (vetrate insonorizzate per esempio).

DPI e DPC sono i dispositivi che un datore di lavoro deve fornire ai suoi dipendenti qualora essi lavorino in un ambiente di lavoro in cui i suoni possono raggiungere la soglia di 80 dB.

Interno

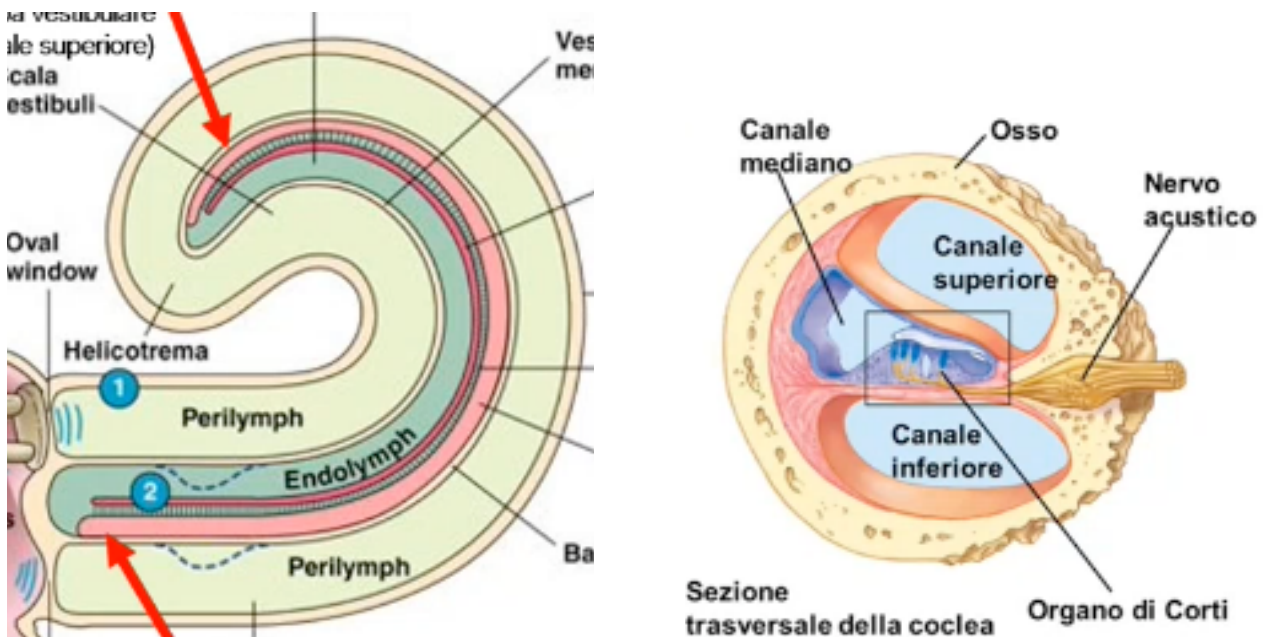
La prima parte dell'orecchio interno che si incontra è la coclea ("coclèa"), essa ha una forma a chiocciola. Come già detto, l'orecchio interno inizia in corrispondenza della finestra ovale, ossia la parte sulla quella la staffa esercita la pressione.

Internamente alla coclea ci sono tre canali l'uno sopra l'altro, essi sono il canale superiore, il canale di mezzo e il canale inferiore, essi sono riempiti con la sostanza oleosa e viscosa, già citata, che prende il nome di **linfa**. La linfa contenuta nel canale superiore e inferiore viene detta perilinfa mentre quella che del canale di mezzo è detta endolinfa. L'onda generata dalla staffa si propaga

lungo il canale superiore, una volta che arriva al centro della chiocciola (i 3 canali sono avvolti su di loro formando una chiocciola, del resto ciò è ovvio essendo la coclea una chiocciola) essa incontra una piccola fessura della **helicotrema**, a partire da questa l'onda viene propagata al canale inferiore e di mezzo.

L'onda propagata nel canale inferiore esce dalla coclea attraverso la **finestra rotonda**, è essenziale che essa esca perché altrimenti l'onda continuerebbe a riecheggiare all'interno della coclea dando vita ad interferenze sonore.

Prima di dire cosa succede all'onda propagata nel canale di mezzo, si descriva la struttura di questo canale essendo questa molto particolare. Il canale mediano ha una sezione di ampiezza più larga da un lato e più tratta dall'altro, l'interno è rivestito da una membrana chiamata **membrana tettoria**, sotto di essa ci sono delle ciglia (circa 20.000) che sono i veri e propri recettori uditivi dell'apparato uditivo. Le ciglia che sono più vicine al centro della chiocciola recepiscono le frequenze più basse, le ciglia che sono più lontane dal centro della chiocciola recepiscono le alte frequenze; il canale di mezzo si può insomma vedere come organizzato in zone ognuna delle quali è in grado di captare un certo range di frequenze.



Le ciglia, quando capano un suono, si irrigidiscono in maniera più o meno lieve a seconda dell'intensità del suono ricevuto.

Proprio per questo fatto che determinate zone del canale di mezzo consentono di ascoltare determinati suoni, si dice che l'orecchio ha un comportamento **tonotopico** ossia, tradotto letteralmente da Greco, "in base a un <tono> si attiva un <luogo> specifico dell'orecchio".

Il canale di mezzo si comporta come un pianoforte, si attivano solo le ciglia relative al suono udito, le altre restano inattive (si comportano come i tasti di un pianoforte).

Quando una zona di ciglia si attiva, cioè recepisce un suono, impiega 0.1s per ripristinarsi e potere recepire i suoni successivi, tuttavia le zone di ciglia che rimangono inattive sono utilizzabili per compiere altre operazioni di ricezione. Quanto appena detto si traduce con il fatto che se l'orecchio riceve 2 suoni di frequenze fra loro diverse, esso li processa allo stesso istante temporale perché vengono recepite da 2 zone di ciglia distinte; i suoni invece di frequenza simile vengono processati l'uno a 0.1 secondi di distanza dall'altro.

A causa di questo comportamento si dice che l'orecchio umano ha una risoluzione sia in frequenza che in tempo; facendo un paragone con i segnali in elettronica, è come se ci fosse un tasso di campionamento minimo del suono nel tempo e nelle frequenze raggiunto il quale due suoni sembrano INDISTINGUIBILI.

L'organo interno al canale di mezzo formato da membrana tettoria e ciglia è cono come **organo di Corti** (dal nome dello studioso che lo scoprì).

Alle persone audiolese viene impiantato un sistema di elettrodi che produce un certo impulso elettrico quando capta una certa frequenza, questo sistema emula il funzionamento dell'organo di Corti.

La coclea, all'estremità della sua spirale, presenta 3 anelli che si intersecano fra loro, questo sistema di anelli costituiscono il **labirinto**. Il labirinto è il sistema di equilibrio del corpo umano, il suo funzionamento è simile a quello di un giroscopio. All'interno del labirinto transitano delle specie di sassolini chiamati **otoriti**, è il movimento di questi all'interno del labirinto a fare capire l'angolo di traiettoria, rispetto la testa, verso la quale l'essere umano si sta muovendo.

Motion-Sickness: è una sensazione di malessere dovuta agli otoriti; quando il corpo umano percepisce movimenti che ne variano il suo angolo rispetto la testa che però sono finti, gli otoriti si dispongono in maniera tale da fare fronte a questo movimento fasullo perciò si percepisce una sensazione di malessere dovuta al fatto che il corpo si è preparato a uno spostamento che poi non è veramente avvenuto. (esempio: se con il visore della realtà virtuale, la visuale di gioco prende un angolo diverso da quello che il giocatore fisico ha nella realtà, si percepisce il malessere dovuto alla motion-sickness).

Frequenza, ampiezza e spettro sono finora sempre state trattate separatamente per definire i suoni, l'obiettivo è trovare una relazione fra di esse

È già stato detto in passato che il fatto che un suono abbia una frequenza udibile dall'uomo, se tuttavia l'intensità di questa non è elevata abbastanza l'orecchio non è comunque in grado di captare il suono.

I suoni complessi, nel dominio delle frequenze, possono essere visualizzati come somma di tanti suoni semplici. La frequenza di Nyquist è molto importante per la digitalizzazione, quando si useranno i filtri passabasso per filtrare i suoni da digitalizzare, il valore della frequenza di N. sarà utile per selezionare la frequenze di taglio.

Due strumenti uguali non suonano 2 suoni uguali, lo spettro, anche se in minima parte, è sempre differente! Quando invece si produce un suono al computer, la sintetizzazione sonora viene fatta con armoniche scelte da colui che sta sintetizzando il suono, perciò si possono produrre suoni perfettamente uguali dato che avranno per certo lo stesso spettro (perché c'è qualcuno che decide come costruirlo).

Sia dato lo spettro delle frequenze di un suono, si ipotizzi che questo suono durante la trasmissione venga inquinato da un rumore molto potente che distorce ogni frequenza. Si prosegue ipotizzando che le frequenze dello spettro vengono distorte in maniera tale da rendere di difficile individuazione la frequenza fondamentale dello spettro, questa difficoltà può essere data da un intervallo in cui la funzione dello spettro presenta molti picchi ammassati fra loro. A causa di questa moltitudine di picchi la frequenza fondamentale non è determinabile con precisione, ma bene o male è noto l'intervallo di frequenze in cui essa è contenuta -> questo fenomeno è chiamato **frequenza fantasma** perché so dove sta la frequenza fondamentale, ma non sono in grado di individuarla con certezza.

Il fenomeno della frequenza fantasma avviene ogni qual volta i suoni oggetti di studio vengono gravemente compromessi.

È già stato osservato il fenomeno dell'effetto Doppler causato dai moti relativi di ascoltatore e sorgente sonora. La luce, come il suono, emette la propria energia luminosa seguendo la legge dell'inverso del quadrato, perciò sorgenti luminose più distanti vengono viste con luminosità attenuata rispetto quelle più vicine.

I dB di tipo SIL (sound intensity level) sono calcolati rispetto l'intensità sonora (ossia l'energia) associata ad un suono.

$$SIL = 10 \log(I/I_0)$$

Dove: I è l'intensità sonora del suono studiato mentre I_0 è la minima intensità sonora udibile dall'orecchio umano.

Ci si vuole ora concentrare sul significato di intensità sonora, a questa grandezza è associato il volume di un suono (non il volume in metri cubi, il volume inteso come intensità del suono, quanto esso è forte o debole!!).

L'unità di misura del volume sonoro che verrà a breve definita è il **phon**.

Il tono puro con frequenza 1000Hz è il suono di riferimento a partire dal quale si costruisce la scala dei phon.

Un suono₁ ha un volume pari a N phon, quando preso come riferimento un suono₂ di 1000Hz questo ha un'intensità pari a N dB di tipo SIL.

Il diagramma di Fletcher-Munson consentono di eseguire questo confronto mediante delle particolari curve chiamate **curve isofoniche**.

I diagrammi di Fletcher-Munson sfruttano sia una scala logaritmica che decimale: la frequenza viene rappresentata in Hz sull'asse delle x (scala decimale) mentre sull'asse delle y viene rappresentata l'intensità sonora in dB SIL (scala logaritmica).

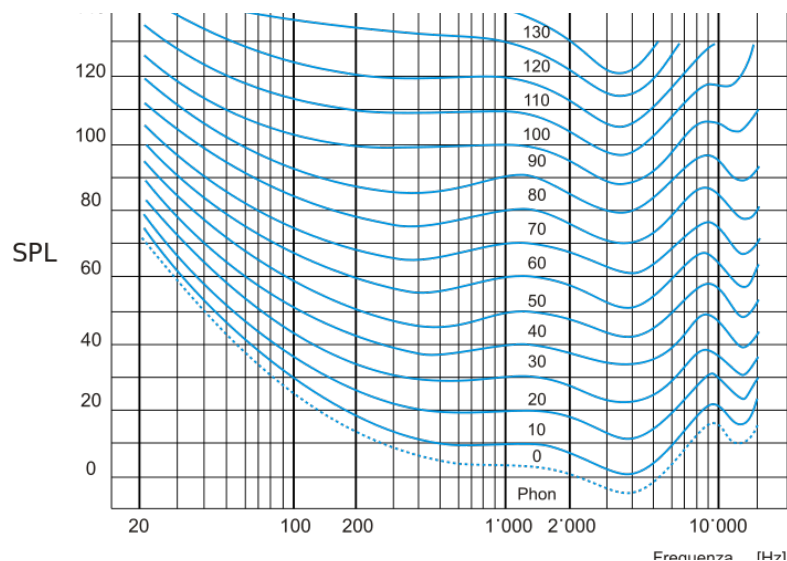
Le frequenze riportate sull'asse delle ascisse vanno da un range da 20 a 20.000; di ultrasuoni e infrasuoni non si tiene conto dato che l'essere umano non può udirli.

Le intensità sonore della scala logaritmica vanno invece da un range che va da un valore minimo negativo (per ora lo si chiami k) e un valore massimo pari a 140 (soglia del dolore dei dB SIL).

Un valore minimo di energia negativo è accettabile perché l'energia in questi grafici è espressa in decibel ossia una scala **adimensionale**; i dB sono numeri puri che mettono a confronto l'energia sonora più bassa udibile con le energie di suoni prodotti; quindi quando avviene che il rapporto del logaritmo è compreso fra 0 e 1, esso darà un numero negativo (l'intensità sonora della camera anecoica produce appunto un valore negativo di dB SIL).

I diagrammi di F-H sono in due dimensioni, se in maniera ipotetica ci fosse una terza dimensione essa individuerrebbe i phon; questa terza dimensione è espressa dalle curve isofoniche che vengono rappresentate nel grafico!!

Il diagramma di F-H è diviso in zone, per ogni zona passa una curva isofonica, la zona alla quale appartiene un punto della curva isofonica individuata da una certa coppia di valori frequenza-intensità individua il valore di phon ai quali corrisponde l'onda sonora di frequenza e intensità sonora quella data.



A parità di energia, ma a frequenze diverse, ci si sposta lungo curve isofoniche diverse, individuando suoni i cui volumi sono più elevati o più bassi.

Lungo le curve isofoniche che il grafico individua, i suoni vengono percepiti allo stesso volume espresso in phon.

Il diagramma di F-M consente di arrivare alla conclusione che anche se dei suoni hanno frequenze e intensità sonore diverse, se essi sono collocati sulla stessa curva isofonica, verranno percepiti dall'uomo allo STESSO volume.

curva isofonica più bassa = suono percepito più potentemente

curva isofonica più alta = suono percepito meno potentemente

i punti che sono su una stessa curva azzurra vengono uditi allo stesso volume!

Soglia del silenzio: curva isofonica più bassa del grafico; questa soglia viene utilizzata per comprare i suoni, in maniera tale che non verranno uditi.

Soglia del dolore: curva isofonica più alta del grafico.

I suoni la cui frequenza è minore, necessitano di intensità sonore maggiori per essere uditi meglio.

Stereofonia: tecnica della localizzazione spaziale delle sorgenti sonore.

Fletcher studiò e confermò che un ascoltatore dotato di 2 recettori sonori è più agevolato a rintracciare i suoni.

Si ipotizzi di avere un suono la cui sorgente è più vicina all'orecchio dx rispetto al sx, ciò significa che l'orecchio dx recepisce prima (in termini temporali) e in maniera più chiara (in termini di intensità sonora) il suono rispetto l'orecchio sx (ciò è sicuramente vero perché l'onda sonora impiega meno tempo a percorrere lo spazio dalla sorgente all'orecchio dx, essendo questo più vicino e sempre a causa della maggiore vicinanza dell'orecchio dx, ad esso arriverà un suono di intensità maggiore per la legge dell'inverso del quadrato).

In termini grafici, il suono che arriva all'orecchio sx è più sfasato e meno ampio rispetto a quello che arriva all'orecchio destro (che parte dall'origine degli assi ed ha un'ampiezza maggiore, si ricordi infatti che l'ampiezza dipende dall'intensità!).

I due suoni arrivano uno al tempo t_1 e l'altro al tempo t_2 , essi quindi giungono ad una distanza l'uno dall'altro pari a $\Delta t = t_2 - t_1$.

Il cervello umano, che ragiona in termini di frequenza, si rende quindi conto della differenza di frequenza fra ogni armonica componente i due suoni (ogni frequenza dista l'una dall'altra di un valore $\partial f = 1/\Delta t$) e riesce quindi a individuare la posizione della sorgente sonora.

I suoni la cui frequenza è molto elevata, sono più difficili da individuare poiché i loro grafici sono molto schiacciati (questo perché le loro componenti oscillano in intervalli di tempo molto piccoli), perciò il Δt di differenza fra i grafici delle componenti del suono giunto all'orecchio destro e quello giunto all'orecchio sinistro è così piccolo da essere quasi trascurabile, quindi $\partial f = 1/t$ è prossima a 0 e il cervello non è più in grado di capire a quale orecchio è giunto per primo il suono.

Sulla base dello stesso ragionamento appena fatto, rilevare le sorgenti sonore sott'acqua è più complicato perché la velocità di propagazione delle onde sonore in acqua è così grande che alle orecchie arriva il suono praticamente in simultanea, per questo motivo il cervello non è in grado di determinare se il suono è arrivato prima ad un orecchio rispetto che ad un altro.

L'area del grafico di Fletcher-Munson dove tutte le curve isofoniche diventano concave (da 2000Hz a 10.000Hz), racchiude i suoni meglio udibili dall'orecchio umano.

Il valore non solo non può essere negativo, ma non può neanche essere pari a 0; un valore di phon pari a 0 equivale all'assenza sonora, ma come è stato detto in una delle prime lezioni l'assenza sonora NON esiste (a meno che non si è nello spazio aperto, ma lì non essendoci un ascoltatore non si può di fatto realizzare un sistema sonoro, perciò in maniera molto sporca si può dire che non esistono i suoni nello spazio).

Il timbro di un suono ne descrive la qualità, mediante un timbro specifico si possono distinguere due suoni nonostante questi abbiano stessa altezza e volume. Quando si è parlato dei due strumenti musicali che nonostante fossero identici producevano due suoni distinti, la componente che li distingueva era proprio il timbro. Il timbro impatta sulla forma d'onda nel tempo.

La voce umana consente di produrre dei suoni che possono essere sostenuti nel tempo, questi sono le vocali!

Sia data una sorgente, se ogni volta che essa riproduce un certo suono lo fa con stessa altezza, stesso volume e stesso timbro, lo spettro di tale suono è detto **formante** (spettro COMPLETO, ossia composto da tutte le armoniche semplici di cui il suono è costituito); le vocali umane, quando riprodotte a parità di altezza e volume, hanno sempre lo stesso timbro, infatti il loro spettro è una formante.

Tremolo: variazione periodica dell'ampiezza di una nota

Vibrato: variazione periodica dell'altezza (frequenza) di una nota

Tremolo e vibrato sono due fenomeni sfruttati dai musicisti per le loro composizioni.

Traducendo questi due fenomeni in termini di trasformazioni della forma d'onda (ossia del grafico in variabile temporale di un suono):

-Tremolo -> la forma d'onda (ossia la rappresentazione nel dominio del tempo del suono) si schiaccia verso l'asse delle ascisse se l'ampiezza viene portata ad un valore minore di quello originale; si allarga allontanandosi dall'asse delle ascisse se l'ampiezza viene portata ad un valore maggiore di quello originale.

-Vibrato -> la forma d'onda si schiaccia verso l'asse delle ordinate se l'altezza (frequenza) viene portata ad un valore minore di quello originale; si allarga allontanandosi dall'asse delle ordinate se l'altezza viene portata ad un valore maggiore di quello originale.

Nelle immagini digitali, più sono i pixel utilizzati per rappresentare un'immagine, migliore è la qualità di visualizzazione di questa, dunque i dettagli contenuti nell'immagine sono più chiari. Un'analogia simile la si può fare anche per i suoni, più sono i campioni dei quali si dispone dello spettro di un suono, più fedelmente si potrà ricostruire la sua forma d'onda.

L'orecchio è tonotopico -> l'attivazione di certe aree di ciglia, consente di captare suoni di certe frequenze. Ma come si definisce l'ampiezza dell'area di ciglia che si attiva quando si riceve un certo suono?

Generalmente le aree di ciglia che si attivano non sono uniformi, nel senso che un'area di ciglia che si occupa di recepire una frequenza f_1 e un'altra area che si occupa di recepire una frequenza f_2 , non hanno la stessa estensione (una magari contiene 1000 ciglia, l'altra 2000 -> esempio casuale, non ha veridicità sperimentale).

Il **mascheramento** è il fenomeno della psicoacustica che consente di identificare l'ampiezza di queste aree di ricezione.

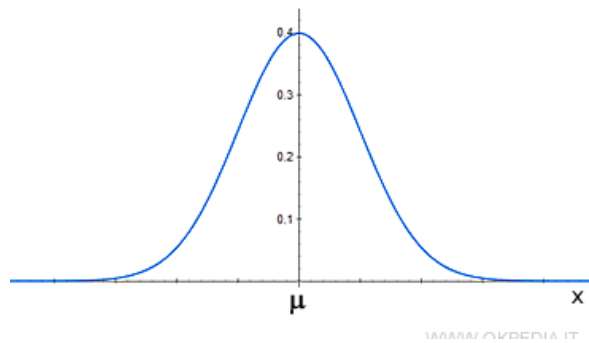
Si parte dal semplice concetto: un segnale forte maschera un segnale debole, ossia le armoniche a frequenza maggiore mascherano la ricezione di quelle a frequenza minore. Le aree di ciglia che si occupano della ricezione sonora si SOVRAPPONGONO (esempio [fatto con numeri casuali, serve solo a capire il mascheramento]: si ipotizzi che l'area di ciglia che riceve i suoni a 2000Hz intersechi per una parte l'area che riceve i suoni a frequenza di 1000Hz e intersechi dall'altra parte l'area che riceve i suoni a 3000Hz; si ipotizzi che arriva un suono che racchiude 2 armoniche, una a 3000Hz e un'altra a 2000Hz, il fenomeno del mascheramento fa sì che prima venga udita l'armonica a frequenza maggiore, ossia quella a 3000, e dopo 0.1 secondi quella a 2000, questa attesa temporale è data dal fatto che parte dell'area di ricezione delle frequenze a 2000Hz è impiegata per ascoltare quella da 3000, dato che le due aree si intersecano, quindi l'orecchio deve aspettare che l'area da 3000 si spenga per potere accendere quella per recepire la frequenza da 2000; poi nei fatti quest'onda da 2000 non verrà mai udita).

A livello fisico ciò che fa il mascheramento è proprio di nascondere il suono mascherato, quindi non lo si sente; occhio alle differenze: "non sentire" non coincide con "scompare", il suono continua ad esistere, solo che l'orecchio non lo percepisce a causa di questo fenomeno fisico!!!

Il mascheramento più banale che può essere fornito come esempio è che se una persona ti parla urlando e una bisbigliando, sentirai ovviamente la persona che urla invece di quella che bisbiglia; questo è un mascheramento dovuto all'intensità sonora.

L'orecchio ha un sistema di **risoluzione in frequenza**, questo parametro indica quanto sono ampie le aree di ciglia all'interno dell'orecchio. Maggiore è la risoluzione in frequenza, migliore è la capacità dell'orecchio di distinguere il suono perché esso, avendo più banda a disposizione, vede ridotte le zone di sovrapposizione fra le aree ciglia e quindi toni a frequenze diverse possono essere uditi contemporaneamente. -> questa caratteristica dell'orecchio è molto simile alla risoluzione in pixel delle immagini.

Banda critica: è la banda MINIMA contenuta all'interno del dominio di frequenze dello spettro di un suono tale che se la parte di spettro del suono in essa contenuto viene sovrapposto allo spettro di un altro suono (come ad esempio un rumore) si osserva un mascheramento DI TUTTO il suono originale.



GENERICA FUNZIONE GAUSSIANA

Per ricavare l'ampiezza della banda critica si va per tentativi, si parte coprendo aree di spettro molto piccole e, ingrandendole progressivamente, si cerca la banda minima per la quale il suono viene totalmente mascherato.

Rumore a banda stretta: è il rumore che se sovrapposto all'onda sonora in corrispondenza della banda critica, maschera per intero il suono.

Si introducano ora le diverse tipologie di mascheramenti; le due più importanti distinzioni fra i mascheramenti sono:

- Non tonali -> quando si maschera un suono mediante un rumore
- Tonalì -> quando si maschera un suono mediante un altro suono

-Mascheramento non tonale dato dal rumore bianco (a banda larga)

Bisogna intanto individuare gli elementi che interverranno in questa analisi: sia il "masking noise" (MN) il rumore che vuole mascherare un certo tono chiamato "test tone" (TT).

In questo primo tipo di mascheramento, il MN maschera abbastanza uniformemente i toni suoi bersagli. Preso un grafico in scala logaritmica, che riporta tutte le curve che mostrano le diverse relazioni fra frequenza ed intensità di ogni tono, si osserva che **qualsiasi** sia la frequenza assunta dal TT, se l'intensità del MT supera di almeno 20dB quella del tono da mascherare, avverrà il fenomeno del mascheramento (la frequenza del TT è nota vedendo quale punto dell'asse delle ordinate individua la curva del tono in corrispondenza della sua frequenza). Per contrastare il mascheramento, in risposta all'aumento di intensità di 10 dB da parte del rumore, il suono deve aumentare la sua intensità di altrettanti dB, il TT deve insomma impedire che la differenza in dB fra lui e il rumore che vuole mascherarlo non sia mai pari a 20dB (supponendo che si voglia fra in modo che il tono risulti essere ancora udibile).

ATTENZIONE: considera il test tone SEMPRE PURO!

-Mascheramento non tonale dato da un rumore a banda stretta

Gli spettri dei rumori a banda stretta sono delle funzioni continue "a campana"(ossia delle gaussiane). Le funzioni gaussiane sono caratterizzate dall'essere 0 ovunque, crescere in un certo intervallo di tempo, raggiungere un picco massimo di altezza e decrescere nuovamente fino a tornare a 0.

Osservando quindi lo spettro del rumore a banda stretta in un grafico che riporta sulle ordinate l'intensità sonora in dB e sulle ascisse le frequenze, indipendentemente dalla potenza sonora che questo rumore possiede, il picco massimo della funzione viene raggiunto in range di frequenze non troppo elevate (di solito il picco della gaussiana è sempre individuato in un intervallo compreso fra i 300 e i 1000 Hz).

Da questa analisi si capisce quindi che questo genere di mascheramento è ottimo per mascherare toni la cui banda critica è contenuta in un range di frequenze basse e, viceversa, è assai meno efficace a mascherare toni di alta frequenza, che conseguentemente avranno una banda critica per alti range di frequenze.

-Mascheramento tonale con un tono puro

Si mascherano i suoni (da assumere come toni complessi) mediante un tono puro specifico.

Lo spettro dei toni puri si ricordi essere un impulso, questo ha una sua intensità ed è individuato per una frequenza specifica.

Aumentando/diminuendo il valore dell'intensità del tono di mascheramento si è in grado di mascherare toni complessi la cui frequenza è più alta/bassa.

Per mascherare toni a basso livello di intensità sonora (fra 0 e 10 dB circa) il MT cresce in modo infinitesimo (esempio: supponendo che il MT abbia un'intensità pari a I , sufficiente per mascherare un tono la cui intensità è di 5dB, supponendo che l'intensità di questo tono cresca fino a 6 dB, basta che I vari di una quantità piccolissima per mascherare questo nuovo tono).

Più cresce l'intensità del suono da mascherare (uscendo dal range da 0 a 10 dB), maggiore è la variazione che si osserva nell'intensità del MT per perpetuare il fenomeno di mascheramento (il prof ha detto che il divario fra mascheramento tra basse e alte frequenze si fa più marcato).

(esempio: se si riprende lo stesso esempio di prima, se I è presa tale da mascherare un suono di intensità 20 dB e poi questo suono passa a 50 dB, I dovrà variare di un valore abbastanza concreto, cioè ci si accorge macroscopicamente della crescita di I).

Insomma, al crescere dell'intensità del TT, la crescita alla quale l'intensità del MT deve essere soggetta affinché il TT sia mascherato NON È LINEARE (come invece era stato visto essere per il mascheramento non tonale a banda larga).

-Mascheramento tonale con un tono complesso

Si mascherano i suoni (da assumere come toni complessi) mediante un tono complesso (cioè composto da più armoniche).

Maggiore è l'intensità del tono mascherante, più sarà semplice mascherare toni la cui frequenza è alta.

curva di Tuning: è una gaussiana ribaltata che rappresenta la robustezza al mascheramento da parte dei testing tone; da essa si evince che un TT: è più difficile da mascherare più per frequenze basse (parte sinistra del grafico); è più facile da mascherare per frequenze di mezzo (parte centrale del grafico); è estremamente semplice da mascherare per alti valori di frequenza (parte destra del grafico).

Bark era uno studioso che si interessò alla suddivisione in aree di frequenza del nostro orecchio; egli individuò 24 bande di frequenza.

Queste 24 bande furono sistemate in un grafico che riporta sulle ordinate una scala logaritmica delle frequenze e sull'asse delle ascisse i numeri da 1 a 24 (che individuano appunto le bande).

Ogni banda individua nel grafico un quadratino blu, questi simboleggiano il range di frequenze che quella particolare banda consente di farci percepire (esempio: la banda 1 si occupa della ricezione dei suoni da 20 a 100 Hz).

Attenzione però, la scala di frequenza del grafico è LOGARITMICA, quindi "area maggiore" non corrisponde a "range di frequenze maggiore" (infatti la banda che consente di ascoltare il maggiore numero di frequenze in assoluto è la 24esima che copre le frequenze da 10.000 a 20.000 Hz, non ci si deve lasciare ingannare da quanto è piccolo il quadratino che essa individua).

Questo grafico riassume perfettamente il motivo per il quale l'orecchio è più bravo a recepire suoni dalle frequenze più basse poiché i suoni le cui frequenze sono maggiori, coinvolgono aree di ciglia che si occupano della ricezione di tanti range di frequenza (per esempio la 24esima banda gestisce la ricezione di ogni frequenza da 10mila a 20mila Hz), il che rende i suoni composti da armoniche di frequenze elevate più soggetti al mascheramento.

Attenzione però: le scale identificate da Bark, NON sono tutte le bande esistenti in natura, esse individuano solamente le bande di frequenza mediante le quali il nostro orecchio lavora.

Grazie a Bark è possibile individuare uno schema ricorrente per l'individuazione delle bande critiche:

-Suoni la cui frequenza è inferiore a 500Hz: banda critica ampia circa 100Hz

-Suoni la cui frequenza è moderatamente elevata (fra i 500 e i 15mila Hz): banda critica ampia circa il 20% della banda di frequenza del suono

-Suoni la cui frequenza è elevata (superiori a 15mila Hz): banda critica ampia circa 6500Hz

La **soglia di mascheramento (S_m)** è il livello di intensità di un tono tale che questo sia ancora udibile nonostante nel sistema di ascolto sia presente anche un suono di mascheramento.

La quantità di energia mascherata è la quantità di energia mascherata del suono originale, è data dalla differenza:

20.000 - Sm

I mascheramenti di cui si è finora parlato sono detti mascheramenti simultanei poiché si è sempre supposto implicitamente che tono mascherante e tono mascherato fossero emessi nello stesso istante di tempo.

Un fenomeno che può avvenire è quello del mascheramento temporale, ossia l'emissione di un tono mascherante quando già da tempo nel sistema era stato emesso un testing tone.

Il mascheramento è uno strumento che viene sfruttato nella compressione dei segnali audio, esso fa sì che dal segnale (ossia l'audio veicolato) vengano eliminate le parti di esso che il nostro apparato uditivo non è in grado di percepire.