

TR-A-0050

音の高さの知覚

--「音響心理と聴知覚」より--

ローラン・ドゥマニ 著
Laurent Demany

上田 和夫 訳
Kazuo Ueda

1989. 5. 25

A T R 視聴覚機構研究所

© (株) A T R 視聴覚機構研究所

音の高さの知覚

——「音響心理と聴知覚」より——

ローラン・ドゥマニ 著
上田 和夫 訳

第1版 1989年5月25日
第2版 1989年7月7日
第3版 1991年10月15日

Laurent Demany
"Perception de la hauteur tonal"

Dans
Psychoacoustique et Perception Auditive
par M. C. Botte, G. Canévet, L. Demany et C. Sorin
Les Editions INSERM, Paris, 1989

目次

序	1
I. 純音の音の高さ：重要な精神物理学的事実	2
I.1. 音の高さの概念	2
I.2. 聴覚系の周波数分析	2
I.3. 純音の音の高さに影響を与える周波数以外の要因	3
a. 音圧レベルの効果	3
b. 両耳複聴	4
c. 同時マスキング音の効果	4
d. その他の効果	5
I.4. 音の高さの二つの尺度	5
I.5. 純音の「クロマ」	6
II. 純音の音の高さの神経対応：データと仮説	8
II.1. 場所と時間：純音の周波数の神経符号化についてのデータ	8
a. 周波数弁別と興奮パタン	8
b. 周波数の時間的な符号化	9
II.2. 音の高さの専ら場所的な符号化説に有利な議論	11
II.3. 時間的符号化説に有利な議論	12
II.4. 思弁的総括	13
III. 複合音の音の高さ：総論	15
III.1. 「基本音の」高さ、「生の」高さ、「スペクトルの」高さ	15
III.2. 周期的複合音の基本音の高さ	16
a. 基本音が欠けていることの問題	16
b. 非調波複合音により生ずる基本音の高さ	19
c. スペクトル優位性の現象	21
d. 位相の要因	23
e. 基本音の高さとスペクトルの高さの関係	24
III.3. 雑音により生じる周期性と音の高さの感覚	25
a. 「凍った」雑音の知覚	25
b. 振幅が周期的に変調された白色雑音の知覚	26
c. スペクトルを制限された雑音の場合	27
d. 両耳系で作られた音の高さの感覚	28

IV. 周期複合音の基本音の高さを抽出するモデル	31
IV.1. 末梢スペクトル分析の役割	31
IV.2. 「スペクトル的」モデル：Terhardt のモデル	32
a. モデルの記述	32
b. モデルについての考察	34
c. スペクトルの高さの神経符号化	34
IV.3. Moore と van Noorden のモデル	36
引用文献	39

序

本論文は4節からなっている。前半の2節は波形が正弦波である音、即ち純音の音の高さに関するものである。第1節では純音の音の高さについて、まず、記述的、あるいは現象的角度から、ついで第2節ではより説明的、神経生理学的角度から考察する。これに続く2節は、純音ではなく、「複合」音によって生ずる音の高さの感覚に関するものである。これらの音についても純音と同様に、まず最初は観察される現象の記述に専心し、現在ある説明仮説についてはその大略だけを示すこととしよう。これが第3節の目的である。この節では、音声の母音部や、従来の楽器が出す音の大部分を近似できる、いわゆる「調波的」複合音について特に注目しよう。これらの音に関しては、音の高さの二つの側面を区別する必要がある。音色の一つである「生の」高さと、メロディーを構成する要素である「基本音の」高さである。本論文では基本音の高さをはっきりと特別なものとして扱う。そして、第4節はこの高さの一つの側面を抽出できる可能性のあるメカニズムのために割かれている。

1. 純音の音の高さ：重要な精神物理学的事実

1.1. 音の高さの概念

音の「高さ (hauteur tonale)」と一般に呼ばれているものについて、はっきりと、一般的に、かつ全く厳密に概念上の定義を与えることはむづかしい。音の高さの知覚とそのメカニズムのために捧げられたおびただしい論文のことを知らない人にとっては、人が音の高さについて話すのを聞いて、それは非常に「単純な」感覚属性である、ときっと思えるであろう。しかし、実はそうではないことは様々な著者が既に強調している通りである（特に Therhardt, 1980 を見よ）。

音の高さの感覚は純音と、ある種の複合音について生じる。しかし定義の問題はその両方の場合において起こる。ANSI (ANSI, 1960) によれば、純音あるいは複合音の音の高さ (pitch) は「音を——音階のように——低から高に伸びる尺度上に配列することができるような聴感覚の属性」である。AFNOR (1977) によれば、それは「周期的な音の周波数と関係が深い聴感覚の性質で、その周波数がより高いか低いかに応じてその音が高いか低いかと言えるようなもの。」である。これら二つの定義は、まず何よりも、「低い」「高い」といったそれ自身を定義することがまだなされていない用語を使っているという理由で不満足なものである。さらに重要なこととして、ANSI と AFNOR の定義は、二つの純音あるいは複合音は、それらの周波数（或はそれらの周期）が近ければ近いほど、常に音の高さも類似しているということを必然的に意味することに注意しなければならない。これらの定義は、従って、オクターヴの間隔をなす二つの純音ないしは複合音はより狭い音程をなす二音よりも音の高さの上ではより類似していると知覚され得る、ということと相容れない。後で見ると、このようなことが除かれてしまっているのは大いに異論の余地があるところである。

音の高さのよりましな定義を——ただし純音の場合に限ってであるが——提案することができる。ここでは、ある一つの純音の音の高さを、周波数を調節することによって、強さの異なる別の純音とその音とを合わせることができる、ということに基づく音の感覚属性である、としよう。そしてまた、それらの音の高さを合わせた後では二つの音は「大きさ」でしか異なっておらず、このことは任意の二つの純音は音の高さと（または）大きさでしか異なることはあり得ないということを必然的に意味する、としよう。ここで、強さの異なる二つの純音について、それらの周波数を一致させておきさえすれば、音の高さが合ったものが即座にいつでも得られる、というわけではないことに注意したい。というのは純音の音の高さは音圧の変化によって少し変化させることができるからである。

このように純音について定義された音の高さの概念は、本当に興味深く、そして有用なものであろうか。多くの精神物理学的事実はこの疑問に対して「イエス」と答えることを許している。例えば周波数を調整して二つの純音の音の高さを正確に合わせることができ、その正確さはこれらの二つの音の間の強さの関係にはほとんど依存しないことは明かである (Verschuure and van Meeteren, 1975; Demany and Lavenant, 1985)。その上、Coninx (1978) は周波数と強さが両方とも（それも極めてわずか）異なる純音を弁別する課題においては、被験者はあたかも音の高さの違いと大きさの違いを全く独立に検知するかのように振舞うと指摘した。さらに、周波数と強さが弁別閾以上に異なる純音の間の主観的距離についての研究から、純音の知覚空間は音の高さと大きさにそれぞれ対応する二つの次元だけを持つ空間で完全に記述し得ることが示されている（例えば Schneider and Bissett, 1981 を見よ）。これら全てのことは、純音の音の高さが「分離可能な」（Garner, 1974 の言った意味で）知覚次元であることを極めて明確に示している。

1.2. 聴覚系の周波数分析

一般に、純音の周波数を少し変化させても、その音の高さは変化するが大きさは変化しない。だが

ら、「周波数分析」と「音の高さの分析」は、入れ換えて使ってもよい表現である。もっともこのことには、例外が生じる場合があることを銘記するべきである。それは特に、強さが非常に小さい場合や (Kemp, 1979) 周波数が聴覚系の通過帯域の限界に近い場合 (Henning, 1966; Coninx, 1978) である。

人間の聴覚系の周波数分析は二つの全く異なった方法で研究されてきた。第一の方法は、純音に連続的に周期的な周波数変調を与えた時、この変調がちょうど検知できる最小の振幅について研究するものである。閾が最小となる変調周波数は Zwicker (1956) によれば 4 Hz であるが、他の研究者は (Shower and Biddulph, 1931; Kay and Matthews, 1972; Demany and Semal, 1989) もっと低いところ、ほぼ 2 Hz であるとしている。第二のやり方は、継時的に提示される定常的な二つの純音がちょうど弁別されるのに必要な最小の周波数差を研究することである。Harris (1952a) によれば二音の間の無音時間は 100 から 1000 ms の間であれば問題にしないでよい。

これら二つの研究法は全く同じ心理学的能力を測定しているのだろうか。そう信じている研究者もいるが (Hartmann and Klein, 1980; Demany and Semal, 1986)、そうでない研究者もいる (Moore, 1976; Moore and Glasberg, 1986)。第二の研究法では音の高さの弁別閾がよりはっきりと測定されるので、問題とする結果だけが示されていることが保証されている。

相対値 (df/f) で表すと、継時的な二つの純音の周波数の弁別閾は基準音の周波数に対応して、最小値が大體 1500 Hz の辺りにくる "U" 字形のカーブを描いて変化する。これは多くの研究 (特に Harris, 1952b; Moore, 1973; Wier et al., 1977; Nelson et al., 1983 を見よ) から引き出せる結論である--論文に報告された平均閾値の厳密な値は、時として互いに大きく異なっている場合もありはするが。ある基準周波数に対して、閾値は最小可聴音圧からほぼ 50 dB SL (最小可聴音圧から 50 dB 上) の間で組織的に減少する。Wier et al. (1977) によれば、高い周波数よりも低い周波数の方が減少がより顕著であるという。しかしながら、Nelson et al. (1983) によれば、弁別閾に対する周波数と音圧の関係は (モノラル受聴の場合) 次式でよく記述できる。

$$\log_{10} df = a \cdot f^{0.5} + b / SL + c \quad (1)$$

ここで df は Hz で表された絶対弁別閾、 f は基準周波数である ($125 \text{ Hz} < f < 8000 \text{ Hz}$)。a, b, c はそれぞれ 0.0214, 5.056, -0.32 に対応する定数で、このときの検出能力 df は信号検出理論の枠組みに沿って (Green and Swets, 1973) 1.0 の d' で表現できる。(1) 式は中位あるいは強い音圧レベルについて、中位の周波数では相対弁別閾は二ないし三千分の一になることを意味している。

(1) 式はよく訓練された被験者にしかあてはまらない、という事実を強調しておこう。周波数の違いについての感受性は適切な訓練によって大いに洗練される可能性があり、多くの被験者において成績がもはや向上しなくなるまでには、少なくとも 10 時間の訓練期間を要する (Wyatt, 1945; Turner and Nelson, 1982; Demany, 1985)。

1.3. 純音の音の高さに影響を与える周波数以外の要因

1.3.a. 音圧レベルの効果

ある周波数の純音の音の高さが、その音圧レベルを変えた時に変化する可能性があることは 19 世紀から知られている。「Stevens の法則」 (Stevens, 1935 により、こう名付けられた) は音圧を上げると高い周波数の音の音の高さは (より鋭くなって) 上がり、中位の周波数の音の音の高さはほとんど変わらず、低い周波数の音の音の高さは減じる、ということをも主張するものである。図 1 にこの法則の正しさを確認した Walliser (1969a) の結果を示す。

Stevens の法則はただ一つの例外を除いては満足 of いくものである--純音の音の高さに対する音圧の影響は被験者によって大きく異なる可能性があり、さらに一人の被験者についてでさえ、ある周波数についての音圧による音の高さの変化はおそらく非単調関係であるようだ (Verschuure and van Meeteren, 1975)。Burns (1982) はさらにこの変化に組織的な違いが生じ得ること--そして方向 (同じ人の一方の

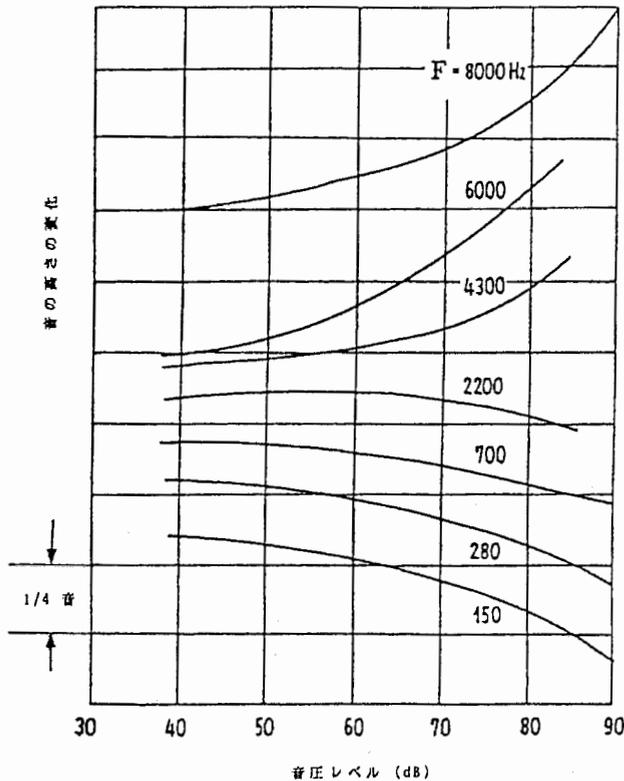


図 1. 純音の音圧レベルが音の高さに与える影響。媒介変数は音の周波数。一音の四分の一（縦軸の単位）は周波数比にしておよそ 1.03 にあたる。Walliser (1969a) による。

耳と他方の耳) についても同様に——を明らかにした。このことは、音の高さに対する音圧の影響が少なくとも部分的には蝸牛に由来することを示唆している。

1.3.b. 両耳複聴

ある周波数、ある音圧の純音の音の高さは、一般的に左の耳で聞いたときと右の耳で聞いたときとでぴったり同じにはならない。「両耳複聴」と呼ばれているのはこの現象のことである。これは聴覚病理学の対象となるような人では特に顕著であるかも知れない。しかしながらこれはどんな人にも、ある程度はあることである。聴力の正常な被験者で、二つの耳に与えられた純音によって生じる音の高さの違いは一音の四分の一（これは 3 % の周波数差に相当する）に達する。その効果は標準音の周波数によって非常に不安定に変化し、変化の仕方の個人差も大きい。van den Brink (1970) によれば、どんな人にも、ある周波数における複聴とその周波数での二つの耳の最小可聴閾の違いとの間に強い相関がある、という。しかしながら、この相関が明らかに強いのは複聴が弱い音圧で測定されたときだけである。1.3.a. 節で引用した Burns (1982) のデータは、ある一人の被験者についての標準音の周波数による複聴の変化は、複聴の測定される音圧に依存していることを思わせるものである。

ある純音が聴力の正常な人の両方の耳に同時に達すると、その人はそれを常にただ一つの音の高さを生じさせる一つの純音として知覚する。このことは中枢神経系が、単耳ではわずかに異なる二つの音の高さをただ一つの音の高さに融合することができる、ということを示唆する。融合の起こる範囲内での二つの音の高さの最大のずれは、聴力の正常な人の最大の複聴の程度をはるかに超えているのである (Perrott and Barry, 1969; van den Brink et al., 1970)。

1.3.c. 同時マスクング音の効果

純音の音の高さは、同側耳にその音を部分的にマスクする（大きさを減じる）別の音を同時に提示することにより、非常にはっきりと変化させることができる。この「誘導」（マスクング）音はノイズであっても別の純音であっても構わない。

テスト純音に白色雑音を付け加えることにより、そのテスト純音の音の高さが上昇する。その音の高

さはテスト純音の周波数が高ければ高いほど、また、純音がノイズでマスクされればされるほど、高くなる。その効果は一音の四分の一にも達し得る(Walliser, 1969a)。白色雑音の代わりに、遮断周波数(音のエネルギーがなくなる周波数)がテスト音の周波数より僅かに低い低域雑音を誘導音として用いても、少なくとも同じくらい大きな効果が得られる(例えばHoutsma, 1981; Stoll, 1985を見よ)。

さらに、純音も他の純音の音の高さを変化させることができる。このことは、誘導音の周波数がテスト音の周波数よりも低い場合には、常に成り立つことが立証されている。二つの音の音圧が等しい場合には、誘導音(周波数の低い方)はテスト音(周波数の高い方)の周波数と近い程、テスト音の音の高さを上昇させる(van den Brink, 1983)。その上、誘導音とテスト音の周波数が単純な比の関係(対2あるいは1対3)にあり、誘導音の音圧がテスト音の音圧より十分大きく、その結果、後者が大きなマスクングを受けるときには、二音間の位相の関係がテスト音の音の高さに大きな影響(半音以上の変化)を持つようである(Plomp, 1967a; Terhardt and Fastl, 1971; Lamoré, 1975)。

ここまで述べてきたのは誘導音とテスト音が同側耳に提示されたときに得られる効果についてであった。Thurlow (1943) と Terhardt (1977) は単耳に提示された純音の高さが、僅かではあるが、反対側耳に提示された音(純音或は雑音)によって影響され得ることを確認した。このことは純音の音の高さに及ぼす同時マスクング効果の原因が、蝸牛だけにあるのではないことを示唆している。

1.3.d. その他の効果

ある純音が他の純音の音の高さに影響を及ぼすためには、二つの音が同時に存在している必要はない。例えば、短い持続時間の純音の高さは、同じ耳に中くらいの音圧で、持続時間が長く、近い周波数の別の純音を先行させることで少し変化する。この効果は反発型である。つまり、テスト音の音の高さは、誘導音の高さがテスト音よりも低ければ上昇し、逆の場合には下降する(Larkin, 1978; Rakowski and Hirsch, 1980)。

さらに、「注意」の効果が純音の音の高さの知覚に影響し得ることも確かである。Houtgast (1976) は、例えば、周波数が f で、ピンクノイズによる強いマスクングを受けた純音が、被験者の注意のサブハーモニック(周波数 f/n 、ここで n は小さな整数)の音の高さに向けられているときには、そのサブハーモニックに対応する音の高さを持つように聞こえることが有り得ることを示した。注意の他の効果は Schouten and Domburg (1966) に述べられている。

1.4. 音の高さの二つの尺度

純音の音の高さ——これは音の主観的な次元なのだが——はどの様にして物理的な決定原理である音の周波数に応じて変化するのだろうか。

音の高さの「制度上の」尺度が存在する。それは西洋音楽の音階であり、それによれば、音の高さは周波数の対数の単純な関数である。音楽家の見地からすると、二つの純音 a と b が、二つの純音 c と d と同じ音程を成すのは(細かいことはここでは無視するとして——Ward, 1954 と Ohgushi, 1983を見よ)、 a と b の周波数比が c と d の周波数比と等しいときである。Attneave and Olson (1971) は、可聴閾の大部分で、この「制度上の」規則は日頃音楽に親しんでいない人についてさえ知覚的実在性を持つことをはっきりと示した。任意の聴取者に移調することを求めると、即ち純音の短い旋律をある周波数帯域から別の帯域に移すことを求めると、元の旋律における一連の周波数の関係は完全に--或は殆ど完全に--移調された旋律にも見られるであろう。その上音楽家は音程を、与えられた周波数の関係に正確に合わせることができ、その正確さは大変なものであると言われている。例えば、ある周波数の純音について、オクターブ音程(オクターブ音程は物理的には周波数がちょうど倍の関係であると定義されている)を続けざまに何度も作るように求めたとすると、音楽家なら、物理的オクターブとの差が1あるいは2%以下という精度で調整できるだろう(Ward, 1954)。

しかしながら、被験者に基準となる純音を与えて別の純音の周波数を調整させ、後の音の高さが基準

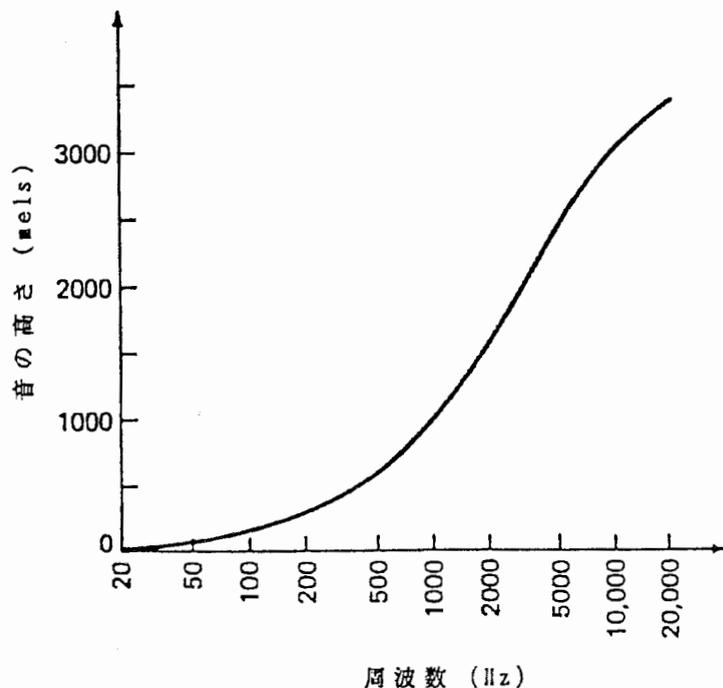


図 2. 周波数の関数として表わした純音の音の高さ、Stevens and Volkman (1940) による。"mel" は二つの音の高さの間のある決まった違いに対応しているとされる。周波数を表わす横軸は対数軸になっている。音の高さを示す関数が直線ではないことに注意、これは Stevens and Volkman によれば、純音の音の高さがその周波数のように対数的に変化しないことを示す。

音の音の高さの「半分に」なるように調整することを求めたとしよう。音楽的に言えばこの課題はあまり意味のあることではない。しかしこのような場合に、個人内で 10% のオーダーの正確さで判断ができることが示されている。基準音の周波数を変化させることで、実験により、音の高さと周波数の間の関係を記述する関数を得ることができる (Stevens and Volkman, 1940)。さて、得られた関数は明かに対数的ではない (図 2)。

音楽での音の高さの尺度と図 2 に示す「メル」尺度とが異なることは、もちろん気になることである。一方の尺度が正しく、他方が間違っていると考えるべきなのであろうか。この問題にはまたすぐに戻ることにして。

1.5. 純音の「クロマ」

純音の音の高さはその周波数を変化させるに従い、「低い—高い」の 1 次元上で変化する。しかしながらこの低い—高いという次元は純音の音の高さの全てを表現しているのであろうか。そう信じている研究者もいるが (Terhardt, 1980; Burns and Ward, 1982)、別の研究者は、実際は純音には二つの音の高さの性質があるのだと主張している。片方は低い—高いという性質であり、そして他方は「クロマ」の性質である。クロマは、周波数の関係が 2.0 となる二つの純音は——そして、従ってオクターブの音程を成す純音は——類似しているかあるいは同じである、という見地からみたときの性質である (特に Bachem, 1950 を見よ)。

大人の、音楽に親しんだ、あるいは親しんでいない被験者について行なわれた多くの実験から、クロマの概念をまじめにとるべきであることがわかる。このことに関して次のような事実を引用しておく。

1. 孤立した純音の音の高さを音名としてほとんど完璧な正確さで同定できる人がいる。しかしそのような人が犯す (稀な) 間違いはたいていオクターブを混同することである。例えば、彼らはファ 3 (174.6 Hz) をシ 3 (246.9 Hz) と取り違えることは決してないが、それに反してときどきファ 3 をファ 4 (349.2 Hz) と取る。これらのことは、彼らが音をその低い—高いという性質の基準よりもそのクロマの基準の上で同定しているかのようである (Lockhead and Byrd, 1981)。

2. ある時間間隔をおいて隔てられた二つの純音の音の高さを比較するような実験条件では、その期間中に「妨害」音 (純音) を提示することにより、成績が低下する。低下の度合は妨害音の音の高さに依

存する。しかしオクターブ離れた二つの妨害音による低下の効果は非常に似ている (Deutsch, 1973)。

3. 純音でつくられたよく知られたメロディーは、メロディーを構成する隣あう音の間の音程をオクターブ、あるいは2オクターブ増やしても容易にそれとわかる。それ以外の音程の増やし方をすると再認はずっと難しくなる (House, 1977; Idson and Massaro, 1978; Kallman and Massaro, 1979)。

4. 優秀な音楽家なら誰でも、二つの継時的な純音を作る音程を、その二つの音が大体60から5000 Hzの間の周波数であれば完璧に同定することができる(3度、4度、という風に)。しかしながら、この周波数の範囲を越えると、純音の音色は変化し続けるにもかかわらず、音程を同定する能力は急に弱くなる (Ward, 1954; Guttman and Pruzanski, 1962; Attneave and Olson, 1971)。この現象を、まず、音程とクロマにおける距離とが知覚的に相関を持つこと、それからクロマの存在する周波数範囲が低い—高いという性質のそれよりも範囲が限られていることを仮定することで説明できる。即ち、後者の領域は可聴周波数範囲全体と同じであるが、クロマを持つのはおよそ60から5000 Hzの範囲の周波数の音だけである、と仮定するのである。

これらの議論を合わせれば、純音のクロマが低い—高いという性質と同様、感覚の「基本的な」性質であることを納得させるに十分であろうか。否、というのはこれらの議論は全て音楽の教育を受けた大人の被験者について得られたデータを根拠としているからである。従って、これらのデータは基本的な感覚の実体よりはむしろ、文化的な条件づけを示しているのだと推論することができる。ここで西洋の音楽においては、そして、他の地域のほかの音楽体系においても、オクターブ音程をなす二つの音は同じ名前(D、レ、...)を持ち、また、機能的には等価であると考えられていることを思いだしてほしい。聴取者の音楽的な教養は心理学実験課題における聴取者の判断と、さらにその成績にも影響を与える可能性があることは全く疑いのないところである(このことについては、例えば、Francés, 1958; Krumhansl, 1979; Shepard and Jordan, 1984を見よ)。実際Allen (1867)のような研究では、二つの純音の類似性を直接(数字によって)判断する課題においては、音楽家は音楽家でない人よりもオクターブをなす二つの音に対して明らかに強い類似性の評価をすることが示されている。一方、Demany and Armand (1984)は3カ月の乳児でさえもオクターブの二つの純音を7度や9度をなす二つの純音よりもより類似したものとして知覚していることを示すことに成功した。この結果を文化的な条件づけの見地から解釈することは理にかなってはいないであろう。従って、純音の音の高さは、低い—高いという性質と同じものではなく、クロマの要素も同様に含む、という考えの正しさが今日ではしっかりと証明されたものと思われる。

van Noorden (1982)が指摘したように、純音の音の高さが単純な知覚上の性質ではなく、実際は独立した二つの性質からなるものだ、と認めることは、とりわけ、音の高さの二つの尺度——メル尺度と音楽的な尺度——の存在を説明することにつながる。このことについて、メル尺度は単に低い—高いという性質からの見方であり、一方音楽的な尺度はとりわけクロマにその基礎を置いているのだと考えることができる。

用語について考えることでこの章を終わろう。純音の音の高さは低い—高いという一つの性質に縮約されないと考えることが正しいのだから、この低い—高いという性質を言い表すためには特殊でより簡単な用語を適用することが有用である。AFNOR (1977)の推奨するところに従い、以後はこの性質を「トニー(tonie)」と呼ぶことにしよう。

II. 純音の音の高さの神経対応：データと仮説

II.1. 場所と時間：純音の周波数の神経符号化についてのデータ

II.1.a. 周波数弁別と興奮パタン

蝸牛の基底膜から(von Békésy, 1960; Khanna and Leonard, 1982)、聴覚皮質に至るまで(例えば Romani et al., 1982 をみよ)、哺乳類の聴覚系全体はトノトピックな機能様式を示している。このことは中耳以降の聴覚系の各段階において、純音によって引き起こされる機械的あるいは電気的な活動の場所が、その音の周波数に応じて組織的、かつ規則的に変化することを意味する。純音の周波数はその音の高さを決定する重要な物理量なので、音の高さの神経対応の性質について調べる際には、もちろんこの事実を無視できない。

ここでは、ヒトの聴覚系の異なる段階における、周波数のトノトピックな符号化の詳細に立ち入るようなことは問題にしない。またその詳細についてはまだよく解っていないのでなおさらである。そうではあるが、聴覚神経線維は基本的にバンドパスフィルターとして働くことを思いだしてほしい。ある音圧レベルにおける聴覚神経の平均発火率(一秒あたりの発火確率)は「特徴周波数」と呼ばれる、ある刺激周波数で最大となり、周波数が離れていくにしたがって低くなっていく(図3のaを見よ)。他方では神経の場所と特徴周波数との間に単純な関係が存在するので、任意の音圧レベルの純音は、聴覚神経のレベルにおいては、「興奮パタン」の位置によってその音の周波数が表現される。純音の音の高さの「究極の」生理学的表現が聴覚神経での周波数の場所的な表現から多かれ少なかれ直接的に生じるも

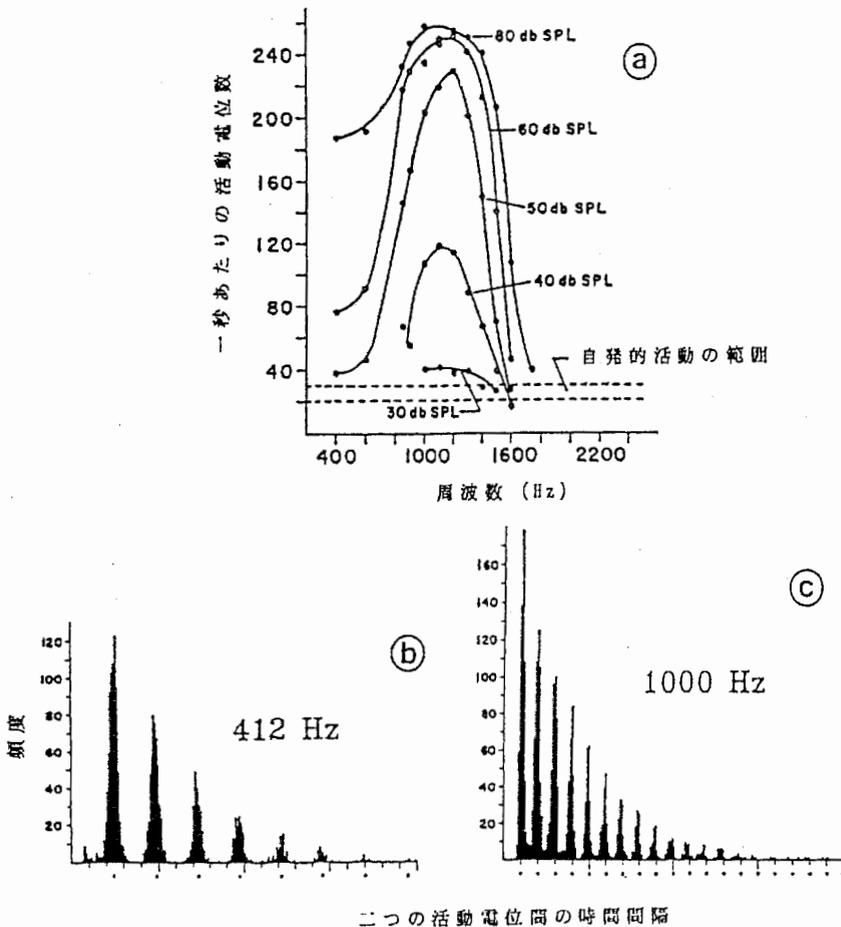


図3. リスザルの聴覚神経線維の純音に対する応答のデータ

(Rose et al., 1967)。グラフa: 音の周波数の関数としての平均発火率(1秒あたりの発火数)、媒介変数は音圧。「自発的(spontaneous)」活動は音刺激の存在しない状態での平均発火率である。グラフb, c: 412 Hzの音(グラフb)と1000 Hzの音(グラフc)に対する応答の微細時間構造を示すヒストグラム。横軸は連続する二つの発火間の時間(軸上の点は音の周期の1, 2, 3, ..., n倍に対応する時間を示す)。各ヒストグラムは、1秒の音について10回記録された結果得られた、異なる時間間隔の頻度を示す。

のだと考えてよい。

以下の点に注意を喚起しておこう——後でこの点にはまた戻るが——それは、ある周波数の純音が引き起こすヒトの聴覚神経内での興奮パタンの正確な位置は、おそらく音の音圧レベルと完全に無関係ではないということである。Evans (1977)と Greenberg et al. (1986) はどちらも、猫について、周波数 f の音の音圧レベルを増大させるとその音によって引き起こされる興奮パターンは、 $f < 1000$ Hzの場合には特徴周波数の低い神経の方向に、 $f > 1000$ Hzの場合には逆に高い特徴周波数の方向に、わずかに移動する傾向があることを示唆するデータを得ている（しかしShamma et al., 1986を見よ）。

II.1.b. 周波数の時間的な符号化

周波数 f の純音を与えたとき、基底膜の異なる場所で観察される振動は、その音と同じ周波数(f)を持ち、各場所で観察される振動は、その場所の変換細胞に見られる(周波数 f の)周期的な活動によって表現される。変換機構は聴覚神経の一本の線維の活動電位の始まりが、その対応する場所の基底膜の振動のある位相と時間的に一致する傾向があるという特性を持っている。その結果、一本の線維に見られる音によって生じた活動電位間の時間間隔は、その音の周期($1/f$)と一致する。この現象のリスザルでの実例が図3(bとc)に示されている。

しかしながら、この聴覚神経における純音の周期の符号化に関して、二つ重要なことを指摘しておかなければならない。

最初に、この符号化がリスザルとネコでは約5000 Hz以下、モルモットとチンチラでは3000 Hz以下に限定されていることである(Rose et al., 1967; Johnson, 1980; Russell and Palmer, 1986) ヒトに関しては、対応する限界は正確には知られていない。

次に、聴覚神経線維は——少なくともリスザルについては——音の周期の時間的な符号化を行なう際に組織的な「誤り」を犯しやすいということである。実際、図3やRose et al.によって公表されたヒストグラムを詳細に調べてみると、これらのヒストグラムの第一、及び第二モードが常に音刺激の正確な周期または倍の周期というふうに表現されているわけではないことに気づく。短い周期については、特に、過大評価のバイアスがあることが明かである。Ohgushi (1978, 1983)はこの現象に大いに興味を持っており、図4は彼が調べたものをまとめたものである。

図3と4に示したデータは聴覚神経の個々の線維の振舞いに関するものである。周波数 f の純音に対するこれらの線維の反応は統計的な性格のものである。というのは、各々の線維についてみれば、連続

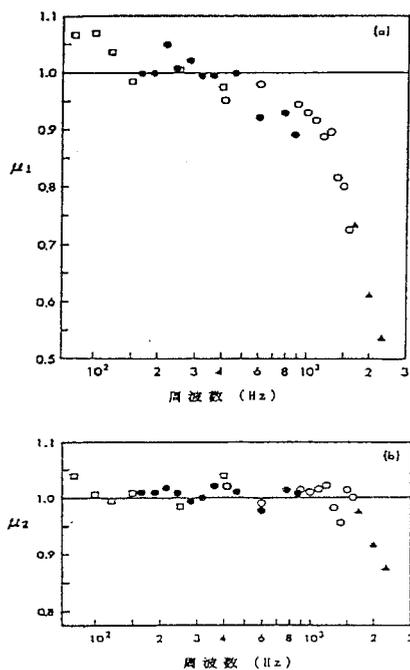


図4. 図3のグラフb, cに示されているものと似た(グラフb, cを含めて)Rose et al. (1967, 1968)のデータをOhgushi (1983)が分析した結果。異なる記号は異なる神経線維を表わす。横軸は音刺激の周波数。 μ_1 (上図)と μ_2 (下図)は商。 μ_1 については、割り算の分子は音の周期で、分母は連続する発火間の時間間隔のヒストグラムの第1モードに対応する時間。 μ_2 については、分子は音の周期の2倍で分母はヒストグラムの第2モードの値。 μ_1 と μ_2 は、もし周期の符号化に誤りが全くなければ常に1.0に等しくなければならない。

する活動電位にみられるほぼ $1/f$, $2/f$, $3/f$ 等に対応する時間間隔の出方は非常に確率的なものであるからである。ところで、他方ではある線維の振舞いはその近くの線維の活動とは比較的独立であるようで、音のある周期に一方の線維が応答しないのに他方は応答するということがある(Johnson and Kiang, 1976)。このことから、純音の周期に近いいくつかの線維が送り出す時間情報を統合することによって、中枢神経系はただ一本の線維から得られる情報よりも文句無しにより豊かな情報を得ることができるのだ、と考えられる(このことに関しては Patterson, 1986 を見よ)。

二種類の音の高さをヒトが区別できることが、これら二つの音の高さの神経対応が空間的に区別されることを意味するのは明らかだと考えて、何人かの研究者は(例えば Whitfield, 1970 を見よ)、純音の音の高さの究極的な生理学的対応は時間的な性質ではあり得ず、場所的なものでしかあり得ないであろうと主張した。この主張の論理ははなはだ曖昧である。とはいえ、いかにして、聴覚神経線維が純音に対する反応として伝達する時間情報が、聴覚中枢において音の高さとの究極的な対応を持つであろう場所的な情報として復号化され得るか、ということを実証する目的で多くの理論的な努力がなされてきた。

このことに関して、Licklider がかなり以前に提案した理論が(Licklider, 1959 を見よ)いまだに影響力を保っている。それは、図 5 に示されるような活動電位の時間的系列の自己相関関数が単純な神経回路によって決定できるという考えに基づいている。この回路では、音の高さの生理学的対応は、同じ型の他のニューロンよりもその音に対してより強く応答する自己相関器ニューロンの位置に相当する、とされている。

別のモデルが Sruлович と Goldstein (1983) によって提案されている。これは Young and Sachs (1979) が初めて述べた考えに基づくものである。しかし、このモデルでは、聴覚系に図 3 のようなヒストグラムの表現を作り上げる能力が要求されるほか、そのようなヒストグラムをある種の周波数スペクトルを持った関数のようなものとして処理する能力が要求される。Pikels (1986) のいうように、神経系に対する要求が過大な原理については、その生理学的な実在性を疑うことが可能であり、そしてこの理由からこのモデルについて、本論文では詳しく扱わないことにする。

Shamma (1985, 1986) は、聴覚系は純音に対する末梢神経の発火によって伝達される場所的情報と時間的情報を同時に利用することができるのではないかと——これは、基底膜は純音に対して非等速度の進行

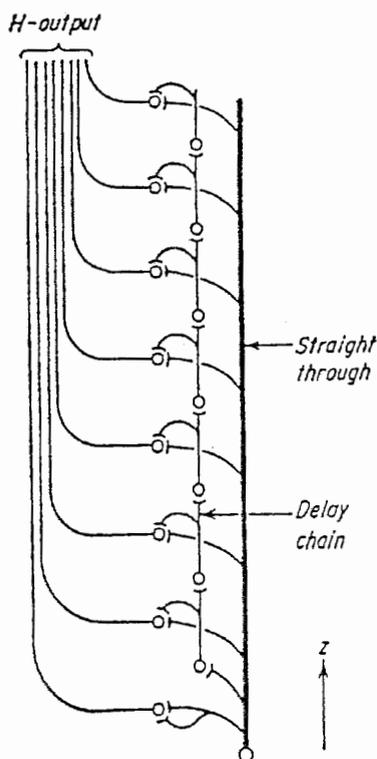


図 5. Licklider (1959) の神経自己相関器。Licklider の想像したシステムでは、聴覚神経線維で運ばれた純音の時間的情報は、シナプス伝達での遅延の結果、いくつかの種類自己遅延と関係を結ぶようになる。もとの情報のある部分は「直通路 ("Straight through")」を通過して運ばれ、そのほかの部分は遅延系列 ("Delay chain") を介する第 2 の通路を通過して運ばれる。二つの通路は自己相関器ニューロン集団のレベルでは一緒になり、そこで各々のニューロンは一方ではもとの情報を、他方ではそれに時間 t の遅延の掛かったものを受け取る。 t は「遅延系列」路上で通過するシナプスの数に依存し、また、図のいちばん左にある自己相関ニューロンのそれぞれについて異なる自己相関ニューロンは t の単調関数にしたがって空間的に組織されていることがわかる。元の情報を受け取った t ms 後で情報を受け取る自己相関ニューロンでは、周期 t の音の発火では二つの入力情報は時間的に一致するが、違った周期の音の発火では一致しない。それ故、時間情報を場所的に再符号化するためには、二つのインパルスを受け取ったときにだけ反応するようなニューロンを仮定するだけで十分である。

波により応答する、という事実によっているのだが——と指摘している。ある周波数の純音によって基底膜の段階で生じる力学的な波は進行波なので、基底膜上の異なる二点における、音によって生じた振動の間には、時間のずれ——位相のずれ——が生ずる。しかし波の速度は一様ではなく、基底膜上で振動の振幅が最大となる領域では速度は最低となる。この結果、振動の振幅が最大となる領域で、基底膜のとなり合う二点における振動の間の位相のずれは最大となる。それ故、振動の相対位相が空間的に最も速く変化する基底膜上の位置がわかれば、刺激純音の周波数を同定するのに役立つことが可能であろう。ところで、位相のこの速い変化の起こる場所は、聴覚神経線維によって符号化できる。なぜなら、神経線維で活動電位がみられるのは、基底膜上の対応する位置での振動のある位相においてであり（これは周波数の時間的符号化と同じ原理である）、問題の位相は少なくとも局部的に定常であるからである。

Shamma のものと似てはいるが、ずっと思弁的な考えがLoeb et al. (1983) によって述べられている。

II.2. 音の高さの専ら場所的な符号化に有利な議論

「純音の音の高さは生理学的にはどの様に符号化されているのだろうか」という疑問に対して、過去にはしばしば、可能な答えは二つしかなく、専ら場所的な符号化の仮説と、専ら時間的な符号化の仮説のどちらかに決着をつけてしまうことが重要であると考えられた。この議論はその後、進展をみた。しかしながら、純音の音の高さの符号化についての「極端な」仮説がどの程度支持され得るのかを調べておくことは有意義である。ここでは、時間的な情報には少しも関わらず、それ故、専ら場所的な音の高さの符号化の仮説に有利なものとして引用できる論拠について調べることにしよう。

最初に考慮する事実は、すでに指摘した生理学的な事実である。それは、聴覚神経線維は、ある周波数以上の——ネコとリスザルについて得られたデータから、人間についても 5000 Hz であろうと推定されるが——純音の周期についての時間的情報を伝送することができないということである。このことは 5000 Hz 以上の音については、音の高さの符号化は場所的なものでしか有り得ないということを示唆する。しかし、それ以下の周波数の音についてはどうなるのであろうか。

I.3. 節において、純音の高さがその周波数だけに依存するのではなく、音圧レベル、音が提示された耳、そして、他の音が存在するか、あるいは存在しないか、といった他の要因にも依存することを示す精神物理学的な相当数の結果について述べた。ずいぶん長い間、このような「例外」は音の高さの場所的な符号化の仮説でしか説明がつかないとされてきた。実際、時間的な符号化のメカニズムが音の周期を表現する際に組織的な誤りを犯し得るということなど思いもよらなかった。既に指摘したように、現在ではもはやそのようには考えられていない。だがこれらの例外のうちの一つに限りそうではない。それは音の高さへの音圧レベルの影響に関する Stevens の法則であって、これは聴覚神経における周波数の場所的な符号化についてのデータと、少なくとも定性的には一致するようだ (Evans, 1977; Greenberg et al., 1986; I.1.a. を見よ)。従って、Stevensの法則は、実のところ音の高さの専ら場所的な符号化の仮説に有利な論拠であると考えることができる。ともかくも、これまで引用されてきた生理学的データには、実際は解釈上の問題があることを強調しておかねばならない (Shamma et al., 1986)。あいにく、I.3. で述べたほかの例外に当てはまるような生理学的な事実はまだ明らかにされていない。

今日場所説に有利であるとして引用できる最も強力な議論が確かに存在する。それは略して「Zwickerの音」と呼ばれているものである。Zwicker (1964) はある周波数領域におよそ 1/2 オクターブの穴のあるスペクトルを持った雑音を長く聞くことで、視覚の領域で知られているものに似た「残効」が生じることを見いだした。後に起こる感覚は、スペクトルの穴の領域の周波数を持つ純音が引き起こすものとよく似ていて、その穴の中心周波数に対応して変化する。Fastl (1986) によれば、Zwicker の音を聞くのに十分有利な条件下では、Zwicker の音による音の高さと現実の純音による音の高さを、純音の周波数を調整して一致させたところ、その精度は 2% であった。Zwicker によって発見された現象は、場所説を強力に支持する、というのは、十中八九、幽霊音を知覚する際の聴覚系内の神経活動には時間的な微細構造がほとんどないからである。この現象はトノトピックに構成された神経全体の中で、スペクトルの穴の領

域にあたる特徴周波数を持つ神経が、幽霊音を引き起こす雑音が終わった直後に、他のニューロンよりも高い平均発火率で発火する（雑音による「疲労」あるいは「順応」）と仮定しない限り説明できないように思われる。

II.3. 時間的符号化説に有利な議論

ここまで、純音の音の高さについての専ら場所的な符号化の仮説を擁護してきたので、今度は反対の仮説、つまり、音の高さを符号化するために聴覚系は聴覚一次神経によって伝達される時間的情報を利用する、という仮説にとって有利なものとして引用できる議論について調べることにしよう。このような時間的情報が場所的な情報の結果として生じるか否かについてはここでは触れないことにする。

定常的、継時的な純音の周波数弁別閾に関して二つの議論を引合いに出すことができる（これらの閾値が音の高さの弁別閾であると考えることがもっともであると言える限りにおいて）。

まず最初は、実際、場所説には不利な議論である。場所説は明らかに、基準の純音と近い周波数の他の純音を弁別する聴覚の能力が、基準の周波数付近での聴覚系の音の周波数弁別力の程度と強い相関を持つことを要求する。被験者の聴覚フィルターが選択的であればあるほど、基準音が引き起こす興奮パターンはより場所的に限定されたものとなるし、そしてそれ故、音の周波数はより精度よく決定されなければならない。ある周波数における、被験者の聴覚系の周波数選択度はマスキングの実験によって精神物理学的に推定可能である（例えば Zwicker, 1974 を見よ）。従って、もし場所説が正しいのであれば、このようなタイプの実験結果と、音の周波数弁別閾の実験結果との間で、この点に関して強い相関が見られなければならない。ところが事實は、そのような相関は實際上皆無である（Tyler et al., 1983; Moore and Glasber, 1986）。

第二の議論は、音の周波数と持続時間との関数としての、周波数弁別閾の変化についての Moore (1973) のデータによるものである。厳密に言えば、音はその波形が正弦波でしかもその持続時間が無限大でなければ、本当の意味で純粋な音（ただ一つの周波数しかスペクトルに含まない）ではない。正弦波の持続時間を短くすれば、物理的現象としてはそのスペクトルが変化する。持続時間が短くなるほど、よりスペクトルは広がり、スペクトル包絡線の傾斜は緩やかになる（例えば Leshowitz, 1978 を見よ）。周波数の場所的な符号化仮説によれば、このことは、正弦波音の周波数弁別閾は、持続時間が短くなれば高くなることを意味する。そして、実際に少なくとも約 200 ms より短い持続時間でこのことが観察できればよい。しかし、Moore はこのことについて注目すべき現象を発見した。200 ms 以下では、持続時間を減少させたことによる閾値が相対的に、周波数が 4000 ないし 5000 Hz から 6000 Hz を少し越えたところで急激に上昇するということである。この現象で特に注目すべきことは、境界の周波数がヒトの聴覚神経で時間的な符号化が可能な周波数限界の推定値と一致していることである。Moore の結果からこの時間的な符号化が有効であると考えられる。というのは、この結果は、短い正弦波音の周波数弁別閾がそのスペクトル包絡によって決まるのは、音の周波数が 5000 Hz を越えるときだけであることを示唆するからである。

これ以外に時間説に有利な議論として、周波数弁別閾には無関係な精神物理学的な二つの事実を挙げることができる。

まず、この文脈で I.5. 節で既に引用した事実を思いだしてほしい。それは音楽家なら誰でも、継時的な二つの純音の音程を同定することは、その二つの純音の周波数が大体 60 から 5000 Hz の間であれば容易なことであるのに、二つの純音のうち一方の周波数が 5000 Hz を越えた場合には非常に難しくなるということである。一般に、このように、純音の音の高さはその周波数が時間的符号化がされないようになると、もはや「音楽的な」ものではなくなってしまうようである。

最後に引用する議論は明らかに最も強力なものである。それは片方の耳、あるいは両方の耳で聴力に異常のある被験者の、単耳による純音の音の高さの知覚に関する研究に基づくものである (Florentine and Houtsma, 1983; Turner et al., 1983)。この二つの研究の被験者は低い周波数で（一側性、あるいは両側性の）聴力損失があった。従って、彼らの異常のある側の耳で、聴力の悪い周波数の上限より少し低い周

波数 f の純音は、一次神経の特徴周波数としては、 f の周波数より上の周波数で検出されるのではないかと予想される。少なくともある被験者においては、この二つの研究以外の研究でこの予想を裏付ける精神物理学的なデータが得られている。問題の被験者については、ようやく聞こえる、そしてその周波数 f が聴力に異常のある領域に含まれる純音 s は、周波数 $2f$ の純音によって、あるいは $3f$ の純音によってすら、 f に近い周波数の純音によるよりも容易にマスクされることがはっきりしている（反対のことが正常な側の耳で観察されている）。ここで、もし場所説が正しいとしたら、このタイプの純音 s の音の高さがどうでなければならぬかについて考えることにしよう。 f よりもずっと上の特徴周波数の末梢神経しか興奮しないのだから、その音の高さは正常な耳にとっての f に対応する音の高さに比べて、よりずっと高いものになるであろう。さて、実際にはそのようなことは全くないことが確認されている。例えば、片方だけに聴力損失がある被験者に、異常のある耳に提示された音 s に、反対側の耳に提示された周波数が調整できる純音を合わせるように教示すると、被験者は平均値としては f と有意な差なしに調整を成し遂げてしまう。ここで引用した二つの研究は、実のところすぐ前に引用したすべての他の結果と同様に、時間説を直接支持するものではない。しかしながら、それは聴覚神経において、純音の音の高さが音によって生じた興奮パタンの場所と形だけで、また単純に、決定されるとする考えに強い疑いを抱かせるものである。

II.4. 思弁的総括

II.2. と II.3. で述べた事実を合わせると、二つの結論を引き出すことができるように思われる。第一に、聴覚神経に与えられる純音の周波数についての時間的情報は（その周波数があまり高すぎなければ）音の高さの知覚に直接的な役割を果たしている可能性が非常に高い、ということである。このことについてのもっとも強力な証明は、二つの純音とその蝸牛における興奮パタンの位置が異なっているにも関わらず同じ高さであるように知覚されるということである。第二に、音の高さの感覚は、それにも関わらず、そのような情報なしでも生じ得るように思われる。というのは、5000 Hz 以上の音にも音の高さの違いはあるし、他方では、「Zwicker の音」が存在するからである。この二つの結論から、聴覚神経のレベルでは、5000 Hz 以下の周波数の純音の音の高さは場所的な様式と時間的な様式で同時に（そして独立に）符号化されていると推論するのが妥当であると考えられる。

この仮説はあまりにも漠然としているため、一見ほとんどなにも言っていないように思われる。しかしながら、純音の音の高さが一つではなく二つあると仮定することは、純音が、一つではなく二つの音の高さの性質を持っていることを認めれば、より魅力的なものとなる。それはトニーの性質と、クロマの性質である（I.5. を参照）。そこで、van Noorden (1982) のように、これらの二つの性質が聴覚神経のレベル以降では、同じ様式で符号化されていないのだと想像することができる。van Noorden の説は、より詳しく言うと、クロマの側面は時間的に符号化されるが、トニーは場所的に符号化される、というものである。

この説は今のところ一つの思弁以外のものとは見なしかねるものであろう。しかしながら、音の高さの独立した二つの側面が、聴覚神経内で独立に符号化されているという考えは、多チャンネルの人工内耳を埋め込まれた被験者についての Tong et al. (1983) の最近の研究結果によって強く支持されている。Tong et al. は人工内耳を用いて、患者の聴覚神経の一つを、次の可能な異なる三通りの周期的電気刺激で刺激した。それは (1) 周期の点で、(2) 刺激される神経の位置の点で（各刺激は神経の限られた範囲しか興奮させない）、(3) 前の二つの両方の点で、異なるものである。種々の刺激の主観的な違いについて、被験者に主観的な強さの面から調整させ、定量的に比較させた結果、彼らは「刺激の周期」という変数と「刺激される神経の位置」という変数が知覚的に全く異なった質的影響を持っていることを明確に示すことに成功した。

上に引用してきたのは、音の高さに聴覚神経で別々に符号化される二つの性質が存在するという考えのものであった。しかし van Noorden の理論はより具体的である、というのはこの二つの性質がトニーとクロマであるとしているし、それぞれの性質に対してははっきりと別の符号を割り当てているからである。場所的符号をトニーに、時間的符号をクロマに割り当てることは容易に正当化される。クロマはオ

クターブの音程にある二つの純音が共有している感覚属性であると定義されているので、オクターブの二つの純音が類似したクロマのコードを持っているとするのは理にかなっている。ところでクロマの符号が時間的なものであるとしたら、それはこの場合にあたる。周波数 f の純音に対応する聴覚神経の発火時間間隔は、ほぼ $1/f, 2/f, 3/f, \text{etc.}$ である。そして同じ間隔が周波数 $f/2$ の純音、つまり f の下のオクターブについても同様に存在する、なぜなら $2/f$ の倍数はすべて $1/f$ の倍数と等しいからである。クロマとは反対に、トニー——低い—高いという性質——は周波数とともに単調に変化する。従って、その符号が周波数とともに明らかに単調に変化する、とすることは理にかなっており、それは、実際の生理学的なデータにより、もしこの符号が場所的なものであったとすればこの場合にあたる。

繰り返すが、ここに述べてきた理論は単に思弁的なものとみなさなければならない。明らかに、純音の音の高さの知覚は今日でもまだよくわからないままである。その生理学的なメカニズムについて、聴覚神経系の末梢以降での純音の音の高さの表現に関する重要なことが不明確なままであるため、まだ我々に知られていないことこそが疑いもなく本質的なことなのである。

III. 複合音の音の高さ：総論

III.1. 「基本音の」高さ、「生の」高さ、「スペクトルの」高さ

ヒトの聴覚系に音の高さの感覚を生じさせることのできるすべての音のなかで、生態学的に最も重要なものは、複合的で周期的（あるいは概周期的）なものである。母音や大部分の楽器が出す音がこれにあたる。フーリエの定理により、周期が $p = 1/f$ の周期的複合音はすべて、周波数の異なる純音、そしてそれぞれの周波数が、総数を n としたとき、 $n \cdot f$ の形となる純音の和として表現できる。ここで周波数 f は周期的複合音の「基本周波数」と呼ばれる。ヒトの音環境中で典型的な周期的複合音は「調波的な」周期的複合音である。基本周波数 f の周期的複合音は「調波的」であるといわれる。それはスペクトルの成分をなす純音がちょうど f の連続した倍数の周波数を持つからである。そのスペクトルは、周波数 $n \cdot f$ （最後のものは除いて）の各成分に周波数 $(n+1) \cdot f$ の成分が続くようになっている。それ以外の全ての周期的複合音は「非調波的」であると言われる。

被験者に周波数の調整できる純音の音の高さを調波的周期的複合音（これはそのスペクトルがある周波数 f の聴取可能な完全な倍数だけで成り立っているもののことである）の高さに合わせるよう求めたとすると、被験者は一般的に、純音の周波数を f に非常に近い値に合わせるであろう。別な言い方をすれば、被験者は一般的に周期的複合音と純音が非常に近い周期になるように調整するだろう。時々、 f ではなしに $f/2$ あるいは $2f$ に調整されることも有り得る（周期のオクターブエラー；Davis et al., 1951 と Ritsma, 1966 を見よ）。しかしながら、被験者に続けて何度かの調整をするように求めると、少数の中央値の周りに調整値が集中してくることが確かめられるであろう。その意味では非常に正確である。このような課題（調波的周期的複合音と純音とを合わせる）での調整の精度は、二つの純音を合わせる課題よりも、確かに少し悪くなる（周期的複合音のスペクトルが、基本周波数成分がその振幅において他の成分よりずっと支配的であるようなものでなければ）。しかし、これは調波的周期的複合音によって引き起こされる音の高さが純音によって引き起こされる音の高さよりも、一般的に、著しくより際だって、あるいは「顕著で (saillante)」ないというわけではない (Fastl and Stoll, 1979)。むしろ被験者は、周期的複合音と純音との音色の違いが課題をより難しくしているようだ、というであろう。

この音色の違いは物理的には二つの音のスペクトルの違いによるものである。純音のスペクトルは一本の線になるが、調波的周期的複合音のそれは一連の線と振幅の包絡——その最大のところは基本周波数ではなしに倍音のところにくることも可能である——からなっている。基本周波数は等しいがスペクトル包絡が異なっている二つの周期的複合音は、音の高さはほぼ等しいのに、音色はやはり異なっている。一方のスペクトルのエネルギーが低周波よりも高周波でより大きく、他方はその逆だとすると、被験者は最初の音は二番目の音よりもより「明るい (brillant)」、あるいはより「鋭い／高い (aigu)」と言うだろう。

このために「低い (grave)」と「高い (aigu)」という言葉を使うことができるのだから、ここで言う音色の属性は、純音について定義した「トニー」のようなものと、明らかに非常に密接な関係にある。Risset (1971) はこの属性を指すのに、「音色」という言葉ではなしに「生の高さ (hauteur brute)」という表現を用いている。我々は、彼が「音色」という言葉は世間一般の言語としては、そして、音響心理学者たちについても同様ののだが、非常に広い意味を持っていると言ったことにしたがうことにしよう。そして周期的複合音の生の高さは、たとえそれを「音色」だといえたとしても、最終的には音の高さの属性と別のものではないのだから、ここでは前に周期的複合音の「音の高さ」と述べたものを指すのに「基本音の高さ (hauteur fondamentale)」というより特殊な表現を提案することにしたい。この表現は英語で "periodicity pitch", "residue pitch", "low pitch", "complex pitch", あるいはさらに "virtual pitch" という言葉が指し示しているものと全く同じものを言い表している。

周期的複合音の基本音の高さはオクターブの曖昧さを示すので、生の高さの性質がトニーの性質と類似しているのと同様に、基本音の高さの性質もクロマの性質と類似している。しかしながら、純音の音の高さが二つの成分——トニーの成分とクロマの成分——を持つことを認めない研究者はいるのに、

生の高さと基本音の高さが周期的複合音の音の高さの全く別の二つの側面であることを否定する人はいない。このことは周期的複合音の生の高さと基本音の高さを、周期的複合音の別々の物理的特徴を変化させることによって全く独立に変化させることが可能であるのだから、全く馬鹿げたことであろう

(Plomp and Steeneken, 1971; Risset, 1971) さらに Risset (1971) と Ueda and Ohgushi (1987) は、周期的複合音の生の高さと基本音の高さの相対的な目立ちやすさが被験者によって異なることを確かめている。

Risset によれば、音楽家は常に基本音の高さに対して特に敏感であるのに、非音楽家のうちのある者は基本音の高さよりも生の高さに対してずっと敏感であるという。しかし、Ueda and Ohgushi (1987) の研究では、被験者の音楽教育の程度と、生の高さ、あるいは基本音の高さに対する相対的なウェイトとの間には強い相関は認められなかった。

ある周期的複合音について、基本音の高さよりも生の高さにより敏感になり得る、あるいはその逆ということがあるものの、生の高さと基本音の高さは周期的複合音により自発的に引き起こされる音の高さの二つの属性であると言ってよい。一般に——我々の日常の環境中で生じる周期的複合音の大部分について——これら二つの属性は、それを理解するための注意力を必要とせずに、意識の中に現われる。

それは Helmholtz (1863) が言うところの「合成的な」聴き方によるものである。しかしながら、Helmholtz は「合成的な」聴き方と「分析的な」聴き方を区別した。それは、真剣に努力して注意を向ければ、いくつかの純音を含む周期的複合音の中から各純音をそれぞれの高さで聞き分けることができる、ということによっている。このことに関する興味深い研究が Plomp (1964), Plomp and Mimpen (1968), それに Martens (1981, 1984) によって行なわれた。これら全部の教訓を数行で書くことは不可能である。しかし基本的には、これらの研究は、振幅が等しく、周波数が $2f, 3f, 4f, \text{etc.}$ (f は大体 60 から 2000 Hz の間に含まれる) の純音からなる周期的複合音の中から、最初の 5 つか 6 つの純音が、すなわち最初の 5 つか 6 つの「倍音」が、分離して聞こえることを示した (繰り返すが、被験者が注意深い態度で臨むなら、である)。複合音のスペクトルの一成分である純音によって生じる音の高さを特に指し示す表現を決めておくべきである。以下で用いるのは「スペクトルの高さ (hauteur spectrale)」という用語である。

複合音のスペクトル成分が分離して知覚されるという事実は、聴覚末梢神経系のトノトピックな構造と関係し、非常に問題となる (II.1.a. 参照)。なぜなら、聴覚系はトノトピックに組織されているので、周波数が非常に離れ、かつレベルの低い二つの純音からなる複合音は、聴覚神経の全く分離した二つの領域を興奮させ、神経線維は実際上二つの純音のどちらか一方にしか反応しない。他方では、振幅が等しく、周波数が $f, 2f, 3f, \text{etc.}$ のスペクトル成分を持つ周期的複合音について、最初の 5 つか 6 つの成分だけが聞こえるという事実は、二つの生理学的事実から理解される。(1) 聴覚神経線維の周波数選択性は無限ではなく、有限である、(2) このレベルにおけるトノトピックな周波数軸は、線形関数より対数関数にずっと近いので、2 対 1 の比率の二つの周波数は、6 対 5 の比率の二つの周波数よりも耳にとっては、はるかに「間隔のあいた」ものになる (Liberman, 1982 を見よ)。

III.2. 周期的複合音の基本音の高さ

III.2.a. 基本音が欠けていることの問題

周期的複合音を「合成的に」聴いたときの音の高さの二つの性質の区別をしたので、これからはこの二つの性質のうちの一つである基本音の高さについてだけ述べることにしよう。基本音の高さの知覚は生の高さよりもずっとよく研究に取り上げられている。これは少なくとも二つの理由による。まず、周期的複合音の基本音の高さは一般的に、生の高さよりもずっと正確に知覚的に定義できることである。次に、そしておそらく何よりも、基本音の高さの知覚の方が生の高さの知覚よりずっと分からないからである。実際、なぜ多数の純音からなる調波的周期的複合音が、スペクトルの高さでは全く異なっているのに、その基本音成分のスペクトルの高さに対応する、あのように明瞭な一つの基本音の高さを (オクターヴの曖昧さを除けば) 引き起こすことになるのだろうか。ここにパラドックスがある。

現象の逆説的な側面は、基本周波数は f であるが、周波数 $2f, 3f, 4f, \text{etc.}$ あるいは周波数 $3f, 4f, 5f, \text{etc.}$ し

かスペクトルに含まない周期的複合音について、周波数 f がそのスペクトルの中に物理的に含まれているときとその値が実際と異なる基本音の高さが生じることが確かめられることに特に明かである。これが「基本音欠如」の謎と言われているものである。

ここで欠如した基本音の知覚が成人に限られたことでないことに注意したい。8カ月の乳児が既にそれを知覚し (Clarkson and Clifton, 1985)、ネコやアカゲザルも同様に聞き取る (Heffner and Whitfield, 1976; Tomlinson and Schwarz, 1988)。しかしながら、人間については、この現象が起こるためには左の聴覚皮質が無傷であることが必要であるようである (Zatorre, 1988)。純音の弁別課題においてはこの皮質が無傷であることは必要ではないようなので、この事実は注目に値する。Whitfield (1980) によれば、ネコについても同様に、欠如した基本音の知覚には聴覚皮質が決定的な働きをしている。

神経学的に異常のない成人については、欠如した基本音を知覚する際の必要な条件は、相対的には少し限定されている。大ざっぱに言って、基本周波数が f でスペクトル成分が $n \cdot f$ の調波構造周期的複合音は、次の場合に f に対応する基本音の高さを生じる。(1) スペクトルに現われる n の最小値が 20 以下であるとき (Ritsma, 1962; Hoekstra, 1979)、(2) 少なくとも 2, 3 個のスペクトル成分の周波数が約 5000 Hz 以下であること (Ritsma, 1962; Fastl and Stoll, 1979)、(3) f が 60 ないし 70 Hz 以上であること (この第 3 の点については後でまた戻る)。二つのスペクトル成分 $n \cdot f$ と $(n + 1) \cdot f$ だけからなる調波的周期的複合音でも f に対応する基本音の高さを生じさせることができはするもの (Smooenburg, 1970; Houtsma and Goldstein, 1972)、物理的に基本音の欠如した周期的複合音の基本音の高さは、その周期的複合音のスペクトル成分が増すにしたがってより顕著となることをつけ加えておこう。

周期的複合音の基本音の高さの顕著さが、基本スペクトル成分の物理的振幅によらないことを明らかにしたのは Seebeck (1841, 1843) が最初であった。Seebeck の仕事以来、基本音欠如の謎は——つまり、実際には基本音の高さの知覚に関する全ての問題は——多くの思弁を引き起こした。Schouten (1970), Wightman and Green (1974), Plomp (1976) そして de Boer (1976) が 70 年代まで、この年代は特に大きな謎の解決にエネルギーが注がれた時期であるが、これらの思弁の活気に満ちた変化について述べてきた。1988 年でも、謎は依然として明瞭な解決がついたわけではないが、例えば Schouten (1970) が彼の章を書いた頃よりはいくらかは明らかになってきている。

歴史的には、欠如した基本音のパラドックスを最初に説明したのは Seebeck (1841, 1843) と Helmholtz (1863) であった。これらの説明は今日では完全に放棄されているので手短かに述べることにしよう。Seebeck は聴覚系は波形の周期を計測することによって周期的複合音の基本音の高さを決定しているのだと想像した。これは欠如した基本音のパラドックスを説明する。なぜなら、周波数 $f, 2f, 3f, 4f, \text{etc.}$ の成分からなるスペクトルを持つ周期的複合音はそのスペクトルの周波数 f を取り除いても依然周期的であり同じ周期 ($1/f$) を保つからである。不幸にして、Seebeck の仮説は聴覚末梢系がスペクトル分析器のように機能すること、すなわち周期的複合音を純音に分離することを無視していた。Helmholtz は彼自身としては、欠如した基本音の現象は単に中耳の非線形性の結果であるという見解を持っていた。この非線形性は差の結合音 ($f_1 - f_2$ の形で) を生み出すことに現われる。彼はこの非線形性の結果、末梢での (物理的な) $2f, 3f, 4f, \text{etc.}$ という形のスペクトルは常に中耳で $f, 2f, 3f, 4f, \text{etc.}$ の形のスペクトルに変換される。このスペクトルの f の成分 (差の結合音) は大きな振幅を持つので知覚される基本音の高さの原因になり得る。Schouten (1940) はこの仮説に見事に反論した。その一連の研究についてはここでは詳しくは述べない。III.2.b. で後に彼が実現し、この仮説の弱点を非常にはっきりと示した別の研究について詳しく述べることにしよう。

Helmholtz の仮説が誤りであることを確信する一方で、基本音の高さを導く現実的なモデルの中では、聴覚末梢系が全ての複合音について実現しているスペクトル分析を考慮にいれなければならないことも意識していたので、Schouten (1940) は図 6 の中程に示すような基本音の高さの理論を提唱した。この図では聴覚末梢系がトノトピックに配列されたバンドパスフィルターバンクとして (左の部分) モデル化されている。周波数軸は対数で、異なるフィルターの中心周波数が左に示されている。もちろん、選ばれた周波数の値は恣意的なものであり、フィルターの数についても同様である。フィルターに関して、このモデルは単に次のことを仮定しているに過ぎない。(1) このフィルターの数は膨大なもので、その中心周波数は、ほぼ対数をなす軸に沿って規則正しく等間隔で並んでいる。(2) フィルターの通過帯域とその中心周波数の間隔がほぼ一定である (この図ではここに示されている全てのフィルターは対数周波数

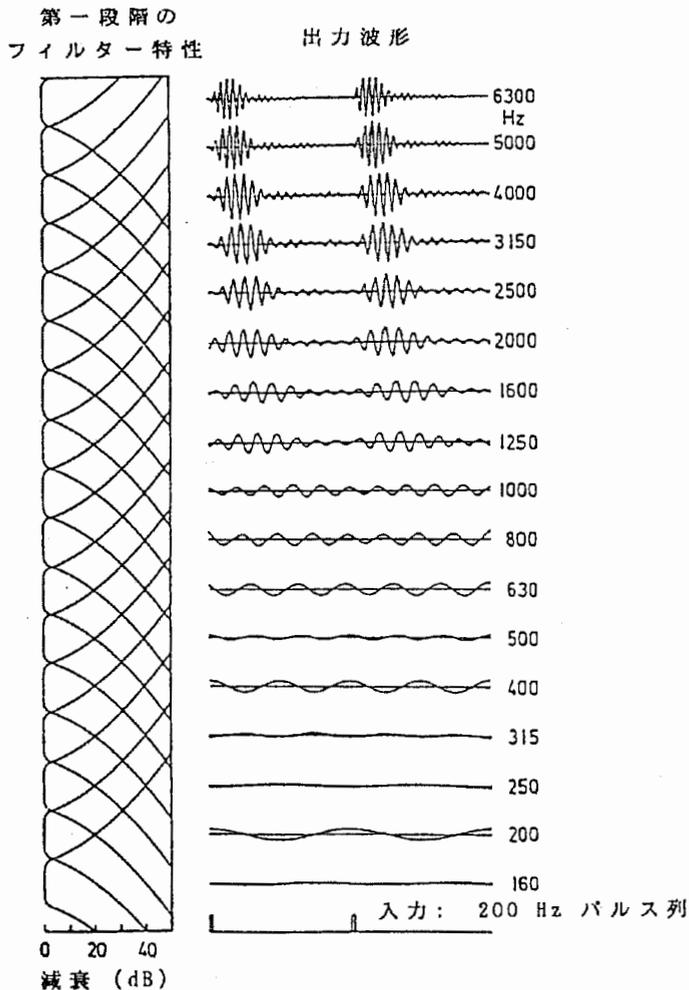


図 6. 聴覚末梢系による単極パルス列のスペクトル分析モデル。
Plomp (1966) と de Boer (1976) による。

軸上で同じ減衰特性を持っているので、その間隔は厳密に一定である)。図の真ん中の部分は周期的な(基本周波数 200 Hz の)単極パルス列の入力信号に対する各フィルターの応答を示している。この信号のスペクトルは等振幅の 200 Hz, 400 Hz, 600 Hz, ... $n \cdot 200$ Hz の周波数を含んでいる。中心周波数の低い(1000 Hz 以下)のフィルターのこの信号に対する応答は、ほぼ平坦な関数となるか(中心周波数 315 Hz のフィルターのようにほとんど何も応答しない)、信号のスペクトル成分のうちの一つをその周波数とする正弦波となるか(例えば、中心周波数が 400 Hz のフィルターがこの場合にあたり、それは第 2 倍音しか「通過させ」ない)、であることがわかる。反対に、フィルターの通過帯域の絶対的な値がより大きい、中心周波数の高いフィルターは、振幅応答の変化がより少なく、とりわけ、個々の倍音を分離できなくなる。各々のフィルターの応答は、連続するいくつかの倍音の波形を組み合わせたものとなり、それ故、入力信号と同じ周期を持つ複合波の形をとる。Schouten はこれらの波形の周期が聴覚系の蝸牛以後の段階で測定され、これが音の基本音の高さの知覚の基礎をなしているのではないかと示唆した。この図で考えている刺激の基本音成分あるいはさらに最初の 3 つないし 4 つまでの倍音を取り去れば、モデルが予測する基本音の高さは基本音成分のスペクトルの高さと同じままであることは明白であるはずであった。

従って、Schouten のモデルでは、周期的複合音の基本音の高さの知覚は、基本的には聴覚末梢系が周波数の近すぎる純音をスペクトル的に分離することができない、ということにより説明されることになる。そしてこのモデルの枠組みの中では、結果的には、全ての周期的複合音について聴覚末梢系が行なっているスペクトル分析は基本音の高さを抽出する妨げとなっているかのように見える。

今日では——なぜか、ということについては後でまとめて調べることにしたいのだが—— Schouten のモデルはどうやら完全に間違ってはいないように見受けられる。しかしながら、一方ではこのモデルが完全に正しいというわけでもないことは確かである。Schouten のモデルに最も損害を与えたのは Hout-

sma and Goldstein (1972) の研究である。彼らは、少なくともある人たちについては、一方は左耳に提示される $n > 1$ で周波数 $n \cdot f$ の、他方は右耳に提示される周波数 $(n + 1) \cdot f$ の、二つの純音からなる刺激によって、 f に対応する基本音の高さが生じることを示したのである。Schouten のモデルでは欠落した基本音を知覚するためにはスペクトル成分の蝸牛での相互作用が必要であると仮定されているので、この現象を説明することはできない。なるほど、Houtsma and Goldstein の刺激音によって生じる基本音の高さは非常に曖昧であり、また、この刺激音は聴覚系にとってきわめて「異常な」ものである（このことに関しては Hall and Sonderquist, 1975 をみよ）。だがしかし、複合音の基本音の高さに関する他のデータも Schouten のモデルと折り合いが悪い。例えば Hall and Peters (1981; Houtsma, 1984 も参照のこと) は f の第 3, 4, 5 倍音を継時的に——時間的に全く重ならず——提示することで f に対応する基本音の高さが生じることを示した。これは倍音が非常に速く続くことが条件である。Schouten のモデルとうまく合わない他のデータについてはもっと後で述べることにしよう。ここでは、Schouten のモデルの理論に真っ向から反対する先駆けとなった理論は、70 年代に陽の目を見た、ということで満足しておこう（けれども全ては de Boer に多くを負っている）。Schouten にとっては蝸牛がスペクトル分析器のように機能するという事実は、周期的複合音の基本音の高さ抽出の際の障害であったが、新しい理論家たちにとっては、この事実は反対に基本音の高さ抽出にとって重要な、そして「積極的な」段階であると考えられている。このことについては後でもっと詳しく述べることにしよう。

III.2.b. 非調波複合音により生ずる基本音の高さ

非調波的周期的複合音も、調波的周期的複合音の基本音の高さと同じ性質、同じ明瞭さの音の高さを生じさせることができる。例えば、 $n \cdot 200$ Hz のスペクトル成分を持ち、200 Hz の極めて近くに対応する基本音の高さを明瞭に生じさせる調波的周期的複合音があったとしよう。この成分のうち一つの周波数を 20 Hz 増やしても、得られた基本周波数 20 Hz の非調波的周期的複合音は、質的には——後でみるように、値の上ではわずかに異なるのだが——元の音の基本音の高さと全くよく似た音の高さを生じさせる。従って、この非調波音の音の高さも基本音の高さと呼ぶことができる。もっとも、その感覚は、この音の物理的な基本周波数である 20 Hz の純音によって生ずるものとは、極めてかけ離れたものではあるが。

何人かの研究者が、ここまで述べてきたような型の非調波的周期的複合音の基本音の高さに興味を持った (Schouten, 1940; de Boer, 1956; Schouten et al., 1962; Smoorenburg, 1970; Patterson, 1973)。Schouten et al. (1962) の研究は特に反響を呼んだものであるし、非常に詳しく書かれるに値するものである。

Schouten et al. は標準刺激として、周波数が $n \cdot 200$ Hz + df の純音の振幅を周波数 200 Hz、変調度 90% で正弦波変調したものをを用いた。各刺激音は 200 Hz 間隔の 3 つのスペクトル成分を含み、各々の周波数は $(n - 1) \cdot 200$ Hz + df , $n \cdot 200$ Hz + df , $(n + 1) \cdot 200$ Hz + df であった。被験者の課題は、各標準刺激の基本音の高さに、基本周波数 f が調整できる調波的周期的複合音をマッチングすることであった。この調波的周期的複合音のスペクトルは、周波数 $(n - 1) \cdot f$, $n \cdot f$, $(n + 1) \cdot f$ (n は標準刺激と同じ値) からなっていた。 n の値は 7 から 12 まで変化した。ただし、異なる値の df を用いる一つの測定系列内では n は一定に保たれた。 df の値は負、ゼロ、正にすることができ、50 Hz ステップで変化された。

実験結果について述べる前に、標準刺激の一般的な波形について考えておこう。合成方法からわかる

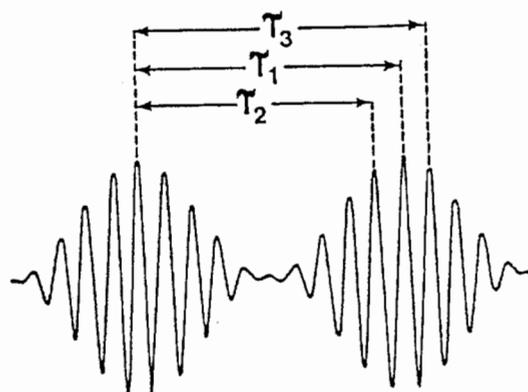


図 7. Schouten et al. (1962) の実験で用いられた刺激の波形。この刺激は 2000 Hz の純音の振幅を 200 Hz の正弦波で変調して得られた。その三つのスペクトル成分は周波数 1800, 2000, 2200 Hz である。 $\tau_1 = 5.0$ ms; $\tau_2 = 4.5$ ms; $\tau_3 = 5.5$ ms。Warren (1982) より複製。

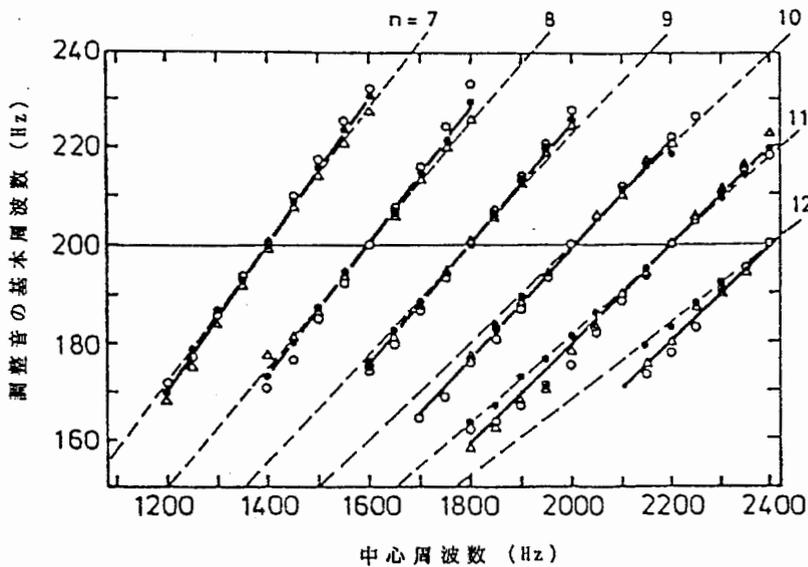


図 8. Schouten et al. (1962) の実験結果。Plomp (1976) による。

ように、全ての標準刺激の振幅包絡は周波数 200 Hz (5 ms 周期) の正弦波であった。しかしながら、波形の微細構造は n と df に依存する。この微細構造にもっと興味を向けてみよう。そして df が 0 (あるいは ± 200 Hz) に近い場合を考えてみよう。標準刺激はこの場合、調波的周期的複合音である。図 7 は $df = 0$, $n = 10$ の標準刺激の波形を示している。微細構造について見てみると、波形はある周期 (τ_1 に対応する) を持ち、それは振幅包絡の周期と等しいことがわかる。次の包絡の「腹」についても、微細構造は変化しない。反対に、 df が 0 (あるいは ± 200 Hz) に等しくないときには、波形の微細構造の周期は包絡の周期よりも長くなる。それ故、この微細構造は次の包絡の腹では変化していて、例えば、前の例のように包絡の各最大点が微細構造の振幅頂点と一致するようなことは、もはや起こらない。

実験結果は図 8 に示されている。この図で、横軸は $n \cdot 200$ Hz + df の値、すなわち、標準刺激の「中心」周波数である。縦軸は調整された f の値、すなわち知覚された基本音の高さを示している。白丸、黒丸、三角は異なる 3 人の被験者のデータに対応している。真ん中の斜めの線はそれぞれ、 n のある値について得られた実験データを示す。この図を調べてみると、3 つのきわめて重要な事実が認められる。

(1) n が一定であれば、調整された f の値は df に依存する。より正確に言えば f の調整値は df の線形関数である。 n の各値について、この線形関数の傾きは df/n に近い (図の鎖線の傾きは df/n とぴったり同じである)。

(2) それにも関わらず、実験で得られた線形関数の傾きは df/n とは異なっている。そしてこの違いは n が増すにつれて大きくなる。即ち、 n が増えれば、傾きは df/n よりもより大きくなる。

(3) n の各値について、200 Hz ちょうどか、あるいはそれを越える df に対する調整値は、直線上にぴったり合ったものが得られる。非調波的な標準刺激に限らず、標準刺激の大部分について、 f の調整値は一つの中心値の周りにではなく、明らかに離れたいくつかの中心値の周りに散布している。例えば、中心周波数 1800 Hz の標準刺激については、4 つの中心値が認められる。200 Hz の調整値は別に驚くべきことではないが、そのほかにも、160 Hz, 175 Hz, 227 Hz の付近にも調整値がみられる。このことは調波的周期的複合音の基本音の高さが、既に指摘したオクターヴの曖昧さ以外の曖昧さを示すことを意味している。しかしながら、この図について、 f の全ての調整値が同じようにはっきりとした基本音の高さに対応していると推論する必要はない。

得られた事実の中で最も重要なのは最初のものである。実際、この事実だけから、周期的複合音の基本音の高さを抽出するメカニズムについての二つの重要な教えを得ることができる。まず、この事実は III.2.a に述べた Helmholtz の仮説が不適當であることを示す。この仮説は、思い出してほしいのだが、基本音欠如のパラドックスを説明するために、聴覚系で差の結合音が生じることを持ち出している。

Schouten et al. の用いた基準音はすべてスペクトル成分の間隔が一定であるので (200 Hz)、この仮説では第一の事実を説明することができない。次に、第一の事実は複合音の基本音の高さに物理的に対応するものが音の振幅包絡の周期ではないことを証明している。もしそうであったなら、実際のとこ

ろ、全ての基準音は同じ振幅包絡を（周期が 5 ms の正弦波）持っているのであるから、df の調整に与える影響もなくなってしまうであろう。

第一の事実と第三の事実を同時に説明するのに、全く異なった二つのやり方が存在する。

その最初のもは、標準刺激において基本音の高さを生じさせる物理的対応物は、音の波形の微細構造の二つの振幅の頂点の間の時間間隔である、とするものである。（このことはもちろん、対応する時間間隔が、なんらかの方法で聴覚系によって測定されることを前提としている。）この解釈をよりよく理解するために図 7 を再検討してみよう。これは中心周波数 2000 Hz の基準音の波形を示しており、振幅頂点の三つの時間間隔が (τ_1, τ_2, τ_3) 示されている。図 8 はこの基準音について得られた f の調整値が中心周波数の三つの値の周囲に分布していることを示している。測定値の直線の傾きはすべて df/n で近似することができるのだから、あるいは、別の言い方をすれば、測定値の傾きを鎖線のように考えるのだから、これらの三つの値はそれぞれ、181.8, 200, 222.2 Hz と推定できる。ところでこれらの周波数は正確に $1/\tau_3, 1/\tau_1, 1/\tau_2$ Hz に対応している。中心周波数 2000 Hz の基準音は調波的周期的複合音である。しかし、この特定の周期的複合音についてと同様に、他の全ての基準音について、 f の調整値の違いと波形の振幅頂点の間の時間間隔との間を関係付けることが可能である。

第一と第三の事実についての第二の可能な説明は、時間領域ではなしに、スペクトル領域から出発している。例として中心周波数 1950 Hz の基準音を取り上げよう。スペクトルとしては周波数 1750, 1950, 2150 Hz の三つの純音の和として表現されるこの基準音を聴く被験者を想像してみよう。さらに被験者は f を調節するために次のような質問に答えようとするのだ、と想像してみよう。「どんな基本周波数なら、自分の聴いている三つの純音が、連続する三つの倍音になるのだろうか」、と。もちろん、基準音は非調波なので完全に満足できる答えはない。しかしながらほとんど完全に満足すべき答えが存在する。例えば、三つの純音を 9, 10, 11 という倍音列であると考えれば、可能な最も良い答えは $f = 195$ Hz である。ところでこの答えが良い答えであるのは 195 Hz の第 9, 10, 11 倍音 (1755, 1950, 2155 Hz) は 1750, 1950, 2150 Hz に非常に近いからである。良さの程度は落ちるが、三つの純音を第 8, 9, 10 倍音であると考えれば、満足すべき答えが同様に見つかる。この仮説のもとで可能な最良の答えは $f = 216.7$ Hz で、そう決めたときに被験者が犯す実際の純音の周波数との誤差は最大約 17 Hz となる（先の場合は 5 Hz であった）。三つの純音を第 10, 11, 12 倍音であると考えれば、答えは $f = 177.3$ Hz となる。従って純音の周波数との誤差はさらに大きくなるが 23 Hz は越えない。図 8 に示された測定値の直線の傾きは df/n に相当するので、中心周波数 1950 Hz の基準音について得られた f の調整値は、これまで述べてきた 177.3 Hz, 216.7 Hz, 195 Hz の三つの f の値に対応しているとこの図から結論できる。中心周波数 1950 Hz の基準音は非調波である。しかし先の論理は調波的な基準音の場合にも同様にあてはまる。例えば中心周波数 2000 Hz の基準音について、 $f = 200$ Hz という答えはこの論理に従えば完璧に満足すべきものである。しかし、181.8 Hz あるいは 222.2 Hz という f の調整値も同様に説明できる。なぜなら、刺激のスペクトル成分 1800, 2000, 2200 Hz は 181.8 Hz の第 10, 11, 12 倍音 (1818, 2000, 2182 Hz) の周波数にも、222.2 Hz の第 8, 9, 10 倍音 (1778, 2000, 2222 Hz) の周波数にも近いからである。

Schouten et al. の結果についての可能な二つの説明については後でまた述べることにしよう。今のところ、さきに上げた三つの事実のうち二番目が解説されずに残っている。なぜ、実験で得られた直線の傾きは n が増えるにしたがって df/n から離れる傾向があるのだろうか。ここまでに述べた二つの説明的モデルはどれも、この疑問に答えてはくれない。なぜなら、いずれのモデルも、経験的に得られた傾きが df/n と厳密に等しいことを要求しているからである。

III.2.c. スペクトル優位性の現象

Plomp (1967b), Ritsma (1967), Smoorenburg (1970), Brookes (1982), そして Moore et al. (1985) の努力のお陰で、今日では第二の事実が明かとなった。この結果は二つの現象——これは各々それ自体が重要なのだが——の結びついた結果として理解されるべきなのである。

第一のものは蝸牛で、ある結合音が生じることである。それは $f_1 - f_2$ の形の音ではなくて、基本的に $2f_1 - f_2$ のタイプの音である。いくつかの研究により（特に Plomp, 1965 を見よ）、同じぐらい弱い音圧（40 あるいは 50 dB SPL）の周波数 f_1 と f_2 をそのスペクトルに含んでいる末梢刺激が出されれば、 $2f_1 -$

f_2 の形の結合音がおそらく聞こえるであろう、ということが示された。例えば周波数 $9f$, $10f$, $11f$ からなる周期的複合音について、この現象により $8f (2 \times 9f - 10f)$ と $7f (2 \times 9f - 11f)$ のスペクトル成分が末梢スペクトルに付け加わることになる。それ故、この現象によりスペクトルの重心が低い周波数の方へ移動することになる。この型の歪は、先に述べた二つの説明モデルのどちらに従っても、測定された傾きを理論的な df/n の傾きよりも「突っ張らせる」に違いない。

さらにより重要な第二の現象は、「スペクトル優位性」の現象と呼ばれている。今日ではいかなる周期的複合音の基本音の高さの抽出においても、音の異なるスペクトル成分に対して聴覚系が与える「重み」はこれらの成分の物理的な振幅のみには依存しないということが知られている。単に特権的な周波数領域に存在しているというだけの理由で、ある成分は特権を持ち、他を圧倒するのである（そして、あるいは、調波的周期的複合音については、その中に含まれる倍音列に特権があるという理由で）。別の言い方をすれば、今日では任意の周期的複合音の基本音の高さの抽出は、刺激そのままについてではなく、その刺激にフィルターが掛けられ、スペクトルが減少したものについて行なわれる傾向があることが知られている。

スペクトル優位性の現象についての興味深い研究が、最近 Moore et al. (1985) によって行なわれた。等しい振幅の、基本周波数 f の最初の 12 の倍音 (f , $2f$, $3f$, ... $12f$) からなる周期的複合音を考えよう。この周期的複合音は調波的で豊富なスペクトルを持っているので、非常に明瞭で、 f にきわめて近い基本音の高さが生じる。しかし、Moore et al. がやったように、そのスペクトル成分のうちの一つ c の周波数を $x\%$ 増やして、「不協和に」したとしよう。もし c が低い倍音で、かつ、 x が大きければ（例えば $x = 15\%$ であれば）、新しい音は純音 (c) と元の音と同じ基本音の高さを持つ周期的複合音を足し合わせたような音に自然と聞こえてくる。即ち、 c はもはや他のスペクトル成分とは溶け合わなくなる。しかしながら、そしてこれこそが何よりもここで重要なのだが、もし x が小さければ ($1-3\%$)、知覚された結果は異なるのである。 c はその場合、元の刺激と同様に他の成分と溶け合い、結果として生じる非調波的な音の基本音の高さは元の調波的な音のそれよりもわずかに高くなる。もっと正確に言うと、基本音の高さの上昇は（周波数の変化したパーセンテージで表わす）、各倍音番号 c について x の一定の割合になる。しかしこの割合の正確な値は c の周波数（そして、あるいは、倍音番号）に依存することが明らかにされている。このようにして、この実験でスペクトル優位性の現象が明らかになった。すなわち、刺激の全てのスペクトル成分は同じ振幅なので、もしスペクトル優位性が存在しなかったとすれば上昇の比率は c とは無関係でなければならない。明らかに、ある成分 c について求められた上昇の比率は基本音の高さの抽出におけるスペクトル優位性の程度を直接指し示すものである。そのずれが基本音の高さに影響を与えれば与えるほど、成分 c は支配的であると考えられるべきである。

ここに述べてきた実験では、Moore et al. は 100, 200, 400 Hz の三つの f の値を用いている。彼らの結果はスペクトル優位性の現象に関して重要な個人差が存在することを示唆している。被験者全体について得られたデータから二つの結論を引き出すことができる。(1) 用いられた周期的複合音全てについて、第 6 倍音以上のスペクトル成分のスペクトル優位度はゼロである。(2) 最も優位なスペクトル成分は第 1 から第 4 倍音で、これらはほぼ同じ程度優位である。

Moore et al. より以前に、違った方法ではあるが、Plomp (1967b) と Ritsma (1967) が既にスペクトル優位性の現象の証拠を示している。これらの研究者の結果は Moore et al. のそれと完全には一致しない。従って、スペクトル優位性の正確な法則は未だにうまく定義されていない。しかしながら、それは調波的で豊富なスペクトルを持つ全ての周期的複合音の中で、比較的周波数の低いスペクトル成分が、そして低い番号の倍音 (7 以下) が、基本音の高さの抽出において他の成分より優位にたつ傾向があることに（一般的に言って、おそらく、基本周波数成分こそが最も優位な成分だというわけではないであろう）明らかに寄与しているのである。このついでにこの法則が示唆する重要なことを指摘しておく。優位な倍音は、「分析的な」聴き方によって純音として聴くことが可能な倍音である、ということである (III.1. 参照)。

III.2.b. で説明した Schouten et al. の実験について、なぜいままで述べてきた規則が二番目の事実を予測できるのかを理解することは難しいことではない。さきに述べた二つの説明モデル——一つは「時間的」なもの、もう一つは「スペクトル的」なものは、基本音の高さの抽出器が元の信号そのままに

対して機能することを前提としている。測定値の傾きが df/n を越えるという事実は、基本音の高さの抽出器が実際にはスペクトル優位性の原理のために、スペクトル成分の周波数が高くなる程、そのスペクトル成分がより減衰していくような信号に対して機能しているのだ、ということ認めれば説明することができる。

III.2.d. 位相の要因

調波的周期的複合音の波形は、そのスペクトル成分間の位相の關係に決定的に依存している。例えば基本周波数 f の、三つの純音の和からなる、次式のような形で記述できる周期的複合音 s を考えてみよう。

$$s = 1/2 \cos 2\pi \cdot 9f \cdot t + \cos 2\pi \cdot 10f \cdot t + 1/2 \cos 2\pi \cdot 11f \cdot t \quad (2)$$

この式は次の式と数学的には等価である。

$$s = \cos 2\pi \cdot 10f \cdot t (1 + \cos 2\pi \cdot f \cdot t) \quad (3)$$

これにより、 s は周波数 $10f$ で、その振幅が周波数 f で正弦波的に変調された純音であると見なせることがわかる。信号 s の実際の波形は図7に示されているようなものである。

今度は s の中心の、周波数 $10f$ の成分の位相を、他の二つの成分より 90 度ずらしたとしよう。すると次のような信号が得られる。

$$s' = 1/2 \cos 2\pi \cdot 9f \cdot t + \cos (2\pi \cdot 10f \cdot t - \pi/2) + 1/2 \cos 2\pi \cdot 11f \cdot t \quad (4)$$

信号 s' の波形は図9に示されている。信号 s に見られる振幅変調が s' では大幅に減少していることがわかる。 s' は実際には周波数が $10f$ で、その周波数と振幅が同時に変調されている純音として記述することが可能である。

周期的複合音のスペクトル成分の位相關係の変化がその波形に影響することを測定するに際して、そのことが同様に、その音によって生じる基本音の高さにも影響するのかどうか、と問うことはもちろん重要である。しかしながら注意しなければならないことがある。位相關係の変化は、波形以外に耳で生じる結合音の振幅にも影響する恐れがあり、それ故、刺激の「内的な」スペクトルにも影響する恐れがある（例えば Goldstein, 1970; Buunen, 1976 を見よ）。位相の影響を波形の影響として解釈できるようにするために、だから、結果をスペクトルからだけでは解釈できないようにしておくことが重要である。

Moore (1977) は $f = 200$ Hz の信号音 s と s' によって生じる基本音の高さの感覚を比較した。スペクトルのアーチファクトを避けるために、彼は信号音に結合音をマスクする低域雑音を付加するという配慮をした。彼の研究は基本的には、 s と s' の基本音の高さに純音の周波数を調整することを課題とするような状況では、 s の基本音の高さは s' の基本音の高さよりもより明確で、より知覚が容易であることを明らかにした。 s については、Moore の被験者はたいていの場合は比較純音を 200 Hz のすぐ近くに調整したが、 s' についてはいかなる基本音の高さも知覚されていないように見えるような調整の仕方がしばしば見られた。これはまさしく III.2.b. で述べた基本音の高さの「時間的」モデルが予想していたことである

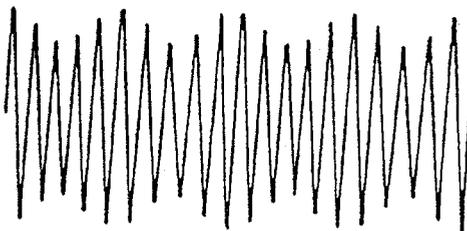


図9. (4)式で記述された刺激 s' 。この図の刺激は実際にはやや図式化されている。振幅包絡が図7の刺激とは明らかに異なることに、何よりも注意せよ。

ことに注意しよう。信号音 s の振幅包絡 (図 7) は特に $1/200 s$ の長さの時間間隔 τ_1 が「突出して」いるのである。同じ振幅頂点間間隔が s' の波形にも存在するが (図 9)、その位置はあまりはっきりしない。

従って Moore の研究は調波的周期的複合音の波形が、それ自体で、この周期的複合音により生じる基本音の高さに影響を与える可能性があることを示唆している。しかしながら s と s' のような周期的複合音については、スペクトル成分は全て基本周波数から高く離れた倍音であり、また、聴覚系にとって近い周波数である (末梢レベルで完全には分離できない) ことに注意しよう。このようなことがあてはまらない周期的複合音については、逆に波形は基本音の高さに基本的に何の影響も及ぼさないことが明らかにされている (Patterson, 1973; Wightman, 1973)。従ってスペクトル優位性の現象の結果 (III.2.c)、そのスペクトル成分が全て 200 Hz に続く倍音であるような周期的複合音は、スペクトル成分間の位相関係とは全く無関係な基本音の高さ (200 Hz に対応する) を生じさせるであろう。その基本音の高さの抽出に対して優位な成分は、実際には低い番号の倍音であって、それゆえ耳にとって互いに比較的離れたものであろう。

III.2.e. 基本音の高さとスペクトルの高さの関係

周期的複合音の基本音の高さがその波形に依存しているのかどうかと問うことは、結局、基本音の高さの抽出において蝸牛でのスペクトル成分の相互作用がどれくらい重要であるか、と問うことになる。同様に、そして逆に、個別に取られた各々のスペクトル成分に担われた聴覚情報の重要性について問われることになる。これは、ある意味で、III.2.c. で述べた Moore et al. (1985) の研究でなされたことである。しかしこの研究は実際には、次のように表現できる根本的な疑問について答えてはいない。周期的複合音の基本音の高さとそのスペクトル成分のスペクトルの高さとの間に、基本音の高さがスペクトルの高さから派生したものであると、別の言い方をすればスペクトルの高さをもとに抽出されたものであると推測することができるような関係があるのだろうか、ということである。この疑問は、先に述べたように、豊富なスペクトルの周期的複合音の優位なスペクトル成分は「分離できる」、即ち純音として独立に聞こえるので、ますます適切なものとなる。

この疑問の意味を明確にしよう。I 節で純音の音の高さはその周波数だけで決定されるものではなく、音圧や、純音が提示される耳や、前後に音が存在するか、あるいは存在しないか、ということによっても決定されることを見てきた。もし周期的複合音の基本音の高さが、分離して聞こえるスペクトル成分のスペクトルの高さから得られるのであれば、この基本音の高さはスペクトル成分の周波数だけに依存するものではない、と予想しなければならないであろう。例えば、ある音圧で周波数 $f = 4f$ の純音 s についてある両耳異聴を示すような、 s を右耳より左耳では少し低いと知覚するような被験者がいたとしよう。この被験者に対して左耳には音 s を、右耳には周波数 $5f$ と $6f$ の二つの純音を提示するものとしよう。このダイコティック周期的複合音の最も顕著な基本音の高さは f に近いであろう。しかしながら、それは s が右耳に提示されたときと正確に同じもの

(値として、別の言い方をすれば「顕著さ」の側面を考慮にいれずに) であろうか。van den Brink (1974) によれば、答えは否である。ダイコティック周期的複合音の基本音の高さは、 s が右耳に提示されればやや低くなり、周期的複合音の基本音の高さが物理的周波数やスペクトル成分よりむしろ、スペクトルの高さに依存していたと予測に従うのである。ついでにいうと、このことから「調波的周期的複合音の基本音の高さ (最も顕著な) とは、その複合音の基本音の音の高さである」と言えば、真実を近似的にしか表現していないことになる (基本周波数と同じ周波数の純音の音の高さはその周波数値にだけ依存しているのではないので、真実はそれ自体相対的なものなのだが)。

ここまで書いてきた実験だけが基本音の高さとスペクトルの高さの関係について捧げられたものではない (Walliser, 1969b; van den Brink, 1975a, 1975b, 1977; Houtsma, 1981; Hall et Soderquist, 1982; Stoll, 1985; Hall, 1986; Houtsma and Rossing, 1987 を見よ)。これら引用した全ての実験結果は完全に一致しているわけではない。しかしながら、今のところ、周期的複合音の分離可能なスペクトル成分が、その物理的周波数よりむしろ、そのスペクトルの高さによって周期的複合音の基本音の高さを決定するのだ、という考え自体はもっともらしく、確実ですらある、と考えることは正しいようである。

III.3. 雑音により生じる周期性と音の高さの感覚

III.3.a. 「凍った」雑音の知覚

1963年に Guttman and Julesz が興味深い聴覚現象を記述している。長く持続する白色雑音を磁気テープに録音し、この録音の任意の断片を切り出し、その初めと終わりを「くっつけて」輪にするものとしよう。この輪をテープレコーダーで再生して得られる刺激は「凍った」雑音、周期的な雑音である。この刺激は実は調波的周期的複合音であり、もしその周期が非常に短ければ（例えば 10 ms 以下）純音の音の高さと容易にマッチングできる基本音の高さを生じさせるであろう。しかしながら、Guttman and Julesz はこの刺激は周期が1秒を越えるまでは真の白色雑音と知覚上区別できることを確かめている。周期が約 200 ms を越えると、先に述べたような音の高さの感覚と共通する特徴のない、周期性の感覚の付け加わった連続雑音によるような知覚が得られる。Limbert and Patterson (1982) によれば、周期の始まり（そして終わり）に知覚上対応する輪の点は被験者によって個人差が生じる余地があり、被験者内では同じ輪をもう一度提示しても変わらなかった。

周期が 200 ms を越える刺激により生じる周期性の感覚は、音の高さの感覚とは明らかに異なるので、このことにはこだわらなくても良いようにも思える。しかし本当に問題なのは、周期が十分に短いときには、音の高さ（あるいはむしろ基本音の高さ）の感覚が十分生じるということ、そして、従って、周期を段々長くしていけば、多少とも連続的に基本音の高さの感覚から異なった感覚へと移っていく、ということである。なんらかの客観的な基準で、それ以下では基本音の高さの感覚が生じ（「音調的

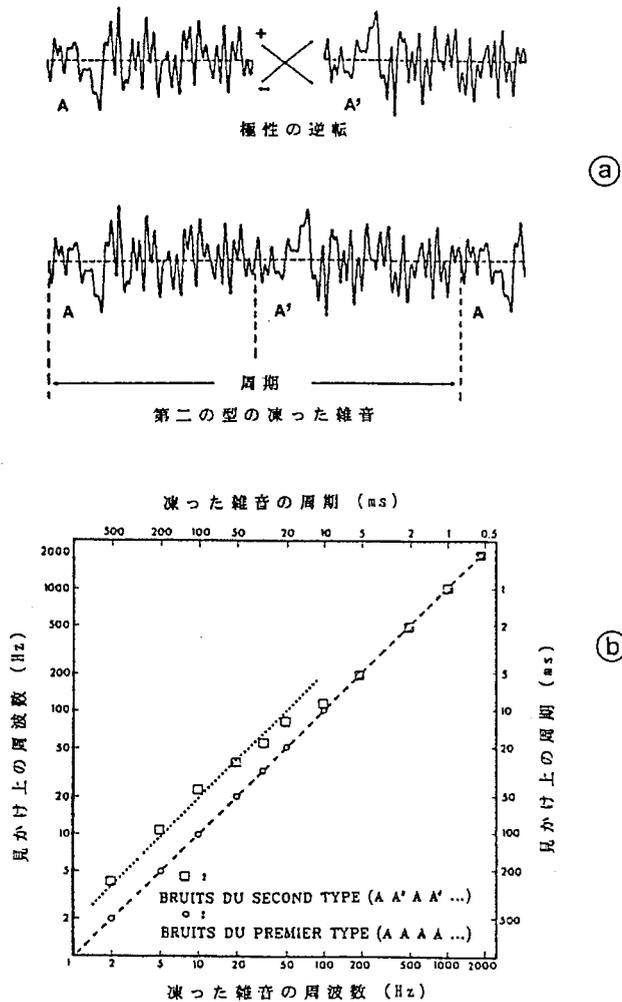


図 10. グラフ a: Warren and Wrightson (1981) の実験の「第 2 の型」の雑音の作り方。グラフ b: 実験結果。Warren (1982) を改変。

(tonale)』である)、それ以上ではもはや基本音の高さの感覚が生じない(「非音調的(atonale)」である)、という境界の周期を定義することが可能だろうか。

この疑問に対する一つの答えを Warren and Wrightson (1981) が述べている。彼らの実験は以前に Flanagan and Guttman (1960) によってなされた実験に着想を得たもので、その結果を一般化したものである。その実験では二つのタイプの凍った雑音が用いられた。第一のタイプの雑音は白色雑音の断片 A それ自身をそのまま繰り返して作られたものである。第二のタイプの雑音では、この断片 A と、図 10 の a に示されているように、A の極性を単純に逆転することによって得られた断片 A' が交替する。従って第二のタイプの刺激の客観的な周期は、第一のタイプの刺激のように A の長さではなく、その二倍の長さとなる。被験者の課題は知覚される周期、あるいは基本音の高さの観点から、各基準雑音と周期を調整できるパルス列をマッチングさせることである。図 10 の b に実験結果が示されている。まず、この図から第一のタイプの雑音(丸印)については、調整された周期は常に客観的な周期に等しいことがわかる。このことは明らかに予想できたことであった。しかしもっと興味深いのは第二のタイプの雑音(四角印)について得られた調整値である。客観的な周期が 20 あるいは 30 ms 以上である限り、知覚された周期は客観的な周期の半分であり、このことは断片 A と A' が互いに同じものとして扱われていることを示している。周期が 20 から 10 ms になり、さらに 10 ms 以下になると、ことは非常に急激に変化し、知覚される周期は客観的な周期と組織的な対応がつくようになる。従ってこの結果は、調波的周期的複合音の周期(周期的な雑音は、思いだしてほしいのだが、調波的周期的複合音なのである)の知覚的な同定のメカニズムが、その周期が約 15 ms (これは約 65 Hz の基本周波数に対応する)以上のときとそれ以下のときとは異なっていることを示唆する。基本周波数が 65 Hz 以上のときに働くメカニズムは、明らかに周期的複合音から基本音の高さの感覚が抽出されるようなメカニズムを考えなければならないので、この実験から、物理的対応が 65 Hz 以下の周波数であるような感覚をさして基本音の高さと言うべきではない、と結論してよい。この考えを正当化する他のデータが Bilsen and Wieman (1980)、同様に Warren et al. (1980) によって得られている。

III.3.b. 振幅が周期的に変調された白色雑音の知覚

「凍った」雑音は実際には、本当の雑音と言うよりは周期的複合音である。しかし、振幅を周期的に変調した(例えば正弦波的に)白色雑音はそうではない。このような刺激(周期的振幅変調雑音)の長時間スペクトルは厳密に平坦な包絡を持ち、それは変調を受けない白色雑音のそれと少しも違わない。このことはしかし、その刺激によって、物理的にはその変調周波数と相関のある感覚が生じることを妨げない。被験者にこの感覚と周波数を調整できる純音の音の高さをマッチングさせることを求めた実験がいくつかある。De Boer (1956) はこの課題は不可能であると主張した。しかしながら、Harris (1963) と Pollack (1969) は 20 から約 1000 Hz の間の変調周波数について比較的正確な調整値を得ることができた。それ故に周期的振幅変調雑音によって、少なくともある範囲内の変調周波数において引き起こされる感覚を音の高さ(あるいは基本音の高さ)と呼んでよいものであろうか。このような音の高さは物理的にスペクトルとは何の関係もないことになるので、この問題は重要である。

周期的振幅変調雑音の変調周波数 f について測られた周波数弁別閾は、少なくとも f が 60 Hz 以上であれば、周波数 f の純音について測定された閾値よりもずっと高い、ということに注意しよう (Hoekstra, 1979; Bilsen and Wieman, 1980; Formby, 1985)。しかしこの事実は我々の疑問には答えていない。なぜなら周期的振幅変調雑音の統計的な性格を考慮にいれれば、このような刺激により生じる感覚は純音の音の高さと比べてより不正確になるであろうことが予め予想されるからである。この感覚がより不正確であるからといって、それが音の高さの感覚とは別のものであると考ねなければならないことにはならない。

Burns and Viemeister (1976, 1981) の研究により、変調周波数が数 100 Hz の周期的振幅変調雑音により生じる感覚の性質が明らかになった。彼らはまず、純音の音の高さと周期的複合音の基本音の高さは非常に特殊な特性を持つものであることを見いだした。絶対判断の課題では、二つの音の高さ(あるいは二つの基本音の高さ)の間の距離は、単独の音の高さ(あるいは基本音の高さ)よりも同定しやすい。音楽家はすべて、純音の周波数が数 100 Hz であれば、継起する二つの純音の音の高さによりできる音程を三度、四度、等という風に、完璧に同定することができる。音楽家は変調周波数の異なる二つの周期

の振幅変調雑音によりできた音程を、同じように三度、四度、等という風に同定できるのだろうか。これが正弦変調された周期的振幅変調雑音を刺激として用いて Burns and Viemeister が問うたことであった。彼らの答えは肯定的なものであって、少なくともある変調周波数範囲内では (60 あるいは 70 Hz 以上)、周期的振幅変調雑音の変調周波数と関係する感覚が音の高さであるとするを妨げる客観的な理由は何もない、という結論を得ている。この結論は重要である。なぜなら、繰り返すが、周期的振幅変調雑音の長時間スペクトルは変調を受けない白色雑音のそれと違わない。Burns and Viemeister の結果は、音の高さは、時間領域では構造を持っているが、スペクトル領域では構造を持っていないような刺激からも抽出され得るということを示しているからである。

III.3.c. スペクトルを制限された雑音の場合

中心周波数を固定したまま (例えば 1 kHz)、遮断周波数を変えることができるバンドパスフィルターを用いて、白色雑音の帯域幅を徐々に減少させていくものとしよう。もし遮断周波数がそれぞれ 20 Hz 以下と 20 kHz 以上であったなら、出力される刺激は、もちろん、無限のスペクトル幅を持つ白色雑音と区別がつかないであろう。しかし遮断周波数が 1 kHz に近づけば、雑音は次第に 1 kHz の純音の音の高さに近い感覚を生じさせるようになる。そして最終的には、帯域幅が大体 10 Hz 以下になれば、生じる感覚は不規則なトレモロの掛かった 1 kHz の純音により生じるようなものになる。このようなことはほとんど驚くべきことではない。特に、中心周波数が 1 kHz でバンド幅が 10 Hz 以下の雑音は、実は、振幅が 10 Hz 以下の周波数の純音の和で変調された 1 kHz の純音と同じものである。バンド幅が 10 Hz を越えた雑音は、振幅だけでなく瞬間の周波数 (あるいは位相とみてもよい) も同様に不規則に、そして帯域幅が増すにつれて次第に速く、変調された 1 kHz の純音として記述できる。

数人の研究者が (Shonle and Horan, 1980; Feth et al., 1982; Iwamiya and Kitamura, 1983) 低い変調周波数で (50 Hz 以下) 周波数だけを、あるいは周波数と振幅を同時に周期的に変調させた純音により生じる感覚について興味を抱いた。厳密に言えばこの刺激は周期的複合音であるが、瞬間の周波数の変化が小さい狭帯域雑音と似ている。そのスペクトル成分は実際には聴覚上分離されず、また基本音の高さを生じさせることもない。このような刺激について行なわれた研究から、少なくともある条件のもとでは、聴覚系は継時的で異なった音の高さの「加重平均」をとって、一つの正確な音の高さを「合成」できることが示唆されている。周波数 f が周波数 6 Hz の三角波関数により $0.97f$ から $1.03f$ の間で変調される正弦波を例に取ろう (Iwamiya and Kitamura, 1983)。この刺激に存在する周波数変調により、音の高さが変調されたように聞こえるであろう。しかしながら、この刺激は一方では、周波数 f の純音の音の高さと等しい、安定で正確な音の高さを生じさせる。今度は周波数変調の上と同じく周波数 6 Hz の三角波による振幅変調を加えて刺激を複雑にしたとしよう。もしこの二種類の変調が同相であれば、別の言い方をすれば、周波数が最大になったときに振幅も最大になり、周波数が最小になったときに振幅も最小になるとすれば、振幅変調度に対して一定の割合で、安定な音の高さが上昇していく。そしてもし逆に二種類の変調が逆相であれば (周波数が最小のとき振幅が最大となる)、やはり振幅変調度に比例して、安定な音の高さは低くなる。いつもビブラートを多用しているオペラ歌手は、彼らのプロとしての活動の中で、この種の現象をきっと考慮にいれているに違いない……。

最後に、高域通過雑音と低域通過雑音の場合について考えておこう。高域通過雑音、低域通過雑音は、フィルターを用いて白色雑音からある周波数より下の、または上のエネルギーを減少させることによって得られる。このことについて注目すべき事実は、それらの雑音を作るのに用いられるフィルターの減衰特性が遮断周波数のところで十分急峻であれば、この雑音は、広いスペクトルを持つ雑音全てが引き起こす感覚にももちろん加えて、純音の音の高さと良く似た音の高さの感覚を引き起こす (von Békésy, 1963; Small and Daniloff, 1967; Fastl, 1980; Klein and Hartmann, 1981)。減衰特性が特に急峻であれば、生じた音の高さは遮断周波数から 5 から 10 % ずれた周波数の純音の音の高さと同じである (Klein and Hartmann, 1981)。von Békésy をはじめ、数人の研究者がこの現象が視覚の「マッハバンド」の現象と似ていることを指摘した。事実、この二つの現象は良く似た解釈ができるように思われる。この聴覚現象はスペクトル対比を強調するための「側方抑制」が末梢過程に存在することで説明されよう

(Arthuret al., 1971; Houtgast, 1972)。その介在の結果、低域通過雑音の内的スペクトル——音の物理的なスペクトルの神経での表現——は、遮断周波数のすぐ下に「瘤」ができるだろう。雑音によって生じた音の高さは、このスペクトルの瘤と関係するのだろう。瘤が遮断周波数のすぐ上にできるであろうことを除けば、高域通過雑音により引き起こされる音の高さについても同様である。

III.3.d. 両耳系で作られた音の高さの感覚

一見したところ最も不適当な刺激の音の高さを聴くことに熱中した挙げ句、音響心理学者はとうとう、単耳での刺激はそれぞれ平凡な（振幅変調のされていない）白色雑音であるようなダイコティック刺激で音の高さを知覚したいと考えた。一見したところ、白色雑音だけでは音の高さは生じないのだから、そのような刺激が音の高さを生じさせることはあり得ないように思われる。しかし、その印象は間違っている。右耳と左耳に達したものを関係付けることによって、両耳系は雑音の音の高さを生じさせることができるようなのである！

両耳系で作られた音の高さが非常に顕著なものであることは決してなく、その上、一般的に、それらを知覚することは難しい。それらが存在する原因となるメカニズムは、音の高さを抽出する「普通の」メカニズム（例えば母音の基本音の高さを抽出するメカニズム）とは関係のないものである、と論じることできる。しかしこのことは確かではないし、それらの音の高さはともかくも、聴覚系の機能的なある側面に関係しているのである。

概して、左耳に提示された雑音の周波数成分（無限に近く隣合った、無数の）と、右耳に提示された雑音の対応する周波数成分との間に特別な位相関係があるときには、両耳系は雑音から音の高さを生み出すことができることが明かにされている。より精確に言えば、両耳間の位相関係は、生じる音の高さに対して、周波数の特別な関数でなければならない。しかしながら、いくつかのタイプの関数で音の高さが生じるので、この意味で、「両耳の高さ」にはいくつかのものがある、といてよい。現在のところ 4 種類の関数が見つまっている。以下では、誘発された音の高さを（独占的に、あるいは主に）発見した研究者の名前を借りて、それらと呼ぶことにしよう。

Huggins の高さ。Cramer and Huggins が 1958 年に報告したもので、最も単純な両耳の高さである。それは音圧が等しく、周波数成分がすべて同じ両耳間位相差 (d) を持つ（ただし、ある狭い周波数領域ではこの位相差は 2π ラジアン [360°] に増加、または減少する）二つの白色雑音を、同時に（ダイコティック）提示することにより得られるものである。 $d=0$ のところが基準状態で、これは図 11 の a に示されている。この状態で誘発される音の高さは、位相差が π ラジアンになっているところの周波数（図の周波数 f ）の純音の音の高さに対応するものである。同じ音の高さが、例えば、 d が低周波数では π ラジアン、高周波数では 3π ラジアン（ $= \pi$ ラジアン）、周波数 f では 2π ラジアン（つまり 0）のときにも生ずる。しかしながら、どの様な場合でも、音の高さが知覚されるためには f はおよそ 1500 Hz 以下でなくてはならない。

Bilsen の高さ。III.2.a. で、周波数がそれぞれ $n \cdot f$ と $(n+1) \cdot f$ (n はすべて 1 よりも大) である二つの純音を同時に、モノラル (Smoorenburg, 1970) あるいはダイコティック (Houtsma and Goldstein, 1972) で提示すると、 f に対応する基本音の高さが生じることについて述べた。Bilsen (1977) はこの二つの純音を二つの「Huggins の音」で置き換えたなら、同じような現象が生じるのではないかと考えた。彼はそこで、両耳間の位相差が連続的に 2π ラジアン分変化する二つの周波数帯域を除いては、両耳間の位相差が一定であるようなスペクトル成分を持つ二つの雑音を、実験の刺激として用いた（図 11, b）。この二つの周波数帯域はそれぞれ 600 Hz と 800 Hz（200 Hz の第 3, 4 倍音）を中心としていた。このような条件で、200 Hz に対応する音の高さ（基本音の高さ）を Bilsen 自身ともう一人の被験者が同様に聞き取った。この音の高さを知覚するのはおそらく極めて難しい。しかしながら、もしも組み合わせる Huggins の音が、ある周波数 f の第 3, 4 倍音にだけでなく、 f に続く全ての倍音に対応するものであれば、 f に対応する音の高さはよりはっきりと聞こえ、 f を変化させることで自然にそれとわかるメロディーを作ることすらできる (Bilsen, 1976)。

Hartmann の高さ。Klein and Hartmann (1981) は周波数 f 以下のスペクトル成分は同相で、この周波数以上のスペクトル成分は逆相（ π ラジアンの位相差）であるような二つのモノラル白色雑音から成る刺

激を実際に作った (図 11, c)。彼らによれば、この刺激は広い範囲の f の値について、しかし特に f が 300 Hz から 800 Hz の間にあるときに、同時には知覚できない二つの音の高さを引き起こす、という。一方は f のおよそ 5% 下の周波数に対応し、他方は f のおよそ 5% 上の周波数に対応する。再び Klein and Hartmann によれば、同じ音の高さ——彼らは「両耳聴の縁の」と言っている——が、もし例えばスペクトル成分が f の下では π ラディアンの位相差があつて、その上では同相であるか、あるいはそれぞれ $-\pi/2$, $\pi/2$ ラディアン位相がずれていれば、知覚できる。ちょうど f までは成分が同相で、 f を越えると成分の位相関係が周波数に対してランダムに変化するようなときには、 f よりも 5% 高いものに対応する音の高さがやはり知覚される (図の d, Hartmann, 1984)。Frijns et al. (1986) が最近、図 11, c に示されている刺激により、Klein and Hartmann の言う二つの音の高さよりむしろ、 f に非常に近い周波数に対応するただ一つの音の高さを生じることを確かめていることに注意しよう。それ故、Hartmann の音の高さは目

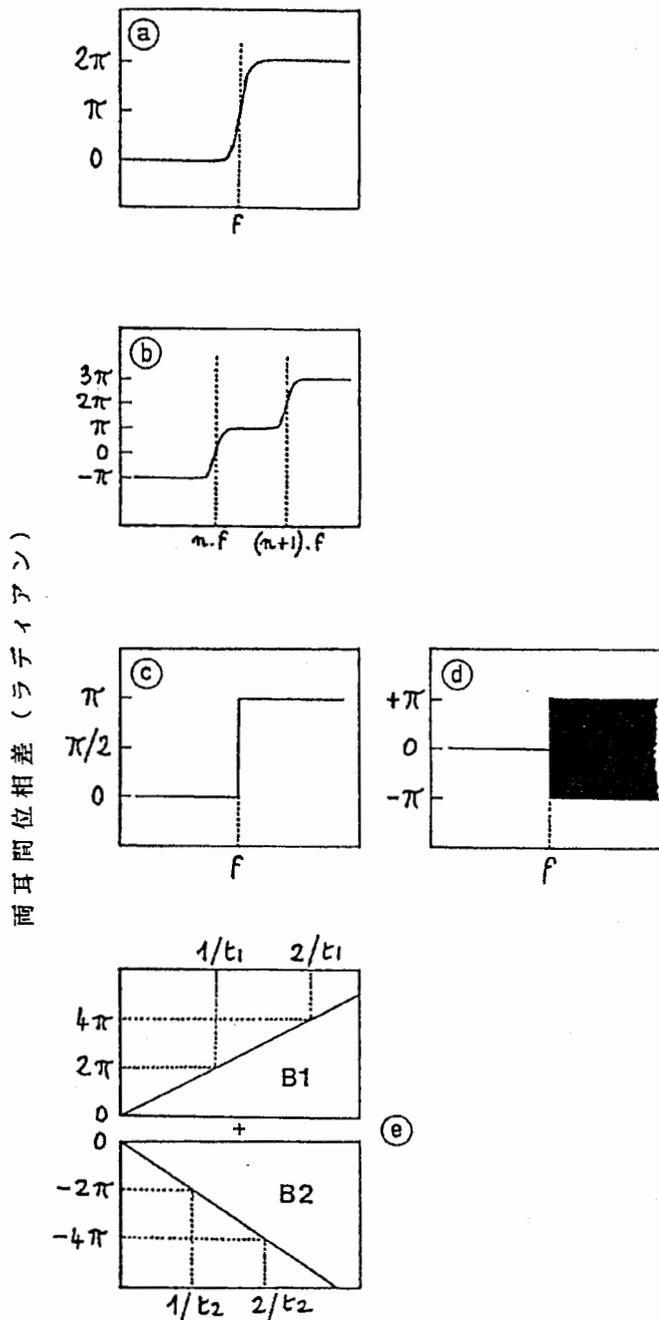


図 11. 音の高さの感覚を引き起こす 5 つのダイコティック雑音。

下のところ論争的となっている。

Fourcin の高さ。 Fourcin (1962, 1970)、それに Bilsen (Bilsen and Goldstein, 1974; Bilsen, 1977 を見よ) の研究により、無相関な二つの白色雑音、B1 と B2 のそれぞれに特定の両耳間遅延 (0 から約 10 ms の間) を与え、それらを足し合わせて作った刺激を二つの耳に同時に提示することにより、音の高さの感覚が生じ得ることが示された。図 11 の e に示されているのは、片方の耳では雑音 B1 が t_1 ms 遅れ、他方の耳では雑音 B2 が t_2 ms (t_2 は t_1 とは異なる) 遅れるときの刺激である。二つの雑音のそれぞれについて、両耳間遅延は周波数の関数としての両耳間位相差の直線的な変化として表わされている。このような刺激により、周波数 $1000 / (t_1 + t_2 + 0.9)$ Hz, $1000 / (t_1 + t_2 - 0.9)$ Hz に対応する (Bilsen, 1977 によれば) 曖昧な音の高さが生じる。Bilsen and Goldstein (1974) によれば、刺激を二つの雑音の一方だけに減らして、つまり、左耳と右耳に、ある両耳間遅延をもって到達する一つの白色雑音だけを残しても音の高さは依然として聞こえる (難しくはなるが)、という。もし遅延 t が約 2 ms 以下であれば、もはやただの白色雑音にしか聞こえなくなり、最初に刺激される側の耳から音が聞こえるように思われる。しかし t が約 2 から 10 ms の間であれば、周波数 $1000 / t$ Hz に対応する音の高さが同様に知覚される。

ここまで述べてきた音の高さはすべて、刺激が各々の耳に白色雑音そのものとして聞こえるような状態でのものなので、聴覚神経のレベルで神経路の交差のところまでで、音の高さの知覚が必ずしもスペクトルの違いとして表現されるものである必要はない、と結論せざるを得ない。この結論はある意味で、周期的振幅変調雑音についての Burns and Viemeister の研究で引用したものと非常に近いものである (III.3.b 参照)。

ただし、特に両耳聴の音の高さに関しては、聴覚系の中枢の段階での、即ち二つの耳からの情報が統合される段階での、スペクトル表現をもとに音の高さの抽出は行なわれる、とすることは考え得ることであるし、さらに一般的に認められてもいる。例えば図 11a の刺激をもう一度考えてみよう。両耳聴覚系が、左耳からと右耳からやってきた神経情報の一方から他方を引きさることができる——これは生理学的に考えられないことではない——と想像してみよう。この引算の結果、周波数成分の位相がずれているスペクトル領域を除いては、二つのモノラル雑音は相殺するであろう。この雑音の平坦な二つのスペクトルは両耳聴覚系によって周波数 f を中心とした「瘤」を持つスペクトルに変換されるであろう。このようにして、 f に対応する音の高さが知覚されると考えられる。さきに述べた、そのほかの両耳聴の音の高さも同様に、両耳聴覚系により、単耳聴スペクトルと機能的に等価な「中枢スペクトル」が形成されるといふ仮説により説明可能である (Bilsen and Goldstein, 1974; Klein and Hartmann, 1981; Blauert, 1983; Raatgever and Bilsen, 1986)。

IV. 周期的複合音の基本音の高さを抽出するモデル

前節の主たる目的は複合音の音の高さ、特に周期的複合音により生じる基本音の高さに関係する精神物理学的事実を示すことにあった (III.2.)。周期的複合音の基本音の高さに関しては、説明的なモデルを取り上げた (III.2.a. と III.2.b.)。しかしそれは「ついでに」取り上げただけで、それを、あるいはその観点を支持する生理学的なデータを引用することはしなかった。この節では、周期的複合音の基本音の高さの抽出メカニズムについての可能な考えを、より詳細に検討する。

IV.1. 末梢スペクトル分析の役割

III.2.b. に述べた、三つのスペクトル成分を持つ調波、非調波周期的複合音が引き起こす基本音の高さについての Schouten et al. (1962) の実験は長らく問題となっていた。この実験を記述した後で、この結果は基本的には基本音の高さ抽出の時間的モデル、即ち聴覚系が音の波形の振幅頂点間の時間間隔を測定することで周期的複合音の基本音の高さを決定する、ということだけのモデルで説明可能であることを指摘しておいた。もちろん、このようなモデルは、例えばバイオリンの低音のように、聞こえるスペクトルが非常に広がった周期的複合音の基本音の高さの抽出にはそのまま適用することはできない。なぜなら、聴覚系は音のスペクトル成分のうち少なくともいくつかは、完全にバラバラに分離するであろうからである。しかしながら、刺激にある種のフィルタリングを (スペクトル成分のうちあるものだけが相互作用を起こすような) 施した波形から基本音の高さが抽出されるような場合には、依然として検討に耐え得るものである。Schouten (1940, 1970) は、この考えを大いに主張した。Schouten の考えは今日でも現実的なものと見なすことができるのだろうか。

Schouten の考えを直接支持する精神物理学的な二つの事実を前に引用した。まず、周期的複合音の波形はそれ自身で、ある条件下では、基本音の高さに影響し得る、という事実である (Moore, 1977; III.2.d. 参照)。次に、周期的振幅変調雑音のように時間領域では構造を持っているが、スペクトル領域では構造を持っていないような刺激が、音の高さを生じさせるという事実である (Burns and Viemeister, 1976, 1981; III.3.b. 参照)。ここで、Schouten の考えを擁護するには、精神物理学的な議論とは別に、現在、生理学的なデータを引用できることを付け加えておこう。

それは調波的、あるいは非調波的な周期的複合音に対するネコヤチンチラの聴覚神経応答についての Evans (1978, 1983, 1986), Javel (1980), Greenberg (1986, 印刷中) の研究である。これらの研究は基本的に、個々の神経線維のレベルでこのような音により生じる活動電位に、ヒトの基本音の高さの感覚と全く一致した時間間隔が見いだされることを示している。例えば、1100 Hz の純音の振幅を周波数 200 Hz で正弦波的に変調して得られる、スペクトル成分が 900, 1100, 1300 Hz の非調波的周期的複合音があるでしょう。この周期的複合音はヒトでは曖昧な基本音の高さを生じさせ、それはたいてい 220, 180、あるいは 100 Hz に対応している。Evans (1978) はネコの聴覚神経線維で、この音で生じた連続する活動電位間の時間間隔の統計的分布を分析した。神経線維の特徴周波数は 1100 Hz であった。彼はこの分布が複数のモードを持ち、そのうちの主なもの三つはそれぞれ 4.6, 5.5, 10.0 ms に対応することを見いだした。ところでこの三つの時間間隔は、周波数が 220, 180, 100 Hz の波形の周期と一致する。Evans は、その分布で 5.0 ms、即ち、振幅変調の周期 (200 Hz の正弦波) と等しい時間間隔のモードを見つけたわけではないことに注意しよう。これはこの刺激では、ヒトでも 200 Hz に対応する基本音の高さは生じないという事実と一致する。

これは Schouten の基本音の高さについての考えを支持すると言われてきた。しかし一方では、III 節にこの考えとうまく合わない精神物理学的结果をいくつか挙げておいた。例えば、Houtsma and Goldstein (1972) が発見した事実、即ち、「欠落した基本音」は刺激のスペクトル成分が蝸牛レベルで少しも相互作用を起こさないのに (III.2.a. 参照) 知覚できる、という事実は明らかに Schouten の考えの面目を失わせるもののようなのである。「Bilsen の」音の高さ (III.3.d. 参照) の存在も同様に、似たような理由でこの

考えの信用を落とすものようだ。周期的複合音の基本音の高さは、物理的な周波数よりむしろ、スペクトル成分によって生じるスペクトルの高さと関係が深いように思われるという事実もまた、Schouten の考えとは一致しない (III.2.e. 参照)。

これらの論拠により、目下のところ Schouten の考えは全く間違っているのではないとしても、少なくとも不十分であると考えざるを得ない。それが全く間違っているという観点から、この 15 年間に何人かの研究者が、周期的複合音の基本音の高さは波形というよりはむしろ スペクトル成分 に関する情報をもとに抽出される、と主張する基本音の高さの理論を提唱してきた。基本音の高さの抽出を支配するスペクトル成分は、調波的で豊富なスペクトルを持つ周期的複合音のなかから、純音として分離して聞こえやすい (III.2.c. 参照)、ということ根拠として、この党派に味方しよう。基本音の高さの「スペクトル的」理論のいくつかを表面的に述べるよりは、ここでは Terhardt の作り上げた一つの理論を十分に長く述べることにしよう。この選択はいくらか恣意的である。よく似た理論である Goldstein (1973; Gerson and Goldstein, 1978; Goldstein et al., 1978) の理論は Demany (1987) にフランス語で詳細に説明されている。

IV.2. 「スペクトル的」モデル：Terhardt のモデル

IV.2.a. モデルの記述

Terhardt (1974, 1979; Terhardt et al., 1982a, 1982b) の作り上げた基本音の高さ抽出のモデルは、以前に Thurlow (1963), Whitfield (1967, 1970), Walliser (1969c) によって形作られた考えをもとにしている。Thurlow and Whitfield は特に、一つの音の印象として周期的複合音のスペクトル成分が知覚的に融合すること、及びこの印象がただ一つの基本音の高さとして特徴付けられることは、連合学習の結果生じたものである可能性があることを指摘した。このような考えが Terhardt のモデルの中心をなしており、彼がそれについて行なった説明は次の通りである。無限の周期的複合音のうちで、調波的周期的複合音はヒト（そして他の哺乳類の多く）のいる環境のなかで非常に特別な役割を演じている。特にヒトの音声の母音はそれらの音の本質をなしている。従ってヒト（あるいは高等な哺乳類）は、小さい頃から調波的周期的複合音を、異なる音の高さをもつ純音の集合としてよりはむしろ、ただ一つの基本音の高さをもつ統合されたものとして知覚することを学習するのだ、とすることはもっともらしい。Terhardt によれば、このような学習は効果的に実現され、聴覚記憶内で純音によって生じた音の高さの感覚間の連合関係の形で表現されるという。周波数 f の純音は、基本周波数が $f/2, f/3, f/4, \text{etc.}$ の周期的複合音の関係の中に組織的に入り込んでいるので、その音の高さは異なる周期的複合音の基本周波数成分からなる純音の音の高さと結びつけられるであろう。純音の音の高さの連合関係がいったん形成されると、どんな周期的複合音の基本音の高さ抽出も、問題の関係に基づいて決定されるであろう。

図 12 は Terhardt のモデルが 220 Hz の第 1 から第 10 までの倍音からなる周期的複合音の基本音の高さ抽出にどういう具合に適用されるかを示したものである。刺激のスペクトルは対数周波数尺で図の上側に示されている。基本音の高さ抽出の最初の段階は、このモデルによればスペクトル成分の末梢での分離である。Terhardt は、このような音については最初の 8 個の成分だけが分離され、第 9, 10 倍音はこの最初の段階を過ぎると基本音の高さ抽出に関わらなくなる、と仮定している。この段階の結果として、Terhardt によれば、スペクトル成分のスペクトルの高さが測定される。各々の成分は他の成分の前後関係の中に入り込んでいるので、スペクトルの高さは成分が単独で提示されたときに生じる音の高さとはわずかに、しかし組織的に異なったものとみなされる。分離された成分のスペクトルの高さは図では黒点で表わされている。スペクトルの高さと周波数の間のずれがスペクトル成分間の反発型の相互作用で表現されていることがわかる。

図で例に挙げた周期的複合音は、Terhardt によれば、統合されたものとして理解することを我々が学ぶ音全体をよく代表しており、この統合を基礎として、結果的に純音の音の高さの連合関係が形成される。これは、例えばこの音の第 4 倍音によって生じるスペクトルの高さ（このスペクトルの高さを h と呼ぼう）は、大人の記憶としては、正確にいえば音の基本音成分（第 1 倍音）により生じるスペクトル

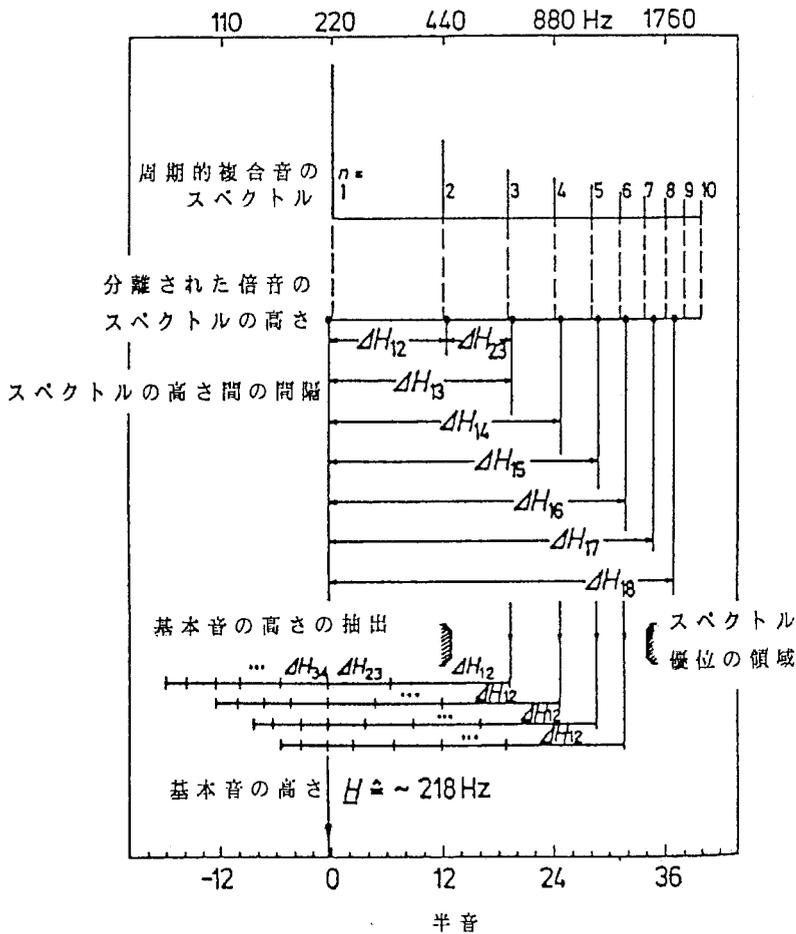


図 12. Terhardt の提案した基本音の高さ知覚に関するモデルの説明。Terhardt(1978) による。

の高さに対応する「潜在的な」—— Terhardt の用いた用語に従えば——音の高さと結びついていることを意味する。記憶上結びつけられているこれら二つの高さの間隔は、周波数の点からみると、正確に 4 の比には対応せず、それより少し大きな比に対応していることに注意しよう（図を見よ）。Terhardt によれば、 h は、記憶の中では同様に、他の周期的複合音の—— h を生じさせるスペクトル成分が第 4 倍音ではなくて他の倍音であるような（第 2, 3, 5, ...8）——基本音成分のスペクトルの高さと結びつけられている。これら他の周期的複合音では、倍音によって形成されるスペクトルの高さの間隔が、図の周期的複合音の倍音によって形成される間隔と似ていることが認められる。このことは h と結びついている「潜在的な」音の高さが、 h との間で図で ΔH_{12} , ΔH_{13} , ΔH_{14} , etc. と名付けられている間隔とほとんど等しい間隔を形成していることを示唆している。もちろん、図の音により生じる他のスペクトルの高さも、それぞれ同様に音の高さの特定の集合と結びつき、この集合の要素と ΔH_{12} , ΔH_{13} , ΔH_{14} , etc. にほぼ等しい間隔を形成するであろう。

Terhardt のモデルは、周期的複合音の基本音の高さは、記憶内で音の支配的なスペクトル成分のスペクトルの高さと結びついている、潜在的音の高さの別の集合を支配する共通の要素を探すことにより抽出される、と主張する。彼によれば、第一次近似として、分離されたスペクトル成分が支配的であるための必要十分条件は、その周波数が 500 から 1500 Hz の間に（このスペクトル領域は周波数弁別閾がよいので、他よりも特権的である）あることである、ということを知ることができる。図 12 の周期的複合音について、500 - 1500 Hz の周波数領域が特権的であるとすれば、基本音の高さの抽出には第 3 から第 6 倍音までだけを考慮に入れることになる。この 4 つの倍音のスペクトルの高さに結びついている潜在的な音の高さの集合が図の下に示されている。4 つの集合に共通なのはただ一つの要素だけで、この共通要素が基本音成分のスペクトルの高さであることがわかる。従って、Terhardt によれば、音の最も目立つ基本音の高さは基本音成分のスペクトルの高さである、という事実が説明できる。しかしながら、各々の集合には現われない他の要素が、それらのうち二つに見いだされることに注意しよう。それは、特に、

基本音のオクターブ上とオクターブ下に対応する要素の場合である。このような部分的な一致が、Terhardt によれば、全ての周期的複合音の基本音の高さにある程度見られる曖昧さの源である。

ここでは Terhardt のモデルの概略を示した。概略に縮めたので、このモデルで III.2.b. で問題としたような、非調波的周期的複合音により生じ得る基本音の高さを説明してはいない。しかしながらこのモデルをそのことを説明できるように洗練することは容易である。このことについてより詳しく知りたい読者は Terhardt et al. (1982a, 1982b) と Moore et al. (1985) にあたるとよい。

IV.2.b. モデルについての考察

Terhardt のモデルの深い正当性を理解するためには、彼がこのモデルによって基本音の高さ知覚の小さな「異常」を、III 節では指摘しなかった異常を、説明したがったことを分かっていなければならない。それは基本周波数 f の、その周波数にエネルギーを持つと持たざるとに関わらずどんな調波的周期的複合音についても、優勢な基本音の高さは周波数 f の純音の音の高さよりもほんの少しだけ低くなる傾向がある、という事実である (Walliser, 1969b; Ohgushi, 1978)。これから述べるように、Terhardt はこの現象を次のように説明する。(1) 調波的周期的複合音の周波数成分のスペクトルの高さは単独で提示されたときにこの成分が引き起こす音の高さとわずかに異なる。(2) この音の成分の形成するスペクトルの高さとその基本音成分との間隔は幼少の頃から学習される。従って、Terhardt の発表した学習仮説は根拠がないわけではなく、同時にいくつかの知覚現象を説明する余地がある。しかしながら Peters et al. (1983) が最近 (1) の見地の現実性に異議を唱えたことに注意しよう。彼らは、平坦なスペクトル包絡の調波的周期的複合音では、一つのスペクトル成分のスペクトルの高さはこのスペクトル成分と物理的に等しい純音の音の高さと違わない、と主張している。Peters et al. の結果と先に述べた小さな「異常」とは、Terhardt のモデルの枠組みの中では相容れないものに思われるので、調波的周期的複合音が引き起こす正確なスペクトルの高さについての別の実験を企てることが役立つであろう。

それ自身についていえば、Terhardt の適用した学習仮説は不条理なものとは考えられない。周期的複合音の基本音の高さ抽出において、聴覚皮質がある程度重要な役割を果たしているようであるという事実は (Zatorre, 1988; III.2.a. 参照)、高さの抽出が結局は、場合によっては記憶痕跡を利用しながら、形の再認を行なう過程である、という考えを支持するものである。更に、Roederer (1975) が示したように、Terhardt の引合いに出した学習は、ニューロン間のシナプス結合によって容易に生理学的なモデル化ができる。一方では周期的複合音の基本音の高さ抽出は、聴覚系の内的な特性だけに依存しているのではないということから、この考えに対してある支援を与える精神物理学的数据も存在する。例えば、Hall and Peters (1982) は非調波的周期的複合音 (1250 + 1450 + 1650 Hz) の最も顕著な基本音の高さが、この音と調波的周期的複合音 (200 + 400 + 600 + 800 + 1000 Hz) との連合の結果、わずかに変化し得ることを確かめた。Hall and Peters によれば、非調波的周期的複合音単独では、大体 210 Hz に対応する基本音の高さが「正常に」引き起こされるのに、この音を調波的周期的複合音と同時に数 10 分間聞かされた後では、この基本音の高さは 204 Hz になり得るのである！

IV.2.c. スペクトルの高さの神経符号化

さて今度は、厳密に生理学的な次元の問題に移ろう。Terhardt のモデルは、全ての基本音の高さの「スペクトル的」モデルと同様、周期的複合音の基本音の高さは刺激のスペクトル成分についての情報をもとに抽出される、と主張する。これから必然的に、どのようにしてそれらの情報が聴覚系で符号化されるのか、ということが疑問となる。

II 節で純音の周波数 (あるいは音の高さ) の符号化についてのデータと現在の仮説を紹介した。周期的複合音のスペクトル成分の周波数 (あるいはスペクトルの高さ) の符号化は比較的是っきりした問題である。しかしながら、純音についてと同様に、一般的な仮説が二つ存在する。最初に、ある周波数のスペクトル成分の存在は、定められた個所に興奮のピークが存在することにより符号化され得る (場所的符号化)。二番目に、その存在は発火系列のリズム構造として符号化され得る (時間的符号化)。

Sachs and Young (1979, 1980; Young and Sachs, 1979) は合成母音 (完全に周期的で、従って調波的周期的

複合音である) の、ネコの聴覚神経線維一本一本における符号化についての研究で、これら二つの仮説を検証した。彼らは平均発火率と、活動の微細時間構造の見地から、基本周波数 128 Hz の音について多数の線維からの反応を調べた。この研究から彼らは二つの大きな結論に達した。(1) 70 dB SPL 以上の音圧——通常の会話のレベル、つまり中くらいの音圧でしかないが——の母音については、一つの倍音の周波数は平均発火率の極大点によってでは符号化できないようである(飽和現象の結果)。(2) 強いレベルを含む、非常に広い範囲の音圧レベルにわたって、神経線維の活動の時間的微細構造が母音の連続スペクトルについての情報を詳細に中枢に送っている可能性がある。

図 13 は第二の結論の根拠となるデータの一例を示している。これらのデータは 57 dB SPL の母音 /a/ によって引き起こされた、特徴周波数が 410, 770, 1130 Hz の 3 本の線維における、活動の微細時間構造に関するものである。3 本の線維の特徴周波数が、それぞれ母音の第 3 (384 Hz)、6 (768 Hz)、9 (1152 Hz) 倍音の周波数の近傍であることに注意しよう。図の a, b, c は各線維について連続する発火間の時間間隔の統計分布を示したものである。

まず、分布 a (最も低い特徴周波数の線維) が図 3 (II 節) に示された分布の形とよく似ていること、そして図 3 は純音に対する神経線維の反応であることに注意しよう。これらの分布のように、分布 a は明瞭に分離されたモードと規則的に減衰していく振幅とを示している。これらのモードは刺激の第 3 倍音の周期の 1, 2, 3, ...n 倍に一致する間隔で得られている。第 3 倍音の周波数はこの線維の特徴周波数と非常に近い周波数である。連続するモード間の差が第 3 倍音の周期と対応していることを確かめるためには、分布のスペクトルを求めればよい。このスペクトルは図の右の欄の上に表示されている。母音の第 3 倍音の周波数がはっきりと際立っていることがわかる。

次に分布 b と c を調べてみると、その包絡線が分布 a のそれとは異なっていることにすぐに気づく。しかしながら、この二つの分布それぞれについても、モードはやはり刺激の倍音の周期に連続する整数を掛けたものと一致していることが明かである。そしてこの倍音の周波数は問題の線維の特徴周波数と非常に近いものである。即ち、分布 b の周波数スペクトルは母音の第 6 倍音の周波数が圧倒的であり、分布 c のスペクトルで最も有力な周波数は第 9 倍音の周波数と等しい周波数である。

Miller and Sachs (1983), Delgutte and Kiang (1984) そして Horst et al. (1986) は Sachs and Young とよく似た

活動電位間の時間間隔の分布 フーリエ変換したもの

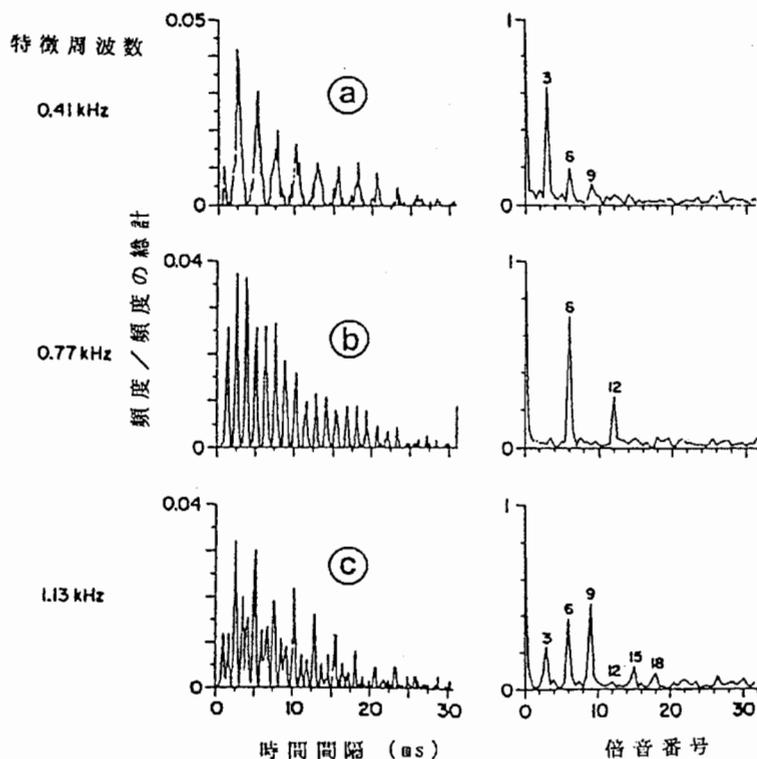


図 13. ネコの聴覚神経で得られた合成母音のスペクトルの時間的符号化についての Sachs and Young (1980) のデータ。

研究方法で、母音のようなタイプの周期的複合音のスペクトルが広いダイナミックレンジにわたって聴覚神経のレベルで時間的に表現されていることを確かめた（少なくともネコについては）。繰り返すが、Sachs and Young のデータは、逆にそのような音のスペクトルは、ネコでは、そしてそれ故おそらくヒトでも、平均発火率の空間的变化によってうまく表現されていないことを示唆している。しかしながら、実際、現在のところ聴覚神経によるスペクトル成分の表現において、平均発火率が果たしている役割については重大な疑問が残されている（Sachs et al., 1986）。

IV.3. Moore と van Noorden のモデル

Terhardt がそうしたように、周期的複合音の基本音の高さはその周期的複合音の個々のスペクトル成分についての情報だけから抽出される、と仮定すれば、周期的振幅変調雑音が生じさせることのできる音の高さを説明することを諦めねばならず（III.3.b. 参照）、周期的複合音の波形がその基本音の高さに対して与え得る影響について説明することも同様に諦めねばならない（III.2.d. 参照）。純粋に「スペクトル的な」基本音の高さのモデルのこの欠陥は重大なものであると考えて、Moore (1977, 1982) と van Noorden (1982) はいわばありとあらゆる周期的複合音の基本音の高さを説明することができる抽出モデルを提案した。このモデルは Terhardt のモデルのように「スペクトル的」でも、Schouten のモデルのように「非スペクトル的」でもない。それは基本音の高さ抽出に関して対立する二つの概念の統合を実現したものである。

Moore と van Noorden のモデルは非常に簡単なものであり、次の三つの命題に要約されるものである。(1) 全ての複合音は、聴覚神経線維のレベルで活動電位の系列によって伝達され、その時間構造は神経線維の特徴周波数の関数として神経線維ごとに異なる。(2) このレベルでは、基本音の高さの感覚と相関を持つのは連続する活動電位間の時間間隔の値である。(3) 周期的複合音の最も顕著な基本音の高さに対応するのは、神経線維全体について（しかし基本的にスペクトル優位な範囲で）、最も頻繁に現われる時間間隔である。

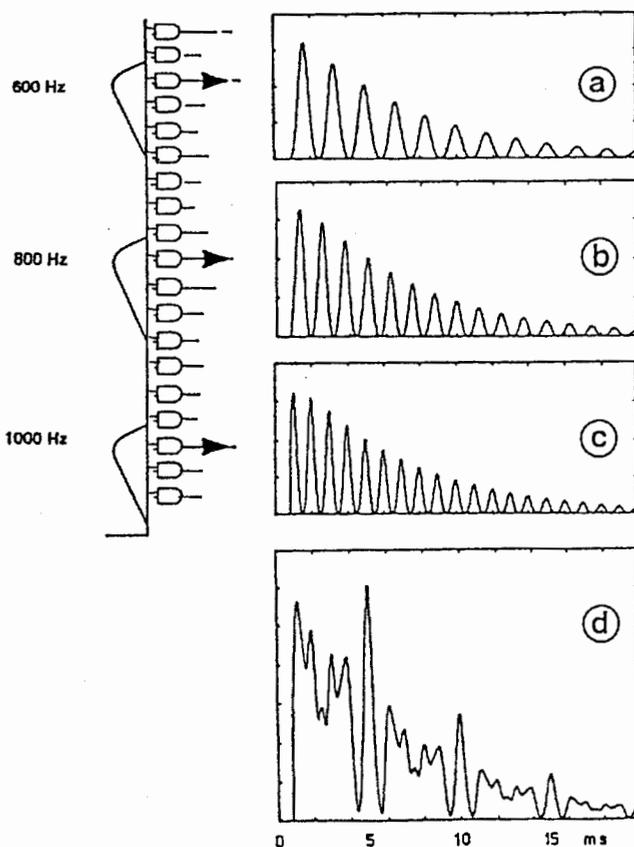


図 14. Moore (1977, 1982) と van Noorden (1982) の提案した基本音の高さ知覚に関するモデルの説明。van Noorden (1982) より複製。

図 14 は音圧が低く、三つのスペクトル成分——周波数 600, 800, 1000 Hz (200 Hz の第 3, 4, 5 倍音)——を持つ周期的複合音の場合についてこのモデルがどのように適用されるのかを示したものである。音の物理的特性と聴覚末梢系の周波数選択性から推測して、音に反応する聴覚神経線維はすべて、実際には特徴周波数に十分近い周波数のスペクトル成分だけにしか応答しないことを認めることができる。ある神経線維について、連続する活動電位間の時間間隔の統計分布は、それ故、純音によって引き起こされた包絡と類似しているであろうし、図の上から三つのグラフ (a, b, c) に示された形を取るであろう。この連続するモードは成分周期の連続する整数倍数と一致している。図の下のグラフは分布 a, b, c の加算を行なって得られた分布 (d) を示している。分布 d が他の三つの分布と非常に異なった分布を持っていること、周期的複合音刺激の周期 5 ms と一致する間隔の値のところに非常に顕著なモードが存在することがわかる。従って、このモデルによれば、この音の主要な基本音の高さが、周期 5 ms の純音のそれとほとんど等しいのは、5 ms の間隔が聴覚神経線維全体で最も頻繁に生じるからである。

図 14 で考えた周期的複合音は 200 Hz の第 3, 4, 5 倍音からなる。もっとスペクトル密度の濃い、例えば 200 Hz の第 5, 6, 7 倍音からなる周期的複合音については、活動している末梢神経線維のうち少なくともいくつかは、複数の倍音のうちの一つだけにではなく、いくつかの倍音に同時に応答しているであろうことを認めねばならない。確かに、線維の特徴周波数が連続する二つの倍音の周波数の間にあるときにはこのようなことが起こるのである。このスペクトル成分の相互作用が、個々の線維のレベルで音の周期と一致する時間間隔が出現する確率を増大させるであろう、と考えるのはもっともなことである (Evans, 1978, 1986; Javel, 1980; Greenberg, 1986; IV.1. 参照)。しかしながら、他方では、聴覚神経の 1 本の線維の活動にみられる周期的複合音の周期の表現は、この線維が非常に高い倍音で刺激されたときには、曖昧なものになるようである。このことについては図 13 の統計的分布を再検討してみよう。分布 c では、これは高い倍音に対する線維の応答から得られたものであるが、音の周期 (この場合 7.8 ms) と一致するモードは、分布 a と比べて隣のモードから離れていない。つまり、このモードに担われる情報は、それによってより曖昧になる。このようにして、Moore と van Noorden のモデルの枠組みの中では、スペクトルの豊富な (母音のような) 周期的複合音については、基本音の高さの抽出において低い倍音が高い倍音を支配する、という事実を説明することができる。

IV.1. で引用した生理学的データは、非調波的周期的複合音により生じる基本音の高さは聴覚神経のレベル以降で、基本音の高さに正確に対応する時間間隔の存在によって符号化されている可能性があることを示唆している。それ故、このモデルの適用範囲は調波的周期的複合音に限られてはいない。このモデルが周期的振幅変調雑音 (III.3.b. 参照) により生じる音の高さも説明できることも同様に明白である。周期的振幅変調雑音の場合には、聴覚神経線維の発火確率は何よりもまず、基底膜の対応する位置における変調関数の瞬間位相に依存するであろう。関数が最大値を取るときに確率も最大となり、関数が最小値を取るときに確率も最小となるであろう。神経線維全体では、結果として、連続する活動電位の時間間隔の統計分布は変調周期と一致するモードしか示さないであろう。しかしながら、刺激の統計的性質を考慮すれば、このモードはあまり「尖って」いないはずであり、だから、変調周期に対する聴覚系の弁別感度が悪いことを説明できる。

結局、周期的複合音の基本音の高さについての十分確立された精神物理学的現象をひっくり返ると、Moore と van Noorden のモデルに基づいて全然説明のつかないことは何もないように思われる。もちろんこのことはこのモデルの強みであるが、しかしこのモデルには、ある不明解さという欠点もあることに同様に注意しなければならない。生理学の領域でまだ正当性が検証されていない命題でモデルを膨らませなければ、このモデルでは、ヒトについて証拠とされる多くの現象を正確に説明することができない。形式が最も練り上げられているという意味で、Terhardt (Terhardt, 1982a, 1982b) のモデルはより精密である。

Moore と van Noorden のモデルは、対する Terhardt のそれとは基本的な問題においてきびしく対立している。それは基本音の高さの知覚における学習過程の介入があるかどうかである。Terhardt は経験論者であるが、Moore と van Noorden は生得論者である。現在のところ、基本音の高さの知覚において生得的な部分と獲得された部分とがよくわからないままになっており、このことが特によりはっきりするような研究を続けなければならない。音の高さの知覚でまだ残っている数多くの問題の中でも、いくつかの問題は神経生理学の領域でしか解決がつけられないが、しかし、それらの問題は精神物理学的方法がなけ

ればそもそも存在し得なかったものなのである。

訳者の謝辞

本論文の翻訳を勤めてくださり、原著者との仲介の労を取られた大串健吾博士、翻訳を許可していただいた原著者の Laurent Demany 博士、及び Les Editions INSERM 社、訳文の改良について御示唆頂いた津崎実氏、宮岡徹先生に深謝致します。

References

- AFNOR (1977). *Recueil des normes françaises de l'acoustique*, Tome 1 (Vocabulaire), NF S 30-107.
- ANSI (1960). *American National Standards, Acoustic Terminology*, ANSI S1.1-1960.
- Allen, D. (1967). Octave discriminability of musical and non-musical subjects, *Psychonom. Sci.*, 7, 421.
- Arthur, R. M., Pfeiffer, R. R., and Suga, N. (1971). Properties of "two-tone inhibition" in primary auditory neurons, *J. Physiol.*, 212, 593.
- Attneave, F. and Olson, R. K. (1971). Pitch as a medium: a new approach to psychophysical scaling, *Am. J. Psychol.*, 84, 147.
- Bachem, A. (1950). Tone height and tone chroma as two different pitch qualities, *Acta Psychol.*, 7, 80.
- Bilsen, F. A. (1976). A pronounced binaural pitch phenomenon, *J. Acoust. Soc. Am.* 59, 467.
- Bilsen, F. A. (1977). Pitch of noise signals: evidence for a "central spectrum", *J. Acoust. Soc. Am.* 61, 150.
- Bilsen, F. A. and Goldstein, J. L. (1974). Pitch of dichotically delayed noise and its possible spectral basis, *J. Acoust. Soc. Am.* 55, 292.
- Bilsen, F. A. and Wieman, J. L. (1980). Atonal periodicity sensation for comb filtered noise signals, in G. van den Brink and F. A. Bilsen (Eds): *Psychophysical, physiological and behavioural studies in hearing*, Delft: Delft University Press.
- Blauert, J. (1983). *Spatial hearing*, Cambridge (USA): M.I.T. Press.
- Brookes, A. J. (1982). On the role of combination tones in the perception of the second effect of pitch shift, *Acustica*, 51, 90.
- Burns, E. M. (1982). Pure-tone pitch anomalies. I. Pitch-intensity effects and diplacusis in normal ears, *J. Acoust. Soc. Am.* 72, 1394.
- Burns, E. M. and Viemeister, N. F. (1976). Nonspectral pitch, *J. Acoust. Soc. Am.* 60, 863.
- Burns, E. M. and Viemeister, N. F. (1981). Played-again SAM: further observations on the pitch of amplitude-modulated noise, *J. Acoust. Soc. A.* 70, 1655.
- Burns, E. M. and Ward, W. D. (1982). Intervals, scales, and tuning, in D. Deutsch (Ed): *The psychology of music*, London: Academic Press.
- Buunen, T. J. F. (1976). *On the perception of phase differences in acoustic signals*, doctoral thesis, Technische Hogeschool de Delft (Netherlands).
- Clarkson, M. G. and Clifton, R. K. (1985). Infant pitch perception: Evidence for responding to pitch categories and the missing fundamental, *J. Acoust. Soc. Am.* 77, 1521.
- Coninx, F. (1978). The detection of combined differences in frequency and intensity, *Acustica*, 39, 137.
- Cramer, E. M. and Huggins, W. H. (1958). Creation of pitch through binaural interaction, *J. Acoust. Soc. Am.* 30, 137.
- Davis, H., Silverman, S. R., and McAuliffe, D. R. (1951). Some observations on pitch and frequency, *J. Acoust. Soc. Am.* 23, 40.
- de Boer, E. (1956). *On the "residue" in hearing*, doctoral thesis, University of Amsterdam.
- de Boer, E. (1976). On the "residue" and auditory pitch perception, in W. Keidel and W. D. Neff (Eds): *Handbook of sensory physiology (Vol. VI/3)*, Berlin: Springer.
- Delgutte, B. and Kiang, N. Y. S. (1984). Speech coding in the auditory nerve: I. Vowel-like sounds, *J. Acoust. Soc. Am.* 75, 866.
- Demany, L. (1985). Perceptual learning in frequency discrimination, *J. Acoust. Soc. Am.* 78, 1118.

- Demany, L. (1987). *La perception de la hauteur tonale*, doctoral thesis of natural science, Université Paris 6.
- Demany, L. and Armand, F. (1984). The perceptual reality of tone chroma in early infancy, *J. Acoust. Soc. Am.* 76, 57.
- Demany, L. and Lavenant, H. (1985). Effet d'une défférence d'intensité sur la précision des ajustements dans une tâche d'égalisation des hauteurs tonales de deux sons purs, *Ann. Psychol.* 85, 329.
- Demany, L. and Semal, C. (1986). On the detection of amplitude modulation and frequency modulation at low modulation frequencies, *Acustica*, 61, 243.
- Demany, L. and Semal, C. (1989). Detection thresholds for sinusoidal frequency modulation. *J. Acoust. Soc. Am.* 85, 1295.
- Deutsch, D. (1973). Octave generalization of specific interference effects in memory for tonal pitch, *Percept. Psychophys.* 11, 411.
- Evans, E. F. (1977). Frequency selectivity at high signal levels of single units in cochlear nerve and cochlear nucleus, in E. F. Evans and J. P. Wilson (Eds): *Psychophysics and physiology of hearing*, London: Academic Press.
- Evans, E. F. (1978). Place and time coding of frequency in the peripheral auditory system: some physiological pros and cons, *Audiology*, 17, 369.
- Evans, E. F. (1983). Pitch and cochlear nerve fibre temporal discharge patterns, in R. Klinke and R. Hartmann (Eds): *Hearing--Physiological bases and psychophysics*, Berlin: Springer.
- Evans, E. F. (1986). Cochlear nerve fiber temporal discharge patterns, cochlear frequency selectivity and the dominant region for pitch, in B. C. J. Moore and R. D. Patterson (Eds): *Auditory frequency selectivity*, New York: Plenum Press.
- Fastl, H. (1980). Pitch strength and masking patterns of low-pass noise, in G. van den Brink and F. A. Bilsen (Eds): *Psychophysical, physiological and behavioural studies in hearing*, Delft: Delft University Press.
- Fastl, H. (1986). Auditory after-images produced by complex tones with a spectral gap, *Proceed. 12th Intern. Congr. Acoust.* (Toronto), B2-5.
- Fastl, H. and Stoll, G. (1979). Scaling of pitch strength. *Hearing res.* 1, 293.
- Feth, L. L., O'Malley, H., and Ramsey, J. (1982). Pitch of unresolved, two-component complex tones, *J. Acoust. Soc. Am.* 72, 1403.
- Flanagan, J. L. and Guttman, N. (1960). On the pitch of periodic pulses, *J. Acoust. Soc. Am.* 32, 1308.
- Florentine, M. and Houtsma, A. J. M. (1983). Tuning curves and pitch matches in a listener with a unilateral, low-frequency hearing loss, *J. Acoust. Soc. Am.* 73, 961.
- Formby, C. (1985). Differential sensitivity to tonal frequency and to the rate of amplitude modulation of broadband noise by normally hearing listeners, *J. Acoust. Soc. Am.* 78, 70.
- Fourcin, A. J. (1962). An aspect of the perception of pitch, *Proceed. 4th Intern. Congr. Phonet. Sci.* (Helsinki). La Haye: Mouton, 355.
- Fourcin, A. J. (1970). Central pitch and auditory lateralization, in R. Plomp and G. F. Smoorenburg (Eds): *Frequency analysis and periodicity detection in hearing*, Leiden: Sijthoff.
- Francés, R. (1958). *La perception de la musique*, Paris: Vrin.
- Frijns, J. H. M., Raatgever, J., and Bilsen, F. A. (1986). A central spectrum theory of binaural processing. The binaural edge pitch revisited, *J. Acoust. Soc. Am.* 80, 442.
- Garner, W. R. (1974). *The processing of information and structure*, Potomac: Lawrence Erlbaum.
- Gerson, A. and Goldstein, J. L. (1978). Evidence for a general template in central optimal processing for pitch of complex tones, *J. Acoust. Soc. Am.* 63, 498.
- Goldstein, J. L. (1970). Aural combination tones, in R. Plomp and G. F. Smoorenburg (Eds): *Frequency analysis and periodicity detection in hearing*, Leiden: Sijthoff.
- Goldstein, J. L. (1973). An optimum processor theory for the central formation of the pitch of complex tones, *J. Acoust. Soc. Am.* 54, 1496.
- Goldstein, J. L., Gerson, A., Srulovicz, P., and Furst, M. (1978). Verification of the optimal probabilistic basis of

- aural processing in pitch of complex tones, *J. Acoust. Soc. Am.* 63, 486.
- Green, D. M. and Swets, J. A. (1973). *Signal detection theory and psychophysics*, Huntington: R. Krieger (2nd ed.).
- Greenberg, S. (1986). Possible role of low and medium spontaneous rate cochlear nerve fibers in the encoding of waveform periodicity, in B. C. J. Moore and R. D. Patterson (Eds): *Auditory frequency selectivity*, New York: Plenum Press.
- Greenberg, S. (in press). Neural temporal coding of lowpitch. II. Responses of single cochlear nerve fibers in the cat to two-component signals.
- Greenberg, S., Geisler, C. D., and Deng, L. (1986). Frequency selectivity of single cochlear-nerve fibers based on the temporal response pattern to two-tone signals, *J. Acoust. Soc. Am.* 79, 1010.
- Guttman, N. and Julesz, B. (1963). Lower limits of auditory periodicity analysis, *J. Acoust. Soc. Am.* 35, 610.
- Guttman, N. and Pruzansky, S. (1962). Lower limits of pitch and musical pitch, *J. Speech Hearing Res.* 5, 207.
- Hall, J. W. (1986). Central adaptation of complex pitch, *J. Acoust. Soc. Am.* 80, 1681.
- Hall, J. W. and Peters, R. W. (1981). Pitch for nonsimultaneous successive harmonics in quiet and noise, *J. Acoust. Soc. Am.* 69, 509.
- Hall, J. W. and Peters, R. W. (1982). Change in the pitch of a complex tone following its association with a second complex tone, *J. Acoust. Soc. Am.* 71, 142.
- Hall, J. W. and Soderquist, D. R. (1975). Encoding and pitch strength of complex tones, *J. Acoust. Soc. Am.* 58, 1257.
- Hall, J. W. and Soderquist, D. R. (1982). Transient complex and pure tone pitch changes by adaptation, *J. Acoust. Soc. Am.* 71, 665.
- Harris, G. G. (1963). Periodicity perception by using gated noise, *J. Acoust. Soc. Am.* 35, 1229.
- Harris, J. D. (1952a). The decline of pitch discrimination with time, *J. Exp. Psychol.* 43, 96.
- Harris, J. D. (1952b). Pitch discrimination, *J. Acoust. Soc. Am.* 24, 750.
- Hartmann, W. M. (1984). Binaural coherence edge pitch, *J. Acoust. Soc. Am.* 75 (Sup. 1), S 22.
- Hartmann, W. M. and Klein, M. A. (1980). Theory of frequency modulation detection for low modulation frequencies, *J. Acoust. Soc. Am.* 67, 935.
- Heffner, H. and Whitfield, I. C. (1976). Perception of the missing fundamental by cats, *J. Acoust. Soc. Am.* 59, 915.
- Helmholtz, H. L. F. (1863). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, Braunschweig: F. Vieweg.
- Henning, G. B. (1966). Frequency discrimination of random-amplitude tones, *J. Acoust. Soc. Am.* 39, 336.
- Hoekstra, A. (1979). *Frequency discrimination and frequency analysis in hearing*, doctoral thesis, Rijksuniversiteit de Groningen (Netherlands).
- Horst, J. W., Javel, E., and Farley, G. R. (1986). Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms, *J. Acoust. Soc. Am.* 79, 398.
- House, W. J. (1977). Octave generalization and the identification of distorted melodies, *Percept. Psychophys.* 21, 586.
- Houtgast, T. (1972). Psychophysical evidence for lateral inhibition in hearing, *J. Acoust. Soc. Am.* 51, 1885.
- Houtgast, T. (1976). Subharmonic pitches of a pure tone at low S/N ratio, *J. Acoust. Soc. Am.* 60, 405.
- Houtsma, A. J. M. (1981). Noise-induced shifts in the pitch of pure and complex tones, *J. Acoust. Soc. Am.* 70, 1661.
- Houtsma, A. J. M. (1984). Pitch salience of various complex sounds, *Music Percept.* 1, 296.
- Houtsma, A. J. M. and Goldstein, J. L. (1972). The central origin of the pitch of complex tones: evidence from musical interval recognition, *J. Acoust. Soc. Am.* 51, 520.
- Houtsma, A. J. M. and Rossing, T. D. (1987). Effects of signal envelope on the pitch of short complex tones, *J. Acoust. Soc. Am.* 81, 439.
- Idson, W. L. and Massaro, D. W. (1978). A bidimensional model of pitch in the recognition of melodies, *Percept. Psychophys.* 24, 551.
- Iwamiya, S. and Kitamura, O. (1983). Localized principal pitch of FM-AM tones as a function of phase difference

- between frequency modulation and amplitude modulation, *Proceed. 11th Intern. Congr. Acoust. (Paris)*, Lanninon: GALF, Vol. 3, 115.
- Javel, E. (1980). Coding of AM tones in the chinchilla auditory nerve: implications for the pitch of complex tones, *J. Acoust. Soc. Am.* 68, 133.
- Johnson, D. H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones, *J. Acoust. Soc. Am.* 68, 1115.
- Johnson, D. H. Kiang, N. Y. S. (1976). Analysis of discharges recorded simultaneously from pairs of auditory nerve fibers, *Biophys. J.* 16, 719.
- Kallman, H. J. and Massaro, D. W. (1979). Tone chroma is functional in melody recognition, *Percept. Psychophys.* 26, 32.
- Kay, R. H. and Matthews, D. R. (1972). On the existence in human auditory pathways of channels selectively tuned to the modulation present in frequency modulated sounds, *J. Physiol.* 225, 657.
- Kemp, D. T. (1979). The evoked cochlear mechanical response and auditory micro-structure--Evidence for a new element in cochlear mechanics, in M. Hoke and E. de Boer (Eds): *Models of the auditory system and related signal processing techniques*, *Scand. Audiol. Suppl.* 9, 35.
- Khanna, S. M. and Leonard, D. G. B. (1982). Basilar membrane tuning in the cat cochlea, *Science*, 215, 305.
- Klein, M. A. and Hartmann, W. M. (1981). Binaural edge pitch, *J. Acoust. Soc. Am.* 70, 51.
- Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context, *Cog. Psychol.* 11, 346.
- Lamoré, P. J. J. (1975). Perception of two-tone octave complexes, *Acustica*, 34, 1.
- Larkin, D. (1978). Pitch shifts following tone adaptation, *Acustica*, 41, 110.
- Leshowitz, B. (1978). Measurement of the auditory stimulus, in E. C. Carterette and M. P. Friedman (Eds): *Handbook of perception (Vol. 4: Hearing)*, New York: Academic Press.
- Liberman, M. C. (1982). The cochlear frequency map for the cat: labeling auditory-nerve fibers of known characteristic frequency, *J. Acoust. Soc. Am.* 72, 1441.
- Licklider, J. C. R. (1959). Three auditory theories, in S. Koch (Ed): *Psychology: a study of a science (Vol. 1)*, New York: McGraw-Hill.
- Limbert, C. and Patterson, R. D. (1982). Tapping to repeated noise, *J. Acoust. Soc. Am.* 71 (Sup. 1), S 38.
- Lockhead, G. R. and Byrd, R. (1981). Practically perfect pitch, *J. Acoust. Soc. Am.* 70, 387.
- Loeb, G. E., White, M. W., and Merzenich, M. M. (1983). Spatial cross-correlation--a proposed mechanism for acoustic pitch perception, *Biol. Cybern.* 47, 149.
- Martens, J. P. (1981). Audibility of harmonics in a periodic complex, *J. Acoust. Soc. Am.* 70, 234.
- Martens, J. P. (1984). Comment on "Algorithm for extraction of pitch and pitch salience from complex tonal signals" [*J. Acoust. Soc. Am.* 71, 679-688 (1982)], *J. Acoust. Soc. Am.* 75, 626.
- Miller, M. I. and Sachs, M. B. (1983). Auditory-nerve representation of voice pitch, in R. Klinke and R. Hartmann (Eds): *Hearing--physiological bases and psychophysics*, Berlin: Springer.
- Moore, B. C. J. (1973). Frequency difference limens for short-duration tones, *J. Acoust. Soc. Am.* 54, 610.
- Moore, B. C. J. (1976). Comparison of frequency DL's for pulsed tones and modulated tones, *Brit. J. Audiol.* 10, 17.
- Moore, B. C. J. (1977). Effects of relative phase of the components on the pitch of three-component complex tones, in E. F. Evans and J. P. Wilson (Eds): *Psychophysics and physiology of hearing*, London: Academic Press.
- Moore, B. C. J. (1982). *An introduction to the psychology of hearing*, London: Academic Press, 2nd ed.
- Moore, B. C. J. and Glasberg, B. R. (1986). The relationship between frequency selectivity and frequency discrimination for subjects with unilateral and bilateral cochlear impairments, in B. C. J. Moore and R. D. Patterson (Eds): *Auditory frequency selectivity*, New York: Plenum Press.
- Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (1985). Relative dominance of individual partials in determining the pitch of complex tones, *J. Acoust. Soc. Am.* 77, 1853.
- Nelson, d. A., Stanton, M. E., and Freyman, R. L. (1983). A general equation describing frequency discrimination as

- a function of frequency and sensation level, *J. Acoust. Soc. Am.* 73, 2117.
- Ohgushi, K. (1978). On the role of spatial and temporal cues in the perception of the pitch of complex tones, *J. Acoust. Soc. Am.* 64, 764.
- Ohgushi, K. (1983). The origin of tonality and a possible explanation of the octave enlargement phenomenon, *J. Acoust. Soc. Am.* 73, 1694.
- Patterson, R. D. (1973). The effects of relative phase and the number of components on residue pitch, *J. Acoust. Soc. Am.* 53, 1565.
- Patterson, R. D. (1986). Spiral detection of periodicity and the spiral form of musical scales, *Psychol. Music* 14, 44.
- Perrott, D. R. and Barry, S. H. (1969). Binaural fusion, *J. Audit. Res.* 3, 263.
- Peters, R. W., Moore, B. C. J., and Glasberg, B. R. (1983). Pitch of components of complex tones, *J. Acoust. Soc. Am.* 73, 924.
- Pickles, J. O. (1986). The neurophysiological basis of frequency selectivity, in B. C. J. Moore (Ed.): *Frequency selectivity in hearing*, London: Academic Press.
- Plomp, R. (1964). The ear as a frequency analyzer, *J. Acoust. Soc. Am.* 36, 1628.
- Plomp, R. (1965). Detectability threshold for combination tones, *J. Acoust. Soc. Am.* 37, 1110.
- Plomp, R. (1966). *Experiments on tone perception*, doctoral thesis, University of Utrecht (Netherlands).
- Plomp, R. (1967a). Beats of mistuned consonances, *J. Acoust. Soc. Am.* 42, 462.
- Plomp, R. (1967b). Pitch of complex tones, *J. Acoust. Soc. Am.* 41, 1526.
- Plomp, R. (1976). *Aspects of tone sensation*, London: Academic Press.
- Plomp, R. and Mimpen, A. M. (1968). The ear as a frequency analyser. II, *J. Acoust. Soc. Am.* 43, 764.
- Plomp, R. and Steeneken, H. J. M. (1971). Pitch versus timbre, *Proceed. 7th Intern. Congr. Acoust.* (Buedapest), Vol. 3, 377.
- Pollack, I. (1969). Periodicity pitch for interrupted white noise--fact or artifact? *J. Acoust. Soc. Am.* 45, 237.
- Raatgever, J. and Bilsen, F. A. (1986). A central spectrum theory of binaural processing. Evidence from dichotic pitch, *J. Acoust. Soc. Am.* 80, 429.
- Rakowski, A. and Hirsh, I. J. (1980). Poststimulatory pitch shifts for pure tones, *J. Acoust. Soc. Am.* 68, 467.
- Risset, J. C. (1968). Sur certains aspects fonctionnels de l'audition, *Ann. Telecom.* 23, 91.
- Risset, J. C. (1971). Paradoxes de hauteur: le concept de hauteur sonore n'est pas le même pour tout le monde, *Proceed. 7th Intern. Congr. Acoust.* (Budapest), 20 S 10.
- Ritsma, R. J. (1962). Existence region of the tonal residue. I, *J. Acoust. Soc. Am.* 34, 1224.
- Ritsma, R. J. (1966). The "octave deafness" of the human ear, *Annual Progress Report*, Instituut voor Perceptie Onderzoek, Eindhoven (Netherlands), 15.
- Ritsma, R. J. (1967). Frequencies dominant in the perception of the pitch of complex sounds, *J. Acoust. Soc. Am.* 42, 191.
- Roederer, J. G. (1975). *Introduction to the physics and psychophysics of music*, New York: Springer, 2nd ed.
- Romani, G. L., Williamson, S. J., Kaufman, L., and Brenner, D. (1982). Characterization of the human auditory cortex by the neuromagnetic method, *Exp. Brain Res.* 47, 381.
- Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey, *J. Neurophysiol.* 30, 769.
- Russell, I. and Palmer, A. (1986). Filtering due to the inner hair-cell membrane properties and its relation to the phase-locking limit in cochlear nerve fibres, in B. C. J. Moore and R. D. Patterson (Eds): *Auditory frequency selectivity*, New York: Plenum Press.
- Sachs, M. B. and Young, E. D. (1979). Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate, *J. Acoust. Soc. Am.* 66, 470.
- Sachs, M. B. and Young, E. D. (1980). Effects of nonlinearities on speech encoding in the auditory nerve, *J. Acoust. Soc. Am.* 68, 858.
- Sachs, M. B., Young, E. D., Winslow, R. L., and Shofner, W. P. (1986). Some aspects of rate coding in the auditory nerve, in B. C. J. Moore and R. D. Patterson (Eds): *Auditory frequency selectivity*, New York: Plenum

Press.

- Schneider, B. A. and Bissett, R. J. (1981). The dimensions of tonal experience: a nonmetric multidimensional scaling approach, *Percept. Psychophys.* 30, 39.
- Schouten, J. F. (1940). *Five articles on the perception of sound (1938-1940)*, Instituut voor Perceptie Onderzoek, Eindhoven (Netherlands).
- Schouten, J. F. (1970). The residue revisited, in R. Plomp and G. F. Smoorenburg (Eds): *Frequency analysis and periodicity detection in hearing*, Leiden: Sijthoff.
- Schouten, J. F. and Domburg, G. (1966). Pitch and anticipation, *Annual Progress Report*, Instituut voor Perceptie Onderzoek, Eindhoven (Netherlands), 35.
- Schouten, J. F., Ritsma, R. J., and Lopes Cardozo, B. (1962). Pitch of the residue, *J. Acoust. Soc. Am.* 34, 1418.
- Seebeck, A. (1841). Beobachtungen über einige Bedingungen der Entstehung von Tönen, *Ann. Phys. Chem.* 53, 417.
- Seebeck, A. (1843). Über die Sirene, *Ann. Phys. Chem.* 60, 499.
- Shamma, S. A. (1985). Speech processing in the auditory system II: Lateral inhibition and the processing of speech evoked activity in the auditory nerve, *J. Acoust. Soc. Am.* 78, 1622.
- Shamma, S. A. (1986). Encoding the acoustic spectrum in the spatio-temporal responses of the auditory nerve, in B. C. J. Moore and R. D. Patterson (Eds): *Auditory frequency selectivity*, New York: Plenum Press.
- Shamma, S. A., Chadwick, r. S., Wilbur, W. J., Morrish, K. A., and Rinzel, J. (1986). A biophysical model of cochlear processing: Intensity dependence of pure tone responses, *J. Acoust. Soc. Am.* 80, 133.
- Shepard, R. N. and Jordan, D. S. (1984). Auditory illusions demonstrating that tones are assimilated to an internalized musical scale, *Science*, 226, 1333.
- Shonle, J. I. and Horan, K. E. (1980). The pitch of vibrato tones, *J. Acoust. Soc. Am.* 67, 246.
- Shower, E. G. and Biddulph, R. (1931). Differential pitch sensitivity of the ear, *J. Acoust. Soc. Am.* 2, 275.
- Small, A. M. and Daniloff, R. G. (1967). Pitch of noise bands, *J. Acoust. Soc. Am.* 41, 506.
- Smoorenburg, G. F. (1970). Pitch perception of two-frequency stimuli, *J. Acoust. Soc. Am.* 48, 924.
- Srulovicz, P. and Goldstein, J. L. (1983). A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum, *J. Acoust. Soc. Am.* 73, 1266.
- Stevens, S. S. (1935). The relation of pitch to intensity, *J. Acoust. Soc. Am.* 6, 150.
- Stevens, S. S. and Volkman, J. (1940). The relation of pitch to frequency: a revised scale, *Amer. J. Psychol.* 53, 329.
- Stoll, G. (1985). Pitch shift of pure and complex tones induced by masking noise, *J. Acoust. Soc. Am.* 77, 188.
- Terhardt, E. (1974). Pitch, consonance, and harmony, *J. Acoust. Soc. Am.* 55, 1061.
- Terhardt, E. (1977). Pitch shift of monaural pure tones caused by contralateral sounds, *Acustica*, 37, 56.
- Terhardt, E. (1978). Psychoacoustic evaluation of musical sounds, *Percept. Psychophys.* 23, 483.
- Terhardt, E. (1979). Calculating virtual pitch, *Hearing Res.* 1, 155.
- Terhardt, E. (1980). Toward understanding pitch perception: problems, concepts, and solutions, in G. van den Brink and F. A. Bilsen (Eds): *Psychophysical, physiological and behavioural studies in hearing*, Delft: Delft University Press.
- Terhardt, E. and Fastl, H. (1971). Zum Einfluss von störtönen und Störgerauschen auf die Tonhöhe von Sinustönen, *Acustica*, 25, 53.
- Terhardt, E., Stoll, G., and Seewann, M. (1982a). Pitch of complex signals according to virtual-pitch theory: tests, examples, and predictions, *J. Acoust. Soc. Am.* 71, 671.
- Terhardt, E., Stoll, G., and Seewann, M. (1982b). Algorithm for extraction of pitch and pitch salience from complex tonal signals, *J. Acoust. Soc. Am.* 71, 679.
- Thurlow, W. R. (1943). Studies in auditory theory. I. Binaural interaction and the perception of pitch, *J. Exp. Psychol.* 32, 17.
- Thurlow, W. R. (1963). Perception of low auditory pitch: a multicue mediation theory, *Psychol. Rev.* 70, 515.
- Tomlinson, R. W. W. and Schwarz, D. W. F. (1988). Perception of the missing fundamental in nonhuman primates,

- J. Acoust. Soc. Am.* 84, 560.
- Tong, Y. C., Blamey, P. J., Dowell, R. C., and Clark, G. M. (1983). Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant, *J. Acoust. Soc. Am.* 74, 73.
- Turner, C. W. and Nelson, D. A. (1982). Frequency discrimination in regions of normal and impaired sensitivity, *J. Speech Hearing Res.* 25, 34.
- Turner, C. W., Burns, E. M., and Nelson, D. A. (1983). Pure tone pitch perception and low-frequency hearing loss, *J. Acoust. Soc. Am.* 73, 966.
- Tyler, R. S., Wood, E. J., and Fernandes, M. (1983). Frequency resolution and discrimination of constant and dynamic tones in normal and hearing-impaired listeners, *J. Acoust. Soc. Am.* 74, 1190.
- Ueda, K. and Ohgushi, K. (1987). Perceptual components of pitch: Spatial representation using a multidimensional scaling technique, *J. Acoust. Soc. Am.* 82, 1193.
- van den Brink, G. (1970). Experiments on binaural diplacusis and tone perception, in R. Plomp and G. F. Smoorenburg (Eds): *Frequency analysis and periodicity detection in hearing*, Leiden: Sijthoff.
- van den Brink, G. (1974). Monotic and dichotic pitch matchings with complex sounds, in E. Zwicker and E. Terhardt (Eds): *Facts and models in hearing*, Berlin: Springer.
- van den Brink, G. (1975a). The relation between binaural diplacusis for pure tones and for complex sounds under normal conditions and with induced monaural pitch shift, *Acustica*, 32, 159.
- van den Brink, G. (1975b). Monaural frequency-pitch relations as the origin of binaural diplacusis for pure tones and residue sounds, *Acustica*, 32, 166.
- van den Brink, G. (1977). Octave and fifth settings for pure tones and residue sounds, in E. F. Evans and J. P. Wilson (Eds): *Psychophysics and physiology of hearing*, London: Academic Press.
- van den Brink, G. (1983). Mutual pitch influence of components in a compound sound, in R. Klinke and R. Hartmann (Eds): *Hearing--Physiological bases and psychophysics*, Berlin: Springer.
- van den Brink, G., Sintnicolaas, K., and van Stam, W. S. (1976). Dichotic pitch fusion, *J. Acoust. Soc. Am.* 59, 1471.
- van Noorden, L. P. A. S. (1982). Two channel pitch perception, in M. Clynes (Ed): *Music, mind, and brain*, New York: Plenum Press.
- Verschuure, J. and van Meeteren, A. A. (1975). The effect of intensity on pitch, *Acustica*, 32, 33.
- von Békésy, G. (1960). *Experiments in hearing*, New York: McGraw-Hill.
- von Békésy, G. (1963). Hearing theories and complex sounds, *J. Acoust. Soc. Am.* 35, 588.
- Walliser, K. (1969a). Über die Abhängigkeit der Tonhöhenempfindung von Sinustönen vom Schallpegel, von überlagertem drosselndem Störschall und von der Darbietungsdauer, *Acustica*, 21, 211.
- Walliser, K. (1969b). Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe, *Acustica*, 21, 319.
- Walliser, K. (1969c). Zur Unterschiedsschwelle der Periodentonhöhe, *Acustica*, 21, 329.
- Ward, W. D. (1954). Subjective musical pitch, *J. Acoust. Soc. Am.* 26, 369.
- Warren, R. M. (1982). *Auditory perception*, New York: Pergamon Press.
- Warren, R. M., Bashford, J. A., and Wrightson, J. M. (1980). Infrapitch echo, *J. Acoust. Soc. Am.* 68, 1301.
- Warren, R. M. and Wrightson, J. M. (1981). Stimuli producing conflicting temporal and spectral cues to frequency, *J. Acoust. Soc. Am.* 70, 1020.
- Whitfield, I. C. (1967). *The auditory pathway*, London: Arnold.
- Whitfield, I. C. (1970). Central nervous processing in relation to spatio-temporal discrimination of auditory patterns, in R. Plomp and G. F. Smoorenburg (Eds): *Frequency analysis and periodicity detection in hearing*, Leiden: Sijthoff.
- Whitfield, I. C. (1980). Auditory cortex and the pitch of complex tones, *J. Acoust. Soc. Am.* 67, 644.
- Wier, C. C., Jesteadt, W., and Green, D. M. (1977). Frequency discrimination as a function of frequency and sensation level, *J. Acoust. Soc. Am.* 61, 178.
- Wightman, F. L. (1973). Pitch and stimulus fine structure, *J. Acoust. Soc. Am.* 54, 397.
- Wightman, F. L. and Green, D. M. (1974). The perception of pitch, *Amer. Sci.* 62, 208.

- Wyatt, R. F. (1945). Improvability of pitch discrimination, *Psychol. Monog.* 58, no. 267.
- Young, E. D. and Sachs, M. B. (1979). Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers, *J. Acoust. Soc. Am.* 66, 1381.
- Zatorre, R. J. (1988). Pitch perception of complex tones and human temporal-lobe function, *J. Acoust. Soc. Am.* 84, 566.
- Zwicker, E. (1956). Die elementaren Grundlagen zur Bestimmung der Informations-kapazität des Gehörs, *Acustica*, 6, 365.
- Zwicker, E. (1964). "Negative afterimage" in hearing, *J. Acoust. Soc. Am.* 36, 2413.
- Zwicker, E. (1974). On a psychoacoustical equivalent of tuning curves, in E. Zwicker and E. Terhardt (Eds): *Facts and models in hearing*, Berlin: Springer.