

Corso di Laurea in LOGOPEDIA

**FISICA ACUSTICA**

# **PERCEZIONE SONORA**

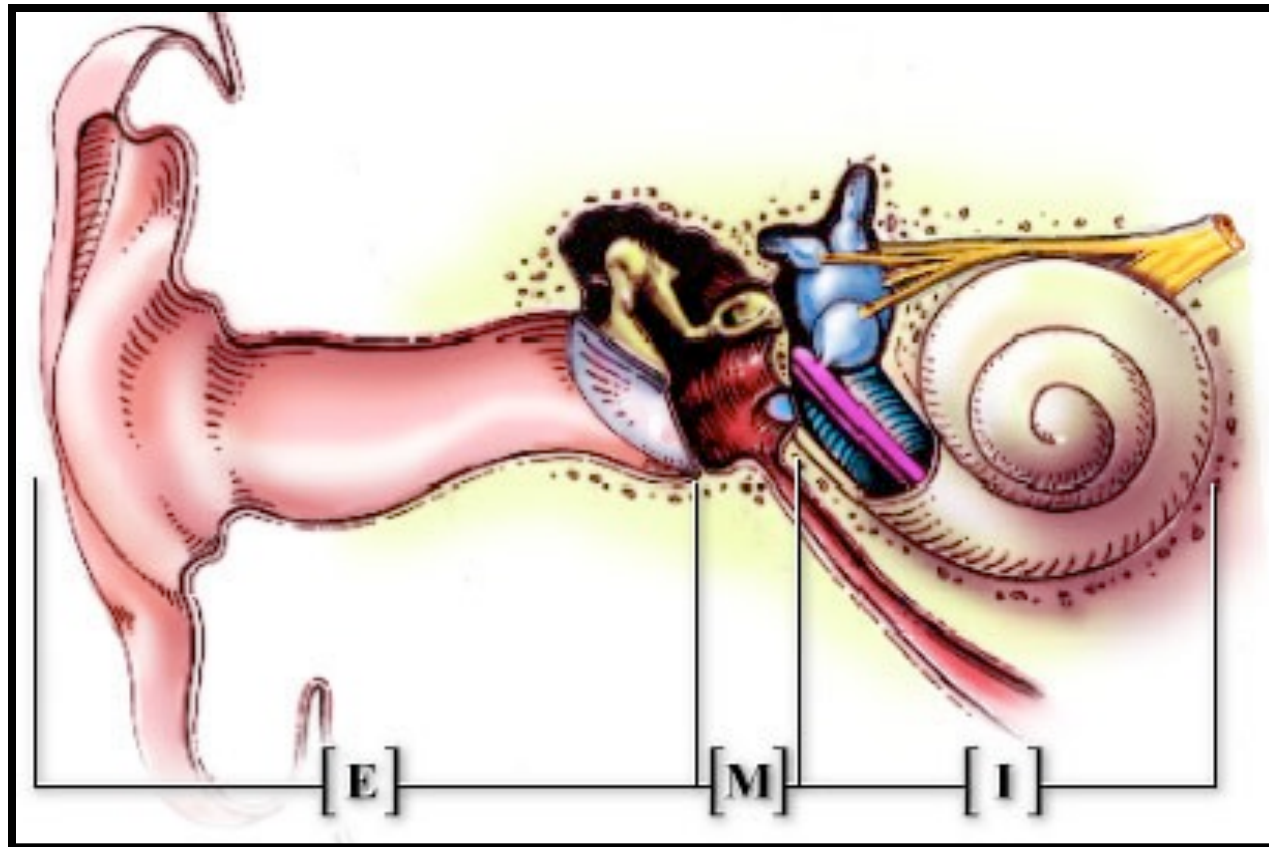
**Fabio Romanelli**

Department of Mathematics & Geosciences

University of Trieste

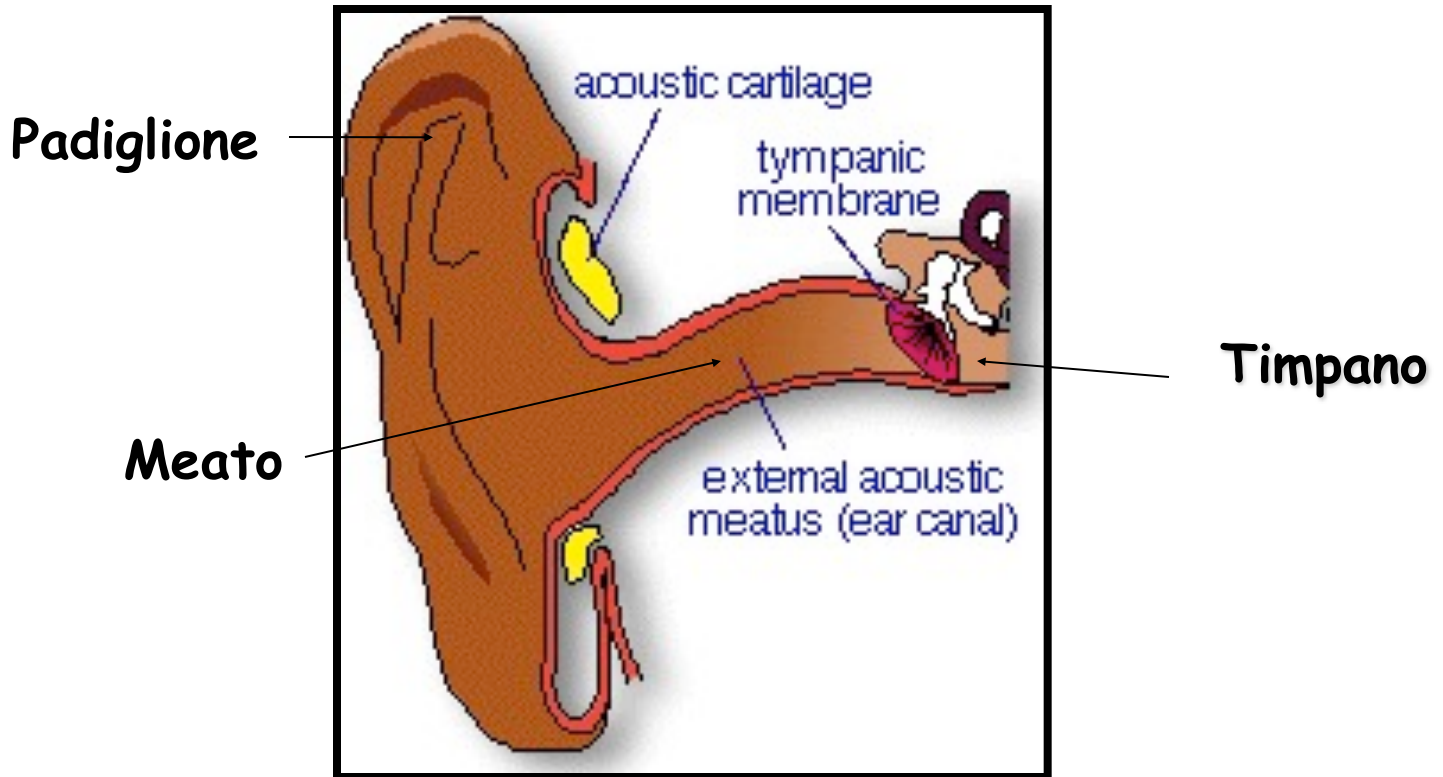
Email: [romanel@units.it](mailto:romanel@units.it)

# L'apparato uditivo umano



**Orecchio esterno (E):** padiglione, meato uditivo. **Orecchio medio (M):** timpano, ossicini, staffa. **Orecchio interno (I):** coclea.

# L'orecchio esterno



**Orecchio esterno:** è costituito dal **padiglione auricolare** e dal **meato acustico** esterno, una galleria dalle pareti lisce, che secerne un cerume protettivo. La sua funzione è quella di convogliare le onde sonore verso la sensibilissima membrana del **timpano**, che separa l'orecchio esterno dall'orecchio medio, facendolo vibrare piano o forte a seconda della velocità e forza dei suoni.



# L'orecchio medio



**Orecchio medio:** permette di trasformare le oscillazioni dell'aria in oscillazioni del fluido presente nell'orecchio interno. Esso è un vero e proprio trasformatore di energia meccanica: gli **ossicini** funzionano come una leva che aumenta la forza trasmessa dal timpano alla staffa; la staffa trasmette il movimento al fluido cocleare mediante un pistoncino membranoso che accede all'interno della coclea.

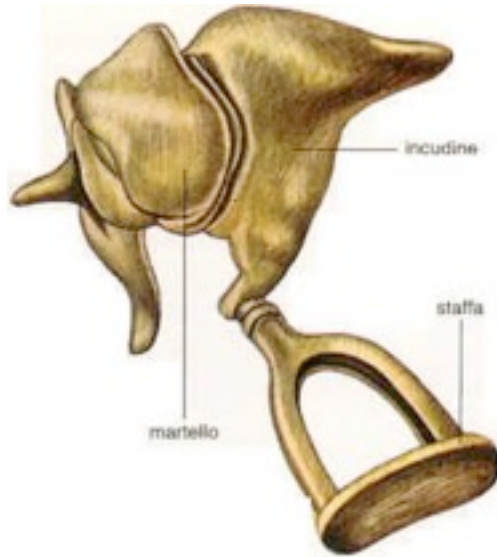
# Ossicini



Il "trasformatore" meccanico (timpano e catena di tre ossicini: martello, incudine e staffa), provvede a **tradurre le vibrazioni sonore raccolte dall'orecchio in variazioni di pressione del fluido cocleare**. Grazie al grande rapporto tra la superficie del timpano e quella della finestrella elastica su cui si appoggia la staffa (circa 35 nell'orecchio umano), all'effetto leva esercitato dal sistema degli ossicini (rapporto = 1.32 circa) e a un terzo effetto dovuto al rigonfiamento del timpano (guadagno = 4 circa), **la pressione applicata alla staffa, in condizioni statiche, risulta circa 185 volte maggiore di quella applicata al timpano**.

In condizioni dinamiche, l'orecchio medio **trasferisce meglio suoni di frequenze prossime a 1000 - 2000 hertz**, e via via meno bene quelli di bassa e alta frequenza, a causa della risonanza generata dalla cavità in cui hanno sede i componenti dell'orecchio medio.

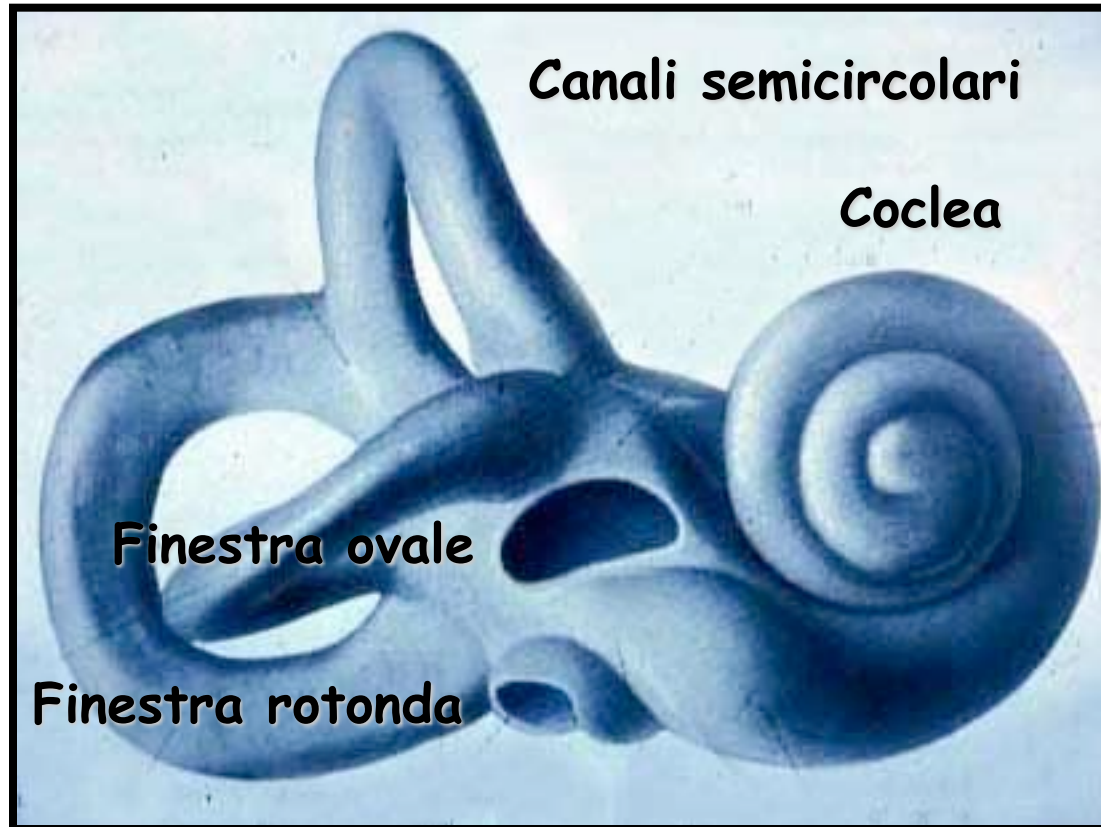
# Ossicini



I tre ossicini provvedono anche a **difendere l'orecchio dai rumori intensi**. Essi, infatti, grazie alla contrazione di due piccoli muscoli, possono modificare la tensione del timpano e quindi la quantità d'energia sonora che viene trasmessa dall'orecchio interno.

L'orecchio medio, comunica con la faringe attraverso la **tromba di Eustachio**, un canale di solito chiuso, ma che si apre a seguito di determinati movimenti muscolari della faringe, come quando si sbadiglia o si deglutisce. La sua funzione è quella di equilibrare la pressione esercitata sulla superficie interna del timpano con quella esercitata sulla superficie esterna.

# L'orecchio interno



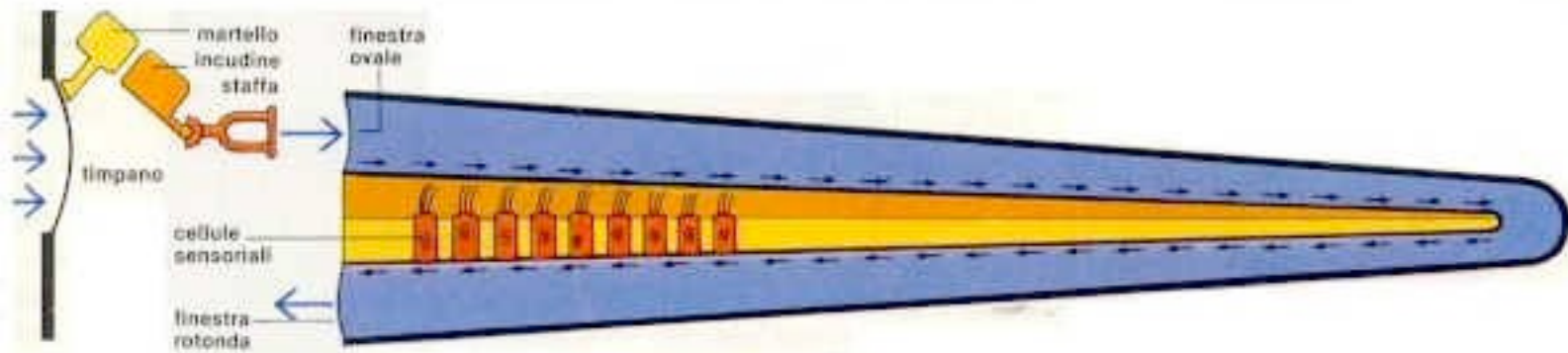
**Orecchio interno:** è la parte più complessa e delicata dell'apparato uditivo. Essa analizza i suoni, separando e facendone risaltare le componenti significative, smorzandone le componenti più intense e accettandone quelle più deboli, favorendo la rivelazione dei dettagli acustici che permettono di individuare le caratteristiche delle sorgenti sonore.

# La coclea



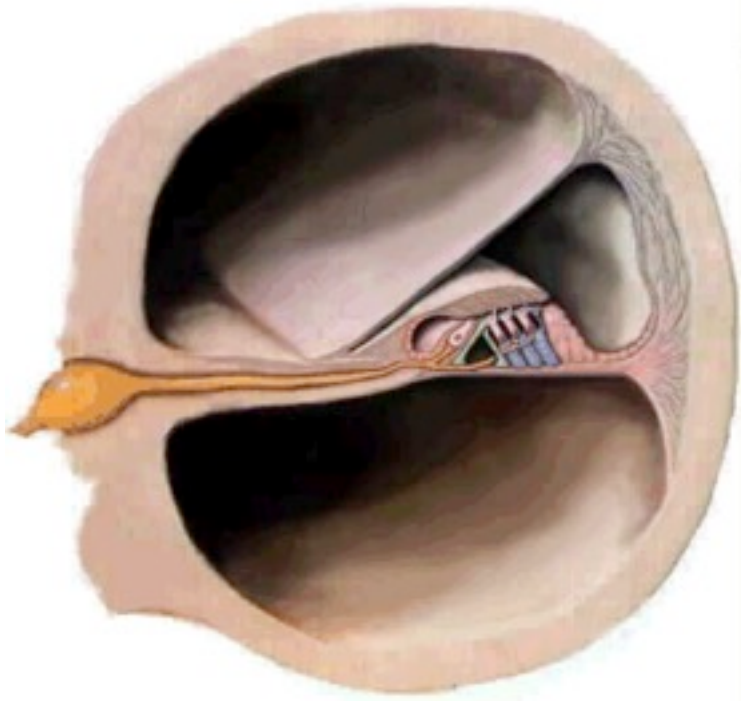
La coclea è interfacciata con l'ambiente esterno mediante due membranelle elastiche: la **finestra ovale**, sulla quale aderisce la **staffa** che trasmette le onde di pressione associate alla perturbazione sonora, e la **finestra rotonda** che permette al fluido (incompressibile) di muoversi anche nella regione inferiore del dotto cocleare, fungendo successivamente da dissipatore dell'energia associata alla perturbazione.





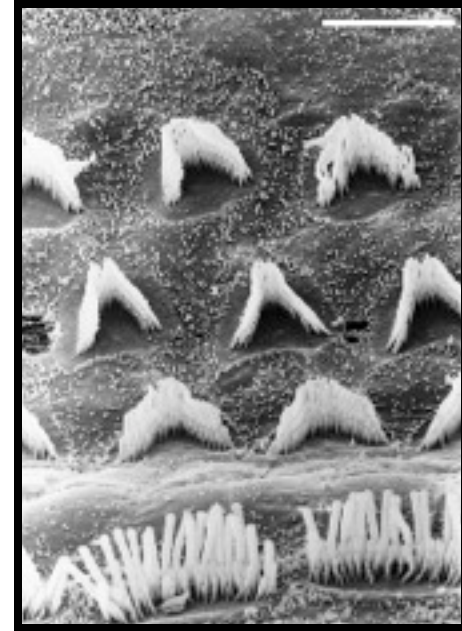
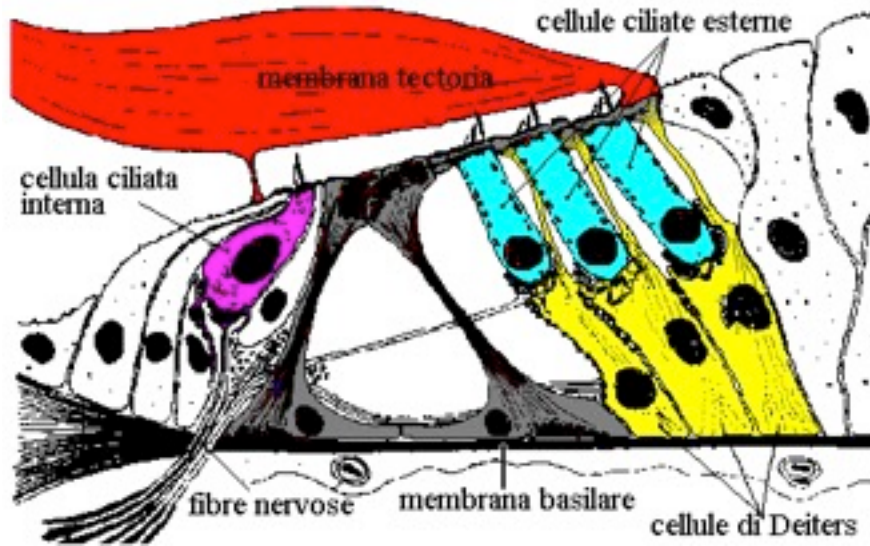
La cavità cocleare srotolata ha la forma di una tromba all'interno della quale sporgono diametralmente due creste ossee che sostengono la membrana basilare). Le due parti dei canali comunicano tra loro solo all'apice, dove la membrana si interrompe lasciando aperto un forame (**elicotrema**) che permette spostamenti di fluido tra le due parti.

L'immersione nel fluido garantisce non solo che le oscillazioni della staffa producano effetti pressoché istantanei sulla membrana basilare, ma anche che ogni porzione di membrana, che si trovi in uno stato di accelerazione, contribuisca a generare un campo di pressione che si estende pressoché istantaneamente su tutta la membrana.

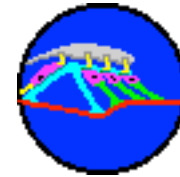
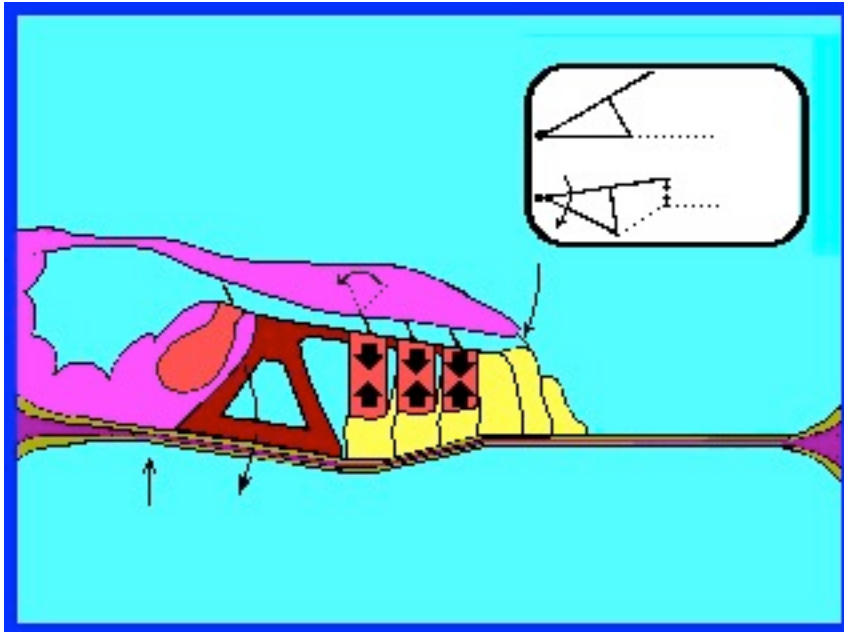


La figura mostra in dettaglio la struttura interna di una sezione cocleare. E' chiaramente riconoscibile l'**organo di Corti** montato sulla **membrana basilare** e la tripartizione del dotto cocleare. La cavità intermedia (**scala media**) è delimitata superiormente dalla membrana di Reissner e inferiormente dalla parte superiore dell'organo di Corti. La cavità superiore (**scala vestibuli**) e quella inferiore (**scala timpani**) comunicano tra loro e pertanto il liquido che le riempie (**perilinf**a) ha la stessa composizione. Il liquido segregato nella scala media (**endolinf**a) ha una composizione elettrochimica diversa dalla perilinf a e si trova a un potenziale elettrico maggiore di questa (circa 80 millivolt). Ciò permette a questo sistema di funzionare come una batteria elettrica che alimenta l'amplificatore elettromeccanico formato dalle triplette di cellule ciliate esterne e che attiva le cellule ciliate interne che costituiscono i sensori delle oscillazioni indotte dai segnali acustici .

# Organo di Corti



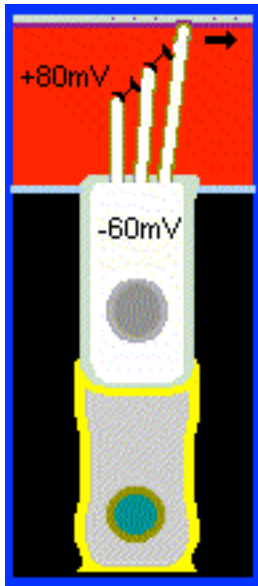
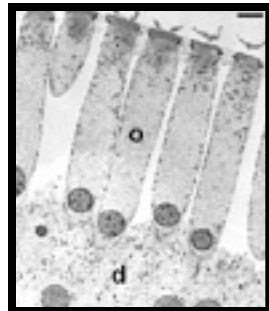
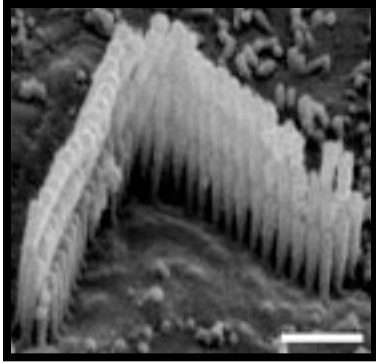
La figura illustra un modulo dell'organo di Corti in sezione. La coclea umana contiene circa 3000 moduli affiancati simili a questo. Ogni modulo ospita una tripletta di cellule motrici (colore ciano), che amplificano le oscillazioni della membrana basilare, e una cellula sensoriale (colore magenta) che trasmette i segnali al nervo acustico. Le cellule di Deiters (colore giallo) collegano le cellule ciliate esterne alla membrana basilare. Esse formano un cuscinetto viscoelastico che trasmette le forze esercitate dalle cellule ciliate.



L'amplificazione è dovuta all'azione motrice delle **cellule ciliate esterne** che si esercita quando vengono stimulate le stereocilia: il caratteristico fascio di tubicini rigidi installati sulla loro sommità cellulare. Quando le stereocilia vengono deflesse verso la regione esterna del dotto cocleare, cioè dalla parte in cui si trova la membrana basilare, le cellule ciliate esterne si contraggono determinando, per una sorta di effetto leva, la rotazione verso il basso dell'organo di Corti.

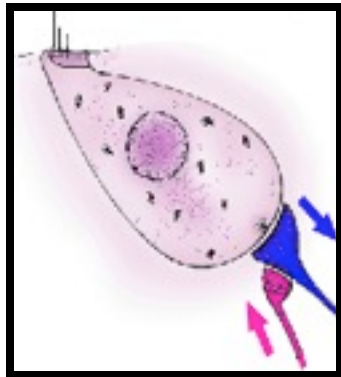
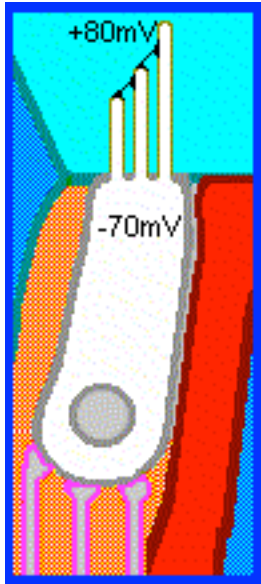
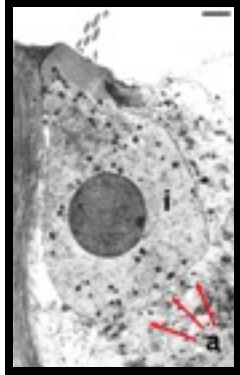


# Cellule ciliate esterne



Organizzate in triplette e localizzate nell'organo di Corti verso l'esterno rispetto all'asse della chiocciola, queste cellule sono i motori elettromeccanici che amplificano le oscillazioni della coclea. In questo modo le oscillazioni elastiche della membrana basilare tendono a persistere, come succede alle corde di una chitarra elettrica. La motilità delle cellule ciliate esterne è regolata dalle **stereocilia**: un fascio di tubicini rigidamente installati su uno zoccolo rigido posto alla sommità cellulare. Le stereocilia si trovano immerse nell'endolinfa, un fluido che possiede un potenziale elettrico di +80 millivolt rispetto a quello della perilinfia che circonda la parte inferiore della cellula. Sempre rispetto a questo, l'interno delle cellule possiede invece un potenziale negativo di -70 millivolt, mantenuto dalle potenti batterie a sodio e potassio delle membrane cellulari.

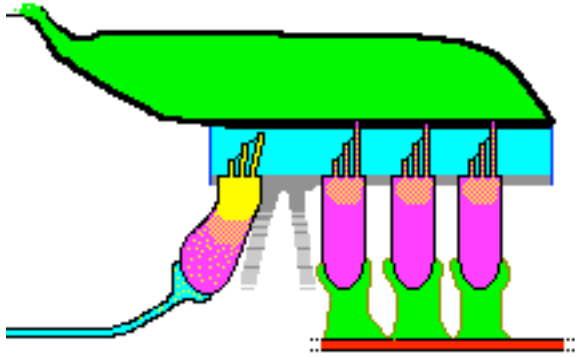
# Cellule ciliate interne



Localizzate nell' organo di Corti verso l'interno rispetto all'asse della chiocciola, queste cellule (non contrattili) sono i sensori che trasmettono i segnali al nervo acustico. Come nel caso delle cellule ciliate esterne, l'apertura dei canali stereociliari determina l'ingresso di una corrente elettrica che fa cadere il potenziale intracellulare. Ciò causa la stimolazione delle terminazione dendritiche, che si trovano alla base della cellula, di neuroni che hanno il corpo cellulare nel ganglio acustico localizzato in prossimità dell'asse cocleare. L'ampiezza del segnale trasmesso è proporzionale alla deflessione delle stereocilia, che funzionano nel modo descritto per le cellule ciliate esterne. Le stereocilia più alte delle cellule ciliate interne, a differenza da quelle delle cellule ciliate esterne, non sono fissate alla membrana tectoria e pertanto la loro deflessione non è proporzionale allo spostamento relativo della membrana tectoria.

# Membrana tectoria

La membrana tectoria è una struttura gelatinosa che sovrasta tutto l'organo di Corti. Essa poggia su una piastrina rigida nella quale sono infisse le punte delle stereocilia più alte delle cellule motrici. A causa della sua viscosità, essa adatta facilmente la sua forma alle lente variazioni di assetto dell'organo di Corti ma risponde rigidamente all'azione più veloce delle vibrazioni sonore. Poiché i fasci di stereocilia sono rigidi e hanno proprietà elastiche differenti in posizioni diverse, le porzioni di membrana tectoria oscillano in modi differenti in corrispondenza delle varie posizioni della sottostante membrana basilare (strisciolina rossa).



Essi formano così un sistema di oscillatori secondari capaci di risuonare alle frequenze caratteristiche delle onde viaggianti che si formano nelle diverse posizioni della membrana basilare. Queste "seconde risonanze" della coclea hanno una notevole importanza non solo perché contribuiscono a migliorare la selettività della coclea ma anche perché in condizioni di risonanza la membrana tectoria attiva le cellule motrici con un ritardo di un quarto del periodo di oscillazione della membrana basilare.

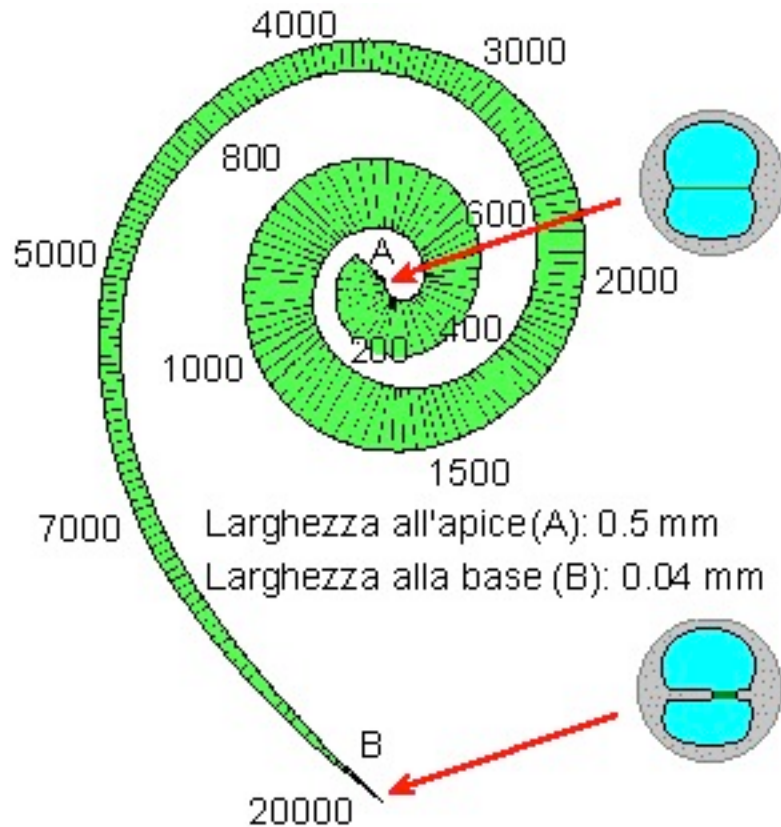


A causa dell'accentuata graduazione di elasticità della membrana basilare, punti diversi di questa rispondono alle variazioni di pressione in modi diversi. Il risultato è che la membrana e il fluido circostante si obbligano reciprocamente a oscillare in un modo complesso (linee chiare nella figura).

In particolare, uno stimolo acustico di data frequenza (**tono**) produce una risposta oscillatoria che presenta un pronunciato picco di ampiezza in una zona precisa della membrana e un ritardo di fase progressivo tale da creare l'illusione di un'onda che "viaggia" dalla base verso l'apice (mentre in realtà il profilo di massima ampiezza rimane fisso). Per questa ragione essa fu chiamata onda viaggiante dal suo scopritore von Békésy (1940).



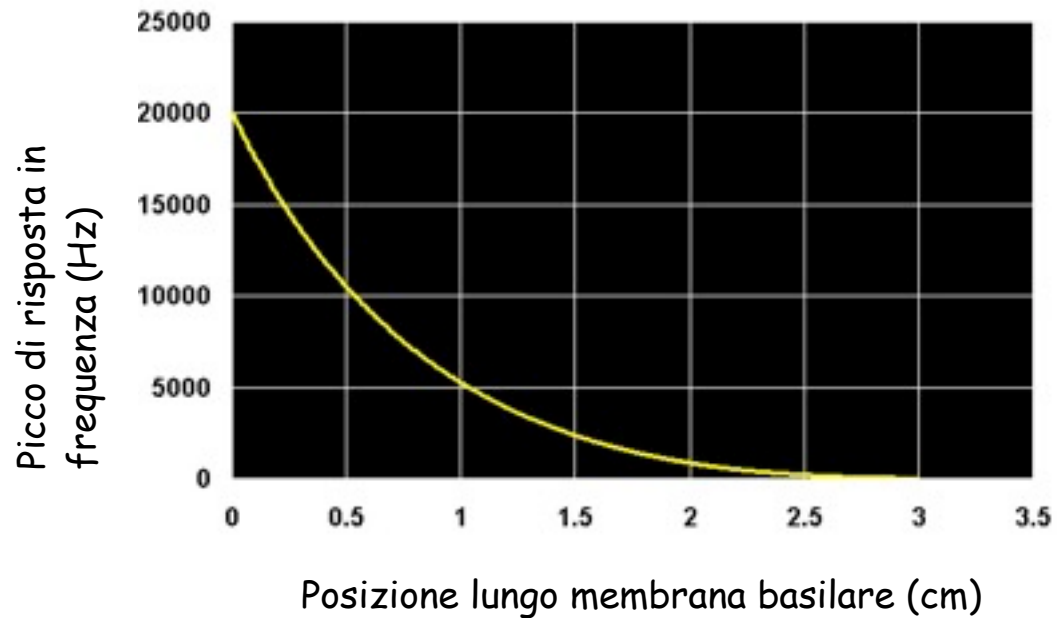
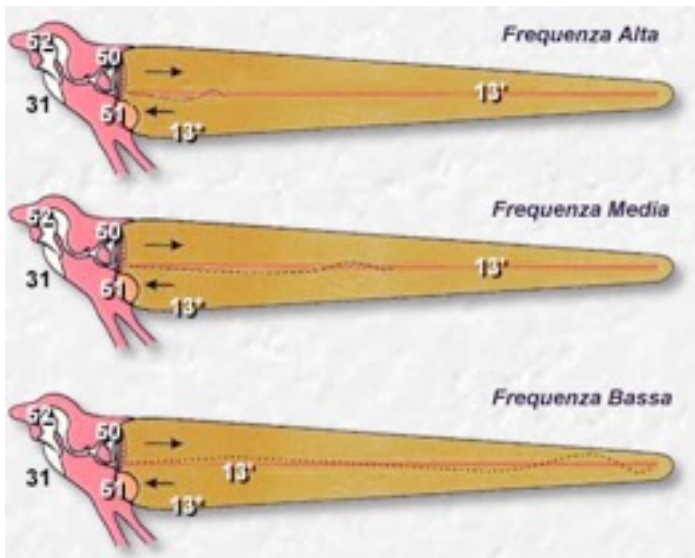
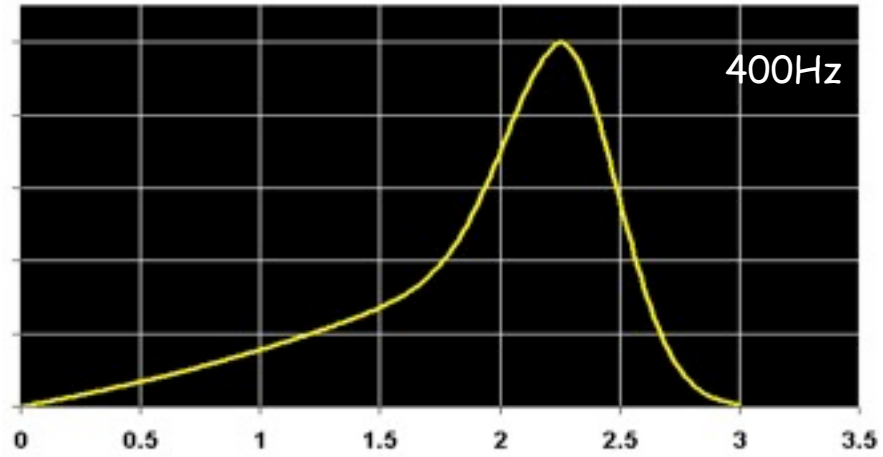
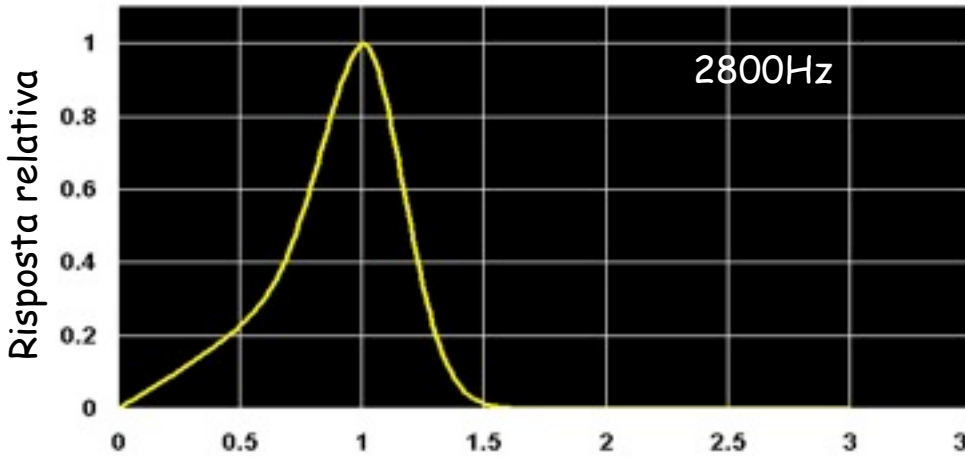
# Membrana basilare



La membrana basilare è formata da una moltitudine di sottili fibre elastiche tese diametralmente tra due creste ossee sporgenti verso l'interno del dotto cocleare. Fitte e corte verso la regione più periferica della coclea (**B**, base), le fibre si diradano e divengono più lunghe mano a mano che si procede verso la regione più interna, dove la chiocciola si arrotola più strettamente (**A**, apice). In tal modo, esse possono vibrare come le corde di uno strumento musicale. Grazie a questa struttura, la rigidità elastica della membrana risulta circa cinquantamila (!) volte più grande alla base che all'apice, diminuendo da un capo all'altro con legge approssimativamente esponenziale. Per questa ragione le oscillazioni della membrana basilare in risposta a suoni di frequenze diverse si formano in punti diversi.

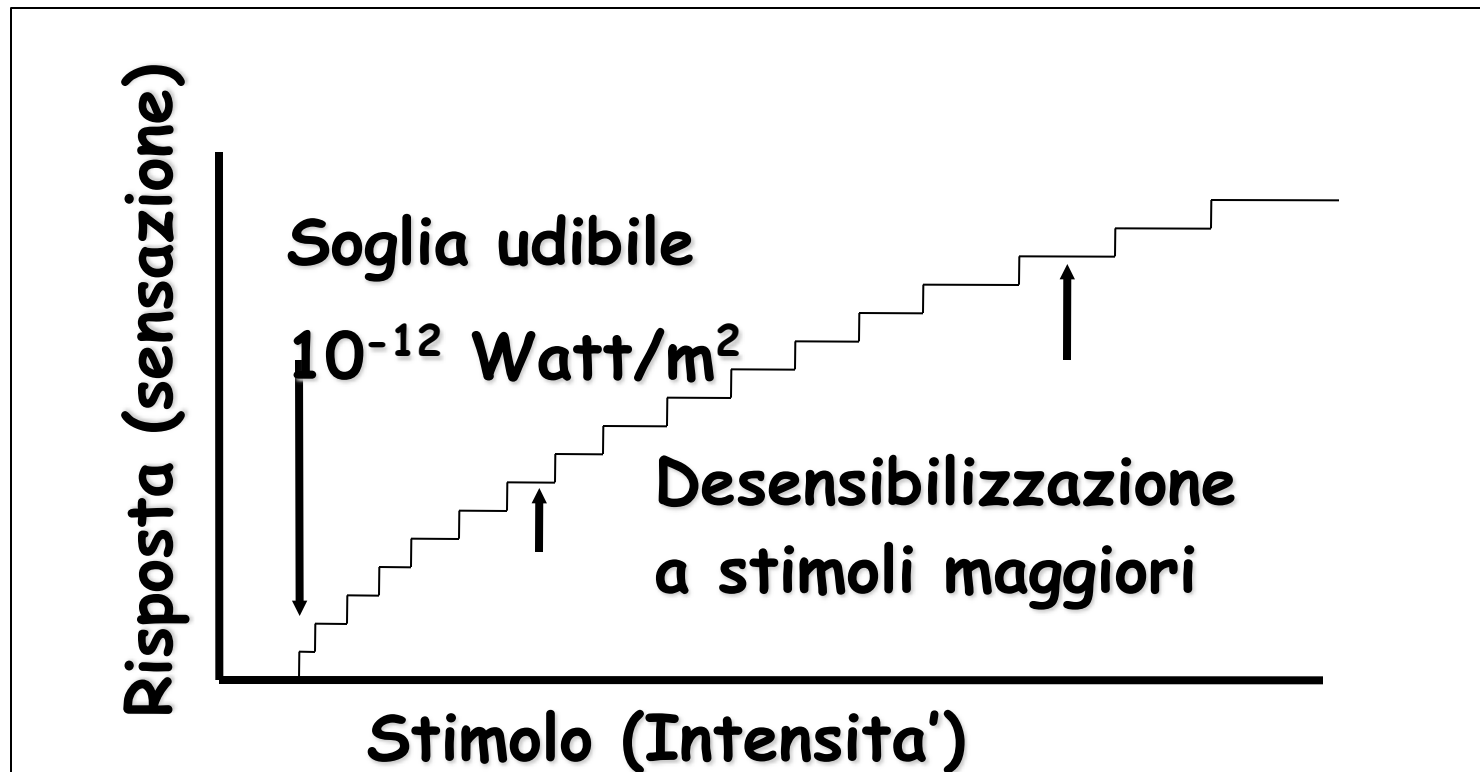
La figura illustra dove si localizzano i picchi di oscillazione in risposta a toni di varie frequenze: da circa 20000 cicli al secondo alla base **B** fino a poche decine di cicli al secondo all'apice **A**. Si noti che la frequenza associata a una certa posizione (**frequenza caratteristica** di quella posizione) diminuisce anch'essa dalla base verso l'apice con legge approssimativamente esponenziale. La risposta a un tono di 1500 cicli al secondo ha luogo a metà circa della membrana basilare.

# Risposta in frequenza

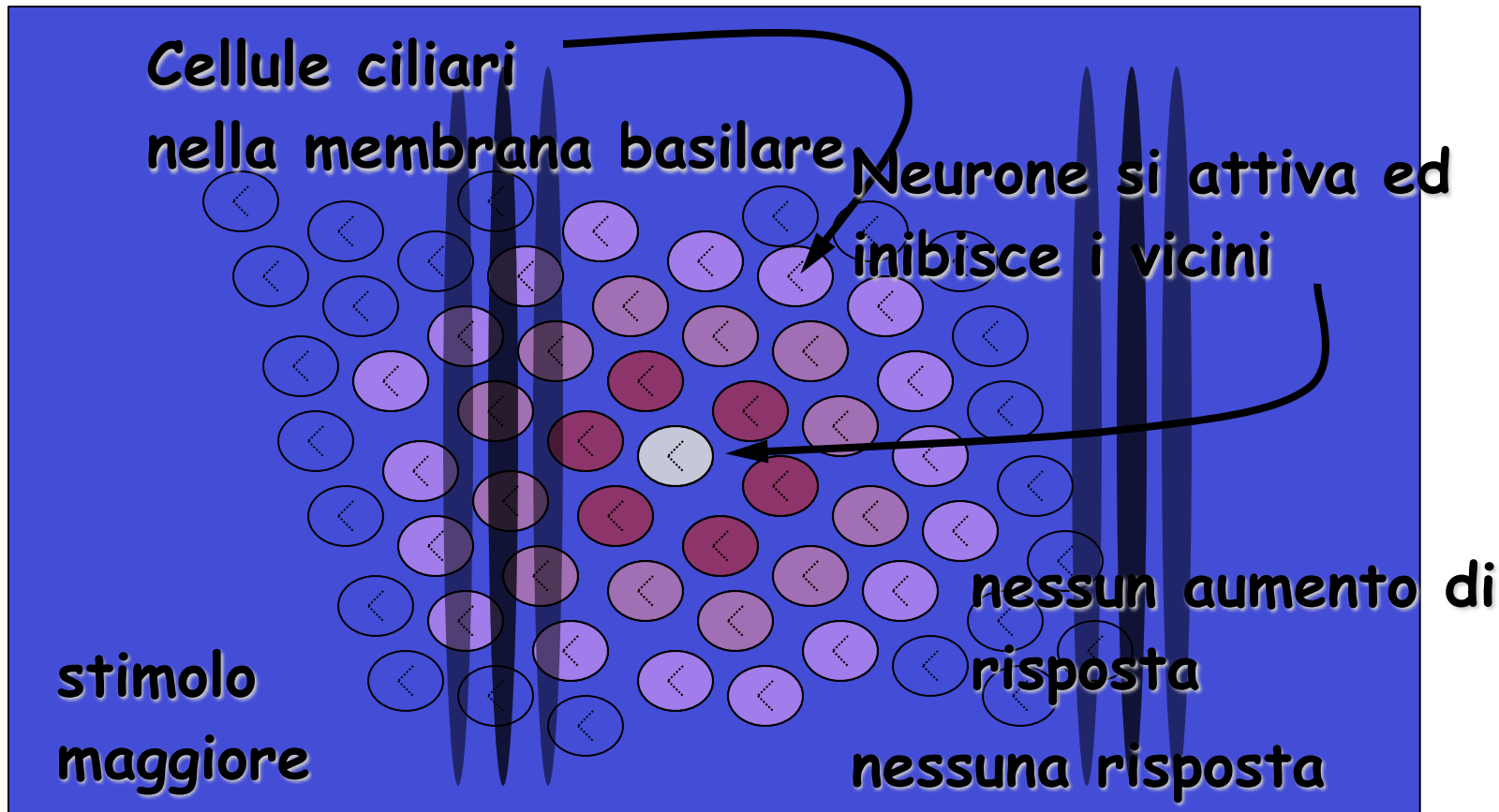


# Risposta neuronale

Le cellule ciliari si comportano come neuroni, idealizzabili come una rete di interruttori. Vi deve essere una **soglia di intensità** affinché si attivino ed inoltre l'attivazione può inibire i loro vicini.

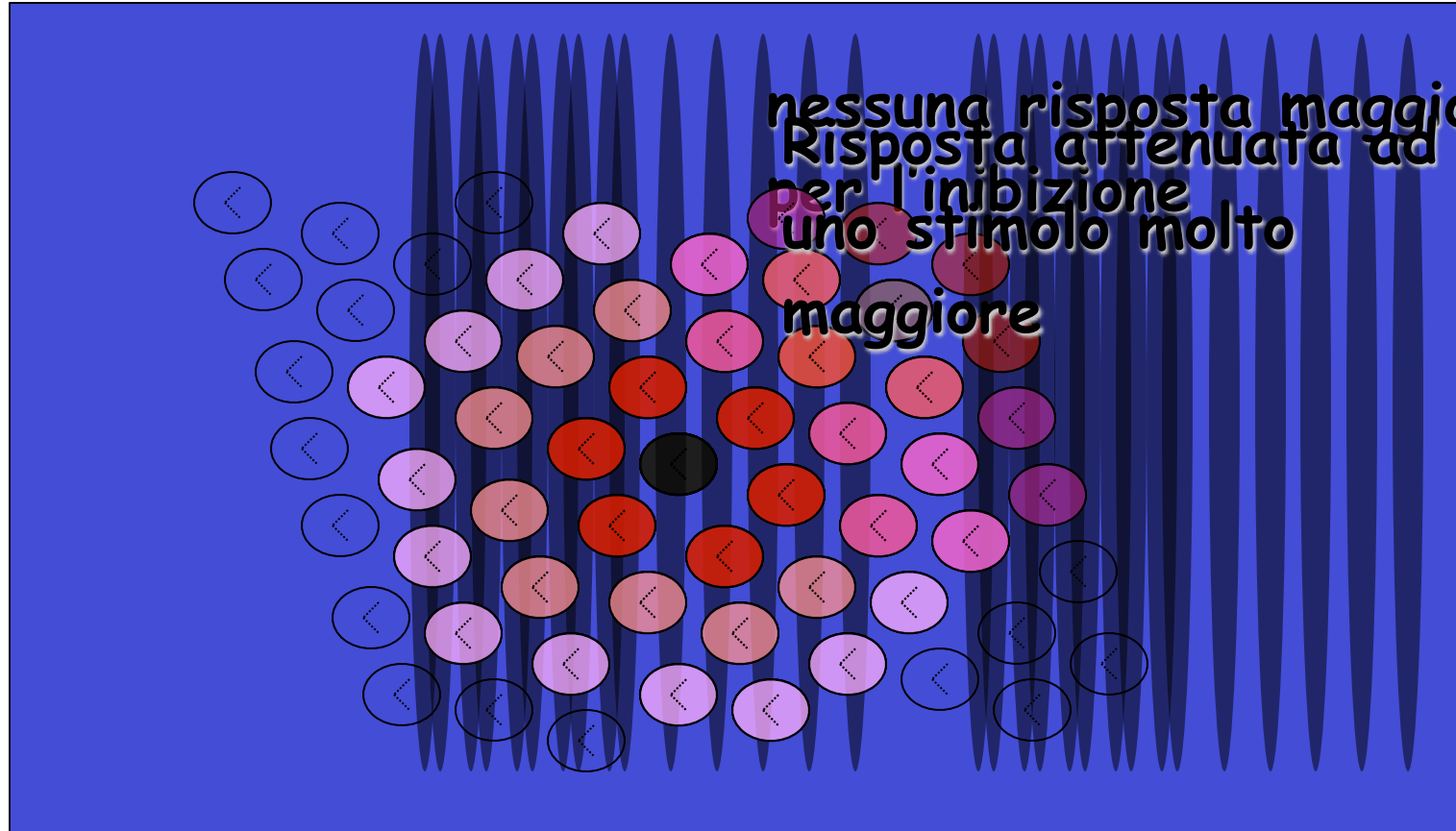


## Risposta neuronale delle cellule ciliari

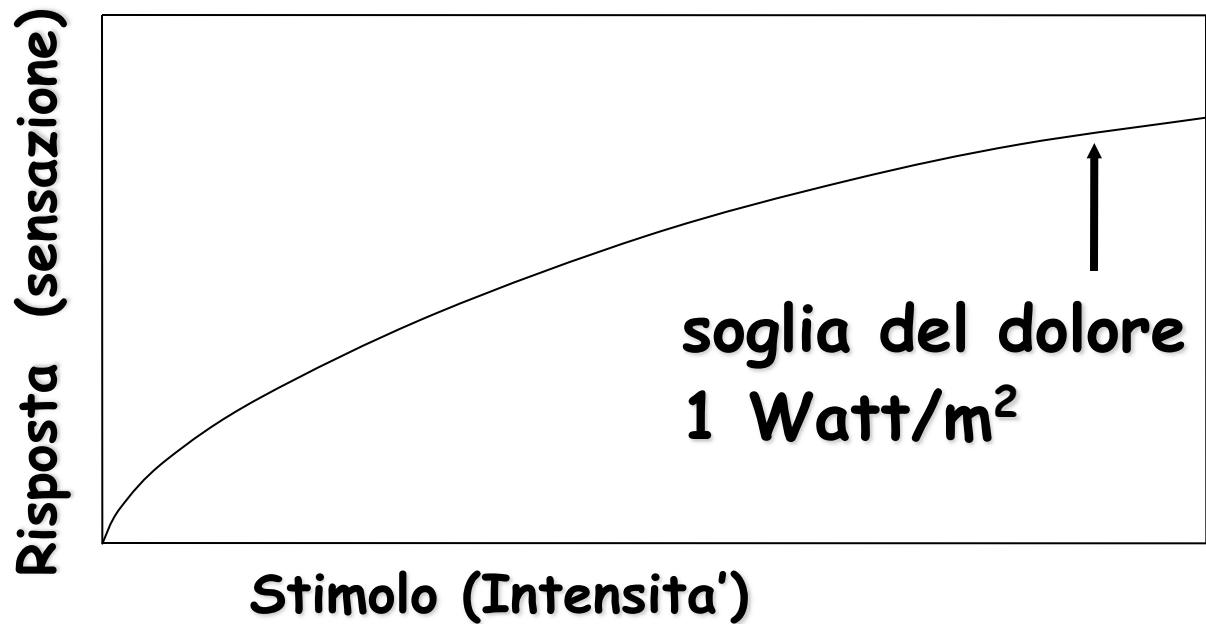




## Risposta neuronale delle cellule ciliari

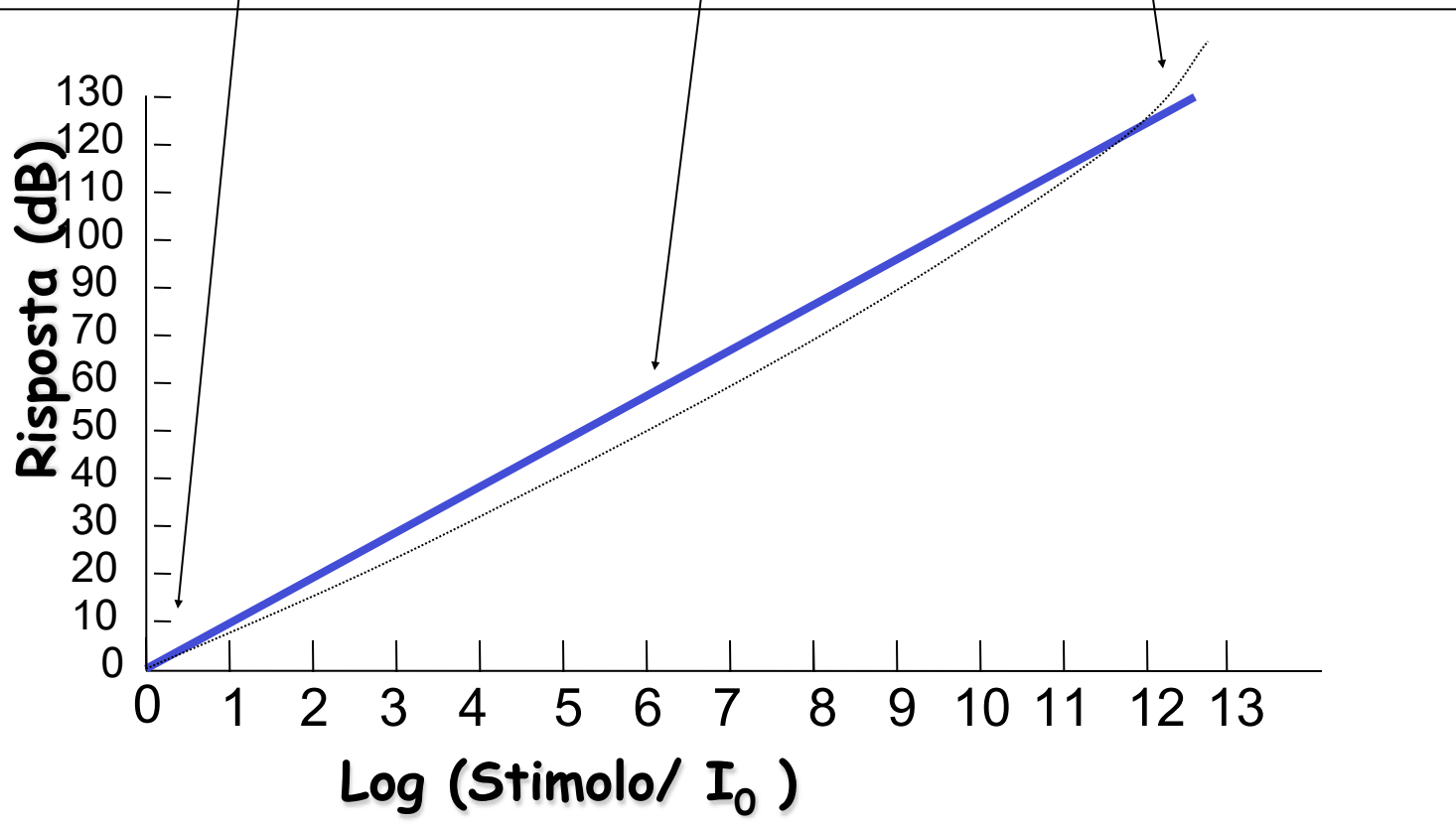


Conseguenza: **Soglia e risposta non lineare**



Logaritmo dello Stimolo in funzione della risposta

$$I_0 = 10^{-12} \text{ W/m}^2 \quad 10 \text{ Log} \left( \frac{I}{I_0} \right) \quad I_0 = 1 \text{ W/m}^2$$



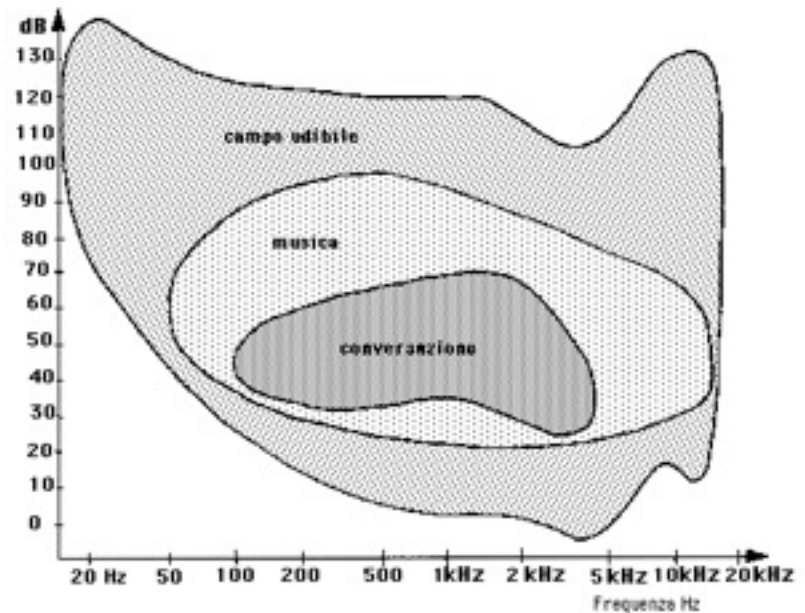
# Livello di Intensità Sonora

Ecco perchè si era definito il **livello di intensità sonora** (LIS) come:

$$I_{\text{dB}} = 10 \log \left( \frac{I}{I_0} \right) \approx 20 \log \left( \frac{p}{p_0} \right) = p_{\text{dB}}$$

dove  $I$  è l'intensità del suono,  $I_0$  è la soglia di udibilità ( $\sim 10^{-12}$   $\text{W m}^{-2}$ ) ed è misurata in decibel (dB).

Il **livello di pressione sonora** coincide con esso in caso di fronti d'onda sferici.





Pratica:

LIS:

$$10 \text{ Log}( 1 ) = 0 \text{ dB}$$

$$10 \text{ Log}( 10 ) = 10 \text{ dB}$$

$$10 \text{ Log}( 100 ) = 20 \text{ dB}$$

$$10 \text{ Log}( 2 ) = 3.0 \text{ dB}$$

$$10 \text{ Log}( 20 ) = 13.0 \text{ dB}$$

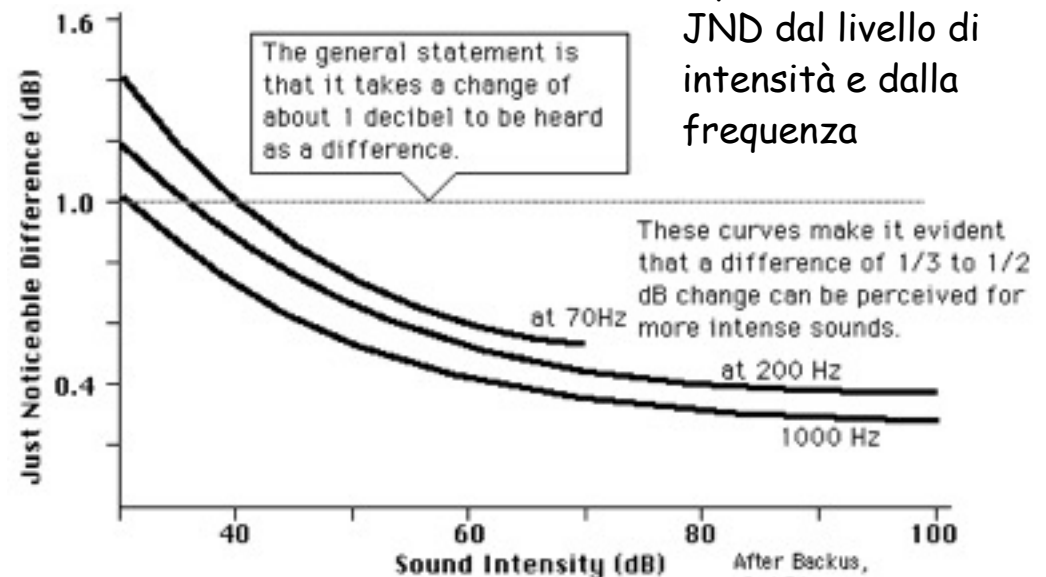
$$10 \text{ Log}( 200 ) = 23.0 \text{ dB}$$

Si supponga di eseguire un test alternativo (vero o falso). In caso di risposta casuale la probabilità di rispondere esattamente è 50%; in caso di conoscenza della risposta è 100%. Il valore di soglia che permette di avere il 75% di risposte esatte è definito JND.

Supponiamo di ascoltare due note con timbro e frequenza uguali, ma leggermente diverse come ampiezza. Man mano che ascoltiamo le due note, rendiamo sempre più piccola la differenza fra le due ampiezze. Arriveremo ad un punto in cui, nonostante ci sia ancora una piccola differenza di ampiezza, l'orecchio sentirà le due note esattamente allo stesso volume.

Si definisce allora la **Just Noticeable Difference** (JND) o "minima differenza percepibile", in acustica, come il rapporto fra le intensità di due suoni pari a 1.26, che corrisponde ad 1 decibel:

$$10 \log(1.26) = 1$$



dependenza della JND dal livello di intensità e dalla frequenza

# Sensazione dell'intensità: SONORITA'

---

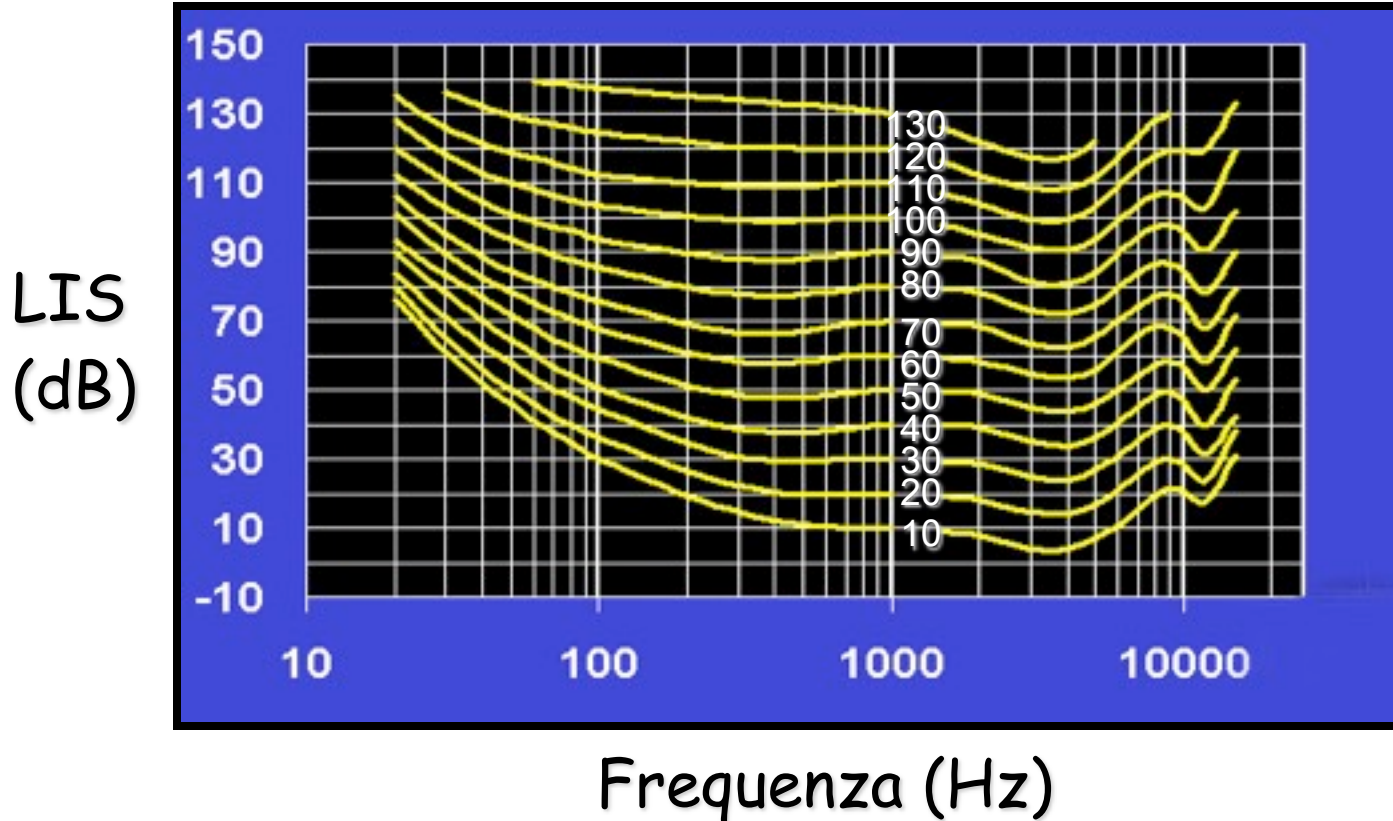
**Legge di Fechner** (1860): se lo stimolo cresce per moltiplicazione la risposta lo fa per addizione. Cioè fra sensazione percettiva ed intensità sonora vi è una relazione **logaritmica**.

**In realtà dipende dal livello di intensità e dalla frequenza!!**

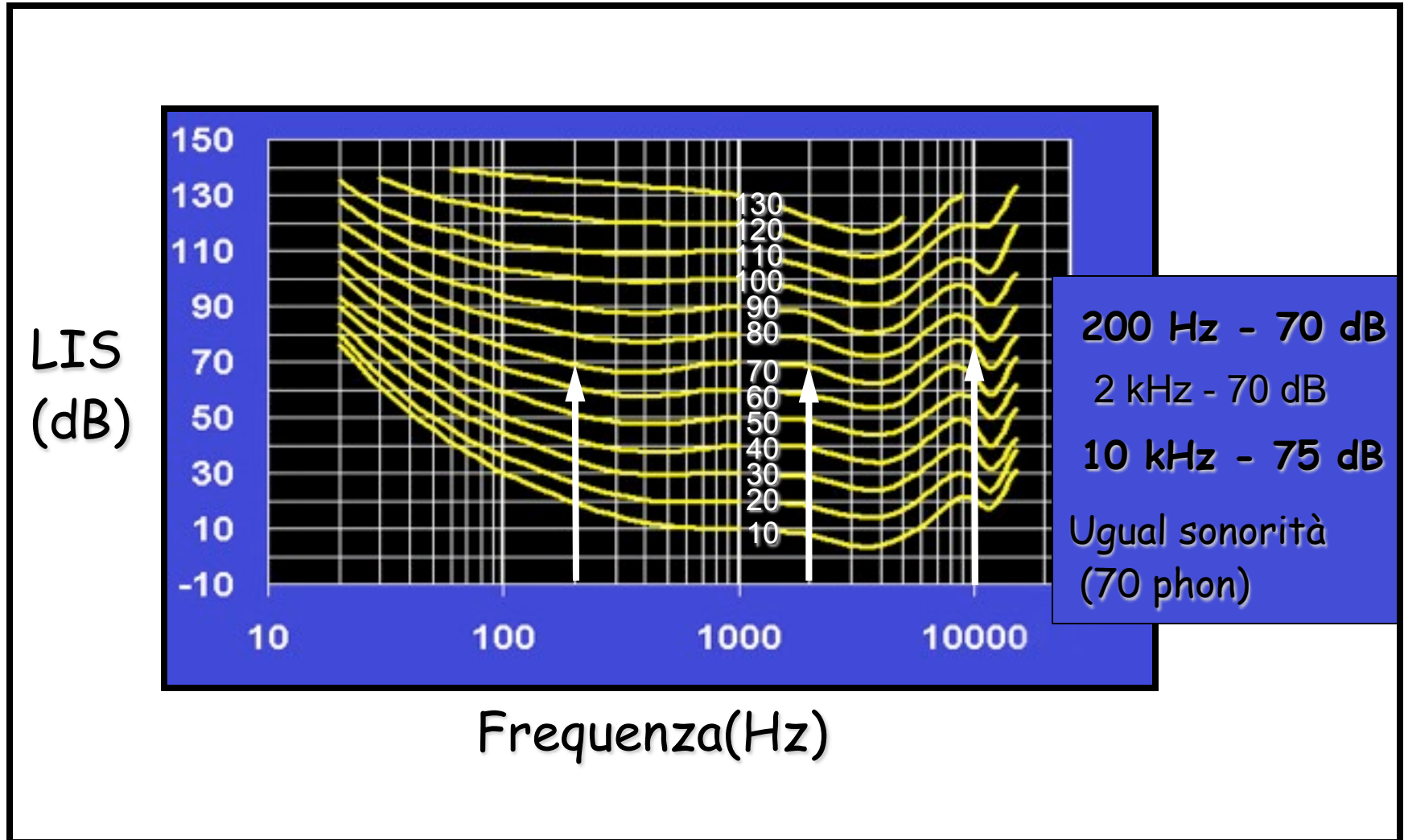
**Stevens** (1933) tentò di definire una grandezza che descrivesse la **sensazione soggettiva di sonorità (loudness)**: Scala di sonorità, la cui unità di misura è

# Curve di Fletcher-Munson

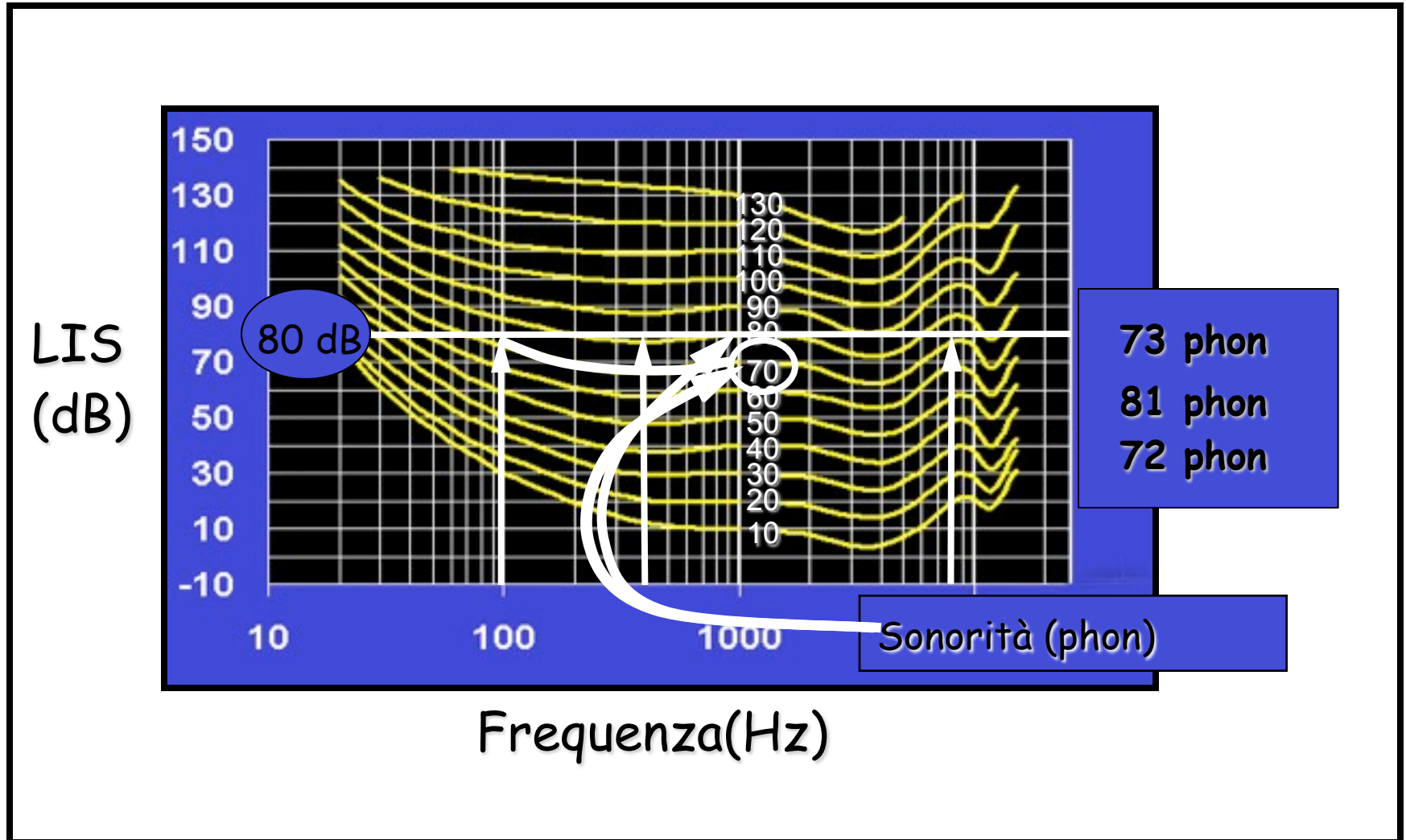
Curve di Fletcher-Munson (1933)



Si determini il LIS necessario per avere un livello di 70 phon a 200 Hz, 2000 Hz e 10,000 Hz.



Si determini il livello di sonorità di un tono di 80 dB a 100, 400 e 8000 Hz.



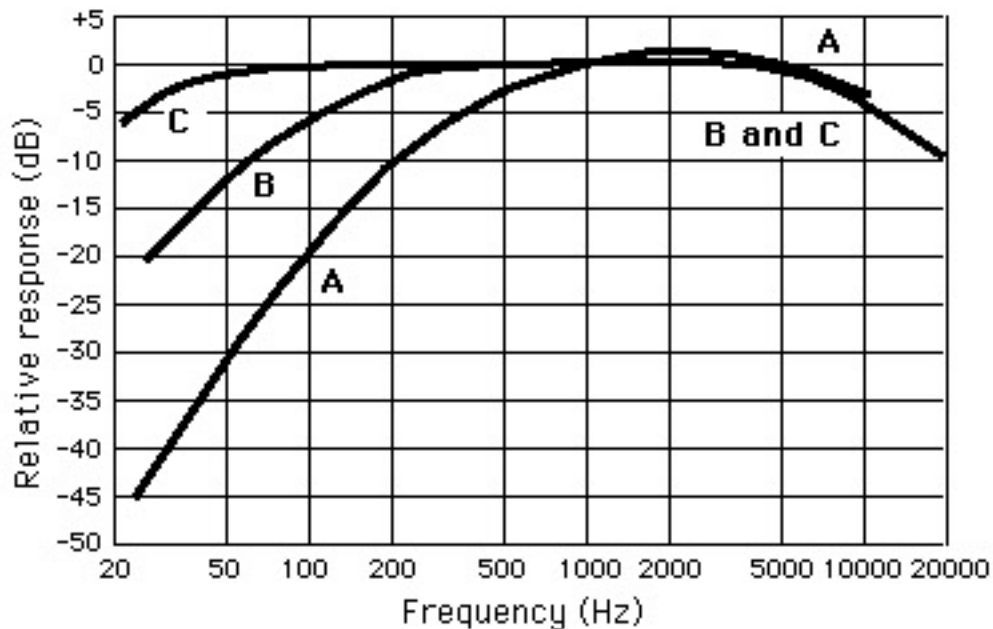


## Dinamica musicale

| Livello dinamico |     | $W/m^2$   | son | phon |    |
|------------------|-----|-----------|-----|------|----|
| Pianissimo:      | pp  | $10^{-8}$ | 1   | 50   |    |
| Piano:           | p   | $10^{-7}$ | 2   | 60   |    |
| Mezzopiano:      | mp  | $10^{-6}$ | 4   | 66   |    |
| Mezzoforte:      | mf  | $10^{-5}$ | 8   | 76   |    |
| Forte:           | f   | $10^{-4}$ |     | 16   | 80 |
| Fortissimo:      | ff  | $10^{-3}$ | 32  | 90   |    |
| Fortississimo:   | fff | $10^{-2}$ | 64  | 100  |    |

# Fonometri

Il fonometro ideale dovrebbe misurare il livello sonoro percepito da un soggetto medio. In realtà, visto il gran numero di fattori in gioco si ricorre a microfoni a cui viene applicato un filtro che simuli la risposta tipica di un orecchio umano. Le diverse curve si applicano a suoni di intensità diversa. La più comunemente usata è la curva A, per suoni attorno a 40 dB (pp) ed il risultato si misura in dBA.

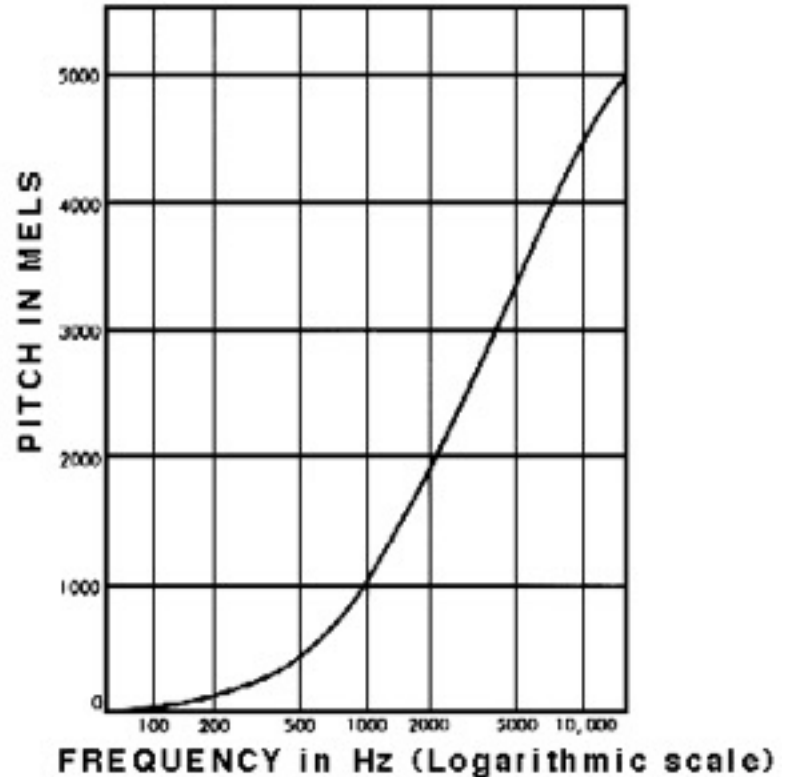


# Sensazione della frequenza: ALTEZZA

La percezione della frequenza di un'onda sonora (**altezza o pitch**) è dipendente dalla risposta dell'orecchio che la percepisce.

La percezione assoluta (fuori da un contesto) dell'altezza di un suono è molto rara (0.01%), mentre la percezione relativa (e.g. intervallo musicale) è piuttosto comune.

Inoltre vi è una corrispondenza **non lineare**: definizione di **mel**.



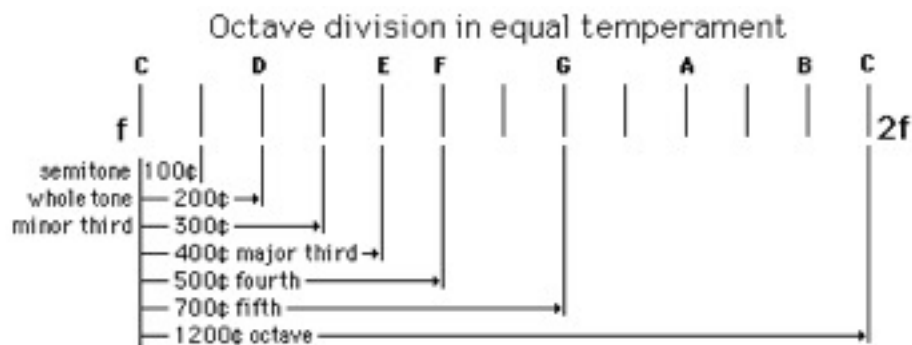
# Intervalli in altezza

Un ugual rapporto di frequenze viene percepito come un' ugual differenza (intervallo) d'altezze.

Un **OTTAVA** è l'intervallo di altezze corrispondente ad un rapporto in frequenza pari a 2/1.

Un **SEMITONO** è 1/12 di un'ottava.

Un **CENT** (¢) è 1/100 di un semitono o 1/1200 di un'ottava. 1 ottava = 1200 ¢



Per trovare un intervallo d'altezze in cent:

$$\frac{f_2}{f_1} = 2^{\frac{\text{¢}}{1200}}$$

$$\text{¢} = \frac{1200}{\log(2)} \log\left(\frac{f_2}{f_1}\right)$$

# JND in altezza

Si supponga di eseguire un test alternativo (vero o falso). In caso di risposta casuale la probabilità di rispondere esattamente è 50%; in caso di conoscenza della risposta è 100%. Il valore di soglia che permette di avere il 75% di risposte esatte è definito JND.

Supponiamo di ascoltare due note con timbro e intensità uguali, ma leggermente diverse come frequenza. Man mano che ascoltiamo le due note, rendiamo sempre più piccola la differenza fra le due frequenze. Arriveremo ad un punto in cui, nonostante ci sia ancora una piccola differenza di frequenza, l'orecchio sentirà le due note esattamente con la stessa altezza.

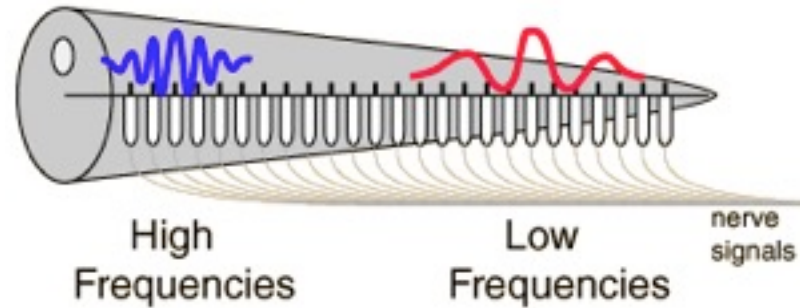
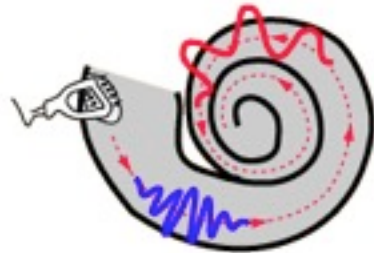
Si è trovato che la **Just Noticeable Difference** (JND) o "minima differenza percepibile", in altezza è il rapporto fra le **frequenze** di due suoni corrispondenti circa a **5¢**. In realtà vari esperimenti hanno mostrato che la JND dipende da: frequenza, livello sonoro, durata etc. In media è **8¢**.

Applicando tale valore al range udibile umano, si ottiene che il numero di frequenze discriminato è:

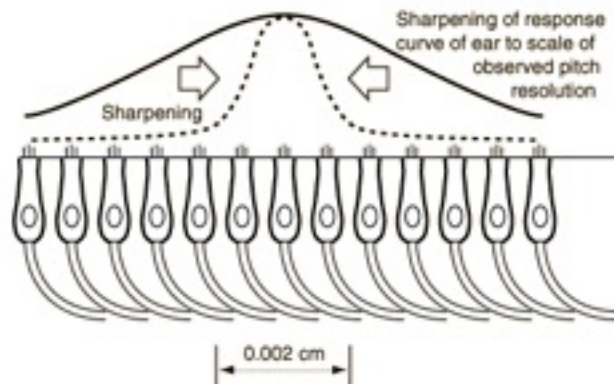
$$\frac{\frac{1200}{\log(2)} \log\left(\frac{20000}{20}\right)}{8} = \frac{3600}{8\log(2)} \approx 1500$$

# Potere risolutivo e teoria del posizionamento

Quindi l'orecchio umano è in grado di risolvere 1500 altezze con 16000-20000 cellule ciliate in 3.5 cm, cioè un'altezza ogni 0.002 cm!



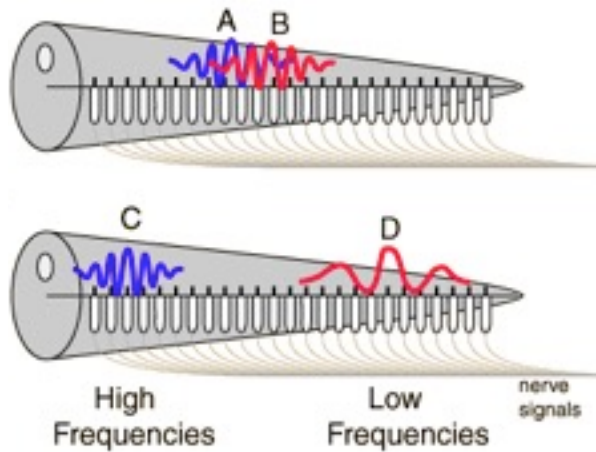
Secondo la teoria del posizionamento, questo alto potere risolutivo implica che solo 4 gruppi di cellule vengono impiegate. Quindi che ci deve essere un meccanismo che rende la risposta dell'organo del Corti più "piccata" alle sollecitazioni della membrana basilare.



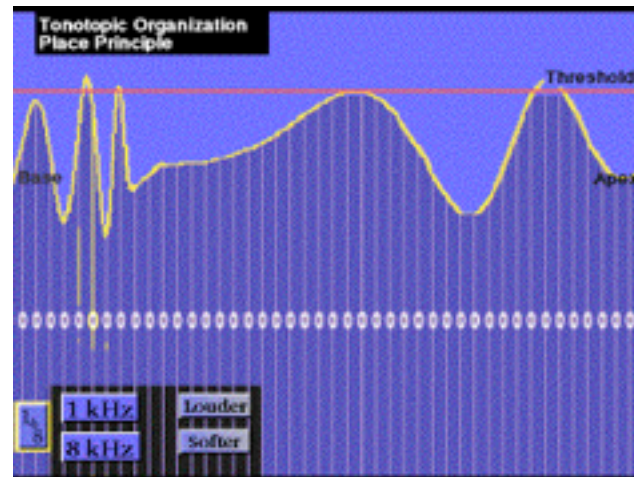
Un meccanismo che può spiegare questo fenomeno è l'inibizione delle cellule nervose vicine



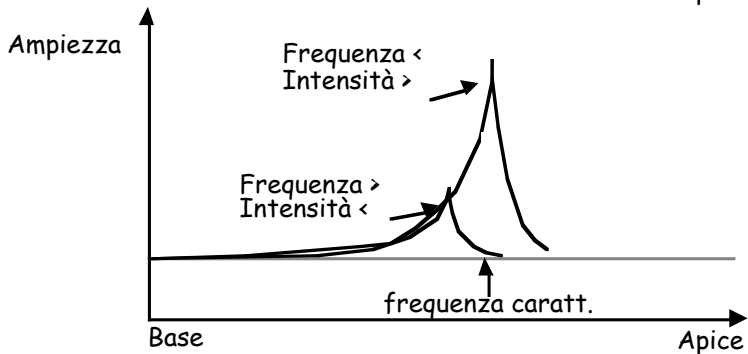
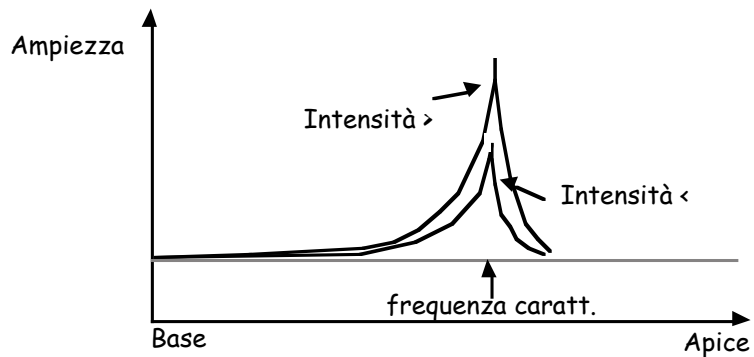
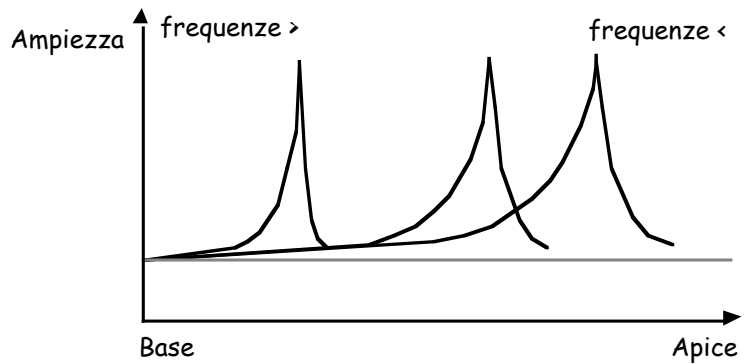
# Altezza e sonorità



Si supponga di sommare un suono ad un altro. La sensazione di sonorità totale dipende dalla separazione in frequenza. C+D avrà una sonorità maggiore, visto che interessa parti diverse della membrana basilare. Si dice che sono fuori dalla **banda critica**, cioè l'intervallo di frequenze corrispondente al settore della membrana basilare, e quindi al numero di cellule nervose, interessato da un suono con una certa altezza.

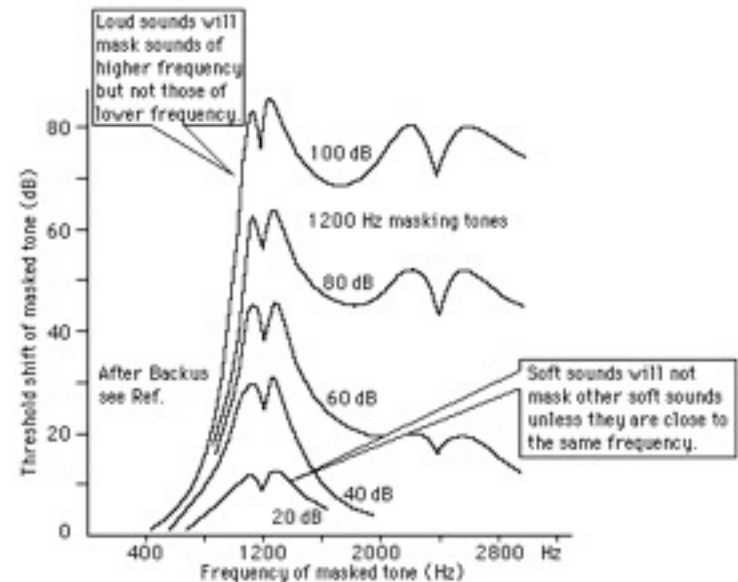


# Mascheramento



A causa dell'asimmetria degli impulsi nervosi che viaggiano sulla membrana basilare, impulsi a bassa frequenza possono mascherare quelli a frequenza più alta

Mascheramento di un suono di 1200 Hz a varie intensità



# Mascheramento

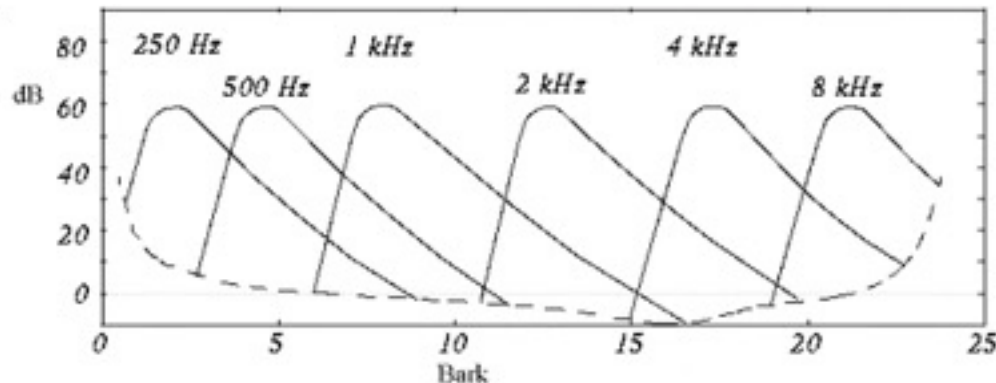
Il mascheramento dipende dalla larghezza della banda critica per le varie frequenze. Si definisce

**1bark = larghezza banda critica**

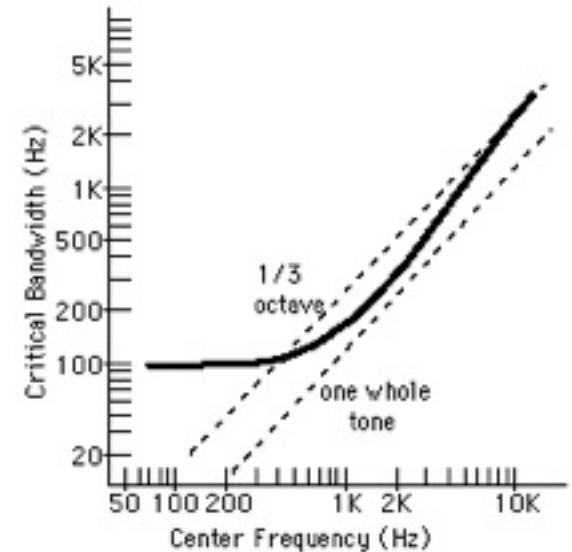
$1\text{bark} \sim f/100$  (per  $f < 500$  Hz)

$1\text{bark} \sim 9 + 4\log(f/100)$  (per  $f > 500$  Hz)

Il mascheramento dipende anche dalla frequenza: nel caso si considerino suoni di 60dB si ottiene il seguente grafico per le curve di mascheramento:

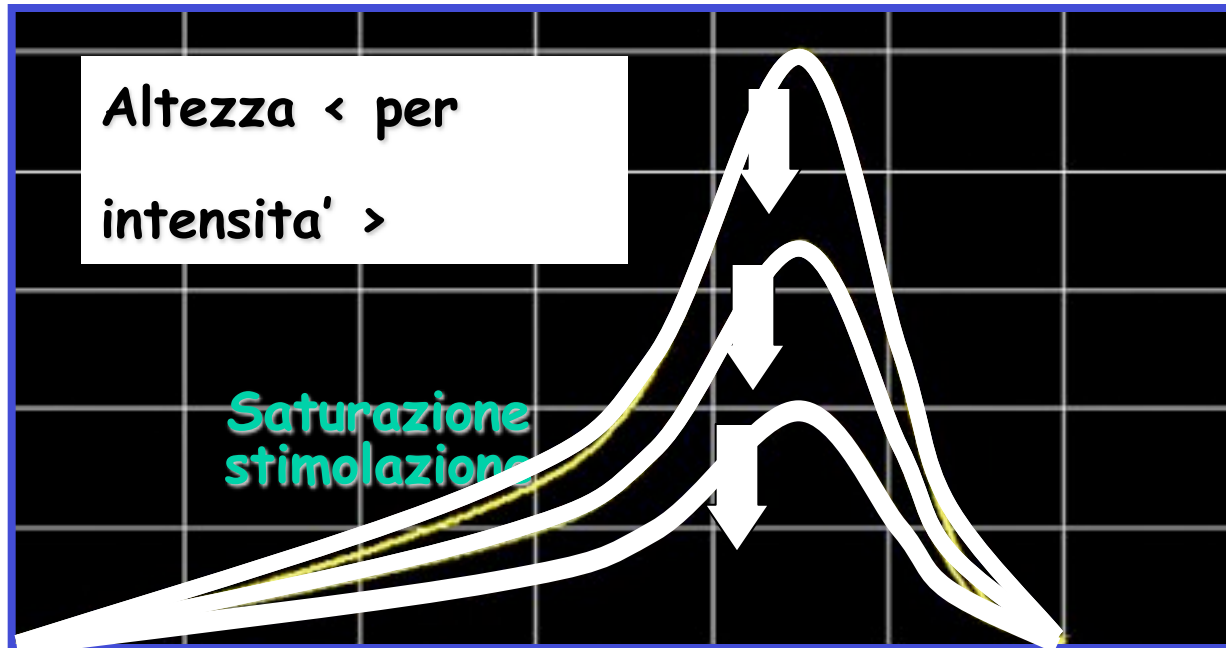


Misura della banda critica



# Dipendenza altezza da sonorità

Cambiamento dell'altezza con l'intensità



Sensazione di altezza >

Sensazione di altezza <

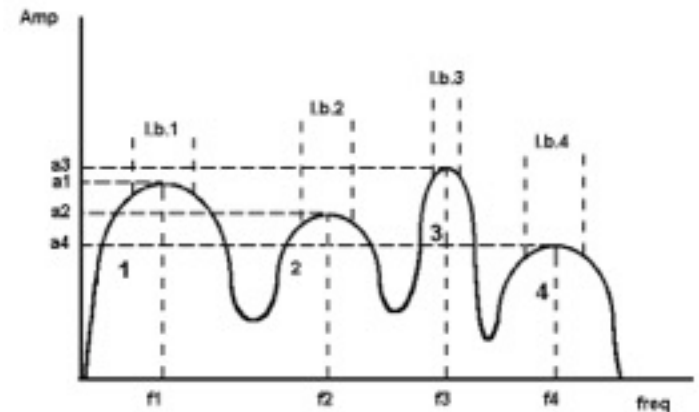


# Suoni complessi: TIMBRO

Il **timbro** (o qualità del suono, o coloritura) descrive le caratteristiche del suono che permettono all'apparato uditivo umano di distinguere suoni con ugual altezza ed intensità. Le qualità del timbro dipendono soprattutto dal **contenuto armonico** del suono: cioè dalla quantità e dall'ampiezza delle sinusoidi che contiene (analisi di Fourier) e dalle **caratteristiche dinamiche** come l'involuppo dell'ampiezza del suono e fenomeni di vibrato.

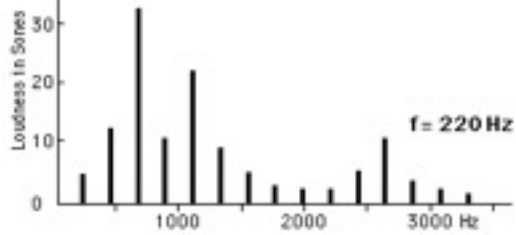
Per la definizione dello spettro e del timbro è possibile fare riferimento non tanto alle singole armoniche e alle loro ampiezze, quanto alla frequenza, all'ampiezza e alla larghezza di banda di alcuni **formanti**, o zone di massima energia sonora.

Sono presenti quattro formanti (numerati da 1 a 4), ciascuno dei quali è caratterizzato da una frequenza centrale ( $f_1, f_2, f_3, f_4$ ), da una ampiezza massima ( $A_1, A_2, A_3, A_4$ ) e da una larghezza di banda (l.b.1, l.b.2, l.b.3, l.b.4). Per quanto riguarda la voce parlata (e in qualche misura anche quella cantata) è sufficiente, per la comprensibilità del linguaggio, definire le caratteristiche di questi quattro formanti, oltre che delle caratteristiche di eccitazione (segnale periodico o non periodico per differenziare fra suoni vocalici o consonantici).

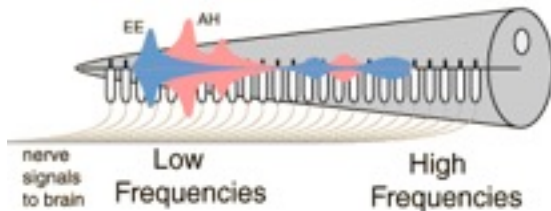
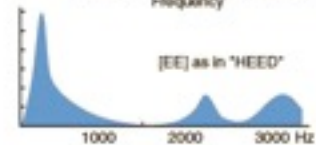
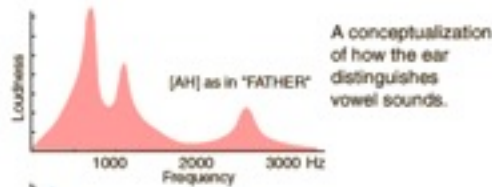
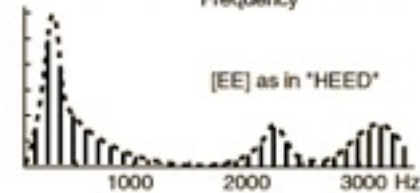
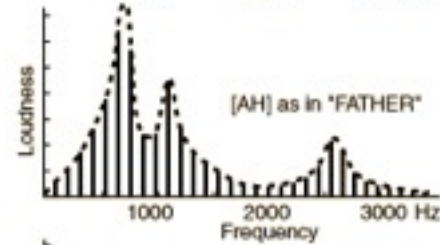
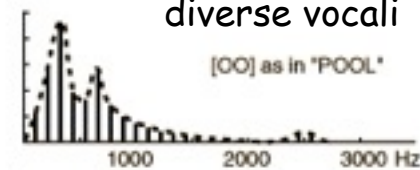


# Vocali

stessa vocale a diverse frequenze

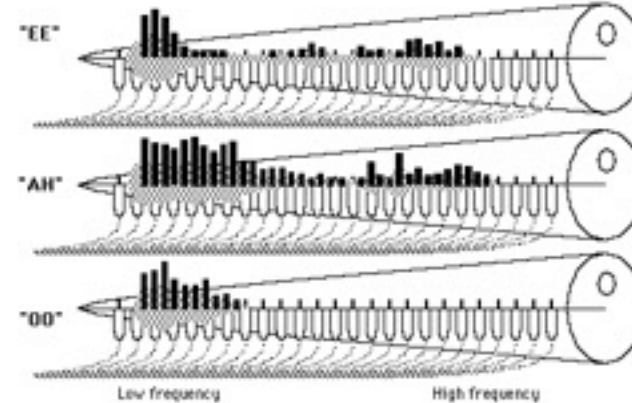


diverse vocali



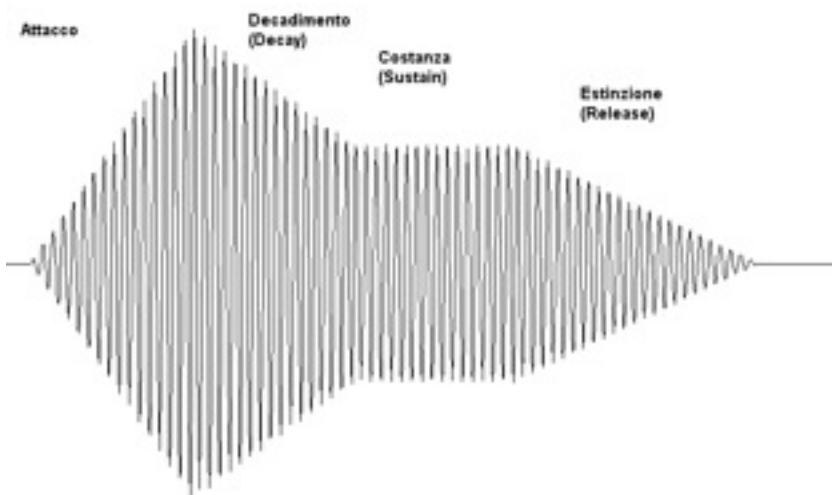
diverse percezioni

Conceptual only!  
No actual scaling to ear was done.





# Inviluppo sonoro

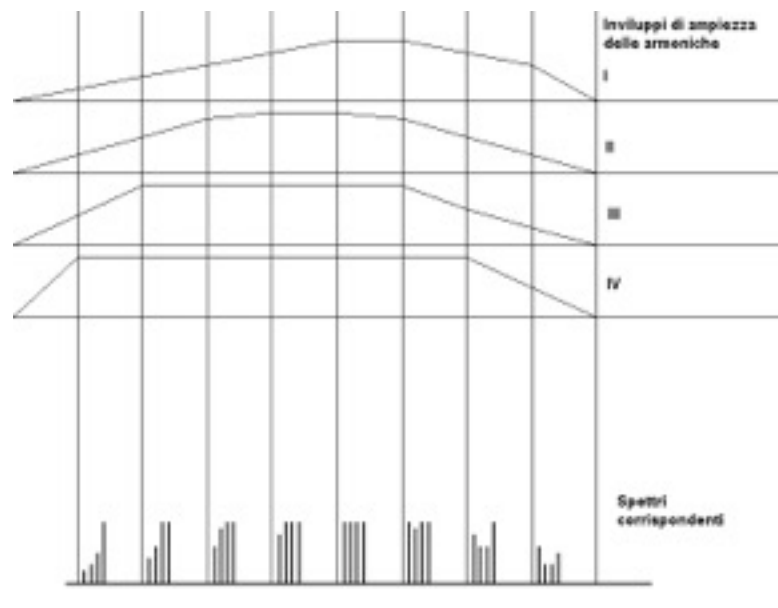
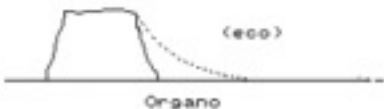
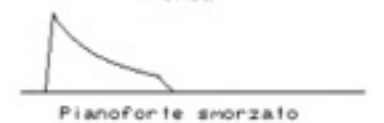
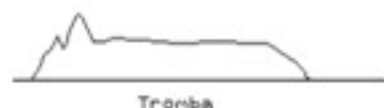
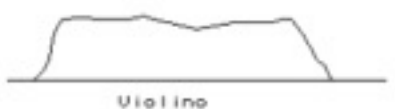
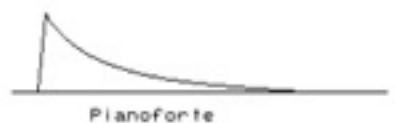
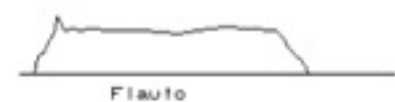


**Attacco:** in cui l'ampiezza gradualmente varia da zero al massimo;

**Decadimento:** in cui l'ampiezza diminuisce fino a un certo livello;

**Costanza:** in cui l'ampiezza si mantiene pressapoco costante;

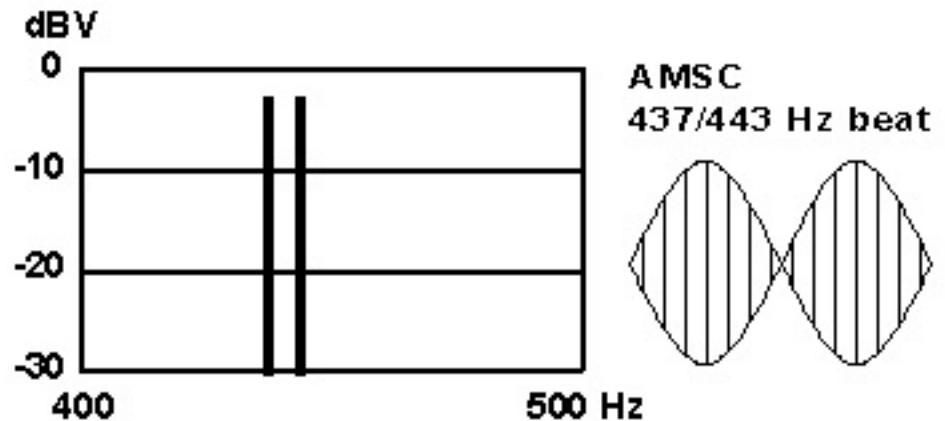
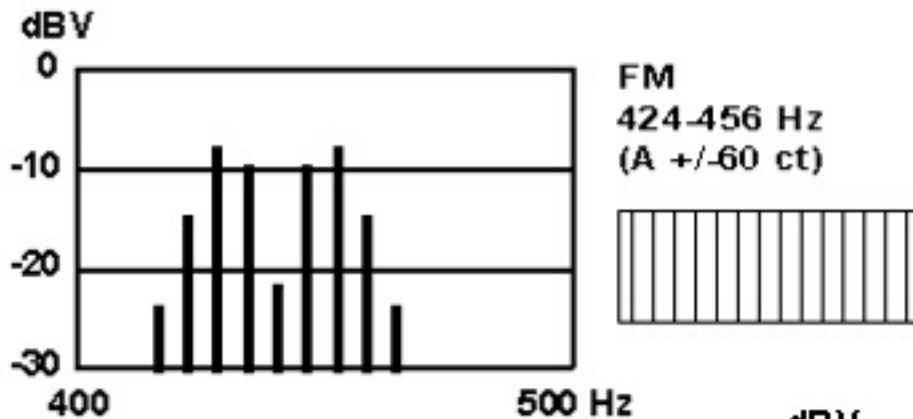
**Estinzione:** in cui l'ampiezza gradualmente diminuisce fino a zero.



# Vibrato e Tremolo

**Vibrato:** cambiamenti periodici nell'**altezza** di un suono (FM)

**Tremolo:** cambiamenti periodici nella **sonorità** di un suono (AM)



# Esempio: Violino

