

TR-H-099

0008

『聴覚による認知』
—聴覚情景分析と関連する知覚現象—

津崎 実

1994. 9. 6

ATR 人間情報通信研究所

〒619-02 京都府相楽郡精華町光台 2-2 ☎07749-5-1011

ATR Human Information Processing Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1011

Facsimile: +81-7749-5-1008

『聴覚による認知』^{注1}

——聴覚情景分析と関連する知覚現象——

ATR人間情報通信研究所

第一研究室

津 崎 実

^{注1} この原稿は東大出版会から1995年刊行予定の「新・認知進学講座 第1巻『運動と知覚』」の第4章『聴覚による認知』の草稿として書かれたものである。

第4章 聴覚による認知

この章の要約

耳に与えられる信号は音が大気中を伝わる圧力の変化であるという性質から、複数の音源からやってくるものが折り重なっている。人間の聴覚系の仕事はこの折り重なった信号を分解し、それを再合成して適切な対象すなわち音源の存在を認知することである。この過程は入力された信号を末梢的レベルでの要素的表現に変換し、この要素の集合の中から関連するものをまとめ上げて一連性のある事象レベルの記述を構築し、それを対象の記述へと結びつける過程として考えることができる。この章ではまず聴覚の末梢系における信号の分解がどのような機構によって行われるかを概観する。その後分解された成分間のどのような関係が事象としてのまとめ上げをする際の有効な手がかりになるかについて、実験例を交えながら紹介する。

1 環境の中における聴覚の位置づけ

聴覚刺激の成立

私たちが通常音として感じる現象は大気中を伝わる圧力の変化、すなわち弾性波である。すべての弾性波が音として感じられるわけではなく、図4.1に示すような周波数、音圧の上限・下限が存在する。聴覚の仕事はこの範囲に与えられる弾性波のパターンを基として周囲の環境で何が生じているかを教えることである。日頃私たちはこのような音の現象から周囲で何が生じているかを知るという行為を当然のこととのように行っているため、この課題、即ち聴覚による情景分析がいかに困難なものであるかということをおそらく意識しない。そこでブレグマン(Bregman, 1990)に倣って以下のようなたとえを持ち出してみたい。

湖につながる川のある場所にハンカチをたらしっているとしよう。湖で何かが生じるとハンカチはそれに応じて動くであろう。聴覚系の仕事はちょうどこのハンカチの動きを通して湖で誰かが泳いでいるのか、ボートに乗っているのか、石を投げ込んでいるのかを言い当てるようなものである。それをたやすい課題だと思う人はそう多くはないであろう。この場合、湖は周囲の音環境、川は外耳道、ハンカチは鼓膜にそれぞれ対応する。

このような聴覚系の仕事は視覚系の仕事とどのように異なるのであろうか？この点については聴覚的刺激と視覚的刺激の発生の過程について考えてみると理解しやすい。基本的に「もの」があれば（その対象に目を向けることによって）視覚刺激は発生する。（さらに厳密を期すならば適切な光源が必要なわけであるが、通常的环境では光源は安定な状態を保つ

ている。)これに対して、聴覚刺激は「もの」が存在するだけでは成立しない。「もの」に何らかの状態の変化、すなわち運動が生じたときに成立する。寺西(1984)はこの聴覚的・視覚的情報源のそれぞれの特徴について、光と音の波長の違いに起因する空間分解能の違いに結びつけて論じており、視覚的情報源が対象(object)であるのに対して、聴覚的情報源は事象(event)であると性格づけている。例えば、バイオリンの音が聞こえるといった場合の知覚像の特徴にはバイオリンという対象だけでなくそれを弓で引く(場合によっては指で弦をはじく)という操作も伴って表現されている。あるいは対象が何であっても「叩いた」音というカテゴリーも成立する。一方バイオリンが見えるといった場合の知覚像の特徴にはそれがどのような光源で照らされているかは本質的ではない。このような考え方は聴覚による情景分析という言葉で定着させたブレグマンの音脈(stream)の概念にも共通している(Bregman, 1990, Pp. 9-11)。

さらに寺西は視覚的情報が空間的であるのに対して、聴覚的情報は時間的であるという対比も行っている。単眼(O)で観察する場合であっても外界に存在する2点X,YはX,Y,Oが直線上に存在しない限り2点として網膜上に投影される(図4.2A)。外界の2点はそれ自身が光源でない限り点滅はしないし、仮にするとしてもそれは別に存在する光源に起因するものであるから時間的には同期して生じる。対して単耳(A)にとって2つの音源X,Yで発生した振動は加算されて入ってくる(図4.2B)。そこには音源の空間内における位置の手がかりは存在しない[注1]。但し、X,Yは音源であるが故に振動の発生を独立に開始、終了しうる。

以上の対比は聴覚と視覚の安易なアナロジーを避け、それぞれの領域、および両領域間の適切な議論を進める上で念頭に置いておくべきである。しかし、聴覚も視覚も実環境下において私たち人間が適切に外界を認知するための機構である以上、極端な対比によって典型的な議論に陥らない注意も必要である。例えば、聴覚的情報源は事象であるといっても、聴覚的には対象の認知に到達しないというわけではない。むしろ一般的には聴覚的事象を何らかの対象、すなわち音源へと結びつけることによって認知は成立する。また、実際の環境では眼も耳も単独ではなく、また観察者、対象も運動する。従って、聴覚系に対しても空間的手がかりが役割を果たす割合は増えてくるし、逆に視覚系に対しても時間的手がかりは重要となってくる。ここで展開した特徴付けはあくまでも、私たちの聴覚系ではどのような情報表現を得意としているかという議論であって、聴覚情報の本質とか限界ということを問うものではない。

聴覚系に対する入力と符号化

音に対する認知が生じるためには当然のことながら、音波を神経信号に変換する過程が必要である。その変換過程を理解することは音の物理的特徴の内どの側面が神経信号として表現されているかを知ることになる。ここではその変換過程を概観する。

4.2.1において、聴覚の仕事を湖につながる川面にたらしめたハンカチのたとえを持ち出して紹介した。このたとえは聴覚がこなすべき課題の困難さをわかりやすくとらえているが、不適切な点が二点ほどある。まず第一に、一本の川しか想定しない点である[注2]。あるひとつの音源から発した音は大気中を伝わり顔の両側面に位置するふたつの耳に到達する。実環境では音源から直接到来する音波に加えて周囲の地面、床、壁などに一旦反射してから到来する成分も存在するが、ここでは直接到来する成分にまず注目してみよう。このとき主に音源と左右の耳の距離の差に応じて左右の耳で受ける音の間に0.2から0.8 ms程度の時間差が生じる(Kuhn, 1977)。さら頭部が作る「陰」の影響による両耳間の強度差も考慮する必要がある。波長が長いほど回り込みにやすいため(回折)、強度差は低い周波数ほど減少する傾向にある。頭の直径を17.5 cmと見積もるとこれに相当する波長を持つ音の周波数は約2 kHzである。これよりも高い周波数に対しては両耳間強度差は最大で10から20 dB程度生じる(Fedderson et al., 1957)。

川面にたらしめたハンカチのたとえのもうひとつの不適切な点は、それが実際の聴覚器官で言えば鼓膜(あるいはアブミ骨)の段階までしか対応していない点である。実際には鼓膜の振動が直接神経信号に変換されるのではない。鼓膜の振動はツチ骨、キヌタ骨、アブミ骨の3つの耳小骨を経由(ここでは機械的なインピーダンスマッチングが行われる)して、蝸牛という構造に伝えられる。蝸牛はまさにその名前の通りカタツムリの殻のように螺旋状の形をしている。基本的にはそれはひとつつながりの管で、前庭階、蝸牛管、鼓室階の3つの部分に仕切られている。前庭階と鼓室階は外リンパ液で、蝸牛管は内リンパ液で満たされている。アブミ骨に伝わった振動は楕円窓から前庭階を蝸牛頂へ向けて伝搬し、そこで折り返して鼓室階を円形窓に向けて伝搬する。その結果として蝸牛管と鼓室階を仕切る基底膜の上に進行波が生じる。基底膜には内有毛細胞という受容細胞が存在しており基底膜の物理的な振動を神経の発火パターンへと変換する。(図4.3参照)

基底膜は物理的に等質ではなく、楕円窓から蝸牛頂に向かってその幅を増加していく(蝸牛管の径はこれと逆の関係になっている)。こうした不等質性から基底膜に生じる進行波の包絡は高い周波数ほど楕円窓の側にピークを持つことになる(von Békésy, 1960)。この発見は聴覚末梢系における周波数分析機能の实在性の裏付けとなるものであった。しかし、観察されたピークの鋭さは実際の人間の周波数弁別能力を説明できるほどではなく、またその後観

察された一次聴ニューロンの周波数選択性はこれよりも急峻な特性を示した。

ベケシー(von Békésy)の観察は屍体から摘出した蝸牛を用いたものであったが、その後メスバウエル法を用いて生きたままの被験体の蝸牛の観察をした結果周波数選択性はベケシーの観察以上に鋭く、また入力音圧に依存して変化することが見いだされた(Sellick et al., 1982; Johnstone et al., 1986)。これらの特性は内有毛細胞とは対峙するように基底膜上に存在する外有毛細胞の能動性と関連すると考えられている(Pickles, 1982, pp.136-148)。この周波数の場所的対応のことを周波数部位再現性(tonotopy)と呼ぶが、この構造は蝸牛神経核、上オリブ核、下丘、内側膝状体、1次聴覚皮質と続く求心性の神経経路の各中継点において確認されている。

基底膜の上下動により内有毛細胞とそれに覆い被さる蓋膜との間に変位が生じ、その結果内有毛細胞の不動毛は動かされる。この動きによってイオンバランスが変化し神経発火につながる。発火の頻度は変位が大きいほど高くなる。さらにこの神経発火は基底膜の振動がおよそ5 kHz以下ならば特定の位相に同期することが知られている(Rose et al., 1967)。不応期があるので個々の細胞が振動のすべての周期で発火するわけではないが、発火間の時間間隔にも入力信号の周期性に関する重要な情報が存在していることになる。

このように音響的信号に含まれる周期性は応答する神経の違い(周波数部位再現性)としても、さらに単一の神経の発火パターンの時間関係(位相同期性)としても符号化されるのである。

聴覚による情景分析の必要性

聴覚系に与えられた入力是最終的には話し声や楽器の音、あるいは身の回りに存在する物の音として認識されることになる。外国語の聞き取りが困難であるということからもこのような最終的認識にいたるまでには学習によって獲得した知識が必要であることは自明であろう。しかし、その学習が日常的な環境で成立するためには認識の対象が引き起こした事象がほかの事象から分離して知覚される必要がある。さもないと、おそろしくヒスノイズのする語学学習テープで覚えた外国語の発音はヒスノイズつきでないと再認できないということになってしまう。この知覚的分離、あるいは知覚的群化にはどのような音響的特徴が活用されるのであろうか?本節ではいくつかの音響的特徴を利用して音の知覚的群化がどのように行われるのかを見ていくことにする。

2 音脈分凝

原始的分凝とスキーマ依存的分凝

2つの音源A,Bがあるとしよう。ある作用が加わることによってそれぞれは周波数a,bで発振する。その結果与えられる音の混合から適当な要素を選び出しその間の結びつきを形成し、音源A,Bに対応するような事象を形成することを音脈分凝(stream segregation)と呼ぶ。この時に形成された音響的事象が音脈である。

通常A,Bには独立に作用が加わるから音a,bの時間的關係も独立である。ここでa,bの時間關係が完全には独立でない場合、例えば、a,bが交替して出現するような場合にどのような知覚が生じるだろうか？まず考えられるのは物理的な条件そのままにa,b,a,b,...という交替としての知覚である。もうひとつの選択肢はaはそれ自身として一定の速さで断続し、aとは別の事象としてbが一定の速さで断続するように聞こえる場合である（このような実験室的な条件下で観察される現象は音脈化効果(streaming effect)という）。ファンノールデン(Van Noorden)はこのような事態において興味深い観察を行った。実際に彼が用いた刺激はa,bの交替ではなくa,b,a,-（ここで-は休止を意味する）という連鎖であった。このような刺激を用いることによってaとbが分かれたか否かはリズムの違いとして知覚される。即ち、分かれなない時はa,b,a-,a,b,a-,...という馬がギャロップする時のようなリズムとして、分かれた場合はa,-, a,-...とそれの倍の周期のb,-,-,-...というリズムが知覚される。彼はこのときにa,bの周波数の分離と、交替の速度を変数とした。その結果は聴き手の聴き方（注意）に依存して2種類の境界が図4.4に示すように現れた。それは、聞き手ができる限りギャロッピング・リズムを聴こうとしても分離してしまう一連性境界(temporal coherence boundary)と、aならばaにのみ注意してそのリズムを聞き出そうとしてもギャロッピング・リズムが聞こえてしまう分裂境界(fission boundary)である。重要な点はこのふたつの境界で交替速度に対する依存性が異なる点である。一連性境界の方は交替速度が遅くなるほど（2つの音のオンセット間隔が長くなるほど）より大きな周波数の違いを分離に必要とする。一方、分裂境界の方は交替速度にはそれほど依存せず半音4つ程度の周波数の隔たりがあれば分離して聴くことができる。従って音脈化効果の中には、(1)一連性境界よりも上の領域で生じる注意に依存しないものと、(2)一連性境界と分裂境界の間の領域で生じる注意による制御が可能なものの2種類がある。Bregmanは前者に代表されるような音脈分凝を原始的音脈分凝(primary auditory segregation)、後者に代表されるものをスキーマ依存型の分凝(schema-based segregation)と呼んで区別した(Bregman, 1990, pp.)。

ファンノールデンの実験結果を見る限り、原始的分凝過程は周波数・時間分解能に関連し

たフィルターモデルを想定し、例えば同一のフィルターを通過すれば同一の音脈に入り、異なるフィルターを通過すれば異なる音脈となるという機構として実現できそうである。しかし、ブレグマンとラドニキイ(Bregman & Rudnicki, 1975)の実験は局所的な周波数と時間の近接関係からだけでは分凝が制御されないことを示した。用いた刺激は図4.5に示すようなものである。聴き手は図中の比較系列内のABの部分と標準系列の順番と同じか違うかを答えることを求められる。A,Bの周波数は2,200, 2,400 Hz、妨害音Xの周波数は1,460 Hzに固定される。捕獲音Cが付かない条件、ならびにCの周波数を590, 1,030, 1,460 Hzとした4条件での成績を比べると、Cがない条件で成績は最も悪く、Cの周波数がXに近づくにつれて向上した。Cのない条件ではABはXとまとまった音脈を形成し順序の判断は困難になるが、Xに周波数の近いCを加えることによりXはCとの音脈を形成し、ABは別の音脈となることにより順序の判断は容易となるのである。この結果は受動的な一段のフィルターのみを用いるのでは分凝過程を実現できないことを示唆する。

音脈分凝が単純なフィルターモデルでは考えられない側面を示すもうひとつの実験結果はその効果が累積的であるというものである(Bregman, 1978a)。この実験では3種類の周波数を持った音、H1, L, H2, L, H1, L, H2, …の系列に対する音脈分凝の形成について調べた。H1, H2, Lの周波数はそれぞれ784, 831, 330 Hzに固定し、聴き手は系列が2つの音脈に分かれて聞こえ始めるようになるまで系列の速度を調整した^[注3]。このとき、4、8、16の音をひとつの「パック」として各パック間に4秒の間隔を空けて提示する。4秒という値は音脈の形成がリセットされるのに十分な時間とみなせる値として予備実験の結果により設定された。実験結果はひとつのパックに含まれる音の数が多いほど遅い速度でも分凝が生じるという傾向を示した。すなわち周波数軸上、時間軸上の近接関係が同等ならば長い系列ほど分凝しやすいということである。

一方のスキーマ依存型の分凝過程については、ファンノールデンの分裂境界の場合以外にも私たちの日常体験がその存在を示唆している。例えば、雑踏の中で自分の名前が呼ばれたときはそれだけが浮かび上がって聞こえるような気がする。

ダウリング(Dowling, 1973)は2つの旋律の音が交替して出現する混合旋律(interleaved melody)を提示して実験ではその過程を示唆する実験例がある。2つの旋律の存在するピッチ範囲が重なっている場合個々の旋律を聞き出すことはたとえそれが既知のものであっても困難であったが、予め旋律の名前を覚えておくことにより聴き手は旋律を聞き出すことができた。(教えられた名前の旋律が含まれない場合も用意されたが、この場合に間違っただけで旋律を聞き出してしまふということはない。)ダウリングはさらに一方の旋律の名前を教えら

れたとしても他方の旋律の聴き出しを促進することはないことも報告している。原始的分凝の場合は2つ（またはそれ以上）の音脈を同時に知覚できるのに対して、この「残差」(residue)を聴くことができないという点はスキーマ依存型分凝の特徴として考えられる。

分凝に対する規則性の役割

スキーマが有効であることは実験的に示すことができても実際に活用されたスキーマがどのようなものであるかを特定することにはならない。先のダウリングの実験を例に取れば聴き出すべき旋律を知っていることにより旋律に含まれる情報のどの側面が活用されたことになるのであろうか？この問題に一步踏み込むために、系列に備わった規則性が分凝に貢献するか否かについての研究が行われている。それは周波数変化の軌跡の追跡に関する研究と時間的規則性に関する研究である。

まず周波数変化の軌跡の追跡に関する研究について紹介する。周波数が一方向へ変化する純音の途中を雑音で置き換えた場合でも雑音が純音に比べて十分に大きければ雑音の背後に純音が途切れずに変化していくように聞こえる(Dannenbring, 1976)。この時に後続する変化音の開始周波数と変化勾配を変えるとどうなるであろうか？軌跡の追跡が完璧に行われているならば先行する変化音から外挿した点から後続部が開始した場合に最も連続的に聞こえるはずである。チオッカとブレグマン(Ciocca & Bregman, 1987)の実験結果によれば外挿点で連続性が高くなるのは先行部と後続部が同じ勾配を持つ場合で、図4.6aに示すように後続部に周波数変化がない場合には先行部が終了した周波数で後続部を開始した場合が最も連続感が強くなった。先行部と後続部の変化勾配の関係に応じて連続性を最ももたらす接続点が異なるということは、軌跡の追跡が常に行われているようなものではなく、後続部をある程度取り込んだ後に逆行的に内挿されていることを示唆する。

次に時間的規則性について見てみよう。先に紹介したブレグマンとラドニキ(1975)の刺激を注意深く見てみると捕獲音(C)と妨害音(X)が時間的に等間隔で並んでいるのに対して、標的音(A,B)はそれらとは外れた時点に存在している。Cが存在しないときにはX,A,B,Xは等間隔となる。従ってCがXと同一の音脈を形成するためにはこの時間的規則性も寄与している可能性がある。ジョーンズら(Jones, Kidd, & Wetzel, 1981)はその可能性を指示する実験結果を報告している。彼女たちはBregmanとRudmickyが用いたものと同様の時間パターンに加えて、A,BはXと等間隔となるがCとは等間隔にならない系列、C,X,A,Bのすべてが等間隔になる系列を用いてA,Bの順番に関する判断を聴き手に求めた。その結果、後の2つの条件では順序判断の成績が低下した。即ち、CによるXの捕獲は生じにくくなった。この時間的

規則性の効果は周期的な注意機構の働きによって分凝を説明しようというジョーンズの立場を支持するものである。しかし、分凝のすべてがこの注意過程によって説明できるわけではない。実際、時間的規則性が分凝に貢献しないという報告もいくつか存在する(Rogers & Bregman, 1993; French-St. George & Bregman, 1989)。ジョーンズらの実験においてもC音とX音の周波数の分離が大きくなるとC,X,A,Bがすべて等間隔で出現しても分凝の効果は現れていた。

音源位置情報の役割

音源位置に対する情報の重要性はチェリー(Cherry)がいわゆるカクテルパーティ問題を提出したときから指摘されている(Cherry, 1953)。実環境では2つの音源から発した音が左右の耳にそれぞれ入るといった単純な状態ではなく[注4]、反射音の影響や音の空間分解能の悪さも考慮すると音源位置情報というものはそれほど頑健なものとは考えられない。事実、2つの音源が、周波数、強度、開始・終了時点を全く同じにして発振した場合は1つの音像を聞くであろう。但し、音源位置情報の違いが他の分凝の手がかりと組み合わせられるとそれは分凝を促進する手がかりとなる。

両耳間マスキング・レベル差(binaural masking level difference: BMLD)の現象は空間情報が音の分離に貢献していることを示す一例である(Carhart et al., 1967)。信号とマスキャーを両耳に同位相で入れた場合に比べて信号のみを両耳間で逆相とすると信号は聞き取りやすくなる。BMLDは前者の条件でのマスキング閾から後者のそれを引いた値である。カハート(Carhart)らの実験では1ないし2音節の音声の同定を4種類のマスキャーが同時に存在する条件で求めた。信号またはマスキャーの両耳間の位相差をかえることによってそれぞれを異なった空間位置に定位させた場合課題は容易となった。但し、最もBMLDの大きかった条件は信号を両耳間で逆相とした条件であった。この条件では信号の音像は空間的に拡散したものとなる。この事実はBMDLに関する現象と空間的な定位とは必ずしも一意に対応しているわけではないことを示す[注5]。しかしながら、空間的情報の担い手として両耳間の位相関係が分凝と関連していることは確かであろう。

キューボビーとハワード(Kubovy & Howard, 1976)は周波数の違いと音源位置情報の違いを組み合わせた興味深い実験を行った。音階上の音名でいうとG, A, B, C, D, Eに相当する1オクターブ内の純音をヘッドフォンからそれぞれの音に異なる空間定位をもたらすような両耳間位相差を設定して提示した。この音のクラスターを一つだけ聞いた場合は全部の音が融合した知覚が生じた。ところが、あるひとつの音の位相差だけを最初のものに変えた第2のクラ

スターを第1のクラスターと短時間(1.5秒以下)のギャップを空けて提示すると位相差に変化を加えた音が他の音から「浮かび上がって」知覚された。

音源位置に関する情報が他の分凝情報と矛盾したような場合はどのような知覚が生じるだろうか? ドイツェ(Deutsch)はオクターブ錯覚というこの問題に関連する興味深い錯覚現象を報告している(Deutsch, 1974)。図4.7Aに示すようにオクターブ離れた音を左右の耳に交替して提示すると図4.7Bに示すような知覚像を得ることが多い。実際にはこの様な刺激に対する刺激は非常に多義的であるので、知覚システムがいつでも図に示すような解に到達するわけではないが、この「収まりの良い」解の特徴は、(1)基本周波数に応じた群が形成され、(2)それぞれの群に異なる空間的位置が割り振られているということである。このような特徴は、異なる群の音を同時に知覚できるという点において違いは持つものの音階錯覚(scale illusion)の場合にも共通するものである(Deutsch, 1975)。

3 連続聴効果と聴覚的誘導

聴覚的誘導 (パルセーション閾)

音脈化効果では交替する音の間の周波数が近接していればそれは単一の音脈を形成するように知覚されると紹介した。しかし、2種類の音の交替と物理的には記述できる条件であっても知覚的には交替とは記述されない場合がある。聴覚的誘導とはこのような現象が起きている場合をいう。

連続聴錯覚と呼ばれる現象はその中の一つである。同一の周波数を持ち振幅の異なる2音が交替する場合を想定してみよう。このような物理的的刺激に対する知覚的な「解」にはどのようなものが考えられるだろうか。まず考えられるのは大きさの異なる2音が交替するという知覚像である。次は矩形上の振幅変調がかかった一続きの音としての知覚像である。そして3番目は連続するひとつの音脈と断続する別の音脈が同時に並行して存在するという知覚像である。振幅の差がある程度以上に存在し、交替の速度が速ければ、3番目の知覚像を得る。このように周波数差やスペクトル領域上での差が存在せずに振幅の違いのみが存在する2音の交替で生じる誘導のことを単音性誘導(homophonic induction)という。

2音間に周波数差やスペクトル領域上での差が存在する場合に生じる連続聴に対しては異音性誘導(heterophonic induction)と呼んでいる。純音と帯域雑音が交替するような場合がこれに相当する。このとき雑音の中心周波数と強度、純音の周波数、交替速度を固定して純音の強度を変数とした場合、純音強度が大きい場合には連続聴は得られない。ある強度以下になると連続聴を感じるようになるが、このときの純音の強度をパルセーション閾と特に呼んでい

る。この閾値は早くから聴覚系の周波数分解能力との関連性が注目されてきた。連続聴が生じるには純音と帯域雑音の間にスペクトル領域上でのオーバーラップがなくてはならない。このことから連続聴を末梢的な神経興奮の連続性によりモデル化する考え方は多くの研究者が採用した（例えば、Houtgast, 1972; 詳しいレビューはWarren, 1984）。神経興奮の連続性モデルは知覚されたひとつの音脈の連続性自体は説明するが、連続聴の現象が生じているときのもう一つの側面、すなわち、聴覚的誘導としての側面を説明しない。その側面とは連続聴が生じているときにもうひとつの断続した音脈も知覚されているということである。ブregmanとダネンブリング(Bregman & Dannenbring, 1977)は純音とそれよりも強度の高い雑音が交替する刺激系列に対する連続聴知覚の実験をした。その中のある条件では大きさの変化が急激に生じないように雑音から純音へ切り替わる時点の純音部に振幅の傾斜を持たせた。その場合、ひとつの音が質を変化（純音から雑音へと）しているように知覚された。このことから聴覚誘導現象が生じるときには刺激系列中に急激な変化が生じる必要があることが示唆される。急激な変化がありかつ周波数領域的にオーバーラップが存在する場合に、聴覚系はその変化分（差分）を新たな（別の）事象の生起に割り当て、共通分に対して連続した知覚像を与えるのである。この差分をとる点、即ちエネルギーの再配分が生じる点は聴覚誘導の興味深い側面である。その証拠は異音性誘導では確認しにくい、単音性誘導では大きさの逆転として確認することができる。音圧差が3 dB以下であると、断続して聞こえる音（本来は振幅の大きい区間に対応している）の大きさはもう一方の連続的な音の大きさよりも小さく聞こえてくる(Warren, 1984)。

聴覚的誘導のもうひとつのタイプは対側性誘導(contralateral induction)である。ヘッドフォンの片方のチャンネルから提示した音声の音像は反対の耳に雑音を入れると雑音を提示した側にずれて知覚される(Egan, 1948)。このような効果は次のような状況を想定すると非常に理にかなっていると考えられる。誰かがあなたの正面で話しており、あなたの右側には非常に大きな音を出す騒音源があるとすると、この時、右耳のS/N比は非常に悪く音声信号はかろうじて左耳からのみ検出できるとすると、音声は左側に定位してしまっても良さそうである。ところが実際には体側性誘導の効果によって音声は完全に左側には定位せず、少しでも「本来の位置」に定位されようとするのである。このような音像の正確な定位感を判断するのは実際のところ困難であるが、ウォレンとバッシュフォード(Warren & Bashford, 1976)は巧みな方法でこの現象にアプローチした。彼らは片方の耳に信号（純音、または音声）、他方の耳に雑音を提示し、提示する耳を0.5秒毎に切り替えた。この時信号の提示レベルが検出閾からある値に達するまでは、中央に動かない信号音の音像が生じる。すなわちこの領域では対側

性誘導が生じている。その領域を超えると信号は両耳の間を交替するように聞こえるようになる。従って聴き手は聞こえの音像の位置そのものを記述するのではなく、信号音の定位位置が動くか動かないかによって判断することができる。

以上の聴覚的誘導の現象そのものは比較的古くから知られていたものが多いが、ウォレンによるマスキングに対する補償効果としての位置づけによって体系づけられ、ブregマンの提唱した「聴覚による情景分析」の発想の中においても重要な位置を占めている。

マスキング可能性の制約

聴覚的誘導現象とマスキングとは、パルセーション閾の測定が臨界帯域幅の推定に使われたことから窺えるように関連性の深い現象である。ウォレンら(Warren et al., 1972)は1000 Hz, 80 dB SPLの純音をマスキャーまたは誘導音として、いくつかの周波数の純音をマスキャーまたは被誘導音を用いて同時マスキング閾と聴覚的誘導閾を測定し両者を比較した。その結果両者に非常によい対応関係があることを見いだした。さらに彼らは1/3オクターブ幅の帯域雑音を誘導音とした場合の誘導の上限値はその中心周波数を中心とした山形になるのに対して、帯域除去雑音を用いた場合は谷型になることも見だし、マスキング特性との類似性を指摘している。但し、このような周波数選択性についてのマスキングと聴覚誘導との対応は両耳性マスキングと対側性聴覚誘導の間には存在せず、対側性聴覚誘導の場合には両耳性マスキング以外の両耳間の交互作用を導入したモデルで説明されるべきであることを示唆する(Warren & Bashford, 1976)。

文脈依存的連結と音韻修復

一連の聴覚的誘導現象を体系づけて論じたウォレンに従えば、聴覚的誘導にはさらにもうひとつ文脈依存的連結(contextual concatenation)と類別されるものがある。これは単音性誘導や異音性誘導では直前(または直後)に存在する音と同じ性質の音をそのまま連続させればいだけであるのに対して、前後の文脈にふさわしい内容をあたかも合成するかのような補完が行われる場合を指す。先に紹介したダネンブリング(Dannenbring, 1976)の周波数変化音に対する知覚的補完はその一例である。これは前後の文脈から存在すべき周波数の変化が決まる例であるが、前後の文脈から存在すべき音の時間的構造が決まる例がロール効果(roll effect)である。周波数が同じ2つの短音(40 ms)が速い速度(立ち上がり間隔で約50 ms)で交替するとき2音間に10 dB程度の音圧差が存在すると、その音は実際のテンポよりも2倍の速さで聞こえてくる。即ち大きい音の背後で小さい音が鳴っているように聞こえるのである。ロール効果は2音間に周波数の差があっても生じうる(Van Noorden, 1975)。

文脈依存的連結の最も劇的な例は音韻修復現象であろう。誰かが話している途中で別の誰かが大きな咳払いをしたとしよう。このようなときでも話し声が途切れて聞こえるということは（話している人が実際にそうしない限り）あまり経験しないはずである。ウォレンの実験では英単語legislaturesの最初のsの区間をとその両側の音韻の一部を含めた120 msを同じ長さの咳払いの音で置き換えたものを文章中に埋め込んで提示した(Warren, 1970)。被験者は咳払いを実際の置換位置に位置づけることはほとんど不可能で、さらに20名中19名の被験者が全ての音声が存在したと報告した。この現象は音声だけでなく旋律を用いても確認されている(Sasaki, 1980)。また、時間構造の錯覚については雑音と音声の間の時間関係のみならず、修復された音声の持続時間の変化も生じている可能性が指摘されている(Tsuzaki, Kato, & Tanaka, 1994)。

サミュエル(Samuel, 1981a)は雑音を音声のある区間に重畳したもの（音声信号は実際に存在する）と雑音置換したものが知覚的に弁別できないことを示してこのような補完が高次の推論過程の結果としてだけでなく知覚的レベルで生じていることを示した。このような知覚レベルでの補完が生じるためには聴覚的誘導に対するマスキング可能性の拘束が適用されることが実験によって確かめられている(Samuel, 1981b; Bashford & Warren, 1987)。また刺激の調音結合による文脈を変化させることによって修復される音韻に系統だった「異聴」傾向が生じることも見いだされている(柏野, 1991)。

但し、音韻修復現象には話し言葉として特有な側面も存在している。例えばウォレンとシャーマン(Warren & Sherman, 1974)の実験結果によれば読み間違えた音韻を雑音置換すると調音結合による手がかりは実際に読まれた音韻を支持する情報を持っているにも拘わらず意味的に適切な音韻を修復する傾向が強い。また、修復が可能な置換区間長の上限值は英文を通常に読んだ場合で300 ms程度であるのに対して逆読みした場合には150 ms程度まで低下した(Bashford & Warren, 1987)。さらにこの値は発話速度が早いほど短くなることが知られている(Bashford, Meyers, Brubaker, & Warren, 1988)。

音韻修復の現象が聴覚的誘導の特殊な例だとするとエネルギーの再配分が生じているべきである。例えば物理的には同じ置換雑音であっても修復されるべき音韻に応じて、「音声にのった雑音」としての聞こえ方が変化するはずである。レップ(Repp, 1992)によるこの問題に取り組んだ意欲的な実験の一部の結果は再配分が生じている可能性を示唆したが明確な答えはまだ得られていない。

4 共通の変化による群化

聴覚系が周波数分解能力を持っているということは、たとえひとつの音源から由来した音であっても末梢的には複数の成分に分解されていることを意味する。ここでは同時に存在する成分を群化するときの手がかりについて概観する。

調波性による群化と周波数変調効果

自然界に存在するいわゆる周期性をもった音は調波的構造を持つことが多い。すなわち、基本周波数に対して整数倍の周波数を持つ成分によって構成されている。聴覚末梢系で時間的な情報（周期性に関する情報）を伝搬しうることからこのような調波的構造は十分利用しうる手がかりとなる。

マーティン(Martens, 1984)は調波複合音のあるひとつの成分を調波的な構造からずらした場合に、その成分に対応したピッチを検出できるようになる音圧が調波的な関係を保っている場合に対して小さくなることを見いだした。例えば、200 Hzの基本周波数、平坦なスペクトル包絡を持つ、10成分の複合音を用いた場合、その第3倍音(600 Hz)に対応したピッチを聞き出すためには16 dBの音圧の増加をしなければならないのに対して、その成分を40 Hzずらした場合(640 Hz)には音圧が-7 dBまで低下させても聞き出すことができた。

このような調波構造に従った群化の存在は2つの母音が同時に存在するときそれを分離するために基本周波数の差が有効に使われることから示唆される(Scheffers, 1982; Assman & Summerfield, 1989; Assman & Summerfield, 1990)。アスマンとサマーフィールドはこの重畳母音を分離するためのモデルとして自己相関を用いて調波構造を検出する機構を提案している。

実際の環境ではひとつの音源から発する音に非調波成分が含まれることもあり得る。このような成分が通常分離してきこえないのは後述する立ち上がり時点・終了時点の同期の存在や共通の振幅変調の存在によるものと考えられるが、これらの手がかりと調波性の手がかりが同時に存在し、その間に拮抗が生じる場合にはどのように知覚されるであろうか？ブレグマンとピンカー(Bregman & Pinker, 1978)は図4.8に示すような刺激を用いて調波的構造の手がかりが他の分凝・統合の手がかり——A, Bの周波数の近接に基づく水平方向の統合、B, C間の立ち上がり時点の時間差——と競合する事態について調べた。その結果では調波性による明確な効果は確認できなかった。

共通の振幅や、立ち上がり時点の同期性はゲシュタルト心理学における「共通運命の法則」に分類できるものである。その中には共通の周波数変調の効果も含まれるが、これと調波構造による統合の間には密接な関連が生じる。即ち、共通の周波数変調の存在により統

合が促されたとしても、それは共通の変動が存在するからではなく、変動を通じて調波的構造が保たれるからであるとみなせる場合が存在する。ブレグマンとドゥーリング(Bregman & Doehring, 1984)はターゲットとなる周波数変化純音を調波関係(2オクターブの距離)にある枠組み音の中に埋め込み、その前にターゲットと同じ変化勾配を持つ捕捉音を提示するという条件の下でターゲットがどの程度捕捉されやすいかについて調べた。このとき独立変数としてターゲットと枠組み音の勾配を変化させた。ターゲットと枠組み音の勾配が一致(周波数対数軸、時間リニア軸で見た場合)してかつその間に調波関係が成立する場合にはターゲットは捕捉されにくくなった。即ち、枠組み音と統合されやすくなった。しかし、勾配が一致していても調波構造が崩れた条件では勾配が一致しない条件とほぼ同等の結果であった。これは調波性の存在により統合化が促進されることを示す。

混在する母音の中からターゲットとする母音を聞き出すときには周波数変調の存在によりその顕在性(prominence)が上昇することが見いだされている(McAdams, 1989)。但し、母音の聴取という課題の場合周波数変調の存在によりホルマント(共振点)のピークの推定の手がかりが増加する可能性(ホルマント・トレーシング; formant tracing)がある(詳細はかこみ「ホルマント構造と周波数変調の関係を参照」)。チャリキアとブレグマン(Chalikia & Bregman, 1989)の実験では2つの重畳母音間で基本周波数の変化が対数周波数軸上で平行である条件(PG条件)を設けてこのトレーシングの効果について検討している。彼らはこれと比較対照する条件として2母音とも基本周波数が変化しないもの(SS条件)、2母音間で周波数変化の方向が逆で互いに交差するもの(CG条件)を設けた。CG条件はトレーシング効果に加えて共通運命の効果が存在するか否かを問うためのものである。さらに各母音の調波構造を変化させ、調波的な条件、調波構造から一定の偏位をした条件(シフト条件)、調波構造からランダムな偏位をした条件(ランダム条件)の3条件について、第3の独立変数として2母音間の基本周波数の差を設定して母音の同定を求めた。その結果、調波構造を持つ母音に対しては2母音間の基本周波数の分離のみが同定成績の向上に寄与し、トレーシング、共通運命ともにさらなる促進効果を生まなかった。シフト条件では2母音間の基本周波数差が6半音になった場合に、SS, PG, CG条件の順に成績が向上し、トレーシング、共通運命の両方とも付加的効果があることが判明した。ランダム条件ではCG条件のみ成績の向上が確認された。

この結果から調波構造から逸脱したものは共通の周波数変調が存在すれば統合される可能性があることが窺える。しかし複合音のピッチに与える影響については共通の周波数変調が貢献しないことが報告されており(Darwin et al., 1994)、共通の周波数変調による統合可能性

については現状では結論に達していない。

同期性の重要性

複数の成分が同じひとつの事象を構成しているということは存在する時間帯を共有しているということであるから、各々の成分の開始と終了の時点が同期するかしないかが統合・分離の重要な手がかりとなることは容易に予想がつく。ブレグマンとピンカー(Bregman & Pinker, 1978)の実験では図4.8に示す成分B, C間に開始時点の(この実験では持続時間が固定されているので同時に終了時点にも)時間差をつけるほどCはBと融合されにくくなり、Cを基本周波数とする音の「豊かさ」が減少することが確認された。

開始時点、終了時点のそれぞれの効果についても調べられている(Dannenbring & Bregman, 1976)。この実験では3成分複合音の中のひとつの成分の開始・終了時点をそれぞれ独立に他の二つの成分からずらした。時間差は他の2成分に対して先行または遅延した。この複合音と交互に、ターゲットとなる成分と同じ周波数の捕捉音が提示された。聴き手はこの捕捉音のリズムを聴くように努め、その聴き易さを評定した。ターゲットが複合音に融合された場合は捕捉音のリズムを聴くことができるが、ターゲットが複合音から分離して捕捉音に捉えられた場合はリズムの速さは倍になってしまうわけである。その結果、時間差の存在によってターゲットは複合音として融合されにくくなること、さらに開始時点の時間差は先行方向、終了時点の時間差は遅延方向に存在する場合にのみ有効であることが示された。ラッシュ(Rasch, 1978)はマスカ、マスキとも複合音を用いたマスキング実験において、開始時点の時間差の効果を測定した。その結果30 msの時間差の存在により40~60 dB相当のマスキング閾の低減効果があることを報告している。この程度の時間差はポリフォニー音楽の演奏で楽譜上同時に複数の楽器が音を出す際に存在している時間差であり、個々の楽器の音色の識別に貢献していると考えられる。

共通の振幅変調の効果

開始・終了時点の同期性はより一般的には共通の振幅変調の存在として考えられる。ブレグマンら(Bregman et al., 1985)は振幅変調による統合の効果についてブレグマンとピンカー(1978; 図4.8参照)が用いた刺激配置と同様の刺激を用いることによって調べた。但し、この実験では各成分に正弦振幅変調を施した。AとBは常に同一のものとして、BとCの間の振幅変調に共通性があるか否かでB,C間の統合の程度が影響を受けるかを調べることに主たる目的であった。但し、このような実験条件では振幅変調の周波数、搬送周波数(振幅変調を施される成分の周波数)の関係に十分注意をする必要がある。例えばB,Cの周波数をそれぞれ

1500, 500 Hzとしてそれぞれを100 Hzで振幅変調した場合、その長時間スペクトルは500 Hzの成分の両側に400, 600 Hzの側帯波が存在し、1500 Hzの両側に1400, 1600 Hzの側帯波が存在するという形になる。即ち、これは100 Hzを基本周波数とする調波構造が存在するということになり、仮にこの条件で統合が促進されたとしてもそれは共通の振幅変調によるものが調波構造の存在によるものか判別できない。そこで彼らは搬送周波数、ならびに変調周波数のいずれかをシフトさせることによって調波構造を崩した刺激を含めることによって調波構造の要因を排除することを試みた^[注6]。その結果、共通の振幅変調の存在により統合が促進されることが確認された。またこの実験の結果では調波構造の有無によって統合は大きく影響されなかった。最近のダーウィンら(Darwin, et al., 1994)による実験結果では共通の振幅変調が非調波的な成分に存在することによって複合音の一部として融合されやすくなり、その結果複合音に対して感じられるピッチが変化することが確認されている。

共通の振幅変調が効果を持つということはその共通性を検出する機構、すなわち周波数チャンネル間の比較をする機構がどこかに存在することを意味する。この存在については以下に説明する共変調マスク解除(comodulation masking release; CMR)や共変調性検出差(comodulation detection defference; CDD)、変調性検出干渉(modulation detection interference; MDI)などの現象が示している。

CMRとはホールら(Hall, Haggard, & Fernandes, 1984)によって最初の例が報告された現象である。古典的なマスクングの実験では定常的な純音信号に対する検出能力を帯域雑音と同時に存在する条件下で求める。このときの雑音はランダム雑音と分類されるもので振幅包絡は平坦である。ホールらは0~10 kHzの帯域雑音と0~50 Hzの低域雑音を乗算することによって各帯域間で振幅の相関をもつ「かけあわせ雑音」を作成し、それをさらにターゲット純音の周波数(1 kHz)を中心とするバンドパスフィルターを通すことによってマスクーとした。マスクーのバンド幅を独立変数としてマスクング閾を測定した結果、「かけあわせ雑音」の帯域幅が0.1 kHzを超えると帯域幅が増加するほど検出閾が低下することが見いだされた。即ち、マスクー(妨害音)の大きさが大きくなるほど検出がしやすくなるという逆説的な結果が得られた。いわゆる古典的な臨界帯域の考え方では、マスクーの帯域幅が臨界帯域幅(この実験の場合は約0.15 kHzと推定される)を超えるまでは検出閾が上昇しそれを超えると頭打ちになることから、臨界帯域の外側の成分は検出に関与しないと考えられていた。しかし、この実験結果はそのようなモデルでは説明できない事態があることを示したのである。

CMRの場合ターゲット音は純音であり、定常的な振幅を持っていた。ターゲットも不規則な振幅の変動を持った帯域雑音である場合について調べた現象がCDDである。マクファデン

(McFadden, 1987)の実験では2.5 kHzの中心周波数、100 Hzの帯域幅をもった帯域雑音をターゲットとし、1.5, 2.0, 2.3, 2.7, 3.0, 3.5 kHzの中心周波数、100 Hzの帯域幅の干渉音が存在する場合のターゲットの検出閾を測定した。このとき、ターゲットと干渉音の間に振幅の相関が存在しない条件と存在する条件を比較すると、相関条件の方が検出閾は増加した。相関があるときに閾値の上昇を招くという点はCMRとは逆の結果のように見えるが、CMRでの相関とは干渉音内の帯域間のものであるのに対して、CDDの場合にはターゲットと干渉音の間の相関であることに注意する必要がある。

次にMDIの課題では聴き手はターゲットの存在の検出をするのではなくターゲットの振幅に正弦変調がかかっているか否かを判断する(Yost, Sheft, & Opie, 1989)。このときに臨界帯域の外に存在する干渉音による影響を調べると、干渉音に振幅変調が存在しない場合は干渉音が全く存在しない条件とほとんど結果は変わらないのに対し、干渉音にターゲットと同じ周波数の変調が加わると課題は困難となった。干渉音に対する変調とターゲットに対する変調の位相差の効果はなかった。但し、変調の周波数の違いは効果を持ち、干渉音に対する変調周波数がターゲットに対するそれよりもずれるに従い干渉効果は低減した(Yost & Sheft, 1989)。

CMR, CDD, MDIなどの知見は振幅変調による共通運命の法則の一例として見ることは確かにできる。ただし、外界に存在する共通運命性を活用した聴覚機能を実現している機構の解明への糸口としての意義を持つ。その機構の候補は振幅変調チャンネルの概念であり、神経生理学的なレベルとの対応も見えつつある(例えば、Schreiner & Urbas, 1988)。

5 むすび

この章では音の体制化の問題を中心に論じてきた。音脈分凝、聴覚的誘導、CRMなどの実験で用いられる刺激はある意味で非常に単純なもの、実験室的なものが多い。それにも拘わらずそこで得られる知見は、私たちが音が媒介として環境を理解するということがどのような問題を解くことで、その時にどのような情報が鍵となるかを教えてくれる。この体制化の機構は日常聴覚という知覚モダリティの主な対象とされている話し声や音楽を聴く上でも重要な役割を果たしている。もちろん話し声や音楽を聴く場合にはその領域特異の問題も生じ、それもまた聴覚による認知の重要な側面ではあり研究の歴史も古い。情景分析そのものも単に音脈がいくつあるかを推論するだけに留まるわけではなく、何らかのオブジェクトを知覚的に構築することが最終的な目標となるはずである。ここでは残念ながらそれらの問題に触れることはできなかったが、他にもよい参考書があるのでそちらを見ていただきたい。

「百聞は一見に如かず」(Seeing is believing)という諺があることから窺われるように、私たちが視覚像から推定するオブジェクトは存在が明確であるのに対して、聴覚像から推定されるオブジェクトは常に曖昧であるようだ。しかしそれは聴覚が視覚に劣っているということをも必ずしも意味しない。役割の違いを意味するといった方が適切であろう。聴覚によって私たちは背後（あるいは死角）から迫る何者かの存在をいち早く知ることができる。このような前注意的なレベルでの処理は聴覚の得意とする分野である。その意味では「一聞は百顧を促す」(Hearing is suspecting)なのである。私たちにとって本当に意味のある環境の変化とは何かを問うことが聴覚による情景分析についてさらに探求するときの糸口となるのである。

囲み4.1

マスキングと聴覚フィルターモデル

マスキングとはある音が別の音の存在により聞こえなくなる（聞こえにくくなる）現象であり、聴覚系をバンドパスフィルターの集合として考える仮説と結びついて長い研究の歴史を持っている。

初期の研究はバンドパスフィルターの形を矩形と仮定し、ある音が別の音に影響を及ぼす範囲、すなわち臨界帯域幅を求めることに主眼が置かれた(Fletcher, 1940)。この矩形フィルターの仮定が不適當であることは生理学的な同調曲線の測定結果と照らし合わせても明らかとなった。生理学的な同調曲線は通常ひとつのニューロンが発火するために必要な音圧レベルをいくつかの周波数毎に求めることによって得る。これと同様の特性を電極を挿入することなく、精神物理学的測定法に則って求めた結果を精神物理学的同調曲線(psychophysical tuning curves)と呼ぶ。

この精神物理学的同調曲線は通常次のような方法で求める。まず検出すべき正弦波信号の周波数と提示レベルを固定する（この時、提示レベルは非常に低い値、例えば10 dB SL程度とする）。マスキングは、正弦波でも狭帯域雑音でも基本的には構わないが、正弦波を用いた場合にはターゲット信号との間で生じるうなりが手がかりとなって検出に影響をするため、狭帯域雑音を使うことが多い。このような条件でマスキングの中心周波数を動かしながらターゲットをマスクするために必要なマスキングのレベルを測定する。この精神物理学的同調曲線の測定方法には次のふたつの問題点がある。まず、測定はかならずターゲット信号とマスキングが同時に入力された状況で行われており、マスキングに対するターゲットからの影響も免れられない。また、このような課題状況では必ずしも単独のフィルターからの出力が常に用いられているわけではない。例えば、ターゲットをそれよりも高い帯域にエネルギーの中心を持つ帯域雑音でマスクする場合、ターゲットと同じ中心周波数を持つフィルターの出力よりもそれよりも低い中心周波数を持つフィルターの出力の方が結果的に高いS/N比を持つことがありうる。そして実際にこのような状況では検出に有利なフィルターの出力を人間が参照していると考えられている（オフ周波数聴取、off-frequency listening）(Johnson-Davies & Patterson, 1979)。

パターソン(Patterson, 1976)はこのオフ周波数聴取の可能性を巧みに抑える測定方法を用いることによって、聴覚フィルターの形状をかなり正確に測定することに成功した。それはマスキングとしてターゲットの周波数に谷間をもつ帯域除去雑音を用いる方法である。「谷」の幅は図に示すようにターゲットの周波数に対して対称に設定する。ここで Δf の値を変えて

ターゲットの検出閾を測定していく。このフィルターを通過するターゲットとマスクのエネルギーの比がある一定の値となるのが検出閾に対応すると仮定し、さらにフィルターは左右に対称の形状であると仮定することによって聴覚フィルターの形状を推定する方法である。このように推定されたフィルター特性には次式で表すようなガンマトーン関数のフィルターの次数を4次とした場合がよく適合するとされている。

$$g(t) = t^{n-1} e^{-bt} \cos(\omega t) u(t)$$

n : フィルターの次数

b : バンド幅に対応する変数

ω : 角周波数

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

囲み4.2

大きさ・高さ・音色

一般に音の3つの知覚的属性として大きさ・高さ・音色があるといわれる。これは今日の音響学の基礎を築いたヘルムホルツ(Helmholtz, 1954)による分類の影響がその後の研究に及んだものと考えられる[注B1]。

従って、音の知覚に対する教科書、ハンドブックには大きさ・高さ・音色の項目が掲げられることが多い。本章ではあえてこのような構成をとらなかった。それはこれらの属性が付与されるためにはそれを付与するための知覚的事象ないし対象を特定する必要があるからで、本章ではむしろ後者がどのような過程を経て行われるかについて紹介することに重点を置くからである。例えば、大きさという概念は「音の持つエネルギーに対する主観的印象」として定義される。音がひとつしかない場合は特に問題はないであろう。ところが、音が二つ以上になったときには与えられたエネルギーの内どの部分を大きさを評価する対象とみなすかという重大な問題が生じる。具体的には部分マスキングの現象を思い浮かべていただきたい。部分マスキングの実験では単独提示された純音の大きさが雑音とともに提示された純音の大きさと比較される。通常耳に到達するエネルギーの総量は雑音を伴う条件の方が多いはずであるが、聴覚系は雑音のエネルギーを評価に含めてしまうということはせず、むしろ純音の大きさは単独で提示されたときよりも小さいと判断されるのである。このような判断ができるということは即ち聴覚系が判断の対象となる純音を雑音から知覚的に分離しているからであり、その分離がどのようにされるかによって大きさの知覚に違いが生ずると見ることが出来る。

従って、大きさ・高さ・音色は2つ（またはそれ以上）の音（音源）を区別するときの知覚的な手がかりである、という表現は適切ではない。私たちが2つ（またはそれ以上）の音源が存在すると知覚する際に、大きさ・高さ・音色さらに定位位置のいずれかまたはその組み合わせのいずれかの間の違いを経験していると表現した方が適切である。

囲み4.3

ホルマント構造と周波数変調の関係

音声のうちいわゆる母音として分類される音は声帯で発せられたパルス列（声帯音源波）が声道を通ることによって成立する。声の高さを一定に保っている場合、声帯音源波は周期性をもち、いわゆる定常的な調波複合音となる。これに対して声帯は音響管であり、その形状（長さ、断面積）によって共振周波数（ホルマント周波数; formant frequency)が変化するフィルターとしてモデル化できる。母音の違いはこのホルマント周波数の違いによって伝えられる。反対に同じ母音を声の高さを違って発声する場合はホルマント構造を保って基本周波数を変化させた場合としてモデル化できる。仮にホルマント位置の推定を行う必要がある場合、単一の基本周波数で駆動されるサンプルよりも複数の基本周波数で駆動されたサンプルを持つ方が推定精度は高くなる。周波数変調がかかった状態はこのような複数のサンプルが与えられた状況に相当する可能性がある。

囲み4.4

音声知覚の調音参照説とカテゴリー知覚

音の知覚・認知と運動との密接な関係を主張したものに音声知覚の調音参照説がある (Lieberman, 1957)。その初期の主張は音声の知覚はその音を自分自身で発するとき用いる調音制御指令を参照することによって成立するというもので、その根拠とされた実験事実がカテゴリー知覚の現象である (Lieberman et al., 1957)。この実験では/be/から/de/を経由して/ge/にいたる刺激連続帯上の14の音を合成し、それぞれの音に対する同定と、各隣接する刺激間の弁別を求めた。その結果は、同定実験で求めたカテゴリー境界の位置に弁別成績が向上する点の存在を示し、音声に対する知覚表象が離散的なものであり、すなわち調音制御指令であるという主張に結びついた。

調音参照説はある意味で非常に直観的に納得しやすい。例えば、外国語の聞き取りが自分がその言葉を話せるようになると非常に改善したように経験された読者も少なくないであろう。しかし調音参照説とカテゴリー知覚に対しては、カテゴリー内の弁別が全く不可能ではないこと、カテゴリー境界が刺激文脈に依存して柔軟に移動しうること、音声以外でも同様の現象が観察されることなどから批判がなされた (例えば、Massaro, 1987)。このような批判やその後の実験事実を受けて調音参照説自体も現在ではその原型とは異なってきた (Lieberman & Mattingly, 1985)。そこでは初期には明確であった離散的な調音制御信号への参照という主張はむしろ一歩後退し、音声処理に対するモジュールの存在が主張され、それゆえ音声は知覚処理過程において「特別扱い」("speech-is-special")されていると考える。モジュール性を仮定する根拠は音声特別扱いされているとみなせる種々の知覚現象の存在(現象の例については(Repp & Liberman, 1987参照)であるが、音声モジュールが音響信号を受け取ったときに具体的にどのような仕事をすべきで、それはどのような上位システムの中で機能するのかについては明確に規定されていない。

「音声知覚について理解を深めるには音声生成に関する種々の制約を理解する必要がある」という調音参照説の主張のひとつの側面は確かに妥当性がある。音源と環境に関する「知識」の活用という点においては聴覚情景分析の考え方も非常に近いものである。その意味では「音声対非音声」として対峙させて考えるよりは環境を構成する聴覚事象のひとつとしてそれぞれを捉えていく方が建設的であろう。

囲み4.5

クックの聴覚情景分析に対する計算論的アプローチ

ブレグマンの提唱した「聴覚による情景分析」の発想法と、マー (Marr) の視覚に対する計算論的アプローチの影響を受けてクック (Cook, 1993)は音の体制化に対する計算論的アプローチを試みている。

音響的信号 (混合した波形) はまず聴覚末梢系の蝸牛基底膜による周波数分析特性をモデル化したガンマ・トーン・フィルター・バンクによって帯域分割される。その後各々のチャンネルについて次の式に従って瞬間周波数 $\nu(t)$ を求める。

$$\nu(t) = \frac{1}{2\pi} \left[\omega + \frac{\mathcal{I}(t) \frac{d}{dt} \mathcal{R}(t) - \mathcal{R}(t) \frac{d}{dt} \mathcal{I}(t)}{\mathcal{I}^2(t) + \mathcal{R}^2(t)} \right]$$

ここで $\mathcal{R}(t)$, $\mathcal{I}(t)$ はそれぞれガンマ・トーン・フィルターの出力の実数部と虚数部である。この瞬間周波数を時間軸方向に平滑化した $\hat{\nu}(t)$ をもとに「場所的まとまり」(place-group)を求める。これは隣接するフィルターがほぼ同位相の出力を出すことに基づいている。より具体的には各フィルターに支配的な周波数はその中心周波数(CF)よりも低い刺激の成分の周波数となるという拘束条件を用いる。「場所的まとまり」 $\{p_i\}$ は次式で与える条件を満たす $[f_i, f_{i+1}]$ の範囲に中心周波数があるチャンネルとなる。

$$\left(\frac{\delta \mathcal{E}}{\delta f} \Big|_{f=f_i, f_{i+1}} = 0 \right), \frac{\delta^2 \mathcal{E}}{\delta f} \Big|_{f=f_i} < 0, \frac{\delta^2 \mathcal{E}}{\delta f} \Big|_{f=f_{i+1}} > 0$$

ここで $\mathcal{E} = \mathcal{E}(t_0, f) = \hat{\nu}(t_0, f) - f$ である。各「場所的まとまり」について周波数、支配性、振幅、振幅変調速度などの属性量を求める。これらの属性量に従ってその後の体制化が行われる。

次の段階では「場所的まとまり」の時間軸方向への集合化が図られ「同相性音縊り」(synchrony strand)が構成される。一連の軌跡上にある「場所的まとまり」はひとつの「同相性音縊り」へとまとめられる。

この「音縊り」による聴覚的情景の初期的表象に対して調波性、共通の振幅変調、共通の周波数変化・遷移、開始・終了時点の同期性などを手がかりとする群化を発見的に探索していく。図の上段は2つの音声刺激を単独提示した場合の「音縊り」による表象である。それらを同時に提示した場合は図中段のように表象される。これに体制化を行い上段左側に相当する群を得たのが下段に示す表象である。この例は分離に比較的成功的な例である。

このモデル自体はそれまでに知られている聴覚末梢系の特性や聴覚的体制化の原理などをなるべく忠実にトレースしようという試みである。それ故モデルから新たに提起される問題点というものは見えにくい。ただし、現象観察に根ざした情景分析規則の発見と信号処理的アプローチとの接点として貴重な位置を占める。

脚注

[注1] 現実の耳の場合は、耳介に存在する凹凸、頭部による回折などから単耳受聴の条件であっても位置に対する手がかりは存在しうる。ここでは音の伝搬経路にそのような干渉物が存在しない場合に条件を単純化している。

[注2] Bregman(1990)のたとえでは川は2本きちんと用意されている。

[注3] 高い音にH1, H2の2種類を設けることにより、分凝したときに高い音の音脈には周波数変調がかかり、低い音の系列は一定の高さを持つように聞こえ方の質的な差異が生まれ、2つの音脈の区別が容易となる。

[注4] これを「単純な状態」とするのは適切ではないかもしれない。私たちは現在ステレオ再生系とヘッドフォンという道具を持つ故に比較的簡単にこのような条件を作り出せるが、自然環境では滅多に存在しないようなこの事態は聴覚系にとっては非常に複雑なものになっている可能性がある。

[注5] 拡散音像も定位感としては「特別な位置」を占めると考えることも可能である。

[注6] 例えば、同じ100 Hzの変調をかける場合でも搬送周波数を528 Hzにすると側帯波は428, 628 Hzに生じ、調波的な構造は持たない。(実際には4 Hzを基本周波数とする調波構造とみなせるが、4 Hzは音としては低すぎる。)

[注B1] ヘルムホルツ自身はまず楽音(周期性のある音)を雑音(周期性のない音)から区別しており、この楽音の間をさらに知覚的に区別する場合に経験する感覚の差として3つの属性を挙げている。

参考文献

- Assman, P. F., & Summerfield, Q. (1989). Modeling the perception of concurrent vowels: vowels with the same fundamental frequency. Journal of Acoustical Society of America, 85, 327-338.
- Assman, P. F., & Summerfield, Q. (1990). Modeling the perception of concurrent vowels: vowels with different fundamental frequencies. Journal of Acoustical Society of America, 88, 680-697.
- Bashford, J. A., Meyers, M. D., Brubaker, B. S., & Warren, R. M. (1988). Illusory continuity of interrupted speech: Speech rate determines durational limits. Journal of the Acoustical Society of America, 84, 1635-1638.
- Bashford, J. A., & Warren, R. M. (1987). Multiple phonemic restorations follow the rules for auditory induction. Perception and Psychophysics, 42, 114-121.
- von Békésy, G. (1960). Experiments in Hearing. New York: McGraw-Hill.
- Bregman, A. S. (1978a). Auditory streaming is cumulative. Journal of Experimental Psychology: Human Perception and Performance, 4, 380-387.
- Bregman, A. S., Abramson, J., Doehring, P., & Darwin, C. J. (1985). Spectral integration based on common amplitude modulation. Perception & Psychophysics, 37, 483-493.
- Bregman, A. S., & Dannenbring, G. L. (1977). Auditory continuity and amplitude edges. Canadian Journal of Psychology, 31, 151-159.
- Bregman, A. S., & Doehring, P. (1984). Fusion of simultaneous tonal glides: The role of parallelness and simple frequency relations. Perception & Psychophysics, 36, 251-256.
- Bregman, A. S., & Pinker, S. (1978). Auditory streaming and the building of timbre. Canadian Journal of Psychology, 32, 19-31.
- Bregman, A. S., & Rudnicky, A. (1975). Auditory segregation: Stream or streams? Journal of Experimental Psychology: Human Perception and Performance, 1, 263-267.
- Carhart, R., Tillman, T. W., & Johnson, K. R. (1967). Release of masking for speech through interaural time delay. Journal of the Acoustical Society of America, 42, 124-138.
- Chalikia, M. H., & Bregman, A. S. (1989). The perceptual segregation of simultaneous auditory signals: Pulse train segregation and vowel segregation. Perception & Psychophysics, 46, 487-497.
- Cherry, E. C. (1953). Some experiments on the recognition of speech with one and with two ears. Journal of the Acoustical Society of America, 25, 975-979.
- Ciocca, V., & Bregman, A. S. (1987). Perceived continuity of gliding and steady-state tones through interrupting noise. Perception & Psychophysics, 42, 476-484.

- Cook, M. P. (1993). Modelling Auditory Processing and Organization. Cambridge: Cambridge University Press.
- Dannenbring, G. L. (1976). Perceived auditory continuity with alternately rising and falling frequency transitions. Canadian Journal of Psychology, 30, 99-114.
- Dannenbring, G. L., & Bregman, A. S. (1976). Stream segregation and the illusion of overlap. Journal of Experimental Psychology: Human Perception and Performance, 2, 544-555.
- Darwin, C. J., Ciocca, V., & Sandell, G. J. (1994). Effects of frequency and amplitude modulation on the pitch of a complex tone with a mistuned harmonic. Journal of the Acoustical Society of America, 95, 2631-2636.
- Deutsch, D. (1974). An auditory illusion. Nature, 251, 307-309.
- Deutsch, D. (1975). Two-channel listening to musical scales. Journal of the Acoustical Society of America, 57, 1156-1160.
- Dowling, W. J. (1973). The perception of interleaved melodies. Cognitive Psychology, 5, 322-337.
- Egan, J. P. (1948). The effect of noise in one ear upon the loudness of speech in the other ear. Journal of the Acoustical Society of America, 20, 58-62.
- Fedderson, W. E., Sandel, T. T., Teas, D. C., & Jeffress, L. A. (1957). Localization of high-frequency tones. Journal of the Acoustical Society of America, 29, 989.
- Fletcher, H. (1940). Auditory patterns. Review of Modern Physics, 12, 47-65.
- French-St.George, M., & Bregman, A. S. (1989). Role of predictability of sequence in auditory stream segregation. Perception and Psychophysics, 46, 384-386.
- Hall, J. W., Haggard, M. P., & Fernandes, M. A. (1984). Detection in noise by spectro-temporal pattern analysis. Journal of the Acoustical Society of America, 76, 50-56.
- Helmholtz, H. (1954(1885)). On the Sensations of Tone (Die Lehre von den Tonempfindungen) (Ellis, A.J., Trans.). New York: Dover.
- Houtgast, T. (1972). Psychophysical evidence for lateral inhibition in hearing. Journal of the Acoustical Society of America, 51, 1885-1894.
- Johnson-Davies, D. B., & Patterson, R. D. (1979). Psychophysical tuning curves: Restricting the listening band to the signal region. Journal of the Acoustical Society of America, 65, 765-770.
- Johnstone, B. M., Patuzzi, R., & Yates, G. K. (1986). Basilar membrane measurements and the travelling wave. Hearing Research, 22, 147-153.
- Jones, M. R., Kidd, G., & Wetzell, R. (1981). Evidence for rhythmic attention. Journal of Experimental

Psychology: Human Perception and Performance, 7, 1059-1073.

柏野牧夫 (1991). 閉鎖区間の前後に分散する手がかりに基づく日本語語中閉鎖子音の知覚. 日本音響学会誌, 48, 76-86.

Kubovy, M., & Howard, F. P. (1976). Persistence of a pitch-segregating echoic memory. Journal of Experimental Psychology: Human Perception and Performance, 2, 531-537.

Kuhn, G. F. (1977). Model for the interaural time differences in the azimuthal plane. Journal of the Acoustical Society of America, 62, 157-167.

Liberman, A. M. (1957). Some results of research on speech perception. Journal of the Acoustical Society of America, 29, 117-123.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology, 54, 358-368.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. Cognition, 21, 1-36.

Martens, J. P. (1984). Comment on "Algorithm for the extraction of pitch and pitch salience from complex tonal signals" [Journal of the Acoustical Society of America, 71, 679-688 (1982)]. Journal of the Acoustical Society of America, 75, 626-628.

Massaro, D. W. (1987). Categorical partition: A fuzzy-logical model of categorization behavior. In S. Harnad (Ed.), Categorical Perception: The Groundwork of Cognition N.Y.: Cambridge Univ. Prs.

McAdams, S. (1989). Segregation of concurrent sounds. I: Effects of frequency modulation coherence. Journal of the Acoustical Society of America, 86, 2148-2159.

McFadden, D. (1987). Comodulation detection difference using noise-band signals. Journal of the Acoustical Society of America, 81, 1519-1527.

Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. Journal of Acoustical Society of America, 59, 640-654.

Rasch, R. A. (1978). The perception of simultaneous notes such as in polyphonic music. Acustica, 40, 21-33.

Repp, B. H. (1992). Perceptual restoration of a "missing" speech sound: Auditory induction or illusion? Perception and Psychophysics, 51, 14-32.

Repp, B. H., & Liberman, A. M. (1987). Phonetic category boundaries are flexible. In S. Harnad (Ed.), Categorical Perception: The Groundwork of Cognition N.Y.: Cambridge Univ. Prs.

- Rogers, W. L., & Bregman, A. S. (1993). An experimental evaluation of three theories of auditory stream segregation. Perception and Psychophysics, 53, 179-189.
- Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. Journal of Neurophysiology, 30, 769-793.
- Samuel, A. G. (1981a). Phonemic restoration: Insights from a new methodology. Journal of Experimental Psychology: General, 110, 474-494.
- Samuel, A. G. (1981b). The role of bottom-up confirmation in the phonemic restoration illusion. Journal of Experimental Psychology: Human Perception and Performance, 7, 1124-1131.
- Sasaki, T. (1980). Sound restoration and temporal localization of noise in speech and music sounds. Tohoku Psychologica Folia, 39, 79-88.
- Scheffers, M. T. M. (1982). The role of pitch in the perceptual separation of simultaneous vowels II (IPO Annual Progress Report No. 17, 41-45). Institute for Perception Research, Eindhoven.
- Schreiner, C. E., & Urbas, J. V. (1988). Representation of amplitude modulation in the auditory cortex of the cat. II. Comparison between cortical fields. Hearing Research, 32, 49-65.
- Sellick, P. M., Patuzzi, R., & Johnstone, B. M. (1982). Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. Journal of the Acoustical Society of America, 72, 131-141.
- 寺西立年 (1984). 聴覚の時間的側面. 難波精一郎 (編), 聴覚ハンドブック (pp. 276-319). ナカニシヤ.
- Tsuzaki, M., Kato, H., & Tanaka, M. (1994). Apparent duration of speech segments interrupted by a noise burst. Journal of the Acoustical Society of Japan (E), 15, 205-206.
- Van Noorden, L. P. A. S. (1975) Temporal Coherence in the Perception of Tone Sequences. Unpublished doctoral dissertation, Eindhoven University of Technology.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. Science, 167, 392-393.
- Warren, R. M. (1984). Perceptual restoration of obliterated sounds. Psychological Bulletin, 96, 371-383.
- Warren, R. M., & Bashford, J. A. (1976). Auditory contralateral induction: An early stage in binaural processing. Perception & Psychophysics, 20, 380-386.
- Warren, R. M., & Sherman, G. L. (1974). Phonemic restorations based on subsequent context. Perception & Psychophysics, 16, 150-156.

- Warren, R. M., Obusek, C. J., & Ackroff, J. M. (1972). Auditory induction: Perceptual synthesis of absent sounds. Science, 176, 1149-1151.
- Yost, W. A., & Sheft, S. (1989). Across-critical-band processing of amplitude-modulated tones. Journal of the Acoustical Society of America, 85, 848-857.
- Yost, W. A., Sheft, S., & Opie, J. (1989). Modulation interference in detection and discrimination of amplitude modulation. Journal of the Acoustical Society of America, 86, 2138-2147.

参考文献

Bregman, A. S. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound.

Massachusetts: MIT Press.

聴覚情景分析の考え方が豊富な実験例とともに紹介されている。多少冗長な点も多いが、様々な観点から問題を発掘するためには貴重な文献である。

Handel, S. (1989). Listening: An introduction to the perception of auditory events. Cambridge: MIT

Press.

音に関係する多岐にわたる事柄が生態学的観点から手際よくまとめられている。紹介される研究についてもなるべくわかりやすい比喻を用いることによって読者の理解を助けるよう工夫されている。

Warren, R. M. (1982). Auditory Perception: A New Synthesis. New York: Pergamon.

古典的な聴覚精神物理学に対して批判的な観点から書かれた独創的な教科書。

Moore, B. C. J. (1989). An Introduction to the Psychology of Hearing . (Third ed.). London:

Academic Press. 大串健吾 (監訳), 聴覚心理学概論 誠信書房 1994

特に聴覚の末梢系に関する実験事実とモデルに関する教科書。古典的な臨界帯域の概念が近代化して聴覚フィルターバンクモデルの成立にいたるまでが明瞭に解説されている。

Pickles, J. O. (1988). An Introduction to the Physiology of Hearing. London: Academic Press.

聴覚の生理学についての教科書。Mooreの教科書と同様に、聴覚末梢系について80年代に確認された事実を蝸牛の能動性、神経発火の同期性を軸に紹介している。

東倉洋一, 津崎実, 山田玲子 (1994). 音声の知覚・認知. 淀川英司(編), ATR先端テクノロジーシリーズ視聴覚情報科学 オーム社.

本章ではあまり触れられなかった音声の知覚を中心とした初心者向けの解説がなされている。

McAdams, S., & Bigand, E. (Ed.). (1993). Thinking in Sound: The Cognitive Psychology of Human Audition. Oxford: Clarendon Press.

事象の知覚からさらに高次の対象の認知へと進む過程の重要性に焦点を当てて各執筆者がそれぞれの専門分野での実験を紹介している。

この他、新編感覚知覚心理学ハンドブック（大山・今井・和気 編、誠信書房、1994）第「聴覚」にも参考となる事例・解説がある。

図キャプション

図4.1 音の存在する範囲。空気中を伝搬する弾性波のうち、私たちの聴覚器官にとって適切な振幅と周波数を持つものが音として知覚される。[Handel, 1989]

図4.2 聴覚的情報源と視覚的情報源の違い。(A)視覚的な2点は互いに重なり合わない限り網膜上でも2点である。(B)聴覚的な2点——空間的位置が異なる音源——からの音は鼓膜上で重なり合う。

図4.3 蝸牛の構造の模式図。螺旋状の管を引き延ばして描いている。前庭階、蝸牛管、鼓室階の3つに仕切られており、蝸牛管と鼓室階の仕切が基底膜である。蝸牛管は蝸牛頂に向かい細くなるが、基底膜の幅は広がっていく。基底膜上には内有毛細胞と外有毛細胞が存在し、その上を蓋膜が被っている。図中蓋膜は断面図のみに描いた。

図4.4 一連性境界と分裂境界。一連性を聴こうとしても分離して聞こえてしまう一連性境界が生じる周波数の分離の程度は交替の速度に大きく依存する。分裂を聴こうとしても一連性が聞こえてしまう分裂境界は速度への依存性があまり大きくない。[van Noorden, 1975]

図4.5 捕捉音の存在の有無による音脈分凝の違い。同じ周波数、時間軸上の分離をもつX,A,B,Xの音列であっても捕捉音Fが存在しない場合は一連の音脈を形成しやすく、捕捉音が存在するとXとA,Bは別の音脈になる。(Bregman & Rudnický, 1975)

図4.6 チオッカとブレグマンの実験で使われた刺激の例。(a)後続部に周波数変化がない場合。(b)先行部と後続部の変化方向の様々な組み合わせ。(c)先行部と後続部で変化方向・変化勾配が等しい場合。(d)変化方向が等しく変化勾配が異なる場合。(Ciocca & Bregman, 1987)

図4.7 オクターブ錯覚。(a)刺激の提示方法。オクターブ離れた2音が同時に左右の耳に交互に入れ替わりながら提示される。(b)最も多く報告される知覚像。(Deutsch, 1974)

図4.8 系列的統合(sequential integration)と同時的統合(simultaneous integration)の競合。同時的統合が優勢な場合、BはCと同じ音脈に取り込まれ、その結果Cの音の「豊かさ」が増加する。(Bregman & Pinker, 1978)

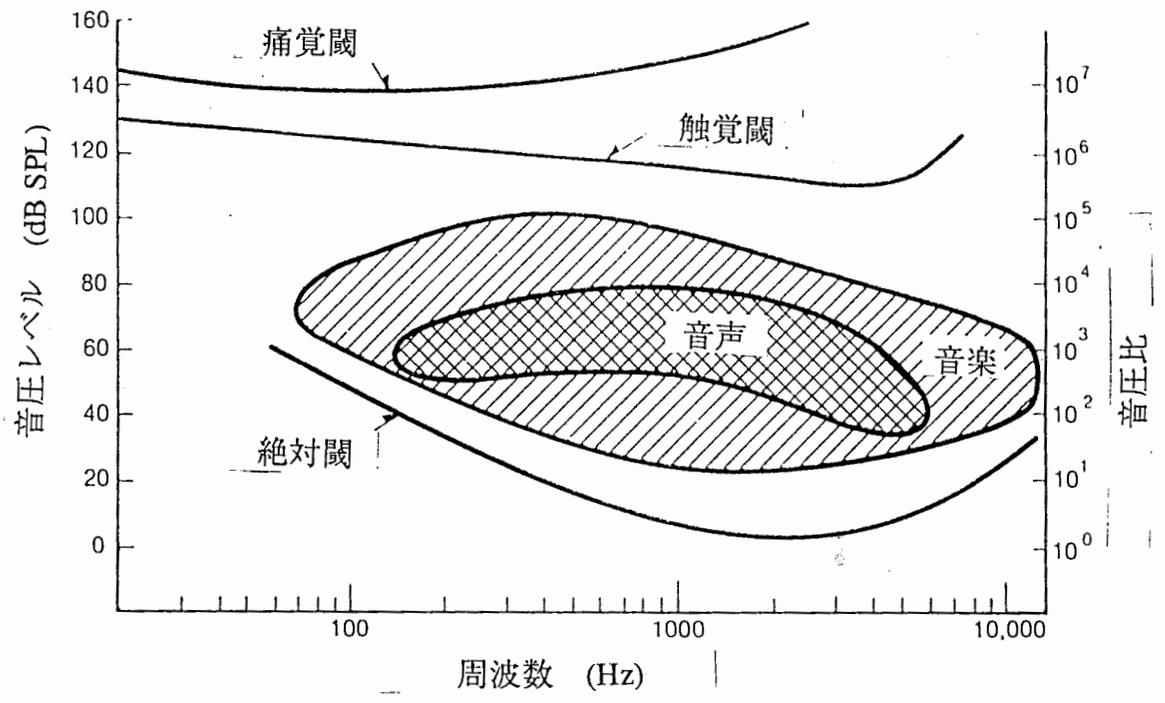
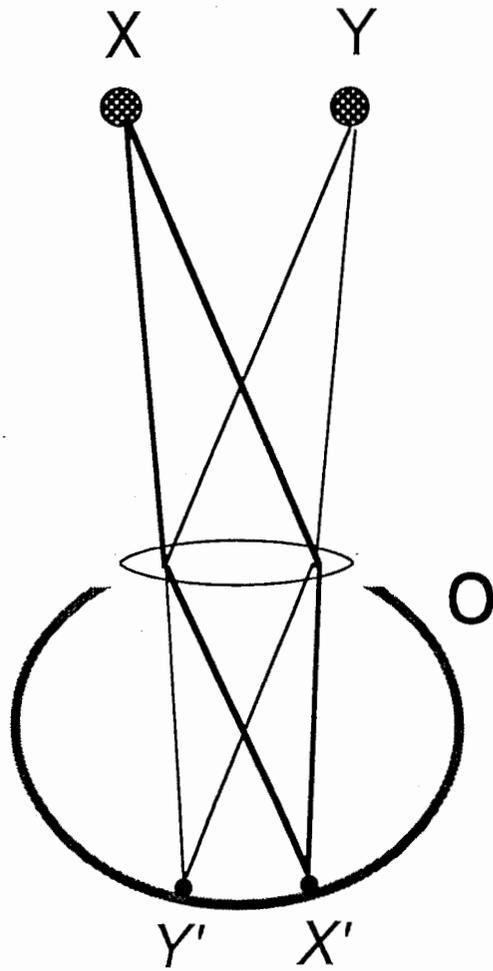


図 4.1

A. 視覚の場合



B. 聴覚の場合

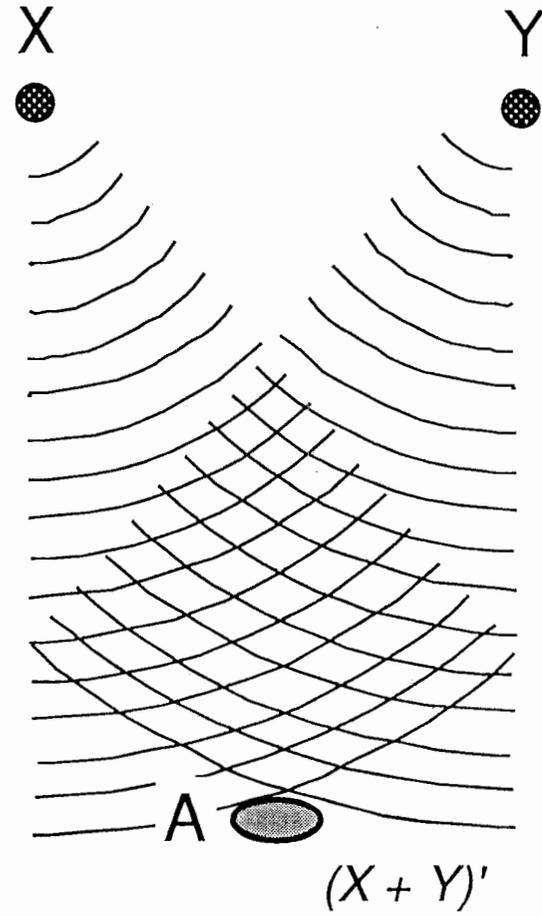


図 4.2

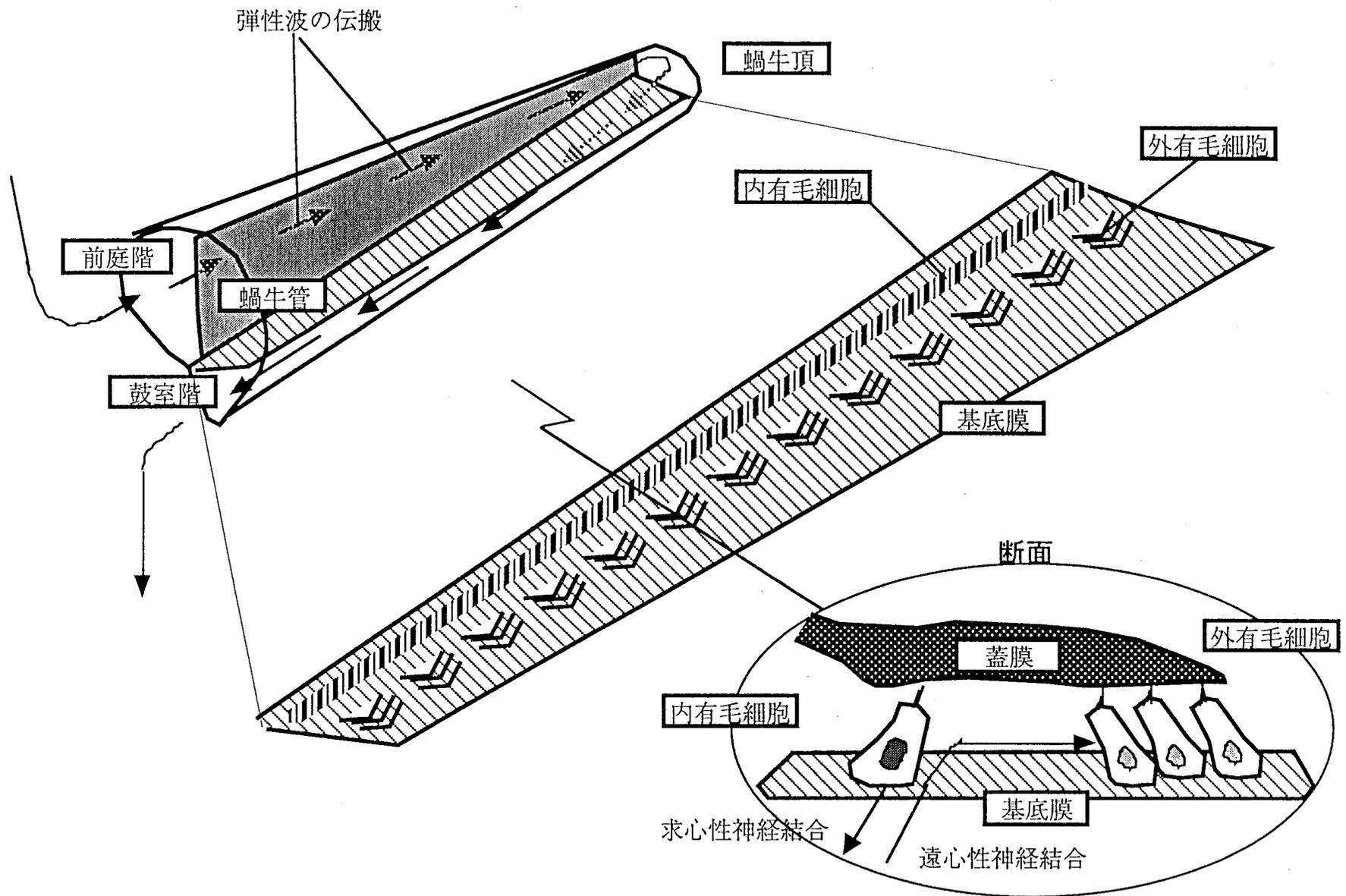


図 4.3

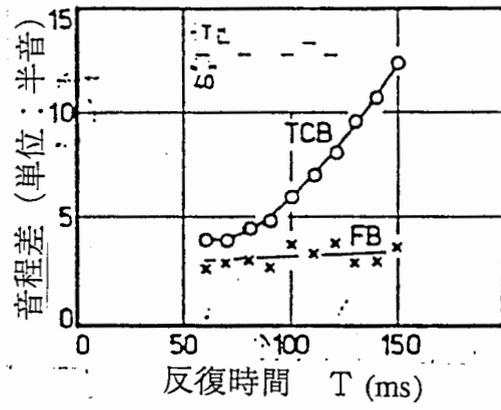


图 4.4

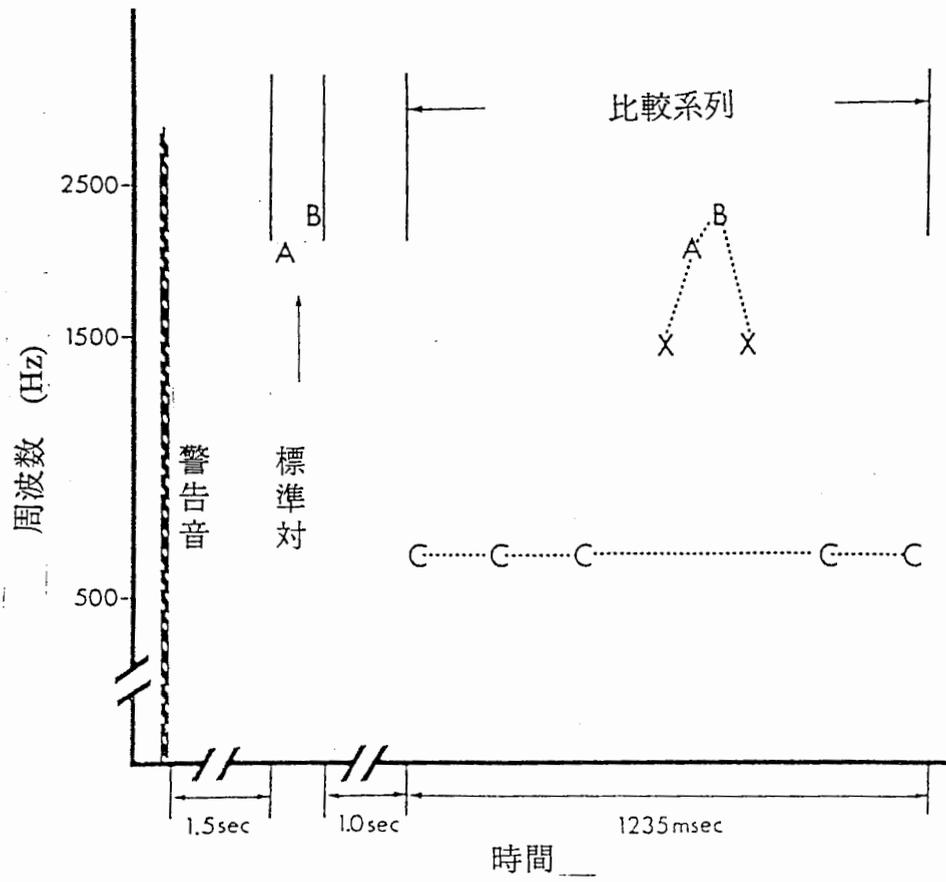


図 4.5

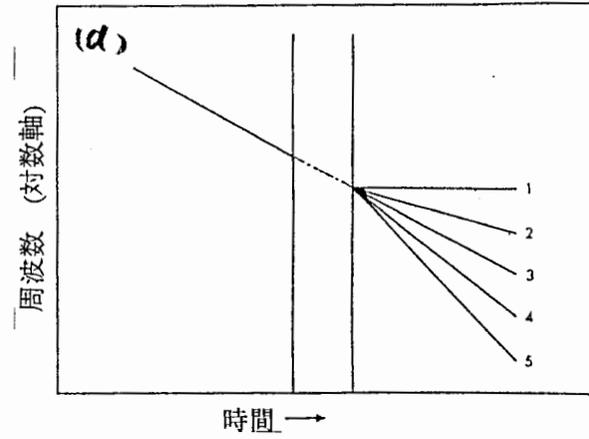
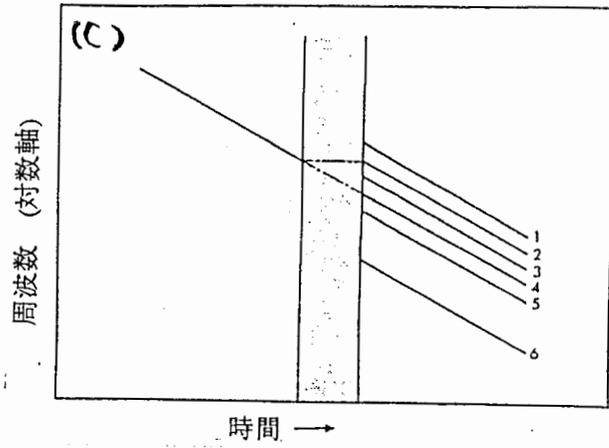
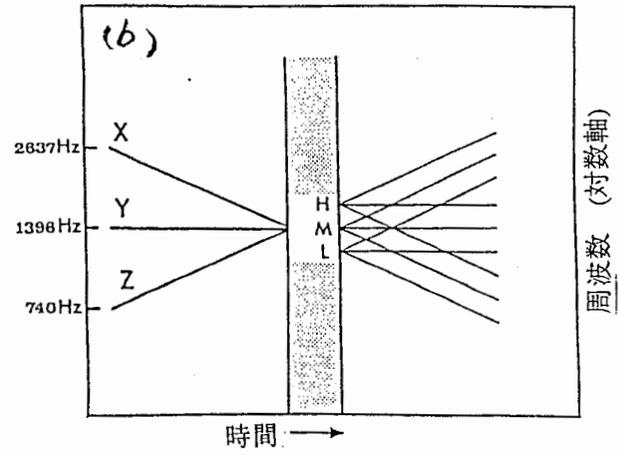
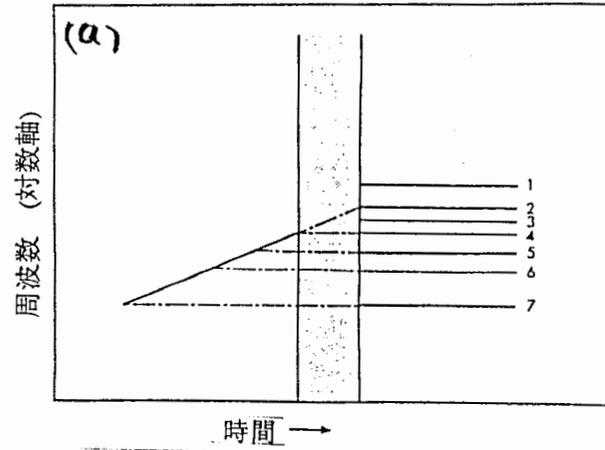


图 4.6

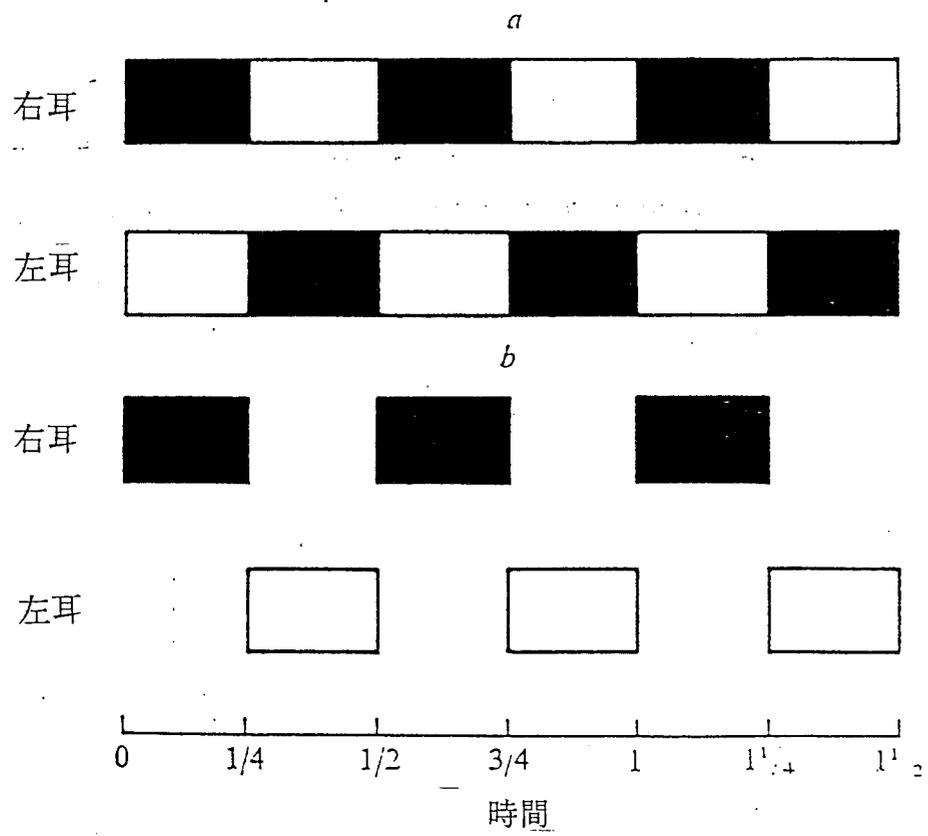


圖4.7

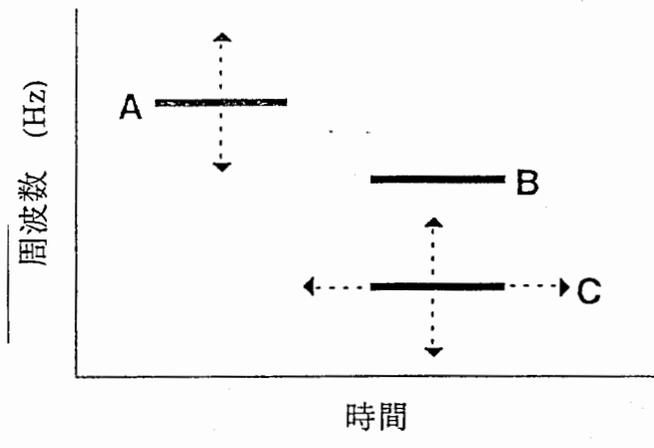


図 4-8

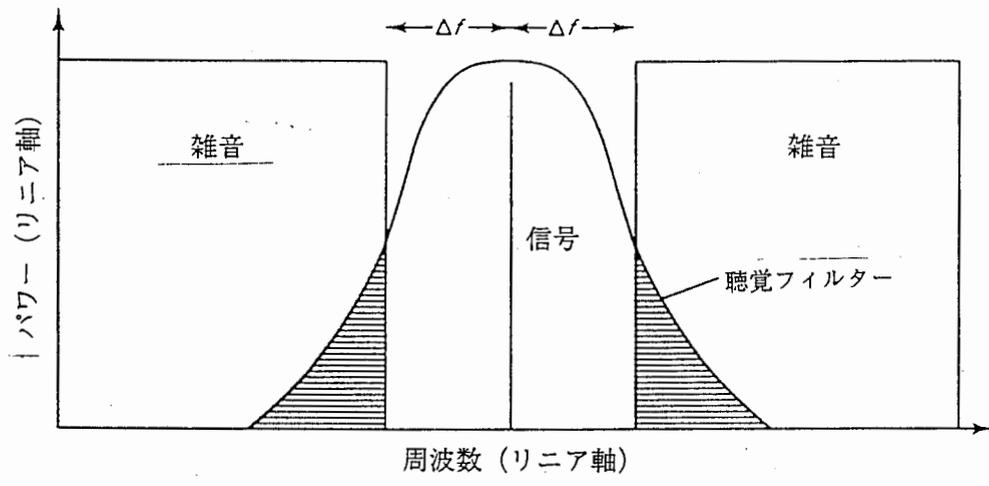


図 4.1

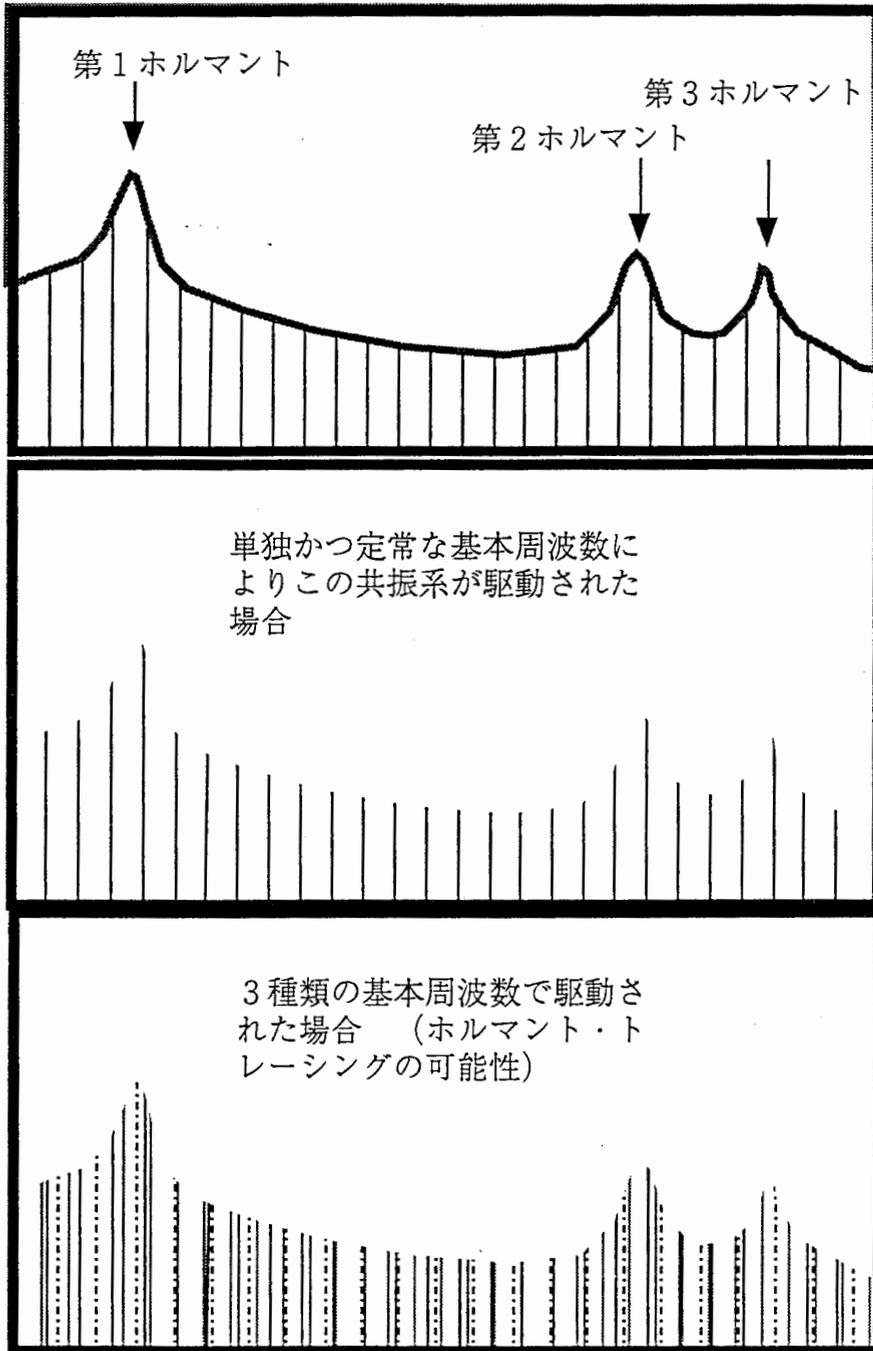


図 図み 4.3

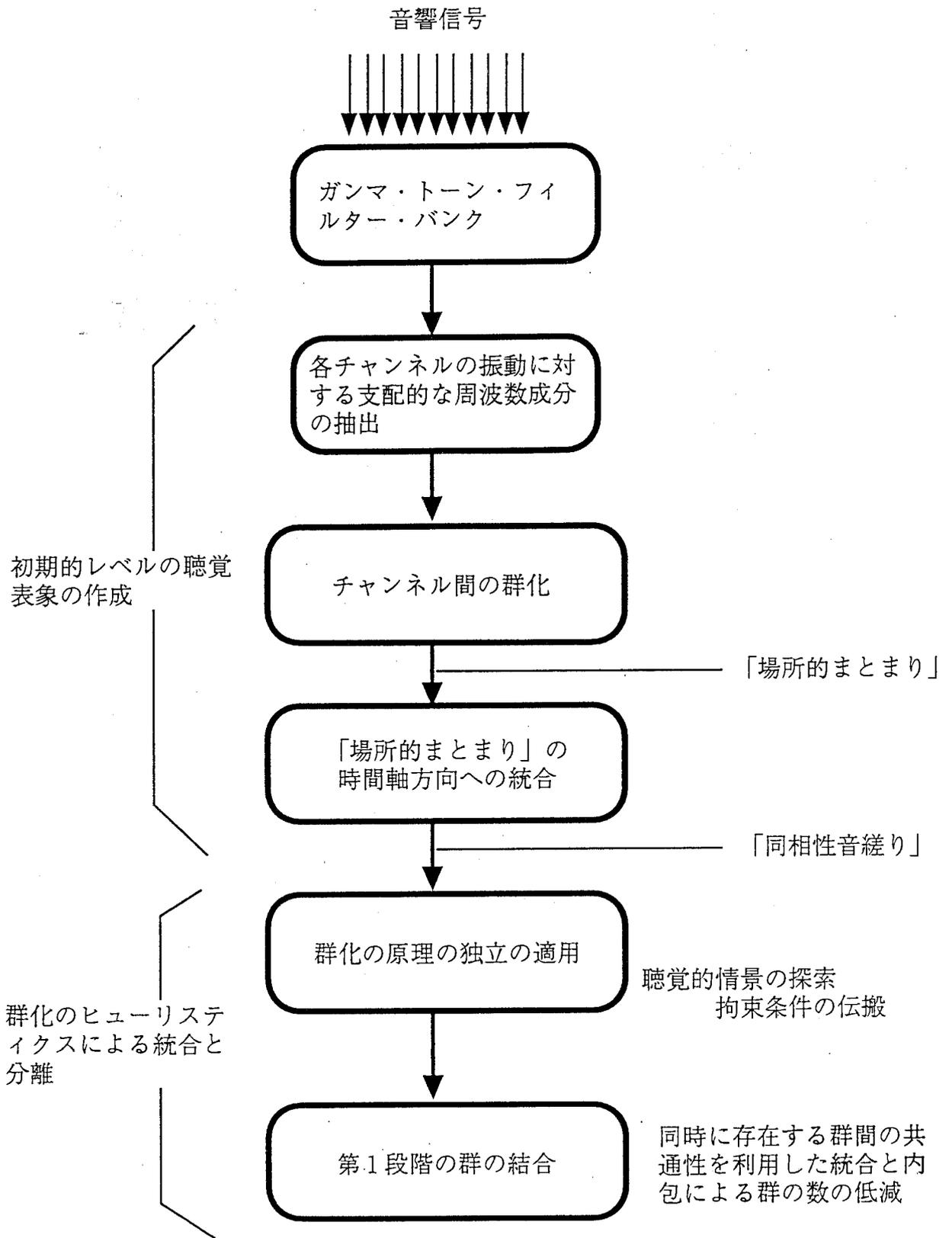


図 図み 4.5 - 1

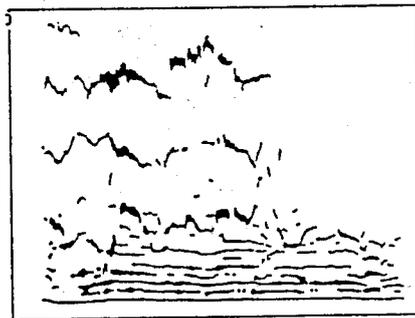
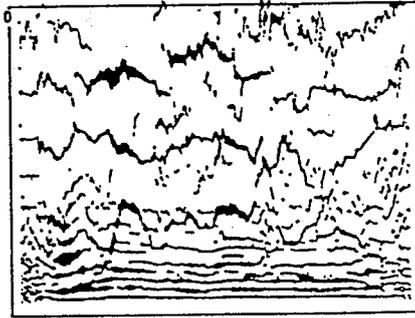
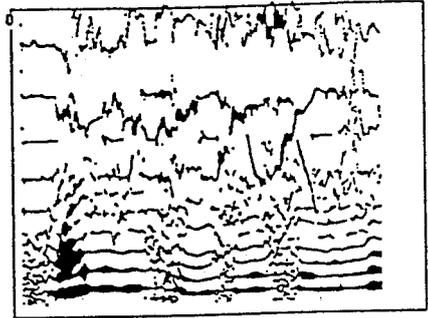
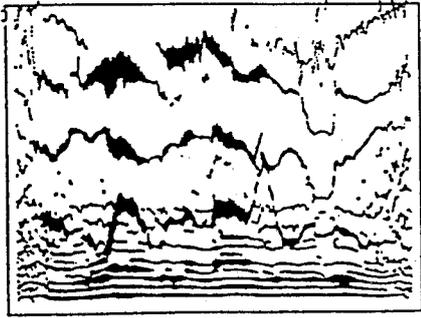


图 图及 45-2