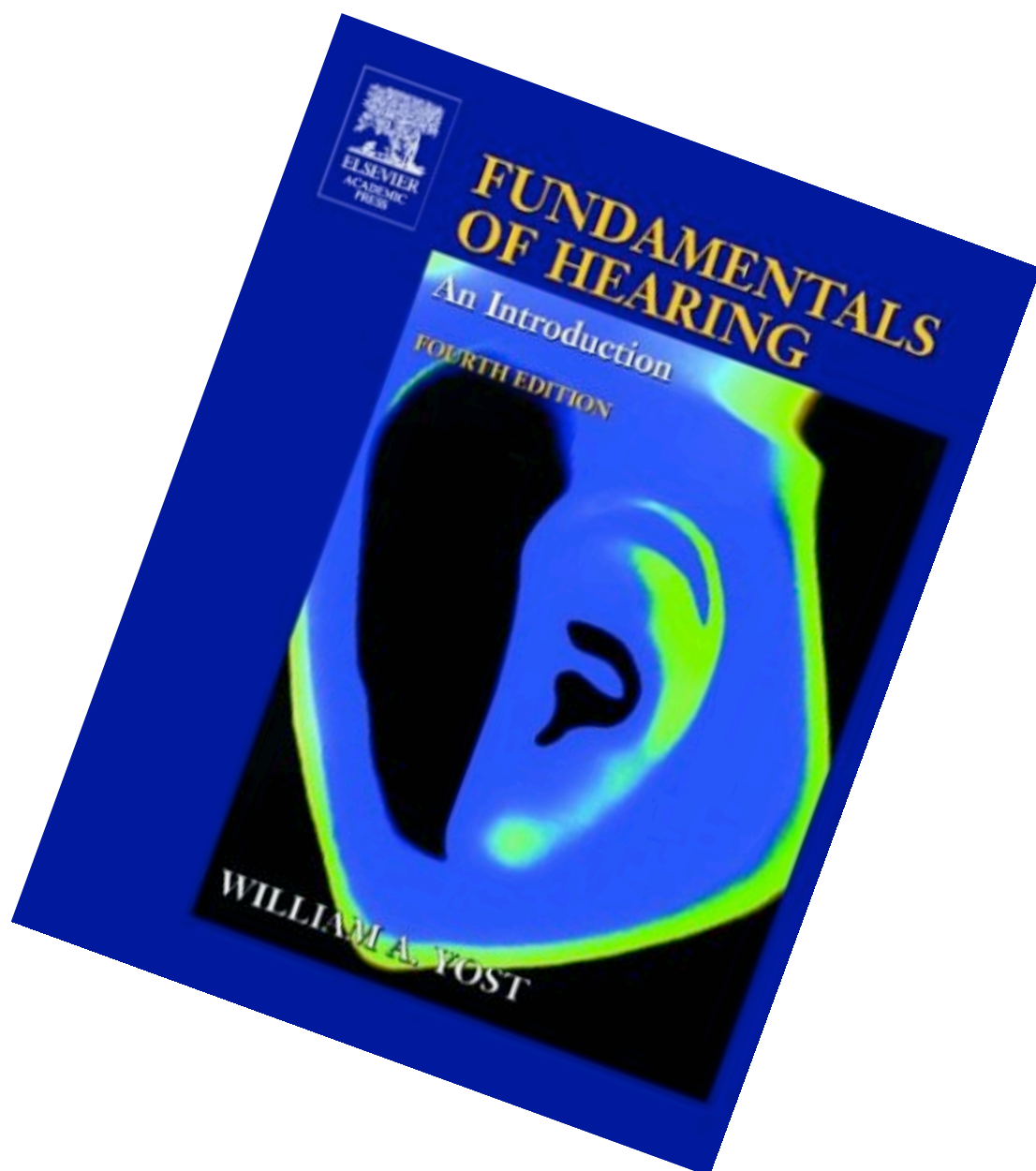


日本音響学会関西支部
聴覚基礎理論談話会

(2004/2/21)

-----**配付資料**-----



聴覚基礎理論談話会

2004年2月21日

於 生駒芸術会館

今回の聴覚基礎理論談話会では昨年に引き続き、Willam Yost著Fundamentals of Hearing: An introduction (2000, Academic Press)に基づいて、聴覚の基礎過程を学ぶために便利な教材の提供と基本知識の提供を目指して発表を行います。今回、紹介する章とその担当者のリストは以下のようになっています。今回は今年のこの談話会に聴衆として参加してくれて、その後に月例の勉強会にも顔を出してくれた「新規」メンバーの人たちが10章以降を担当してくれています。

昨年度は各章の内容をどちらかというと淡々と紹介していきましたが、今回は各担当がそれぞれの章から「目玉」になる題材を採り上げて、PowerPointなどの機能を活かした効果的な説明を試みます。その為に各章の具体的な内容については配付される資料に譲ります。実際に、このような教材で聴覚を学ぶ立場になった入門者の目になって、質問、コメントなどをしていただけると幸いです。

Chap. 7 Structure of the Inner Ear and Its Mechanical Response

栗栖清浩 TOA 株式会社

Chap. 8 Peripheral Auditory Nervous System and Hair Cells

高野佐代子 ATR-HIS

Chap. 9 The Neural Response and the Auditory Code

白土保 CRL

Chap. 10. Auditory Sensitivity

中川誠司 産業技術総合研究所 ライフエレクトロニクス研究ラボ

Chap. 11. Masking

藤坂洋一 産業技術総合研究所 ライフエレクトロニクス研究ラボ

Chap. 12. Sound Localization and Binaural Hearing

岡本洋輔 産業技術総合研究所 ライフエレクトロニクス研究ラボ

Chap. 13. Loudness and Pitch

添田喜治 産業技術総合研究所 ライフエレクトロニクス研究ラボ

この談話会は年に1回のこのような公開の研究発表会以外に、月例の勉強会を行っております。月例会のご案内は一般には行っておりませんが、こちらに対しても関心のある方は以下の連絡先までご連絡下さい。

津崎実(ATR)

minoru.tsuzaki@atr.jp

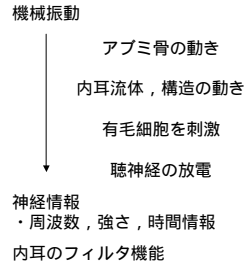
Fundamentals of Hearing:
An Introduction, 4th. ed.
by William A. YOST
(Academic Press 2000)

Chapter 7
Structure of the Inner Ear and its Mechanical Response

栗栖清浩
TOA株式会社
April 12, May 24, 2003(月例会)
Feb. 23rd., 2004(公開談話会)

7章の内容

6章 外界から内耳への音刺激の伝播経路について
7章 内耳の解剖学的構造と振動刺激との関係



Summary 1/2

- ・内耳は3つの部分に分かれ、そのうち聴覚を担当するのは蝸牛．
- ・蝸牛は基底膜とライスネル膜により3つに分けられている：
前庭階，中央階，鼓室階．
- ・前庭階と鼓室階は外リンパ，中央階は内リンパが満たされている．
- ・基底膜の蝸牛回転軸側は骨ラセン板で，血管条側はラセン韧带で支持されている．
- ・基底膜は頂端部で幅広，柔らか，張力がかかっていないが，蝸牛基部では狭く硬く，おそらくいくらか張力がかかっている．
- ・中央階にはラセン器があり，コルチンネルより内側(蝸牛回転軸側)に内有毛細胞が，反対側に外有毛細胞がある．
- ・内外有毛細胞の繊毛は蓋膜と接している．
- ・有毛細胞と神経線維は支持細胞で支えられている．

Summary 2/2

- ・アブミ骨の駆動により基底膜上に進行波が生じる．
- ・進行波の変位は刺激が低周波数のとき頂端部で最大となる(頂端部に達するまでの時間遅延がある)．
- ・進行波の変位は刺激が高周波数のとき基部で最大となる．
- ・刺激のレベルが変化すると基底膜振動変位も変化する．
- ・基底膜振動の時間パターンは刺激の時間パターンに従う．
- ・基底膜の動きはCF付近で特に非線形である．

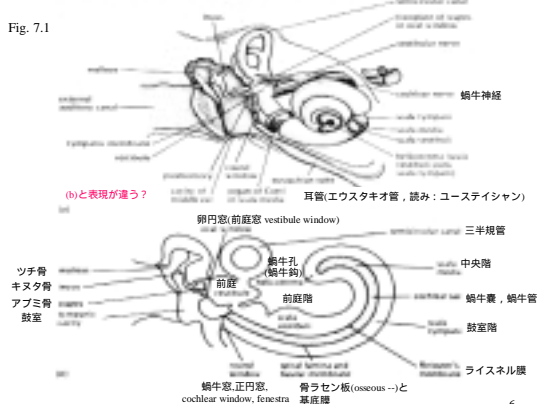
注意：
本資料において「音響用語辞典」から引用した図を用いて補足説明しているが，
これらは「新版音響用語辞典(コロナ社2003.7発行)」ではなく，
月例会(2003年4-5月)当時の「音響用語辞典(コロナ社1988初版)」からの引用である．

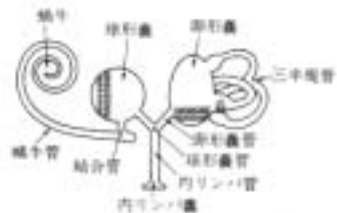
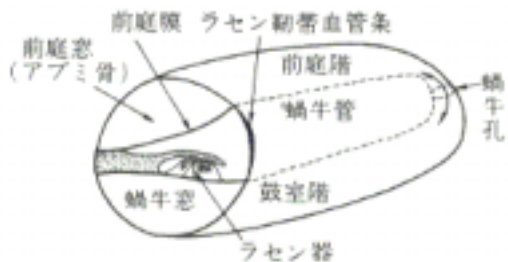
音響用語辞典・聴覚器官



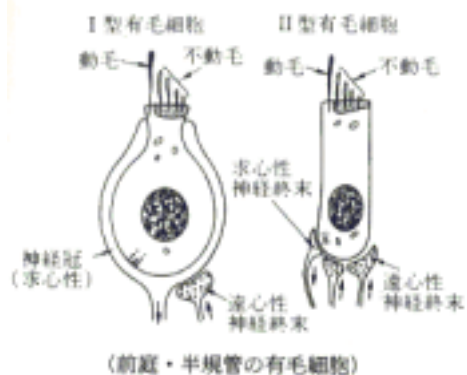
(音響用語辞典「聴覚器官」より)

Fig. 7.1





(音響用語辞典: 「前庭」より)



(前庭・半規管の有毛細胞)



(音響用語辞典「内耳」より)

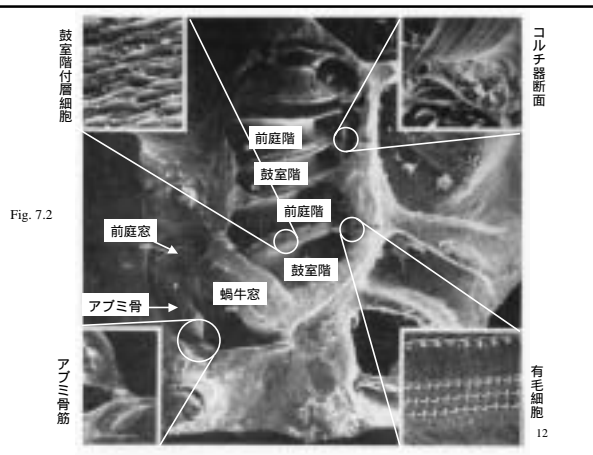
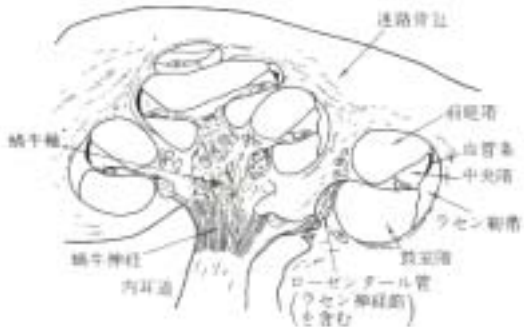
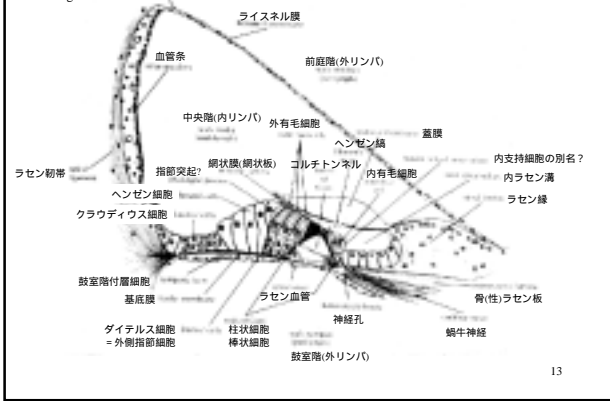
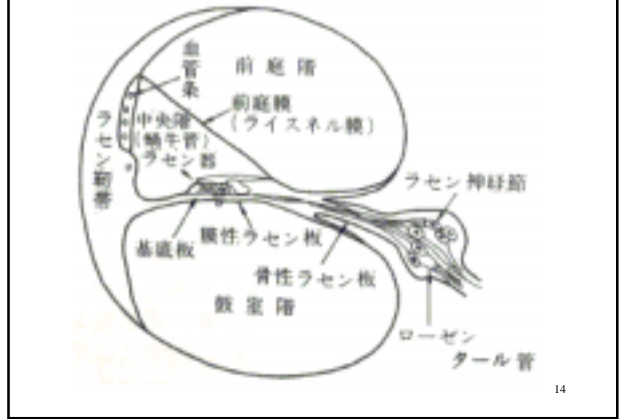


Fig. 7.2

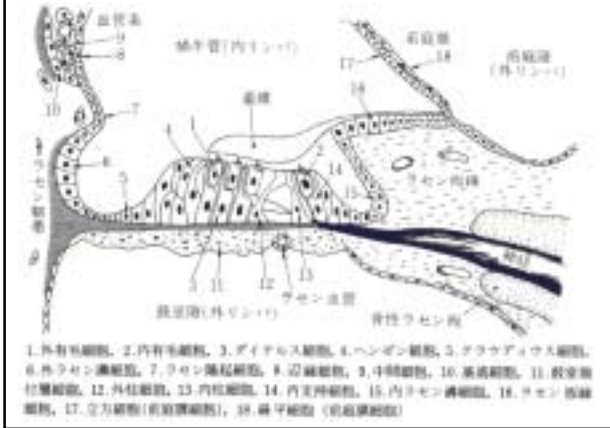
Fig. 7.3



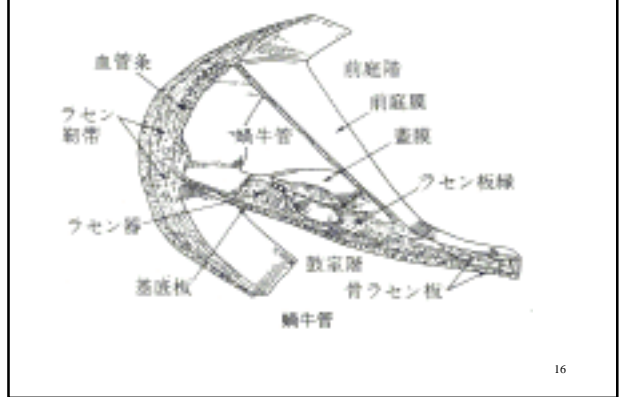
音響用語辞典:蝸牛管



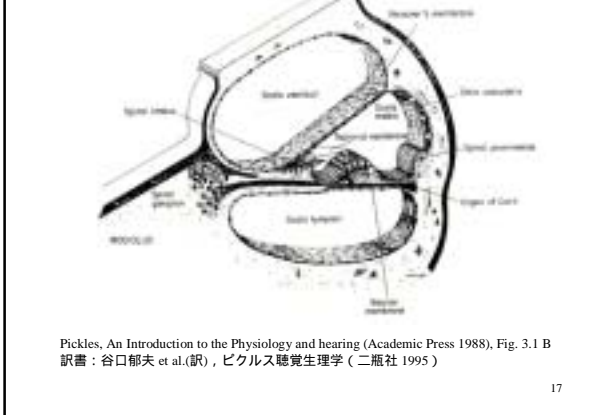
音響用語辞典:ラセン器



音響用語辞典:感音器 1/2



Pickles: 蝸牛管断面



Pickles, An Introduction to the Physiology and hearing (Academic Press 1988), Fig. 3.1 B 訳書: 谷口郁夫 et al.(訳), ピクルス聴覚生理学 (二瓶社 1995)

Fig. 7.4

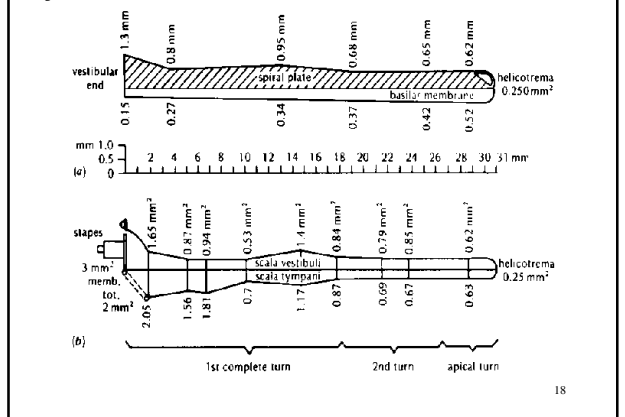
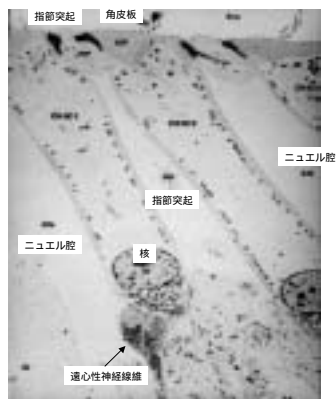


Fig. 7.7a 内毛細胞と外毛細胞

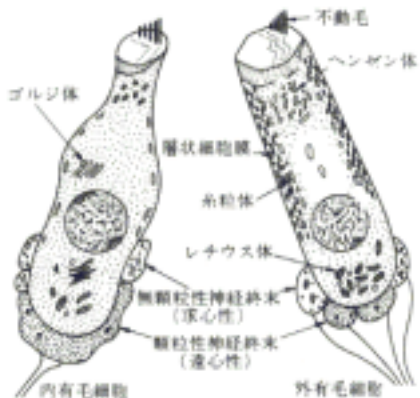


Fig. 7.7b

外毛細胞



音響用語辞典: 内毛細胞



話題提供: 器官や組織の名称について

同じものに対し異なる名称が使用されている。これらは研究者の名に因んだもの、機能的名称などであり、分野(生理学や解剖学)によって異なる名称で呼ばれているようだ。初學者や学際領域の研究者のためにも用語集などではその辺の事情を詳しく解説して欲しい。

日本語表記	英語表記
前庭窓, 卵円窓	vestibular window, oval window
蝸牛窓[4,5,6], 正円窓[3], 円形窓[8], 鼓室窓[7]	cochlear window[4,5], round window, fenestera rotunda
基底膜[3,6,7], 基底板[4, 5]	basilar membrane
ラセン器[4], コルチ器	spiral organ[4], organ of Corti
内柱細胞 / 外柱細胞[4], コルチの柱	inner/outer pillars, rods of Corti[1]
鼓室階	scala tympani, tympanic canal[8]
ライスネル膜, (ライスネルの)前庭膜[8]	Reisner's membrane, vestibular membrane (of Reisner)
中央階, 蝸牛管	scala media, cochlear duct, cochlear sac[1]

出典:

- [1] YOST W.A. Fundamentals of Hearing: an introduction 4th. ed.(Academic Press 2000)
- [2] PICKLES J.O. An Introduction to the Physiology of Hearing 2nd. ed.(Academic Press 1988)
- [3] 谷口郁雄 et al. (訳), ビクルス聴覚生理学 (二瓶社 1995)
- [4] 日本音響学会 (編), 音響用語辞典 (コロナ社 1988)
- [5] 日本音響学会 (編), 新版音響用語辞典 (コロナ社 2003)
- [6] 日本音響学会 (編), 聴覚と音響心理 (コロナ社 1978)
- [7] 電子情報通信学会 (編), 新版聴覚と音声 (コロナ社 1980)
- [8] マクローヒル科学技術用語大辞典 (日刊工業新聞社 1979)

Fig. 7.8 蓋膜を取り除いたコルチ器上部

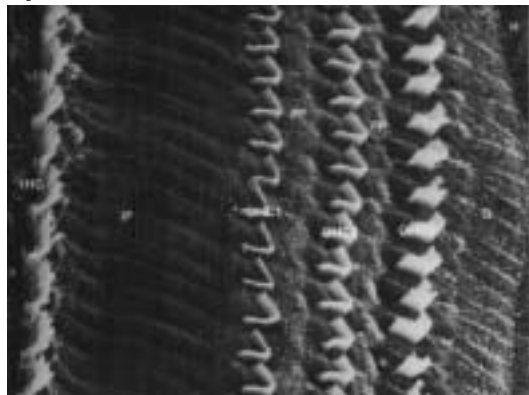


Fig. 7.9a

内毛細胞

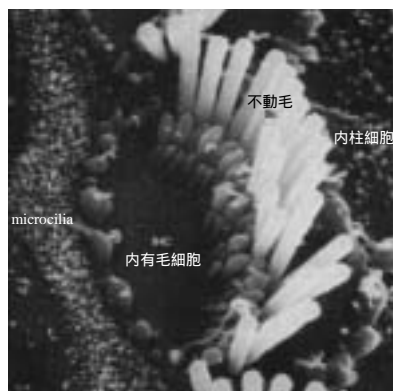
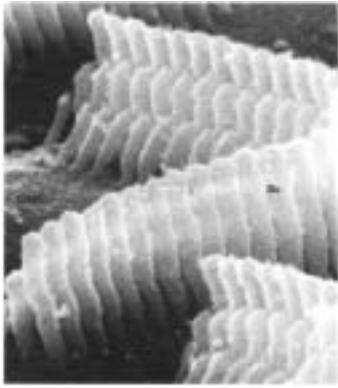
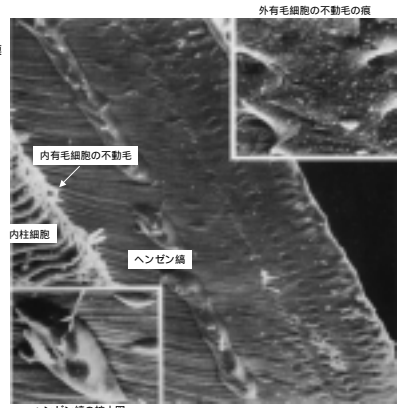


Fig. 7.9b
外有毛細胞



31

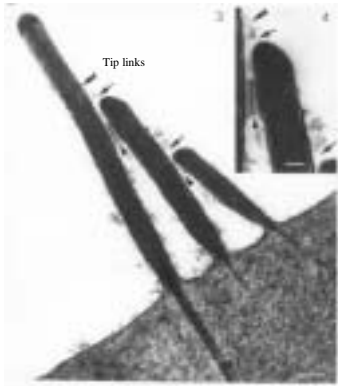
Fig. 7.10
有毛細胞側から見た蓋膜



ヘンゼン橋の拡大図

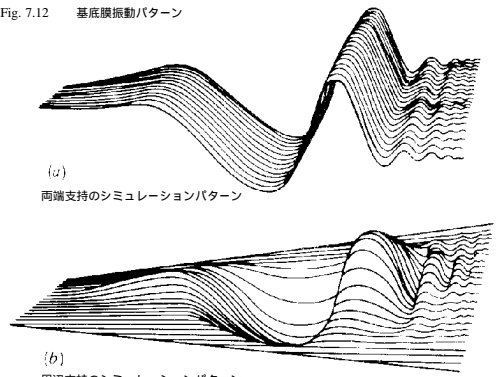
32

Fig. 7.11
不動毛の拡大写真



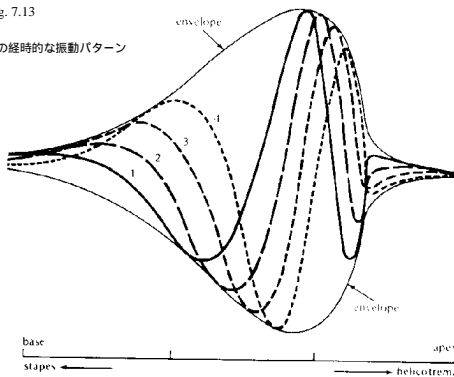
33

Fig. 7.12 基底膜振動パターン



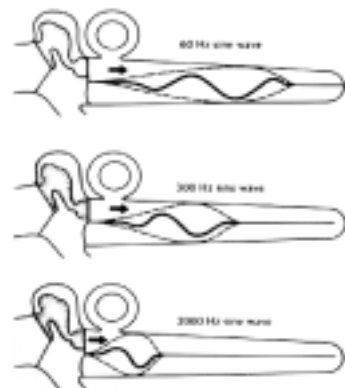
34

Fig. 7.13
基底膜の経時的な振動パターン



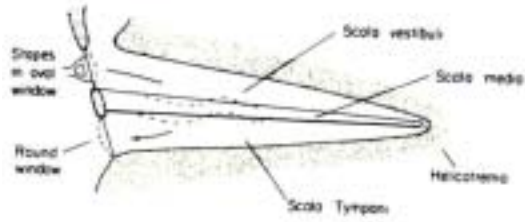
35

Fig. 7.14 異なる周波数における基底膜上の進行波とその包絡線



36

Pickles: 中央階の振動



Pickles, An Introduction to the Physiology and hearing (Academic Press 1988), Fig. 3.1 C
訳書：谷口郁夫 et al.(訳)，ピクルス聴覚生理学（二瓶社 1995）

Fig. 7.15 50[Hz], 100[Hz]を独立に提示した時の基底膜振動と、同時に提示した時の基底膜振動

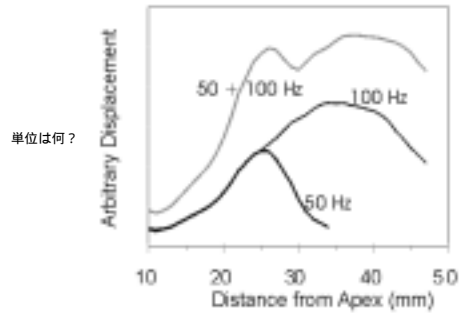
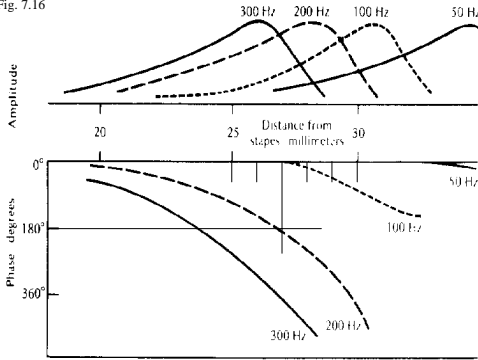
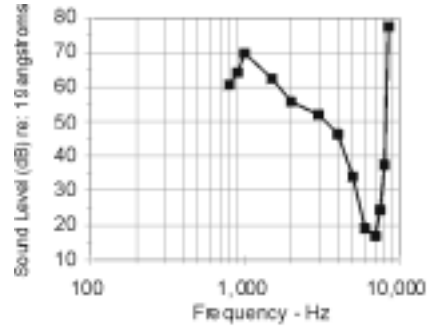


Fig. 7.16



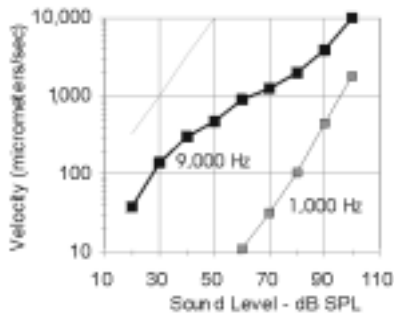
アブミ骨振動を基準とした基底膜上の位相遅れ, e.g., 200[Hz]が提示されたとき基底膜上27[mm]位置では0[mm]位置(アブミ骨)と比べてほぼ逆相で振動している。

Fig. 7.17



基底膜上のある一点を振幅 1.9×10^{-8} [cm]で振動させるのに必要な音圧レベル

Fig. 7.18 基底膜振動の非線形性



基底膜上で9000[Hz]で一番よく振動する位置(CF=9000[Hz])における振動速度 vs. 音圧レベル。提示周波数がCF=9000[Hz]の時、振動速度は音圧レベルと(ログスケールで)直線的な関係ではない。提示周波数がCFでないとき(ここでは1000[Hz])、直線的に変化する。

Chapt. 8. Peripheral Auditory Nervous System and Hair Cells

Introduction

●蝸牛電位

cochlear potentials

●生物学的な変換器としての毛有細胞

hair cells

●毛有細胞とつながって、脳幹へ信号を送ったり受け取ったりするための聴覚神経の解剖

auditory nerve

Cochlear Potentials

●静止電位 (resting potentials)

直流。聴覚刺激なしに存在する。

●累加電位 (summing potentials)

直流。聴覚刺激があるときのみ存在する。

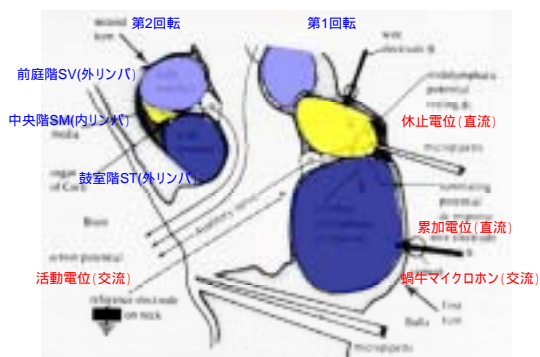
●蝸牛マイクロフォン電位 (cochlear microphonic)

交流。聴覚刺激があるときのみ存在する。

●活動電位 (action potentials)

交流。内耳の構造ではなく、神経によって生成される。

蝸牛軸部分の図



静止電位 (resting potentials)

●内リンパ電位(SM) +80mv

endocochlear / endolymphatic potentials

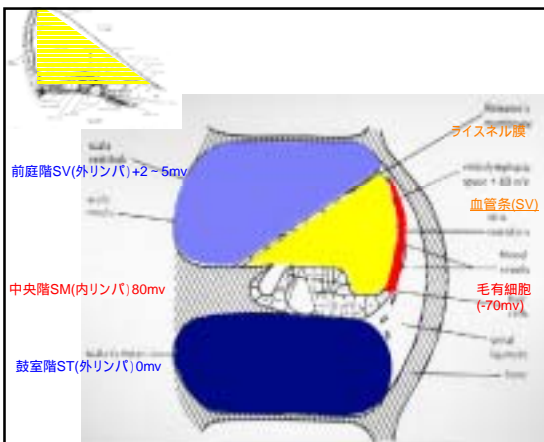
体内で最も高い正電位

源は血管条 (strial vascularis, Fig. 7-3左上)

毛有細胞の内側の静止電位は-70mvなので、最大で150 mvの電位差がある。効率的。損傷により、聴力低下を引き起こす。

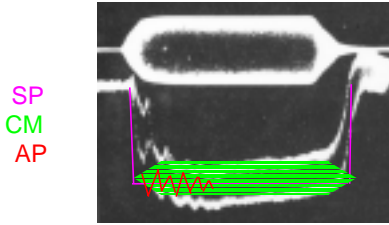
●外リンパ電位(SV) (+)2-5mv

●外リンパ電位(ST) 0mv



累加電位 Summating Potential

- 累加電位は蝸牛で記録される直流の電氣的反応である。



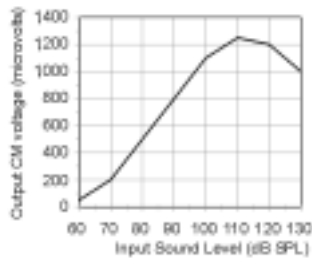
蝸牛マイクロフォン電位

- 聴覚刺激があるときのみ生じる交流の電位
- 500Hzの刺激ならば500Hzの応答
- 刺激の大きさはある程度比例 Fig.8-4
- 信号源は外毛有細胞
(繊毛支持点 cilia-bearing? の終わり)
- CMの大きさは基底膜の変位と比例関係 Fig.8-5
von Bekesyの結果と一致

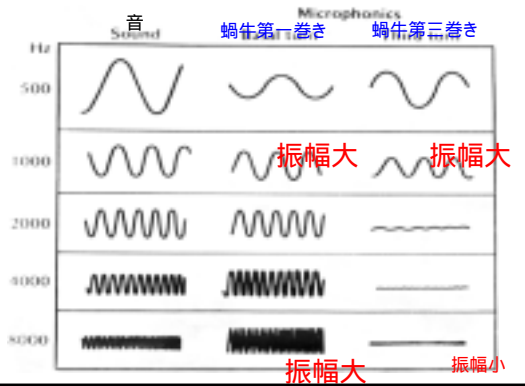


蝸牛マイクロフォン電位の応答

- 100dB SPL程度まではほぼ線形



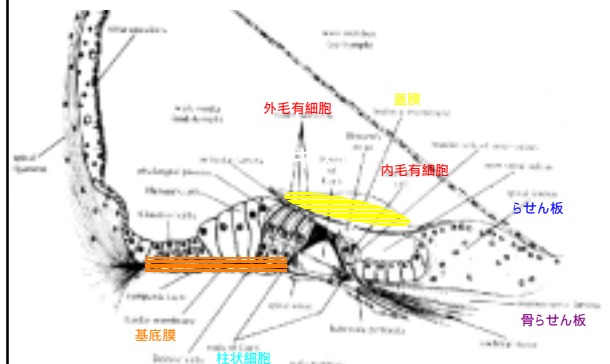
蝸牛マイクロフォン電位



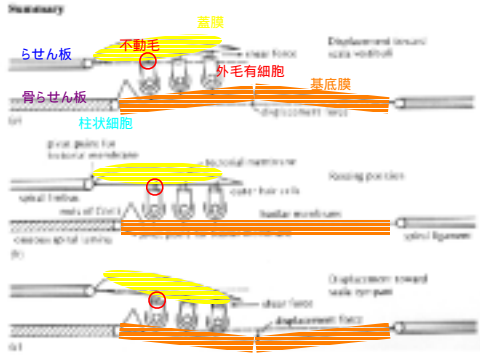
活動電位

- 交流。内耳の構造ではなく、神経によって生成される。
- 聴覚神経束とほぼ同期して発火するたくさんの個々の聴覚ニューロンの活動の和である(chapt9)。
- 個々のニューロンの発火は基底膜の振動によって生じる。
- 蝸牛の底を神経でつなくニューロンは頂上よりは下方の方が早く発火する。
- 時間の違いは進行波が底から頂上まで移動するための時間である。(種や計測方法によって異なり、2.5 ms ~ 4msである)。

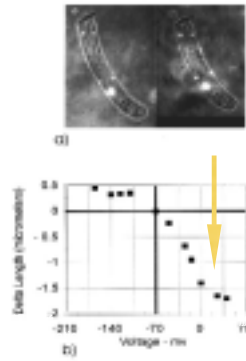
有毛細胞と不動毛



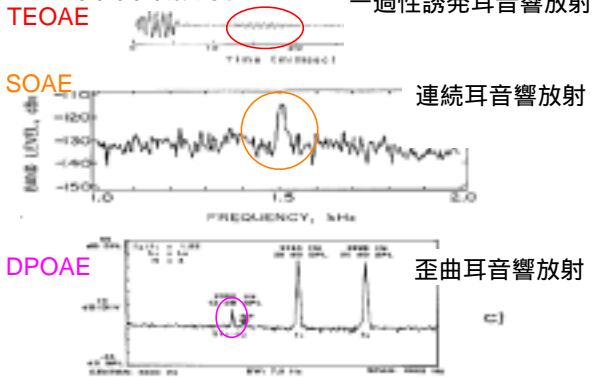
基底膜の位置変化によるねじれの力



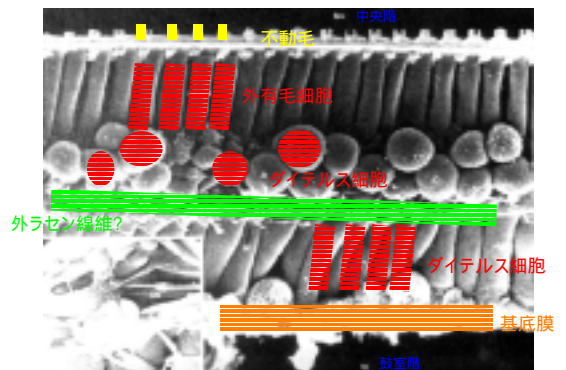
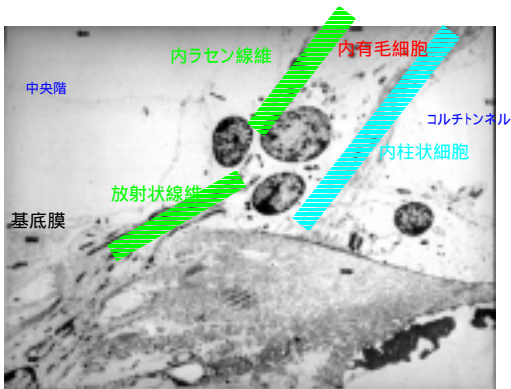
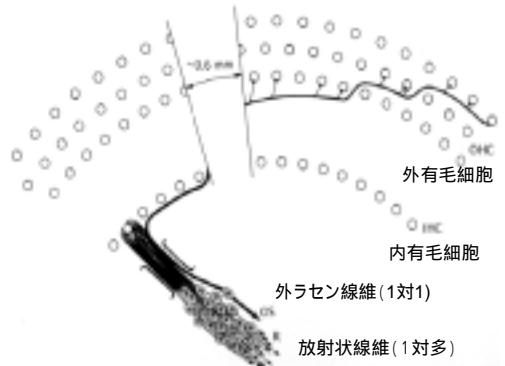
外毛有細胞の長さ変化



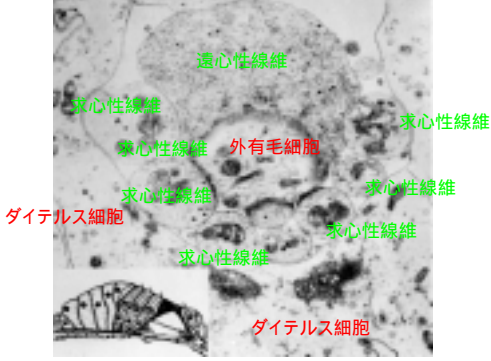
耳音響放射



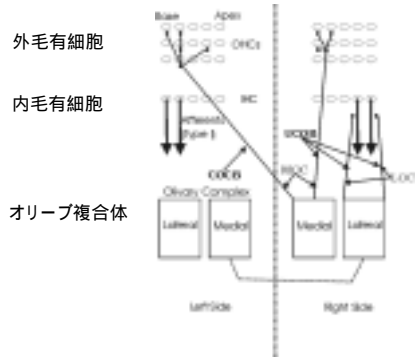
聴覚神経の構成



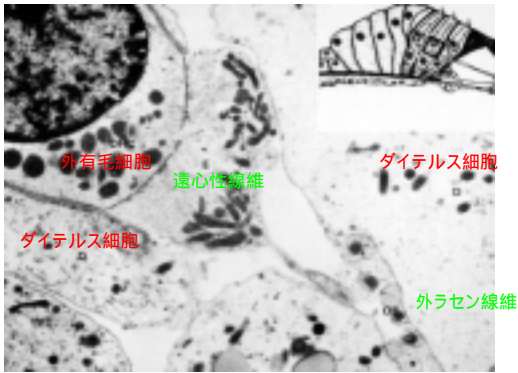
外毛有細胞をとりまく求心性線維



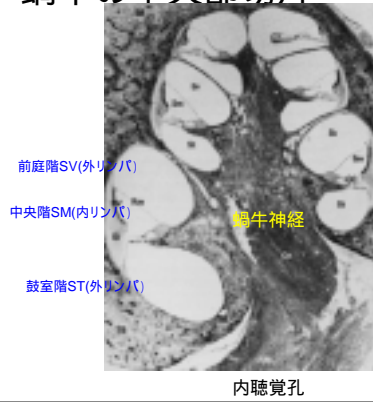
脳と内耳の神経接続



外毛有細胞下部の求心性線維



蝸牛の中央部切片



FUNDAMENTALS OF HEARING

Chapter 9: The Neural Response and the Auditory Code

- ・本章の内容: 第VIII神経束における求心性、および遠心性の聴神経。
- ・刺激音の周波数、レベル、時間的特徴、及び遠心性聴神経が、有毛細胞に交絡する求心性聴神経の出力をどのように変化させるか? について学ぶ。
- ・章末では、聴覚末梢系での音の符号化について述べる。

{pp.144-145 SUPPLEMENT}: 神経の反応を表す言葉として、以下をほぼ同義語として用いる: *discharge, firing, spike, response (rate)*

FUNCTION OF THE AFFERENT AUDITORY NERVE

- ・有毛細胞のところから聴覚プロセスの神経の部分。
- ・(このプロセスは以下の順):
 - ・不動毛の機械的変形 graded electrical potential (刺激の大きさに比例した電位) であり受容器電位 (即ちCM,SP) と呼ばれる。 有毛細胞が化学伝達物質放出
 - ・有毛細胞の根元に交絡している神経線維上にgraded electrical potential発生
 - ・神経線維を伝播 神経孔 神経孔でスパイク発生 蝸牛神経核
- ・generator potential: 髄鞘のない神経線維で作られるgraded electrical potentialのことであり、スパイクを生成する元と考えられるためこう呼ばれる。
- ・スパイク発生率は基底膜の動きの速度に比例。
- ・(SUPPLEMENT p.144): Sellickら(1980)、天竺ネズミにおける有毛細胞の受容器電位と鼓室階のCMの比較 100-200Hz以下では有毛細胞の反応は基底膜の速度に対応、より高い周波数では、基底膜の変位に対応。
- ・放電パターン: 図E.4(p.291-292)参照。
- ・不応期(点A-Dの時間、約1ms) = 絶対不応期(点B-C) + 相対不応期(点C-D)
- ・絶対不応期: どんなに強い(後続の)刺激に対しても反応しない期間
- ・相対不応期: (後続する)刺激の強さに依存して反応する期間
- ・個々のニューロンの活動電位(スパイク)は1.0。即ち、放電強度は刺激のレベルと共に変化はしない。
- ・APは多数の聴神経の組み合わせ(放電の和)なので、APの大きさは刺激のレベルと共に変化。

Pickles 聴覚生理学

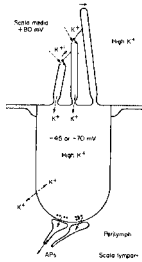


図 9.17 Davis の電位理論の現在版は、有毛細胞の先端部にある変換器チャンネルが可変抵抗として働くというものである。内リンパ電位と細胞内電位からなる電位で駆動されたイオンが、細胞内に流れ込む。細胞内電位が弱分極すると伝達物質が放出され、聴神経が活動する。有毛細胞が活動すると、有毛細胞の膜電位が減少し、鼓室階の電位が反対方向に変化する。この図では、変換器チャンネルが低い電位(不動毛の先端)にあるように描かれているが、これはまだ解明の域を出ない。Pickles ら (1984) の用語に言って、不動毛の先端を引っ張らないで、基底膜よりチャンネルが開いた状態を示している。基底膜の先端の方向に不動毛を動かすと興奮(脱分極)が起こる。外有毛細胞と内毛細胞の静息電位はそれぞれ -45 mV と -75 mV である。APs: 活動電位。Perilymph: 外リンパ。Scala media: 中央階。Scala tympani: 鼓室階。Pickles (1987, 図 2.16) より改変。

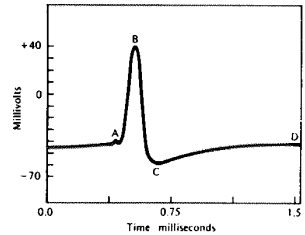


FIGURE E.4 Voltage changes during stimulation of a nerve. From the resting potential there is an increase (depolarization), A, to a positive voltage, B, then a return to negative voltage (hyperpolarization) with a negative undershoot, C, and back to the resting potential, D. This is the voltage change that exists at the point of nerve discharge shown in Figure E.3. The time between A and D is the total refractory period of the neural discharge.

FIGURE E.4

Spontaneous Activity and Neural Thresholds

- ・同義語: *neuron, unit*.
- ・単一ニューロンの放電パターンを研究する方法のひとつとして、刺激のないときのニューロン活動(自発性活動)を刺激のあるときのニューロン活動と比較する方法がある。
- ・単一ニューロンの自発性放電率は、低いものでは0/sec, 高いものでは100spikes/sec以上。
- ・絶対不応期が1msなので理論上は単一ニューロンの最大放電率は1,000spikes/sec。殆どの聴神経線維の最大放電率は500spikes/sec以下だが。
- ・単一ニューロンの(最小)閾値: 自発性放電率より高い放電率を生じさせる最小の刺激値。ただし、自発性放電率は変動するため、一般には、放電率平均の統計的变化を以って決める。
- ・図9.1は、自発性放電率が異なる3種のニューロン群に対して測定された(最小)閾値の分布。
- ・図の横軸は、最適(特徴)周波数、縦軸は特徴周波数における閾値。
- ・自発性放電率Low: <0.5 spikes/sec, Medium: 0.5-18 spikes/sec, High: > 18 spikes/sec
- ・閾値は、自発性放電率High < Medium < Low の(ニューロン群の)順。
- ・閾値は1,000から10,000 Hzで最小。これは、ネコの耳が最もよく聞こえる周波数領域 (p.151 図10.2: 音の絶対閾) にあたる。

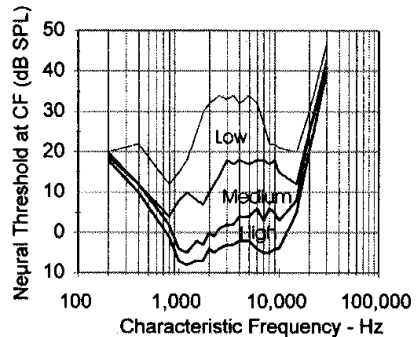


FIGURE 9.1 (ネコ)

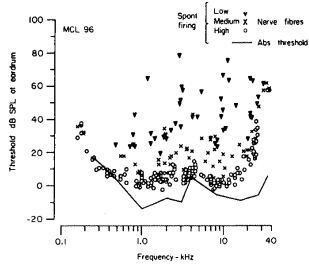


図 4.4 1匹のネコにおける聴神経線維の最小閾値の分布。自発放電頻度の高い線維 (○、18/s以上) は閾値が低く、自発放電頻度の低い線維 (▽、0.5/s未満) は閾値が高い。中間の自発放電頻度の線維 (×) はその中間の閾値を持つ。行動学的に得られた、鼓膜での音圧強度で表されたネコの絶対閾値は、聴神経線維の最も低い閾値のすぐ下にある。神経のデータは Liberman と Kiang (1978, 図 2)、行動データは Elliott ら (1960) より。縦軸: 鼓膜音圧 (dB SPL) で表わした閾値。Abs threshold: 絶対閾値、Nerve fibres: 神経線維、Low: 低、Medium: 中、High: 高、Spont firing: 自発放電。

Rate-Level Functions

・刺激音のレベルに対する単一ニューロンの放電率のこと。以下は *rate-level* 関数の同義語: *intensity* 関数, *input-output* 関数。
 ・図9.2は、3つのニューロン (自発性放電率 Sr: high/medium/low) の *rate-level* 関数。
 ・基本的に、(図) 全ての聴神経反応は、閾値を超える 20-50 dB の刺激レベルの範囲 (*neuronal dynamic range* と呼ばれる) では増加し、それより大きいレベルに対しては一定あるいは 僅かに減少 (閾値を超える 30-40dB で特徴的)。
 ・自発性放電率が異なる線維の間では、*input-output* 関数 (実線) と自発性放電率 (点線) が交わる点の刺激レベルが異なる。

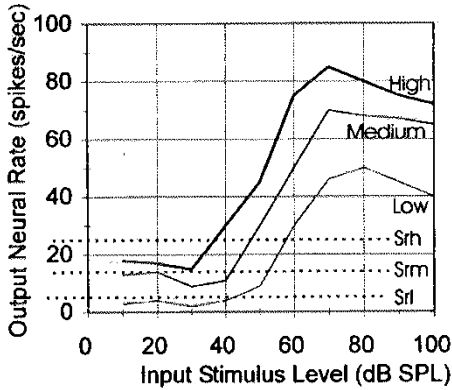


FIGURE 9.2

Response Areas and Tuning Curves

・図9.3は、3種類の刺激周波数に対する単一ニューロンの *rate-level* 関数。
 ・どのニューロンも幅広い周波数に反応する。しかし、周波数依存性がある。
 ・図9.3aを再プロットした図が、図9.3b、及び図9.3c。
 ・図9.3bは、固定した刺激レベル (この図では 50 dB SPL) に対する放電率であり、等レベル曲線、等強度曲線、あるいは応答野と呼ばれる。
 ・図9.3bでは 2,000 Hz に対する放電率が最大。この周波数を特徴周波数 (CF) と呼ぶ (これが図9.1の横軸 (図9.1での定義は同調曲線の CF と整合))。
 ・図9.4は、4レベル (captionの "six" は間違い) の刺激音に対する単一ニューロンの反応曲線 (応答野)。
 ・図9.3cは、閾値に対応する放電率 (この図では 15 spikes/sec) を引き起こす刺激レベルであり、等発火率曲線、あるいは同調曲線と呼ばれる。
 ・図9.3cでは 2,000 Hz に対する刺激レベルが最小。この周波数を特徴周波数 (CF) と呼ぶ (これが図9.1の横軸)。

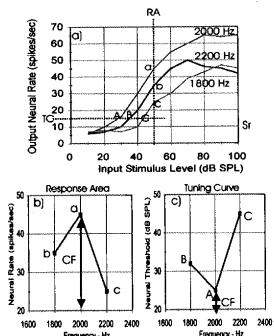


FIGURE 9.3

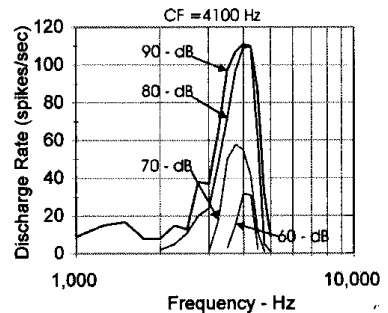


FIGURE 9.4 (リスザル)

Response Areas and Tuning Curves 続き

- ・図9.5は、異なるCFを持つ6つのニューロンの同調曲線(左上から右下に向かってCFが高くなる)。
- ・同調曲線の横軸は通常、対数周波数。
- ・どの同調曲線も、CFより高い周波数での傾きは急であり、より低い周波数での傾きは緩い。
- ・図の右(本文"bottom")は間違いない3つの同調曲線では、低い周波数での長いtailが顕著に見られ、tip&tailの差はいずれもほぼ40-50 dB。
- ・全てのニューロンで言える傾向として、tip&tailの差はいずれもほぼ40-50 dB
- ・以上のことは、CFより高い周波数範囲の刺激に対しては反応しにくく、より低い周波数の刺激に対しては、CFの閾値よりほぼ40-50dB大きい刺激レベルには反応することを意味する。
- ・図9.4、図9.5は、聴神経線維の周波数選択性が高いこと、即ち各ニューロンはそれぞれ異なる周波数範囲(CF)で最もよく発火することを示す。
- ・このようなI型神経線維は、基底膜に沿って並びつあるいくつかの内有毛細胞にシナプス結合しているので(8章)、放電率パターンが、基底膜の各点で得られるのと同様の周波数選択性を反映していることは驚きではない。つまり、聴神経は、進行波の動きによって基底膜に沿って得られた周波数選択性を保存する。
- ・8章で述べたように、外側の神経束は蝸牛基部側(周波数高)の有毛細胞に交絡しているが、この神経線維のCFは高い。一方、内側の神経束は蝸牛頂側(周波数低)の有毛細胞に交絡しているが、この神経線維のCFは低い。このように、周波数選択性に関し、蝸牛と神経線維の対応がある。
- ・以上の周波数選択性は、しばしばバンドパスフィルタリングの概念を用いて記述される

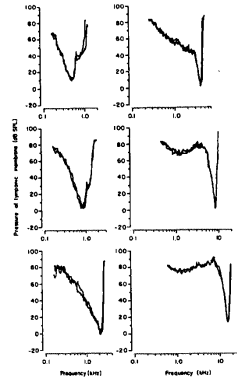


FIGURE 9.5

Histograms

- ・放電パターンの持つ特徴の内、周波数の知覚に重要と考えられるもの。
- ・CFが4.5kHz以下の神経線維は、1 kHz以下の純音刺激に対してはほぼ毎周期1回ずつ反応する(図9.6)。この反応は、時間固定、位相固定、あるいは*following response*と呼ばれる。
- ・1 kHz以上の刺激に対しては、絶対不応期のため単一のニューロンは毎周期1回ずつの反応はできないが、2周期毎に1回、3周期毎に1回...などのように反応する。
- ・図9.6では、正のピークのときのみ発火することに注目。つまり、不動毛が一方にずれた(せん断された)時のみ放電する。これを整流と呼ぶ。
- ・位相固定の性質を見るため、ここではヒストグラムという解析手法を用いる。この縦軸は反応頻度、反応率、あるいは放電数そのものである。横軸は時間、ただし、刺激音開始からの経過時間であったり、隣り合う放電間の時間の場合がある。

Function of the Afferent Auditory Nerve

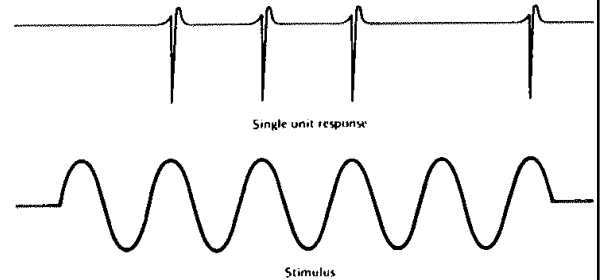


FIGURE 9.6

Poststimulus time histograms(PSTs)

- ・一定時間内における総反応数のヒストグラムであり、(同じ)刺激&各時刻での反応数の測定、を何度も行なうことによって得られる。
- ・ヒストグラムの開始時刻は通常、刺激音の立ち上がり直前、例えば5ms前にとる。
- ・図9.7は、トーンバーストに対する典型的なPSTヒストグラムを示す。
- ・刺激音立ち上がり時に放電率最大("on" effect)、その後急速に減少(順応)し刺激音呈示の間は一定、刺激音終了直後、放電率は0かほぼ0(自発性放電率より小)。その直後、自発性放電率まで回復。
- ・図9.7を図9.6と比べれば明らかのように、PSTヒストグラムは、各刺激に対する単一のニューロンの放電率を測定することでは見えなかった反応パターンが明らかになる。

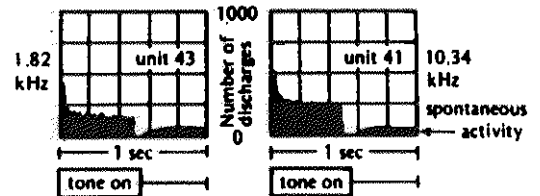
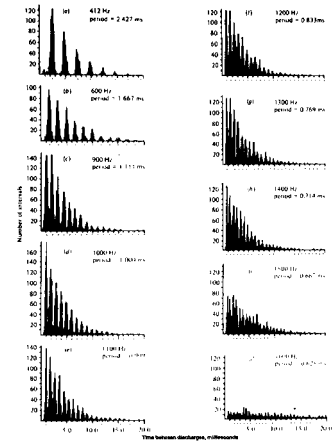


FIGURE 9.7(ネコ)

Interval histograms {ISI histograms}

- ISIヒストグラムの横軸は、連続した放電 (spike)間の時間。縦軸は、その頻度。
- 図9.8は、いろいろな周波数の純音 (持続長1 sec, 80 dB SPL)それぞれ10回ずつ測定して得たISIヒストグラム。
- 絶対不応期があるため、1msより短いISIは期待されない。
- 図9.8dでは、1msのISIの頻度が最大だが、その他、2から10msのISIにもピークがあり、このピークは、刺激音周期 (=1ms)の整数倍のISIのところ。他の周波数でも同様。
- 以上は、低い周波数を持つ刺激音に対し、各々のニューロンの放電は刺激音の cycles (1周期毎とは限らない)に固定する、というアイデアを支持する。
- 周期に対する何らかの固定は、ほぼ5 kHzくらいまで見られる。

FIGURE 9.8
(サル)



Phase-locked PST histograms (time-locked PST histograms, period histograms)

- PSTヒストグラムと同様、ただし刺激音の同じ位相に対する放電数のカウント。
- しばしば、刺激音1周期以上に対してこのヒストグラムを計算。
- 図9.9を見てわかるように、放電数は刺激音の特定の位相 (正の最大振幅の時の位相)で最大 (図9.9は、2つの同時的純音 (レベル9通り) に対するヒストグラムだが、ここでの説明上は冗長)
- neural synchrony (coefficient of synchrony): 放電パターンが刺激音波形の周期にどの程度一致しているかを表す指標。
- 1周期以上呈示された (正弦波的な) 刺激に対し、ヒストグラム上の放電パターンが正弦波的变化 (即ち、あたかも刺激の正弦波の上半分、あるいは整流された正弦波に追従) するときには coefficient of synchrony はほぼ1である (図9.9はこの例)。一方もし放電パターンが刺激位相に關して均一に分布しているならば coefficient of synchrony はほぼ0である。
- 刺激レベルに対する coefficient of synchrony の関数は、synchrony-level 関数と呼ばれる。
- 図9.2および図9.3に示したように、rate-level 関数は、刺激レベル30-50 dB以上では増加しないが、coefficient of synchrony はより大きい刺激レベルでも増加する。

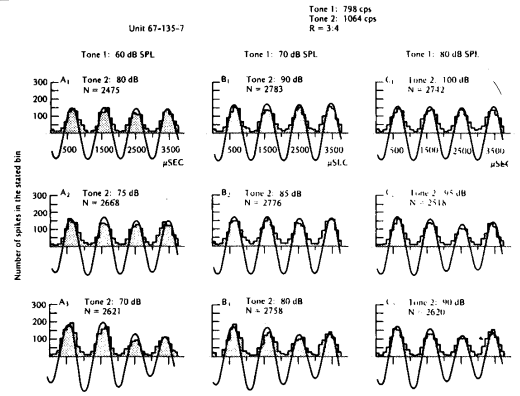


FIGURE 9.9

Phase-locked PST histograms 続き: クリック音に対する反応

- クリック刺激に対する神経線維の反応は、神経の変換に関するいくつかの重要なポイントを明らかにしてくれる。
- 図9.10a, bはそれぞれ疎クリック、密クリックに対するヒストグラムを示す。
- 図9.10cはそれぞれ、疎クリック、密クリックの、バンドパスフィルタによるフィルタリング前、フィルタリング後、の時間波形を示す。
- フィルタリング後の波形の振動周期は、フィルタ中心周波数の逆数に等しい。
- もし神経線維がこのフィルタリングの出力波形のピークで放電するならば、ヒストグラム上の主なピークはこの神経線維のCFの逆数になるはずであり、図9.10a, bはそうなっている。このことは、内耳の biomechanics があたかも刺激音をバンドパスフィルタリングしているかのように働いていることの根拠である。
- 疎クリック、密クリックに対するフィルタ出力波形は互いに逆位相であるが、もし不動毛が一方に動くときのみ放電が起こるならば、密クリックのヒストグラムのピーク群は疎クリックのピーク群より半周期遅くなる。図9.10a, bはこれを示す。

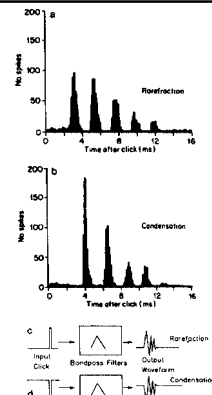
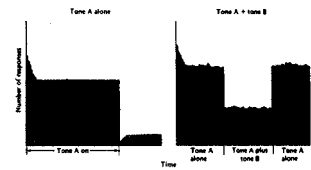


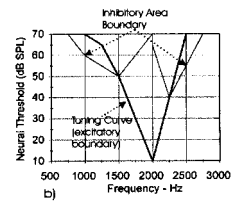
FIGURE 9.10

TWO-TONE SUPPRESSION AND OTHER NONLINEAR NEURAL RESPONSES

- 複雑な刺激(複合音など)に対する神経活動は、単純な刺激(ひとつの純音など)に対する活動に関する知識に基づきしばしば説明可能であるが、2つの正弦波が同時に表示された際、これでは説明できない現象が起こることがある。
- ニューロンにひとつの純音 (tone A) が与えられた際、特に純音の周波数がCFに近い時(かつ一定のレベル以上の時)には通常、ニューロンの放電率は自発性放電率より大きくなる。ここで、特定の周波数とレベルを持つふたつ目(tone B)の純音を与えられた時、先のニューロンの放電率が減少することがある(図9.11-a)これは、tone Aによる神経活動をtone Bが抑制あるいは抑圧するような働きであるので、2音抑圧と呼ばれる。
- 図9.11-bは、tone Aに対する放電率を減少させるようなtone Bの周波数と強度を表す。濃い線はtone Aに対する同調曲線、薄い線で挟まれた領域、抑圧野はtone Aを抑圧するtone Bの周波数とレベルを表す。つまり抑圧野は同調曲線のすぐ上と下の領域。
- 2音の組み合わせによってはneural synchronyの低下も生じる。これをrate suppressionあるいはsynchrony suppressionと呼ぶ。



a)



b)

FIGURE 9.11

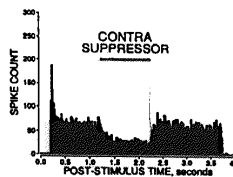
TWO-TONE SUPPRESSION AND OTHER NONLINEAR NEURAL RESPONSES 続き

- 2音抑圧は、音刺激に対する蝸牛の非線形反応に起因する現象のひとつ。他には本書の5章(差音(pp.58-60)); 7章(入力レベル - 基底膜速度, p.101), 14章でも紹介。高い刺激音レベルにおいて、高調波ひずみや差音、特に3次差音(2f1-f2)に関連するひずみが、聴神経の反応において計測されている。
- 2音抑圧や他の非線形現象は神経放電だけでなく、biomechanical motionや蝸牛マイクロフォン電位でも見られる。それゆえ、これらの非線形現象は神経だけに起因して起こる現象ではないと思われる。
- 2つ以上の音の間での興奮や抑圧は同時的な音でも継時的な音でも起こる。

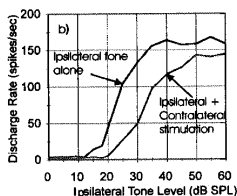
FUNCTION OF THE EFFERENT SYSTEM

- 遠心系の正確な機能は未だ良く分かってはいないが、電気生理学的証拠からはこの系が本質的には抑制的に働くことが示唆される。
- 交差オリブ蝸牛束へ電気あるいは音刺激を与えた際、単一の聴神経の放電、AP、及びCMに影響を及ぼすことが分かった。
- オリブ蝸牛束に関する多くの研究は、(測定する側と)同側の耳に音刺激を与えた上で、対側の耳に音刺激を加える、あるいは対側の交差オリブ蝸牛束(通常、第4脳室)に短い電気ショックを与えることにより行なう。
- 図9.12aは、同側の耳に4.37kHzの音(3.5sec)を与え、その途中で対側の耳に広帯域ノイズ(1sec)を加えたときのPSTヒストグラムを示す。対側の耳にノイズが加えられている間、同側の耳の刺激音に対する反応率が抑圧されていることがわかる。それゆえ、対側の耳に加えられたノイズが交差オリブ蝸牛束を刺激し、それが同側の耳への音に対する反応を抑制したかのように思われる。
- 図9.12bは、対側の耳への音がある場合(薄い線)となし場合(濃い線)のrate-level関数を示しており、どの(同側の耳の音の)レベルにおいても、前者が後者より放電率が低いことがわかる。交差オリブ蝸牛束に電気刺激を与えた場合も同様。
- SUPPLEMENT末: 図9.12a(第2音は対側)は図9.11a(第2音は同側)と似ている。対側の耳への刺激に際しては、実験者はその音が頭を通して漏れて同側の耳への刺激となってしまう可能性を排除するよう注意しなければならない。もしそのような漏れが生じるならば、計測された抑圧は交差オリブ蝸牛束の活動よりは、2音抑圧による結果を見ていくことになるかもしれない。

FIGURE 9.12



a)



FUNCTION OF THE EFFERENT SYSTEM 続き

- 抑制効果は、耳にダメージを与えるようなノイズに対し神経系を保護するため、強いノイズが与えられた状況における神経活動を低下させるのに役立つのかもしれない(16章) また、背景音による効果を減らして所望の音を検知するのに役立っているのかもしれない(11章)。
- {workbookのhints: 上記には時間がかかるので、大きな音に対する耳の保護に関しては、遠心系の抑制機能は、時間遅れ(時間的余裕)があるときのみ役に立つと思われる。}

ENCODING OF FREQUENCY, INTENSITY, AND TIME

(p.140右段1行目: "simulates" "stimulates")
 前述のように、聴神経は周波数局在的に構成されているので、入力音の周波数は、聴神経線維のどの位置の線維の放電率が最大であるかによって決めることができる。この考え方を、場所説と呼ぶ。
 図9.6, ISIヒストグラム、ピリオドヒストグラムのデータで示したように、聴神経線維はまた、ほぼ5,000Hz以下の音に対しては、周期に調和(同期)して放電する。従って、放電率の周期性が入力音の周波数を決めるのに使われ得る。この考え方を、周波数符号化の時間説と呼ぶ。
 音の強度は放電率の増加として符号化されると思われる。このことは、各々の聴神経線維の放電率が増加することを意味する。しかし図9.2に示したように、単一ニューロンの放電率は、比較的狭い入力レベルの範囲(通常、35 dB未満)に対してのみ増加する。そのため、単一ニューロンのダイナミックレンジでは、我々が持つダイナミックレンジ: 140dB(10章)を説明できない。これは、たくさん神経線維における放電率の増加によって音の強度が決められていることを示唆する。例えば、図9.3(図9.2の関連1)に示す3つ(閾値: H.M.L)の神経線維の出力を組み合わせることで、より広いダイナミックレンジの音のレベルの符号化ができる。
 刺激音周波数から離れたCFを持つ線維の出力を組み合わせることもによって、ダイナミックレンジの広さの説明の助けになる(例えば、図9.5のtail)。
 脳幹における処理も(15章)140dBに及ぶダイナミックレンジ(動物の多く)の符号化を助けているかもしれない。

ENCODING OF FREQUENCY, INTENSITY, AND TIME 続き

神経の放電が刺激音の位相に同期していることが、音の位相や動的なタイミングの符号化の元を与えていることは明らかだが、この同期はほぼ5,000Hz(周期0.2ms)までに限られている。
 図9.13aは、"bet"の合成音声における母音e(図9.13bは振幅スペクトル)に対する、数百の神経線維放電率の平均(最大放電率で正規化)を示す。図の横軸は各線維のCF。2レベルの刺激音声 が用いられた。
 図9.13bに現れている主なピークはホルムント周波数と呼ばれる、これは声帯で生成された音が、口、舌、のど等でフィルタリングされた結果生じる(14章)。ホルムント周波数は(人による)音声認識の鍵となる音声上の特徴であるので、ホルムントピークが聴覚末相系の出力において適切に表現されていることは重要である。図9.13aでは、低い刺激レベルに対してはOKだが、高いレベルでは多くの情報が失われている。
 音声のレベルが変化しても認識能力はそれほど変わらないので、刺激レベルが高いときも低いときもスペクトル情報を保存しているはずである。

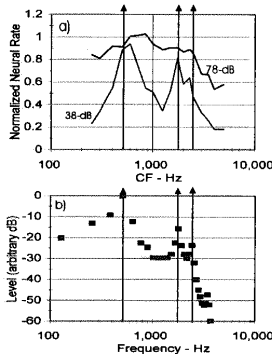


FIGURE 9.13 (ネコ)

ENCODING OF FREQUENCY, INTENSITY, AND TIME 続き

図9.14に、図9.13bで示した母音eに対して計算した、ALSR(Average Localized Synchronized Rate)を示す。ALSRは、刺激音の周期に対するニューロンの同期と放電率を考慮した指標である。図を見ると、低いレベルだけでなく高いレベルに対してもスペクトル情報を保存していることがわかる。
 聴神経は放電率が飽和する程度の高い入力レベル(図8.4: CMのinput-output関数、図9.2、図9.3: rate-level関数)においても同期するので、このような比較的高いレベルの音声に対するスペクトル情報の神経的な符号は、放電率の情報と同期の情報(統合した情報)として保持される。
 聴神経の集団としての反応は、特に低いレベルの刺激音に対し、聴神経がその音の物理的特性を非常に忠実に符号化していることを示唆する。脳幹において更なる先鋭化が行なわれる。

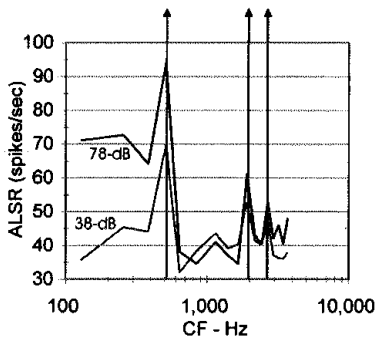


FIGURE 9.14

Modeling the Auditory Periphery

聴覚末相系の働きをシミュレートする計算モデルを開発できる程度には、内耳と神経線維の働きは明らかになっている。
 計算モデルは以下2つの段階の手続きからなる。
 1 段目: 進行波に対する基底膜の周波数選択性のシミュレート
 2 段目: 1 段目で得られた振動出力を神経放電パターンに変換する。内毛細胞のシミュレート
 図9.15は、図9.13bに示した母音eに対する出力のシミュレーションを示す。
 縦軸左は各聴神経線維のCF、縦軸右はあぶみ骨からの距離。両尺度はほぼ対数的に対応。
 横の各線は、各線維の0-50msのPSTヒストグラム。
 図に見られる低い周波数チャンネルでの時間遅れは、進行波が蝸牛頂まで到達する時間(CFに対応する基底膜上のそれぞれの位置に到達するまでの時間)に対応する。
 図は、550Hz付近のCFを持つ線維のPSTの活動が高いことを示すが、これは図9.13bで示したeの第1ホルムント周波数(500-550Hzの周波数領域)に対応。
 各線の周期性(10ms、つまり100Hz)は、母音eの時間領域における周期性(14章音声生成)を表している。(男性の声帯振動の基本周波数程度)。
 それゆえ、この図はeが発音(を聞いた)ときの、神経核まで至る神経線維の活動のシミュレーションを表す("stimulation" "simulation")。

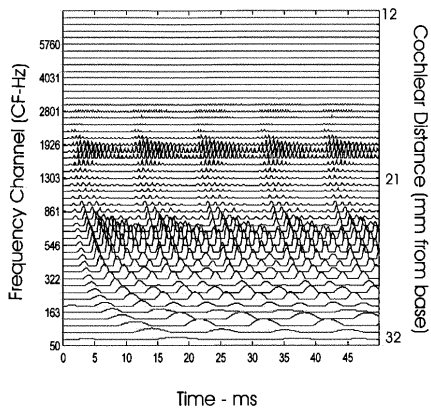


FIGURE 9.15

SUMMARY

- ・自発性放電率が小さい聴神経の閾値は大きく、自発性放電率が大きい聴神経の閾値は小さい。
- ・*rate-level*関数で示したように、閾値より20-50dB大きいレベルにおいては、刺激音のレベルが大きくなるほどニューロンの放電率が大きくなる。
- ・各々のニューロンは広い周波数に反応するが、ひとつの周波数に対し最もよく反応する。この周波数を特徴周波数 (CF) と呼ぶ。CFは応答野や同調曲線の測定で明らかにされた。
- ・高いCFのニューロンは低いCFのニューロンより広い周波数範囲に反応。
- ・CFが5kHz以下の線維は低い周波数の刺激音の周期に反応 (同期) することが、ISIヒストグラムとビリオドヒストグラムの測定で明らかにされた。
- ・2音抑圧は、複合音刺激に対して生じる非線形な相互作用の一種。
- ・交差オリーブ蝸牛束 (遠心性) を刺激すると、単一のニューロンの放電率が減少する。
- ・音の周波数の符号化の説明には、場所説、及び時間説がある。
- ・強度は、放電数として符号化されていると考えられる。また、5 kHzまでは位相、及び動的なタイミングの情報は、音の位相に対する聴神経放電の同期として保存される。
- ・聴覚末相の神経出力をシミュレートできる計算モデルも作られている。

Chapter 10 Auditory Sensitivity

独立行政法人 産業技術総合研究所
中川 誠司

Summary

- 聴力閾 (聴覚閾, 可聴閾, 検知閾, thresholds of audibility) は音の強さと周波数に対する聴覚系の感度を表す
- 聴力閾は 最小可聴音場 (Minimum Audible Field: MAF) もしくは 最小可聴音圧 (Minimal Audible Pressure: MAP) によって計測される。それぞれにキャリブレーションを必要とする
- その際の純音の duration は重要
 - duration が 10 ms から 300 ms の場合, 検知閾はエネルギーに依存
 - duration > 300 ms では, 検知閾はパワーに依存
 - duration < 10 ms では, 周波数スペクトルが広がってしまうため, より多くのエネルギーが必要
- このような duration の効果は, 聴覚系の“時間的積分 (Temporal Integration)”で説明できる
- 周波数弁別閾は, Weber-Fechner 則に従う。 $\Delta f / f = 0.002$ (0.2%)
- 強さの弁別閾も (広帯域雑音は) Weber-Fechner 則に従う。 $\Delta I = 0.5 \sim 1.0$ dB
- 時間弁別閾の推定は, duration の変化に伴う検知閾, 周波数スペクトルの影響を受けるために困難
- TMTFs (時間的変調度伝達関数) は, 変調周波数の違いによる振幅変調の検知閾の変化を記述するために用いられる
 - 音の振幅変化 (m) の検知閾を, その変化の速さ (Fm) の関数として計測したものを

Thresholds of Audibility (1/6)

- はじめに
 - 聴覚の理解のためには, **周波数, 振幅, 初期位相**に関する聴覚閾 (可聴閾, 検知閾)を理解することが重要
 - すべての刺激音は正弦波の組み合わせで表現可能
正弦波に対する聴覚感度から, 聴覚感度の全容が推測可能
- 初期位相
 - 片耳に純音が示されただけでは, その初期位相を感知不可能.
 - しかし, 聴覚が位相を感知できないというわけではない.
 - 周波数の異なる二つの正弦波が示された場合, 音質(?)は周波数が変化した始めた初期位相に依存して変化
 - 両耳間の位相差が変化すれば, 音像も変化(Ch.2 および Ch.12 参照)

Thresholds of Audibility (2/6)

- 周波数と振幅
 - 正弦波の知覚に必要な強さを周波数ごとに調べることで, 周波数と強さに関する聴覚感度を計測することが可能
 - どんなに強い正弦波を与えても知覚できない周波数帯域が存在
= 周波数に対する感度限界
 - ある周波数の純音の最小可聴音圧: 聴力閾 (thresholds of audibility)
 - 聴力閾の計測方法 (Fig 10-1)
 - 精神物理学的手続きによって, psychometric (精神物理)関数を作成
 - あるパフォーマンスレベル(正答率75%が一般的)を選択
 - 高感度 = 低聴力閾

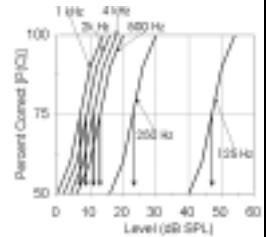


Fig. 10-1

Thresholds of Audibility (3/6)

- Minimum Audible Field (MAF)
 - 純音の最小可聴音圧
 - 自由音場
 - マイクロナンを被験者の位置 (両耳の間点) に設置してキャリブレーション
 - 被験者は音源から 1 m の距離で, 音源に正対
 - 両耳聴
- Minimal Audible Pressure (MAP)
 - 鼓膜における最小可聴音圧
 - イヤホン
 - 様々な手法で(?)鼓膜における音圧を計測
 - 鼓膜の位置における音圧のキャリブレーションは困難
 - テストカブラー
 - 健聴耳の音響特性
 - マイクロナンを内蔵し, 鼓膜の位置での音圧を推定可能
 - 等価参照音圧閾 (Reference Equivalent Threshold Sound Pressure Levels: RETSP_Ls) - マイクロナンとテストカブラーの組み合わせに存在

Thresholds of Audibility (4/6)

- 3種類のイヤホン

名称	装着方法	キャリブレーション方法
Supra-aural phone	耳介を覆う	6 cc カブラー (成人のイヤホン-鼓膜間の外耳道の容積)
Circum-aural phone	耳介および周囲を覆う	Artificial ear
Insert earphone	外耳道に挿入	Occluded ear-simulator / 2 cc カブラー (イヤホン先端-鼓膜間の外耳道の容積)
- MAF thresholds は MAP thresholds より小さい (Fig 10-2)
 - 頭部による回折
 - 外耳道による共鳴

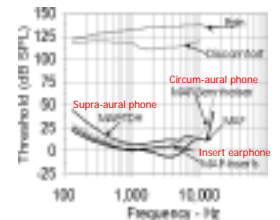


Fig. 10-2

Thresholds of Audibility (5/6)

- 1000Hz以下、4000Hz以上での感度低下 (Fig 10-2)
 - 1000 Hz 以下: 6 dB/Oct
 - 4000 Hz 以上: 24 dB/Oct (ただし、8000 Hz 以上の検知閾計測は困難であるため、高周波側の減衰勾配の推定はより困難)
 - 外耳、中耳の音響関数を左右する要因を示唆
 - 中耳のゲインの周波数特性 (Fig 6-8)

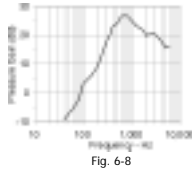


Fig. 6-8

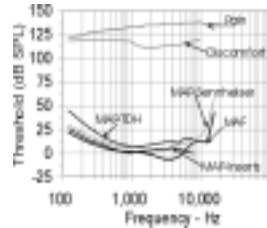


Fig. 10-2

Thresholds of Audibility (6/6)

- 20 Hz ~ 20,000 Hz の音をすべて聞くには、高性能のスピーカが必要
- Sound Pressure Level
 - 0 dB SPL = 20 μ Pa
 - このときの鼓膜の変位は水素原子の直径程度
- Hearing Level
 - 健聴者の平均値 = 0 dB HL
 - Supra-aural phone、1000 Hz の場合、30 dB HL = 37.5 dB SPL
- 聴覚系が耐えられる最大の音圧 (最大可聴音圧レベル)
 - 被験者が“音を(皮膚で)体感する”、“痛みを感じる”、“ちくちくするような感じがある”場合の音圧
 - 不快になる音圧 (不快音圧レベル)
 - 120~130 dB SPL
 - 周波数によって、あまり変化しない
- ダイナミックレンジ
 - 最大可聴音圧と最小可聴音圧の差
 - 1000 Hz: 125-135 dB、100 Hz: 80-90 dB

Duration (1/4)

- duration が大きいほど、簡単に知覚できる (検知閾が低下する)
 - 時間 T, 強さ(パワー) P, エネルギー E
 - $P = E/T$ and $E=PT$

duration (T) を変化させている場合、P および E の両方を一定に保つのは不可能

- Fig 10-3
 - duration が大きいほど、小さなパワーで検知可能
 - duration が 250-500 ms 以上では、あまり変化しない
 - 250 ms 以下では、変化が大きい
 - duration が 1/10 検知閾が 8-10 dB 上昇
 - つまり、信号のエネルギーはほぼ一定

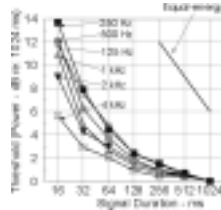


Fig. 10-3

Duration (2/4)

- 時間 T, 強さ(パワー) P, エネルギー E
 - $E = PT$
 - $10 \log E = 10 \log P + 10 \log T$ ($T < 1 \text{ sec}$)
- $T < 1 \text{ sec}$ では、duration が小さいほど、検知するために大きなパワーが必要
 - 例: 500 ms トーンのパワーが 80 dB のとき、そのエネルギーは
 - $E \text{ (dB)} = P \text{ (dB)} + 10 \log T$
 - $= 80 \text{ dB} + 10 \log (500 \text{ ms} / 1000 \text{ ms})$
 - $= 80 \text{ dB} + 10 \log 0.5$
 - $= 80 \text{ dB} - 3 \text{ dB}$
 - $= 77 \text{ dB}$
 - duration が 500 ms から 50 ms に減少した場合、エネルギーは 77 dB で一定に保たれるので、
 - $77 \text{ dB} = P \text{ (dB)} + 10 \log (50 \text{ ms} / 1000 \text{ ms})$
 - $= P \text{ (dB)} + 10 \log 0.1$
 - $= P \text{ (dB)} - 10 \text{ dB}$
 - or $P \text{ (dB)} = 87 \text{ dB}$

Duration (3/4)

- duration が 10 ms から 300~500 ms の場合、エネルギーはほぼ一定。しかし、この法則は“およそ”のものであって、周波数に依存。
- $T < 10 \text{ ms}$ では、エネルギーを一定に保つ以上のパワーが必要。
 - ON/OFF を伴う純音のスペクトルは広帯域になる (Ch.4 参照, 図4-3)
 - 純音の duration が短いほど、スペクトルは広がる
 - ごく小さな duration ($T < 10 \text{ ms}$) では、スペクトルはかなり広がる。可聴周波数帯域外のエネルギーは、検知に貢献しなくなる
 - 純音の周波数が低い場合は、検知に貢献しない周波数帯域のエネルギーの割合が減るため、検知されやすくなる

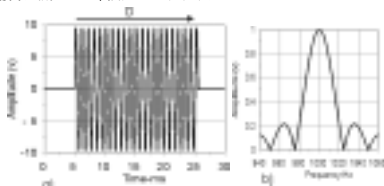


Fig. 4-3

Duration (4/4)

- 聴取閾を決定する際の信号の duration はとても重要
 - 300 ms ~: 聴取閾は強さ(パワー)に依存
 - 10 ~ 300 ms: 聴取閾はエネルギーに依存
 - ~ 10 ms: スペクトルの広がり、検知閾の決定を困難にする

Temporal Integration

- Temporal Integration (時間的積分, Fig 10-4)
 - a) 検知に必要なエネルギー
 - b) a) より duration が小さい分, パワーが増加
 - c) duration がある値より長くなると, エネルギーは積算されない
- エネルギーの積分は 300 ms (the integration time, 積分時間) まで
 - Duration が 300 ms 以下ならば, 検知に必要な信号パワーは増加する必要あり
 - Duration が 300 ms 以上でも, 検知に必要な信号パワーは減少せずに一定
- Integration time は刺激条件によって変化
 - 純音: 300 ms
 - クリック: 数 ms
 - 1~2 ms のこともあれば, 500 ms を超えることもある

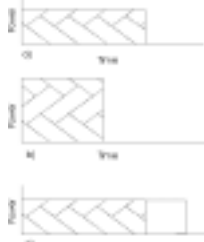


Fig. 10-4

Differential Sensitivity (1/2)

- 聴力閾 (thresholds of audibility) は, 刺激の強さや周波数の“変化”に対する感度 (弁別感度) を表しているわけではない
- Weber の観察
 - 1 ボンドと 2 ボンドを区別するのは容易
 - 100 ボンドと 101 ボンドを区別するのは困難
 - つまり, 二つの物体を区別するために必要な重さの違いは, 小さいほうの重さの値に比例
 - 1 ボンドの重さに対する Just-noticeable difference (jnd = differential limen, DL, 丁度可知差異, 弁別閾) は 0.1 ボンド
 - 100 ボンドの重さに対する jnd は 10 ボンド
- Weber's law (Weber-Fechner law)
 - $\Delta S/S = \text{constant}$ $\Delta S: \text{jnd}, S: \text{比較される小さいほうの値}$
 - $\Delta S/S: \text{the Weber fraction (ウェーバー比)}$

Differential Sensitivity (2/2)

- 弁別感度の推定には注意が必要
 - 純音が短い場合, または突然 ON/OFF された場合の周波数スペクトルは広く分布 (energy splatter)
 - このようなエネルギーの分散は, 周波数, 強さ, 位相が急峻に変化した場合にも起こる クリックが発生
 - 強さ, 周波数, 位相の変化を被験者に申告させても, 単純に弁別感度のみを調べることができない
- 多くの研究では, 二つの純音を使用した方法がとられる
 - Standard tone と Comparison tone
 - 二つの純音は, 強さ, 周波数, 位相などをわずかに変化させる

Frequency Discrimination

- Fig. 10-5
 - 周波数 f の純音が与えられたとき, 周波数変化を検知できる弁別閾 Δf
 - Δf は f に伴って増加
 - 中域の周波数では, Weber 比がほぼ一定に保たれている (Fig. 10-6)
 - $\Delta f / f = 0.002$ (0.2%)
 - すなわち, 低周波域においては, Δf が 1 Hz 程度でも弁別可能
 - when $f = 800 \text{ Hz}$
 - $\Delta f = f * 0.02 = 1.6 \text{ Hz}$

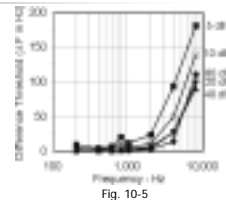


Fig. 10-5

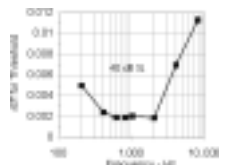


Fig. 10-6

Intensity Discrimination (1/3)

- Fig. 10-7
 - 強さ I の純音が与えられたとき, 音圧変化を検知できる弁別閾 ΔI
 - ΔI は広い周波数にわたって 0.5 ~ 1.0 dB 程度
 - すなわち, 音圧, パワー, エネルギーのどれに関しても, $\Delta I/I$ はおよそ一定 (?)

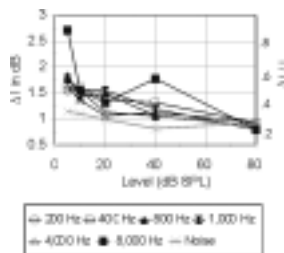


Fig. 10-7

Intensity Discrimination (2/3)

- Weber 比を dB 表記
 - $\Delta I/I = c$ (a constant)
 - $\Delta I/I + 1 = c + 1 = K$ (another constant)
- すなわち
 - $\Delta I/I + 1 = (\Delta I + I)/I = K$
- dB で表すと,
 - $10 \log [(\Delta I + I)/I] = 10 \log (\Delta I + I) - 10 \log (I) = 10 \log K = C$

Intensity Discrimination (3/3)

Fig. 10-7 をよくみると

- 広帯域ノイズに対しては、 ΔI (dB) は一定
- 純音の場合は、 I の増加に伴って ΔI (dB) がわずかに減少。つまり、レベルが高くなると、弁別能が上昇。
- 1000 Hz 純音では
 $\Delta I = 1.5$ dB when $I = 20$ dB
 $\Delta I = 0.7$ dB when $I = 40$ dB
 $\Delta I = 0.3$ dB when $I = 80$ dB
- ウェーバー比へのニアミス (Near miss to the Weber fraction)

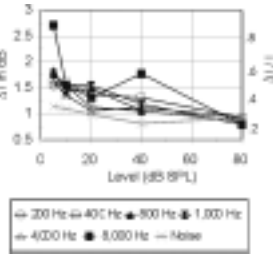


Fig. 10-7

Temporal Discrimination

- duration の変化は、duration 以外の変数 (検知閾、周波数スペクトル) にも影響する - これらの変数が時間弁別能の研究を困難にしている

Acoustic Markers を使った実験方法

- 被験者には、standard stimulus, comparison stimulus を呈示し、両者におけるマーカーの間隔を比較させる

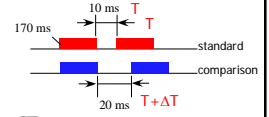


Fig. 10-8

- T (temporal separation) が増加するに従って、弁別閾 ΔT も増加
- しかし、Weber 比 ($\Delta T/T$) は一定でない
 - T = 1 ms のとき $\Delta T/T = 2$
 - T = 300 ms のとき $\Delta T/T = 0.1$
- 弁別に必要な時間差 (ΔT) は、T の増加に伴って増加するが、厳密な時間特性は様々な刺激条件に依存する

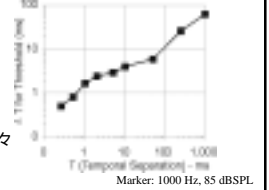


Fig. 10-8

Temporal Modulation Transfer Functions (1/4)

- Fig. 10-8 の実験は、時間特性のうちの一つだけを変化。しかし、実世界の音はもっと激しく変動している

- Sinusoidally amplitude-modulated (SAM) wideband noise (正弦波状に振幅変調した広帯域雑音) を用いて、時間変動 (変動する時間変化?) への感度を計測

- SAM 定式化 (Fig. 4-16, Equation 4.6)

$$x(t) = [1 + m \sin(2\pi F_m t)] n(t)$$

where m: 変調深度 (modulation depth)

F_m : 変調周波数

(modulation rate, modulation frequency)

$n(t)$: ノイズ搬送波 (noise carrier)

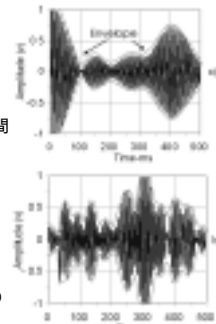


Fig. 4-16

Temporal Modulation Transfer Functions (2/4)

Fig. 10-9

- 被験者は a), b) のどちらが正弦波変調されているかを判断
- m (depth of modulation) を変化させて正答率等を観察
- 変調深度閾値 (Threshold Modulation Depth) を F_m の関数として記述

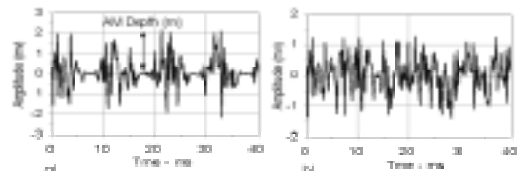


Fig. 10-9

Temporal Modulation Transfer Functions (3/4)

Fig. 10-10

- $F_m < 50$ Hz では、m がほぼ一定
- $F_m > 50$ Hz では、 F_m の上昇に連れて m も上昇 (閾値は下降?)
- 変調周波数が増加する (変調波の振幅のピーク間の時間差が短くなる) と、変調を検知するのに必要な変調深度は増大する

- 曲線はカットオフ 50 Hz のローパスフィルタのような形状

- カットオフ以上の周波数で変調深度を減衰させ、高変調周波数における振幅変調の検知を困難にする作用を持つ
- このようなローパスフィルタは、聴覚系が如何にして振幅変調信号を変換しているかを示す
- これらのデータによって TMTFs (時間的変調度伝達関数) が決定づけられる

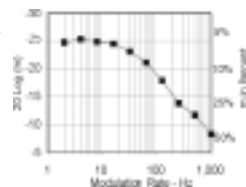


Fig. 10-10

Temporal Modulation Transfer Functions (4/4)

TMTFs (時間的変調度伝達関数)

- 音の振幅変化 (m) の検知閾を、その変化の速さ (F_m) の関数として計測
- もっとも一般的には、白色雑音を正弦波状に変調し、その検知閾を変調深度と変調周波数の関数として計測
 - 白色雑音は、変調しても長時間振幅スペクトルが変化しない

Summary

- 聴力閾 (聴覚閾, 可聴閾, 検知閾, thresholds of audibility) は音の強さと周波数に対する聴覚系の感度を表す
- 聴力閾は 最小可聴音場 (Minimum Audible Field: MAF) もしくは 最小可聴音圧 (Minimal Audible Pressure: MAP) によって計測される。それぞれにキャリブレーションを必要とする
- その際の純音の duration は重要
 - duration が 10 ms から 300 ms の場合, 検知閾はエネルギーに依存
 - duration > 300 ms では, 検知閾はパワーに依存
 - duration < 10 ms では, 周波数スペクトルが広がってしまうため, より多くのエネルギーが必要
- このような duration の効果は, 聴覚系の“時間的積分 (Temporal Integration)”で説明できる
- 周波数弁別閾は, Weber-Fechner 則に従う。 $\Delta f / f = 0.002$ (0.2%)
- 強さの弁別閾も (広帯域雑音は) Weber-Fechner 則に従う。 $\Delta I = 0.5 \sim 1.0$ dB
- 時間弁別閾の推定は, duration の変化に伴う検知閾, 周波数スペクトルの影響を受けるために困難
- TMTFs (時間的変調度伝達関数) は, 変調周波数の違いによる振幅変調の検知閾の変化を記述するために用いられる
 - 音の振幅変化 (m) の検知閾を, その変化の速さ (Fm) の関数として計測したもの

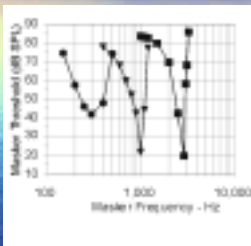
Masking

- 音は単独ではめったに存在しない
- マスキング研究は音の相互作用に関係
- 1つの刺激が他の刺激の知覚で引き起こされる干渉量に関心がある
- 閾値の変化は干渉量の典型的な度合い

Tonal Masking(1)

- リスナーに1000Hz, 5dB SLのトーン信号を検出させる
- マスキングトーン(周波数は通常, 信号とは異なる)は信号と同時に与える
- 信号の閾値検出に通じるマスキャーレベルはマスキング量の尺度として使用される

Tonal Masking(2)



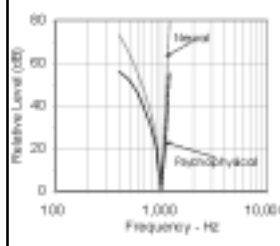
信号 + マスキャーに引き続いて, マスキャー
マスキャーに引き続いて, 信号 + マスキャー
のどちらかが呈示される

そして1回目か2回目どちらに信号があったかを2者択一(強制選択)

信号周波数に近いときマスキングとして有効でない

第8章(9章? Fig. 9.5?)での同調曲線を考慮する

Tonal Masking(3)



刺激の周波数がbest frequencyと異なるとき, より高いレベルの刺激がニューロンを駆動するために必要

マスキングによるものは, 神経の同調曲線と区別して「心理物理的な同調曲線」という

心理物理的と神経の同調曲線は中心, テスト周波数で鋭い頂点を有して, 高周波側は低周波側よりも急

右図は, 心理物理的と神経の同調曲線の比較

Tonal Masking(4)

周波数の近い2つのトーンがともにあるとき, 時間波形には"beating"が起こる
音の大きさに関して, $(f_1 + f_2)/2$ でトーンを聞く(各周波数を f_1, f_2 として)

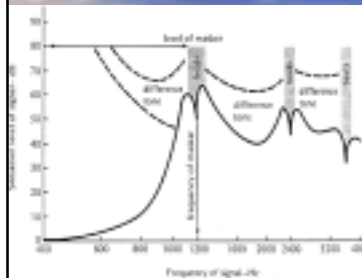
通常最も聞こえるビートは3 ~ 5Hzの速度

マスキング実験において, 信号とマスキャーが周波数において近接しているときリスナーがビートを聞くことは可能

マスキング実験においてビートの感覚をなくす方法のひとつは, 短いデューレーションの信号を呈示すること

3 ~ 5 Hzのビートは信号強弱周期が200 ~ 333 msec

Tonal Masking(5)



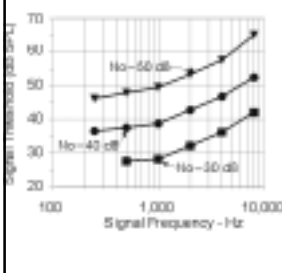
1200Hz, 80dB SLのマスキャー
リスナーは信号 + マスキャーを聞き
マスキャーと異なる周波数信号を検出

マスキャーが低い周波数よりも
高い周波数信号にマスクをかける
ので, "上方に広がるマスキング"
がある (Fig. 11.3)

影の部分はリスナーが
ビートを聞いた信号周波数
マスキャーの倍音でも起こることに注意

1200Hzのマスキャーは2400, 3600Hzの
聴覚倍音を生成した

Noise Masking(1)

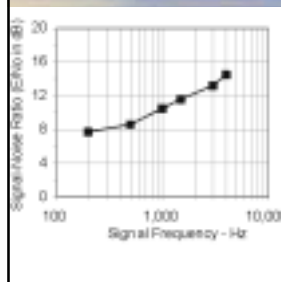


広帯域の白色Gaussian雑音は異なる周波数のトーン信号にマスクをかけるのに用いられる
NOの雑音呈示で、トーン周波数に大きく依存する

しかし、雑音のバックグラウンドレベルが上昇しても、トーンに関するマスクングの閾値は周波数にさほど依存しない

なので、レベルが上昇したら信号を検出を維持するために同じ量増加しなければならない

Noise Masking(2)

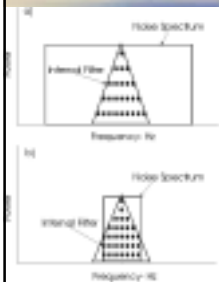


信号エネルギーは雑音スペクトルレベルよりも5~15dB強くないと、信号は検出できない

信号エネルギー(E)/雑音スペクトルレベル(N0)(S/N)はマスクングの閾値を表現するのに用いられる

E/N0は信号を検出するために5~15dBにしなければならない

Critical Band and the Internal Filter(1)



狭帯域雑音なら、その帯域の中心にあるトーン周波数の検出が簡単になるかな

逆にマスキングに信号周波数のエネルギーが無いかもしれないので、信号を非常に検出しやすいかな

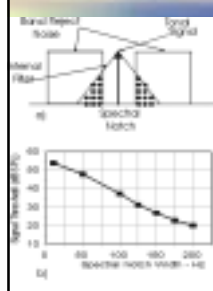
トーン信号の検出は狭帯域で存在しているトータルパワーの大きさに従う

図11.6a: 雑音スペクトルが内部フィルタの通過帯域より広いので、マスキングの最大量が生じる

図11.6b: 最大量の雑音パワーより小さいものが通過するので、11.6aよりも検出しやすい

Fletcherは内部フィルタを“臨界帯域”と呼んだ

Critical Band and the Internal Filter(2)

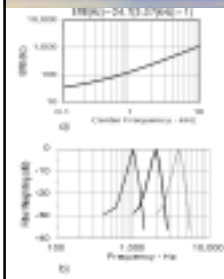


帯域排除フィルタをかけられた雑音は、臨界帯域の形状を推定するのに有効なマスキング

帯域排除雑音におけるトーン信号の検出閾値は、スペクトル幅の関数として与えられる

スペクトルの幅が広くなるにつれて、フィルタを通るパワーは小さくなり、図11.7bにみられるようにマスキング効果は小さくなる

Critical Band and the Internal Filter(3)



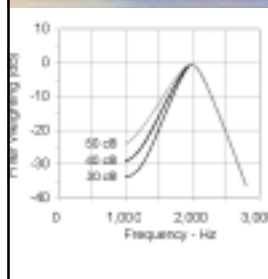
周波数に対するERBは、臨界帯域幅がその中心周波数に比例している

図11.8下は異なる中心周波数における聴覚フィルタの推定

フィルタ形状は変わらないが、帯域幅は変わる

$ERB = 24.74(4.37f + 1)$

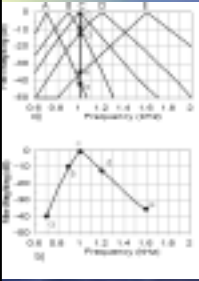
Critical Band and the Internal Filter(4)



臨界帯域幅は信号レベルを増加させることで増加する

第8章と16章で論ずる蝸牛変換の非線形性を表すかもしれない

Relationship between Excitation Patterns and Critical Bands



Best frequencyがどのくらいマスキャーに近いかが、とマスキャーのトータルレベルに依存する

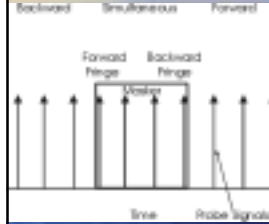
マスキングトーンは聴神経で興奮パターンをセットアップする

マスキャーによって引き起こされた興奮は周波数においてマスキャーの高低に位置する臨界帯域に広がる

図11.10は臨界帯域の内部フィルタから興奮パターンをどう得るかを説明している

11.10bの発火パターンと11.3のマスキングパターンの形の類似性に注意(どこが合ってるのかわからない)

Temporal Masking(1)



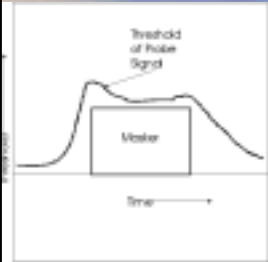
さっきまでは、マスキャーと信号は同時に発生

信号が時間内にマスキャーに先行するとき、逆向性マスキング、信号がマスキャーに続くとき、順向性マスキング

図11.11は継時マスキングの状態の図解

長方形の領域がマスキャー

Temporal Masking(2)



同時よりも、フリッジ状態の方がマスキングが起こる

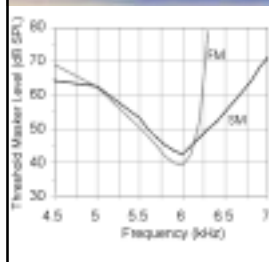
順向性の方が逆向性よりマスクされる(逆向性の方が、順向性や同時マスキングより学習によって効果が減少するらしい、音響用語辞典)

順向性は継時差 75 ~ 100 msec

逆向性は 50 msecまで
でマスキングは生じる

逆向性マスキング量は、順向性マスキングより継時分離を増加させる関数として急速に減退する

Tonal-Temporal Masking(1)



FM: 信号はマスキャーが止められた直後に示すこの図(Fig. 11.13)は、FMとSMに関する心理物理的な同調曲線の直接比較

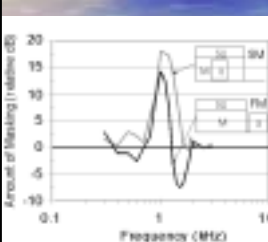
・FMの結果は信号がオーバーラップしていないとき、マスキャーが信号検出に影響を及ぼす。

・効果としては、SMとほぼ同じ。(Fig. 11.1)

・FMの方が鋭いので、FMの方がSMより良く検出できる。

Forward masking : FM
Simultaneous masking : SM

Tonal-Temporal Masking(2)



M (Masker): 1000Hz 40dB SPL

S (Signal): 1000Hz レベル変動

SU (Second tone):

周波数変動

Mより20dB高い、つまり 60dB SPL

・この図(Fig. 11.14)の0はベースラインを示すので、Mによるマスキングに加えてSUがマスキングを提供するのなら、閾値は0dBより上になる。

・SUがMと周波数が異なるにしたがってベースラインより上での閾値変化は減少。

・ちょっとSUの方がMより高いとSは検出しやすいし、しばしば、SUを抑圧トーンと呼び、その効果をunmaskingと呼ぶ

Summary

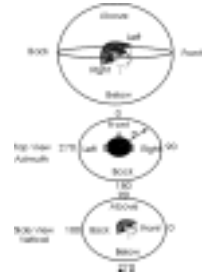
- ・低周波トーンが高周波信号をよりマスクする。
- ・ビートは、聴覚システムが限られた周波数解決力を持っているが、それは入力刺激の振幅に従う。
- ・白色雑音がトーン信号にマスキャーとして用いられるとき、信号エネルギー(E)対マスキャースペクトルレベル(N0)の比率はおおよそ5 ~ 15dB.
- ・臨界帯域とexcitation/パターンは、聴覚システムの周波数決定能力の帯域幅特性を推定するのに用いられる。
- ・フリッジマスキングと順向性、逆向性マスキングは、同時に表示されない音の相互作用を決定するために用いられる。
- ・順向性マスキングで得られる同調曲線は、同時マスキングより鋭い
- ・マスキャーが2つ以上の周波数からなると、心理物理的な抑圧もしくはunmaskingが起こりうる。

CHAPTER 12

Sound Localization and Binaural Hearing

岡本 洋輔

LOCALIZATION (方向定位)



3つの空間的次元

- (1) 水平面 (左右)
- (2) 正中面 (上下)
- (3) 距離 (遠近)

空間のある一点で観測される音信号そのものは空間的次元を持たない



よって音の定位は、音源の物理的特性(周波数, レベル, 時間/位相)を処理する聴覚システムによるものである

Fig. 12.1

Localization in Azimuth (水平面の方向定位)

音は右耳に先に到達する
→ 両耳間時間差 (interaural time difference : ITD)

音速は周波数によらず一定
周波数によって両耳の時間的な違いは同じ
しかし、両耳間位相差 (interaural phase difference : IPD) は周波数によって変わってくる

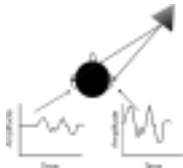


Fig. 12.2

frequency (Hz)	ITD (msec)	IPD(°)
1000	0.5	180
500	0.5	90

Localization in Azimuth (水平面の方向定位)

音は右耳に先に到達する
音の強さは右耳の方が強い (逆2乗則) $I \propto P/(4\pi r^2)$
→ 両耳間レベル差 (interaural level difference : ILD)

しかし逆2乗則による両耳間レベル差はとても小さい
→ 聴取者から音源までの距離に対して頭の大きさ (両耳間の距離) は小さいから

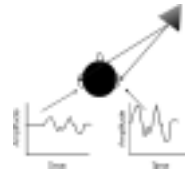


Fig. 12.2

音が高くなる (波長が短くなる) ほど、「音の影」は大きくなる
→ 大きな両耳のレベル差は高い周波数でできる

Localization in Azimuth (水平面の方向定位)

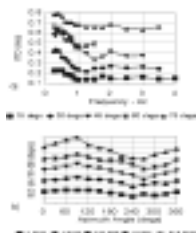


Fig. 12.3

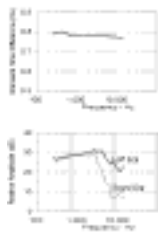


Fig. 6.5

ITD: 周波数の関数になっており、音源の位置によって0から0.8msecの値を取る
ILD: 高い周波数のとき、かなり変わる

両耳の聴覚システムは右耳がはじめて音を受け、そして右耳の刺激がより強いことを知ることで音が右側から来たことを決定できる

Localization in Azimuth (水平面の方向定位)

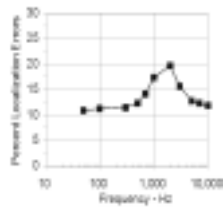


Fig. 12.4

刺激: 周波数の異なる正弦波
2000 Hz 付近 (mid-frequency region) で誤答率が

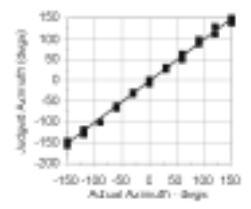


Fig. 12.5

刺激: 広帯域雑音

Localization in Azimuth (水平面の方向定位)

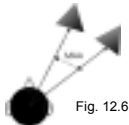


Fig. 12.6

最小可聴角度: MAA (minimum audible angles)
2つのスピーカーの位置を区別できる最小の角度
スピーカーが正面から離れるほど、区別するのにより大きな角度差が必要となる

middle-frequency で間違いが多くなる
middle-frequencyは正確に方向定位をするには両耳の時間とレベル差が小さすぎる

両耳の時間差 低い周波数
両耳のレベル差 高い周波数

→ 方向定位の二重理論 (duplex theory of localization)

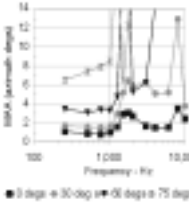


Fig. 12.7

Localization in Azimuth (水平面の方向定位)

なぜ低い周波数において両耳間のレベル差が小さいのか、どうしてそれを手がかりとして用いるのに問題があるのか
→ 頭の音の影が引き起こす両耳間のレベル差は周波数が下がるにしたがって減るから

なぜ両耳間の時間差は低い周波数においてのみ方向定位に関する情報を与えるのか

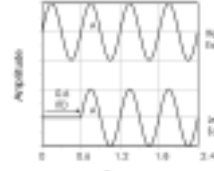


Fig. 12.8

周期 (ピーク間隔) 0.6 msec 1666 Hz
片耳からもう片方の耳へ音が伝わるのにかかる最大の時間

1番目のピーク以後 (A), 両耳の信号波形は同じ
→ 聴取者は正面からの音だと思うかもしれない

このような誤解は1666 Hz以下では起こらず、1666 Hzとそれより高い周波数でおこる

このように周波数が高いと両耳の時間差がもたらす方向定位に関する情報はあいまいになる

Localization in the Vertical Plane (正中面の方向定位)

時間とレベルについて同じ両耳間の差をもたらす位置はいくつかある
例えば、正面と真後ろ頭上 (正中縦断面)

→ 混同の円錐 (cone of confusion)

この円錐の表面上にある音源は同一の両耳間差をもたらす

混同の円錐上の音の方向定位ができる

→ 正中面に関しては両耳差のほかには何か手がかりがある

→ 頭部伝達関数 (HRTF) によって引き出されるスペクトルの手がかり (spectral cues)

頭と体の表面の部分、特に耳介が小さな音の影として作用する 高い周波数

雑音のような複合音の場合、耳介、鼻、頬などの大きさによって、違う周波数成分が、違う量だけ、減らされ、運らされる。

この減衰や遅れの量は音がやってくる方向による
例えば、耳介は前から来る音よりも、後ろから来る音をより減衰させる

→ 複合音のHRTFの形は音源の位置による

1つの耳だけで振幅スペクトルの変化は高周波の複合音の方向定位を可能にする (HRTFの主な変化は高周波数で起きるから)

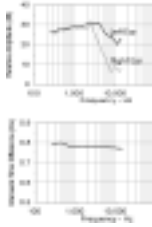


Fig. 6.5

Localization in the Vertical Plane (正中面の方向定位)

10,000 Hz 付近でHRTF にスペクトルの谷が見られる

この谷の位置や形は正中面の音源の位置によって変わる

→ HRTF の形は特に正中面において音源の位置についての情報をもたらす

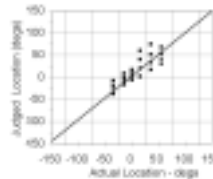


Fig. 12.9

正中面の方向定位

刺激: 広帯域雑音

結果: 水平面の方向定位ほど良くない

Localization as a Function of Distance (距離に関する定位)

音源の距離 (distance) や範囲 (range) を判断するキューはあまり知られていない

ラウドネスやレベルは強力な手がかりとなる

すぐ近く的面から来る、早く到達する音源の反射音も重要な手がかりである

PRECEDENCE-LOCALIZATION IN REVERBERANT SPACE (残響空間における先行音定位)

耳に到達する第一波が刺激の定位に重要である

聴覚システムは音源の定位には、第一波面を処理し、残響の後続波面は抑圧する

直接音はどの反射音よりも早く両耳に到達するので、直接音は音源の情報を含んでいる
これを第一波面の法則 (law of the first wave front) または先行音効果 (precedence effect)

(音が急になったり、止んだりすることに關係して、正弦波はエネルギーの拡散を減らすために音がゆっくり来たり去ったりしなければならない)

LATERALIZATION (頭内定位)

自由音場において、両耳の時間差やレベル差、HRTFによるスペクトルの違い等をコントロールするのは不可能である

より正確にこれらをコントロールできる方法の1つにヘッドフォンによる提示方法がある

多くの場合、音像は頭内にはできる → 頭内定位 (lateralization): 音像が頭内

ヘッドフォンによる両耳への提示したときの音像は時々 fused image と呼ばれる
(理由: 両耳に到達する音像が融合されたように1つの音像が知覚される)

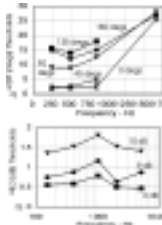
両耳の時間差や周波数が大きく違えば、fused image は知覚されない

頭内定位における fused image は最初に、またはより強い音を聞いた耳の方にあらわれる (頭外定位と同様)

頭内定位の手順は音源を定位する聴覚システムの能力における両耳の時間差とレベル差の効果を学ぶのに適していると思われる

LATERALIZATION (頭内定位)

周波数の異なる音、両耳間の位相とレベル差の関数



位相差 < 180°, fused imageは先に音を聞いた耳のほうに位置する(時間的に先)
 位相差 > 180°, fused imageは反対の耳のほうに位置する(時間的に後)
 位相差 = 360°, その像は頭の中心に位置する
 両耳間に差のない音は、頭の中央または正面に知覚される

ΔIPD fused imageの位置が変わったと知覚できる両耳の位相差の量
 ILD fused imageの位置が変わったと知覚できる両耳のレベル差の量
 (1) IPD 900Hzまで一定、その後増加 → 900Hz以上では両耳間の位相は両耳間の時間差を処理する手がかりとしては弱い

(2) 音像が片側に向かって動くにつれて(IPDが180°に近い、またはILDが大きいの)、IPDとILDも増加する

Fig. 12.10

ヘッドフォンを使った実験で、ILDが全周波数でほぼ一定ということ、自由音場で方向定位をするのに低い周波数で両耳間のレベルが使えないことを意味するわけではない

つまり、ヘッドフォンで200Hzで両耳のレベルが2dBの変化が分かっていても、自由音場でこの差は起きない自由音場で知覚できるほど大きな両耳間レベル差は起きない

LATERALIZATION (頭内定位)

両耳システムは低い周波数の刺激に対して両耳間の時間を処理する(複合音は別)

波形のエンベロープが低い周波数の繰り返し成分を持つ高周波数を提示すると、低い周波数の正弦波と同様に両耳間時間差に敏感になる

3900 Hzの搬送波を300 Hzの正弦波で振幅変調した信号は、300 Hzの同じように両耳間の時間差を聞き分けることができる

高周波だけでなく(時間領域で低い周波数の繰り返し成分を持つ音は両耳間時間差に基づいて区別される)

- ・両耳間のレベルは高い音の定位の手がかり
- ・両耳間の時間差は、低い周波数の音、または時間領域の波形で低い周波数の繰り返し成分をもつ高い周波数の音の定位の手がかりである

聴取者は両耳間の位相が3°変化したことが分かる

これは1000 Hzの音では0.01 msecの両耳間の時間の変化に対応する

聴覚システムは10 × 百万分の1秒に値する時間差に敏感である

両耳間の時間差が低い周波数や低い周波数の繰り返し成分を持つ高い音の手がかりであることは、神経システムがどのように両耳間の時間情報を処理しているかを理解するのに位相固定が重要であることを示している

LOCALIZATION VERSUS LATERALIZATION (頭外定位 対 頭内定位)

ヘッドフォンを通した音は、「頭外」よりも「頭内」に定位する

頭部伝達関数(HRTF)の振幅と位相スペクトルを反映したフィルターをかけた複合音をヘッドフォンで提示すると、鼓膜に達する音は頭や胴を通した音と同様の性質をもつ

→ 多くの聴取者はHRTFでフィルターしていない音を聞いたときよりも、外の空間で音がなっている様に聞こえる

方向判断の比較(相関係数)

Listener	Actual	Simulated
S1	0.99	0.96
S2	0.97	0.83

Actual - 聴取者が聞こえたと判断した方向と実際の音源の比較

Simulated - 聴取者が聞こえたと判断した方向とHRTFフィルターとの比較

BINAURAL MASKING (両耳マスキング)

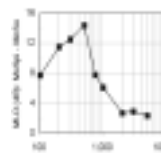


Fig. 12.11

はじめに両耳にノイズと信号を提示してマスク閾をきめる

その後、信号を片方の耳から取り去る(ノイズは両耳に、信号は片耳に)

→ 信号が知覚しやすくなり、マスク閾は下がる

モノティック(monotic) - 片耳だけに刺激が提示

ダイオティック(diotic) - 両耳にまったく同じ刺激が提示

ダイコティック(dichotic) - 両耳に異なる刺激が提示

モノティックとダイオティックでは信号のマスク閾は同じ

ダイコティックでは信号はより低い閾値を持つ

マスキングレベル差(masking-level difference: MLD): $[M_{n,S_n}]$ (モノティック)において検出するのに必要な信号のレベル - [ダイオティックやダイコティックにおいて検出するのに必要な信号のレベル]

周波数が高くなるにしたがってMLDは小さくなる → MLDは両耳間の時間とレベル差に関係あるかもしれない

ダイコティックでは、両耳に到達する音の間には両耳間の時間とレベルの差があるがダイオティックではその差がない

MLDとカクテルパーティー効果は関係があると考えられる

音を定位する能力は、音の位置を決定することだけでなく、周りの多くの音源からある音に注意をむけることにも役立つ

SUMMARY

・水平方向の方向定位において

両耳間の時間は、低い周波数と低い周波数の繰り返し成分を持つ複合音の定位の手がかりとなる

両耳間のレベル差は高周波数の音の定位の手がかりとなる

・頭部伝達関数(HRTF)がもたらすスペクトルの違いは垂直面の定位の手がかりに使われる

・ラウドネスと反射面からの初期反射音は距離に関する定位の手がかりかもしれない

・先行音効果は音の定位を決定するのに第一音が重要であることを強調する

・ヘッドフォン提示による頭内定位の実験においては、fused imageの定位は低い周波数の音における両耳間の時間とすべての周波数における両耳間のレベルに依存している

・頭部伝達関数(HRTF)によるスペクトルの違いを用いることでヘッドフォンから提示された音でも自然に起こっている音のように判断される

・ダイコティックで提示された場合、ダイオティックやモノティックよりも低いMLDとなる

SUPPLEMENT

・動く音源についての研究

ある位置にある1つのスピーカーから瞬間に音が出て、次の瞬間に違う場所にある1つのスピーカーから音が出る音が動いたように聞こえる

スピーカーから出る音の間隔を変えて動きの速さを変えることができる

・両耳ビート

わずかに異なる周波数の音をそれぞれ違う耳に提示すると両耳ビートが起こる音像が両耳間で動いているように感じる

例えば1つの耳に500 Hz、もう1つの耳に505 Hzを提示すると音像が両耳間を1秒間に5回動く

・MLDモデル

equalization-cancellation (EC) モデル: MLDの実験で、両耳のマスキングの波形は同等化され、その後、1つの耳の神経情報はもう片方のものから引かれる

MoSoで両耳に同じ波形となるので完全に打ち消される

MoSではマスキングだけが完全に打ち消される

打ち消す過程はMoSnにおいて検知される大きな信号をもたらす、MoSoにおいては検知されるものは何もなくなる → MoSnにおいて信号の閾値が低くなるのが予測される

CHAPTER 13

Loudness and Pitch

この章では強さ (Intensity) と周波数にかかわる主観的な属性であるラウドネスとピッチを扱う

産業技術総合研究所 ライフエレクトロニクス研究ラボ 添田寛治

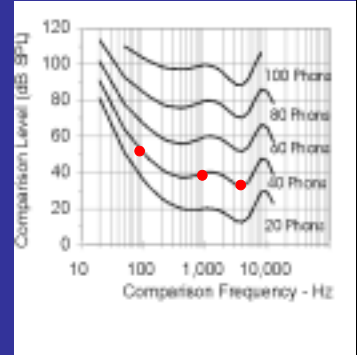
Loudness (1)

ラウドネスを研究するために、スケールリングやマッチングのような精神物理的測定法が使われる。

右図はラウドネスマッチング実験結果を示している。

リスナーは比較音と基準音のラウドネスが、知覚上等しくなるよう調節することを求められた。

それぞれの曲線上にあるすべての音は等しい大きさに感じられるので、等ラウドネス曲線と呼ばれている。



Loudness (2)

二つの用語が、刺激のラウドネスの測定や記述に用いられる。

フォンは、1000 Hzの純音と等しい大きさのdBスケール(SPL)レベルである。

1000 Hz で40 dB SPLの純音と等しいラウドネスと判断されるすべての純音は、ラウドネスレベルが40フォンである。

1ソンは40 dB SPLで提示される1000 Hzの純音のラウドネスである。

1ソンは40フォンに等しい。

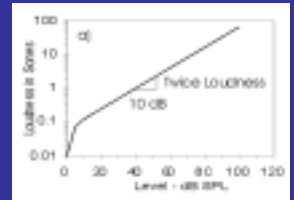
nソンの大きさの刺激は、1ソンのn倍と判断される。

Loudness (3)

右図は、1000 Hzの純音のソン対その音のdB SPLの図である。

データは、約30フォン以上で、おおそ直線回帰される。

この直線の傾きから、およそ10 dBレベルの増加ごとに、ラウドネスは倍増する。



ソンスケールは、ratio-scaling techniqueで得られる。

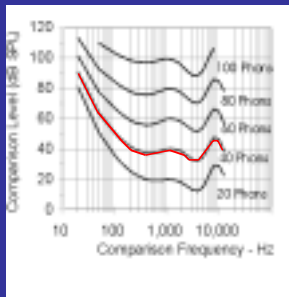
リスナーは、1ソンの基準音(1000 Hz, 40 dB SPL)の半分あるいは二倍の大きさに聞こえるよう、比較音のレベルを調節することを求められた。

Loudness (4)

40 dBの等ラウドネス曲線は、A特性で重み付けされたデシベル、すなわちdBAと呼ばれるレベル測定に用いられる。

dBAで総ノイズパワーが計算される際には、500-5000 Hzの領域の周波数成分が最も強調される。

我々が最も敏感な周波数成分のみが、最もdBA測定へ寄与する。



Loudness (4)

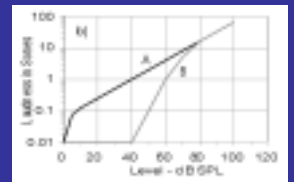
右の曲線Bは30 dBスペクトルレベルの広帯域ノイズでマスクされた純音のラウドネスを表している。

マスクされた音の閾値は、マスクされていない音の閾値よりも40 dB高い

両方の音が約80 dB SPLの時には、等しい大きさであると判断される。

このことは、マスクされた音(曲線B)のラウドネスが、マスクされていない音よりも速く増加することを意味している。

このラウドネスの上昇は、ラウドネス補充現象と呼ばれる。



Loudness (5)

音響心理学的尺度として閾値を用いて研究された大部分の相互作用は、ラウドネスを用いて研究できる。

ラウドネス尺度は時間的積分や臨界帯域幅を求めるのに使われる。

ラウドネス順応(あるいは、一過性疲労)は、長時間(数秒あるいは数分のオーダー)順応刺激にさらされる間に起こる。

刺激のレベルが一定に保たれたとしても、周波数、継続時間、強さ、帯域幅の変化すべてが、刺激のラウドネス知覚に影響を及ぼす。

強さが振動の大きさの客観的な尺度であるのに対し、ラウドネスは音の主観的な評価である。

Pitch (1)

ピッチは周波数と高い相関がある。

よく議論される周波数とピッチの関係は、蝸牛、聴神経、あるいはもっと一般的に聴覚システムの部位的構成や場所説によっておこなわれる周波数の符号化である。

リスナーはエネルギーが存在しない周波数に刺激のピッチを報告するという事実、ピッチが聴覚の場所説に関わるメカニズムに加えて、神経発火の時間的な機構を基に符号化されていることを意味している。

ピッチ実験では、通常、基準刺激(正弦波)が比較刺激のピッチ調節の基準として使われる。

刺激のピッチ変化は、大きい小さいあるいは多い少ないといった一次的に変化しているように感じられない。

Pitch (2)

一般的に、音階間の関係は、オクターブあたり12間隔のように配列される。

等平均律スケールでは、オクターブは半音と呼ばれるログスケールで12の間隔で分けられ、それぞれの間隔はセントと呼ばれる100の等しいログ間隔(音程)に分けられる。

1半音は、100セントである。従って、1オクターブは1200セントもち、それぞれの間隔は100セント、あるいは1半音である。

この表では音Aが、440 Hzの逆数となる振動の基本周期を持つと仮定されている。

Musical tone	End notation	Start frequency (Hz)	End frequency (Hz)	Periodicity (cycles)
1	1	440	440	1
2	2	220	440	2
3	3	146.7	440	3
4	4	110	440	4
5	5	88	440	5
6	6	73.3	440	6
7	7	63.4	440	7
8	8	55	440	8
9	9	48.9	440	9
10	10	44	440	10

Pitch (3)

ラウドネスは主観的な尺度であるが強さと結びつけられるように、ピッチは周波数と深い関連がある。

他の変数、例えば、レベル、波形の繰り返しや帯域幅により波形のピッチは変わらう。従って、ピッチも主観的な尺度である。

Tonality (調性)は、観測者がピッチの存在を検知していることを暗示している。

Von Békésyは、1000 Hz以下の周波数の音は、もし明確なピッチを持つのであれば、3~9周期に等しい継続時間を持つ、とした。

1000 Hz以上では、このピッチあるいは調性の境界継続時間は、その周波数に関わらず10 msである、とした。

Pitch (4)

ノイズのような広帯域刺激もまた、エネルギーがある周波数帯域に集中していれば、ピッチのような質をもつ。

調整音の周波数で表される示される知覚されたピッチは、最もエネルギーが集中する周波数領域に関連している。

しかしながら、リスナーは構成音700, 800, 900, 1000 Hzからなる刺激を、100 Hzに全くエネルギーが存在しないにもかかわらず、ピッチは100 Hzであると報告する。

複合刺激の基本周波数がスペクトル上に存在しなくても、ピッチはその基本周波数と関連するという観測結果は、ミッシングファンダメンタル現象と呼ばれる。

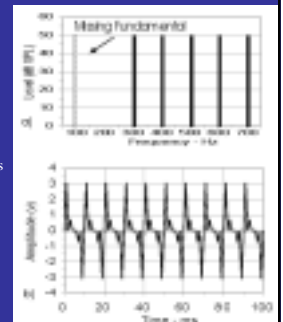
Pitch (5)

右図は、複合音(300, 400, 500, 600, 700 Hz)の振幅スペクトルと時間波形である。

複合音の時間波形の主要なピーク間には、10 msの間隔がある。

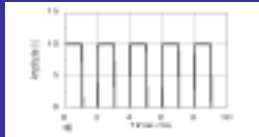
10 ms周期に対する周波数は100 Hzなので、10 msの周期性により、聴覚システムが100 Hzのピッチを知覚したと考えられる。

多くの周期的時間波形をもつ刺激は、“周期性ピッチ”と呼ばれる周期の逆数に等しいピッチを持つ。



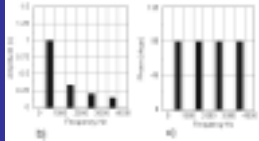
Pitch (6)

右図に示す矩形波の時間波形は、2 msの周期で発生する矩形波で構成されるので、この時間波形は500 Hzの周期性ピッチを持つ。



この場合、振幅スペクトルの500 Hzにもエネルギーがある。

複合音がピッチ周波数にエネルギーをもつ時(矩形波)や、ピッチの周期にエネルギーをもたない時(ミッシングファンダメンタル複合音)でも、周期性ピッチを知覚する。

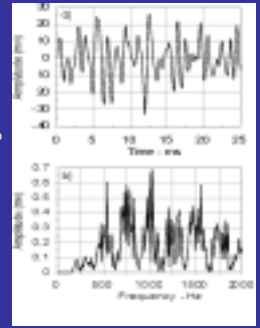


Pitch (7)

右図に示すようなスペクトラムや時間領域の波形を持つ音は、250 Hzのピッチを生み出す。

この音のスペクトラムは連続で、750、1000、1250、1500、1750、2000 Hzにノイズのピークを持つ。従って、スペクトラルのピークは、ミッシングファンダメンタルピッチの純音成分のようである。

報告されたピッチ(250 Hz)にエネルギーはなく、ノイズの波形のエンベロープに周期性は存在しない。



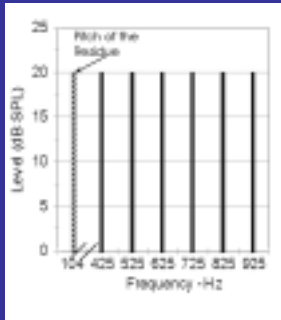
Pitch (8)

右図は、複合ピッチのもうひとつの側面を示している。

すべての周波数は100 Hzの等間隔であるが、100 Hzはミッシングファンダメンタルではない。

25 Hzがこの複合音の最も高い基本周波数であるが、リスナーはこの音のピッチを104 Hzと判断した。

このピッチはレジドゥーのピッチ移動と呼ばれる。



Nonlinear tones (1)

二次と三次の高調波(倍音)、差音、三次の差音は、最も良く知覚される非線形音である。

いくつかの正弦波からなる複合音が、特に高いレベルで提示された時、聴者は刺激の周波数に対応するピッチに加えていくつかのピッチを知覚する。

これらのピッチは、聴覚末梢の非線形特性によって生み出される倍音、差音と関連している。

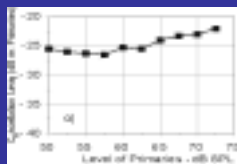
特に三次の差音は、非線形性によって生み出される最も知覚されやすい音であるため、とても興味深い。

Nonlinear tones (2)

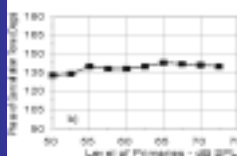
非線形音を記述するためには、周波数と同様にレベルと位相も明確にしなければならぬ。Cancellation methodは、非線形音のレベルと位相を推定するのに使われる。

キャンセレーションは、二つの音(キャンセル音と三次の差音)のレベルが等しく、位相が180°ずれている時に起こるはずである。

右図の曲線(a)は、三次の差音を生み出す二つの主音(840 Hzと1000 Hz)全体の関数として、三次の差音をキャンセルするのに必要な680 Hzの純音レベルを示している。



右図の曲線(b)は、三次の差音をキャンセルするのに必要な680 Hzの純音の位相を示している。



Other subjective attributes of sound

複合音はピッチとラウドネスに加えて、主観的な属性を持っている。

音色は、同じピッチ、ラウドネス、継続時間を持つ、ふたつあるいはそれ以上の音を異なると判断するときの、その相違に対する感覚の属性であると定義されている。

音色に加えて、音楽家は複合音の協和や不協和のことを言及する。

Stevensは、音は密度と容積の属性をもつとした。音の密度は、周波数あるいは強さとともに増加する。容積の増加は、一般に周波数とレベルの減少と結びついている。

Summary

等ラウドネス曲線とラウドネス尺度化実験から、我々はラウドネスのゾン、フォン尺度を構成できる。

ピッチの音階には、オクターブ、間隔、半音、セントがある。ピッチのメルスケールは、ピッチ尺度化実験から構成される。

ピッチとラウドネスは、物理的な対照物である周波数や強さとは完全には結びつかない。

非線形音のレベルと位相(特に、三次の差音)は、Cancellation methodによって測定される。

複合音は、音色、協和、不協和、ビート、flutter、roughness、密度、容積等の他の主観的屬性を持つ。