

きる患者もいる。ある程度のレベルの背景雑音がある状態でも無理なく語音を理解できる人が、少数ながらいる。複数電極システムが、単一電極システムに比べて良いという結果が増えてきているが、どのタイプの符号化システムが最も良いかという問いについては、これらの実験結果からはまだ明確に回答することはできない (Gantz *et al.*, 1987; Tyler *et al.*, 1989; Spillman and Dillier, 1989)。人工内耳は全聾の治療にますます広く用いられていくものと思われる。

### 3 ハイファイ装置を選ぶにあたっての音響心理学的考察

ここ数年の間、音響再生機器の販売は、かなり伸びてきた。同時にメーカーは、「より良い」アンプ、スピーカ、コンパクトディスク (CD) などを生産しようとお互いに競いあってきた。あいにく、「良い」についての基準は、必ずしもはっきりと定義されているわけではない。そして、たいていの場合、どの程度技術的な仕様を良くすれば、実際に聞いてわかるほど再生音を良くできるかについて検討することには、ほとんど努力は払われて来なかったのである。ここでは、高忠実度再生の目的を、レコードの制作者や演奏家の意図にできるかぎり近い音の再生と定義する。これゆえ、再生の忠実度や精度といった点を強調する。

どんなに高忠実度 (hi-fi) の再生システムであっても、数多くのコンポーネントから成り立っており、これらの各々がさまざまに信号を変えてしまったり歪ませたりする可能性がある。普通のシステムコンポーネントには次のようなものがある。レコードプレーヤ (カートリッジ, アーム, ターンテーブル), CD プレーヤ, カセットデッキ, アンプ, スピーカ (たいていは二つ) である。これらはそれぞれ、単純にあるいは複雑に信号をひずませる。そして、どんな種類のひずみがあったとしても、再生系全体としての性能は最も弱い部分の性能によって決まる。次の項では、ステレオコンポーネントの性能を表示する方法から考察を始め、これらの表示と知覚に関連する指針をいくつかあげる。次に、再生システムの主要な各コンポーネントについて考察する。ここでは、そのシステムの再生に最も重要であると思われる性能の側面を強調する。

## A 物理特性の解釈

一般によくもち出される物理特性の一つに周波数応答特性がある。これは、振幅が一定で、周波数が異なる正弦波をコンポーネントへ入力することによって測定される。コンポーネントの出力が、周波数とともに変化しないで、すべての周波数が同じレベルで再生されることが理想的である。現実には、コンポーネントが再生できる周波数範囲には限界がある。そして、応答特性が平坦にならない可能性もある。理想的な周波数特性と実際の特性的間のどの程度隔たりがあるかは、ある特定の周波数（多くは1kHz）についての出力を基準レベルとして表すことができる。他の周波数の出力は、このレベルとの関係で表され、総合的な特性は、出力レベルの変動が特定の範囲内に収まっている周波数の帯域幅として表される。たとえば、周波数応答特性が $50\sim 15000\text{ Hz}\pm 5\text{ dB}$ であれば、手ごろなスピーカといえよう。これは、50 Hz から 15000 Hz の間の範囲のいかなる周波数に対しても、音圧レベルは10 dB 以上変わらないということの意味する。

周波数応答特性の二つの側面が知覚と関連する。一つは、全体的な周波数範囲である。応答特性を約30 Hz 以下にまで広げる意味はほとんどない。なぜなら、音楽にはこれより低い周波数のエネルギーはほとんどないからである。そして、いずれにせよ、一般の家庭で30 Hz 以下の周波数を再生することは非現実的である。高周波数端については、20 kHz あるいはそれ以上まで聞くことのできる人もいるが、ほとんどの人は、16 kHz になると語音や音楽に低域通過フィルタを通したとしても、その影響を検知することはできない (Ohgushi, 1984)。もしその違いを聞き分けられたとしても、低域通過フィルタを通したからといって音質が劣下することは決してない。したがって、周波数応答特性が30 Hz から 16000 Hz あれば高忠実度再生には十分である。

周波数応答特性のもう一つの重要な側面は、その平坦さである。もし、周波数応答特性に大きなピークやディップがあれば、再生音を「色づけ」するような影響を音色に与える。音響心理学的な研究によれば、理想的な状況において被験者は、ある周波数領域に比べて、別の周波数領域のレベルが1~2 dB 高くあるいは低くなっただけで、スペクトルの形状の変化を検知することができることが示唆されている (Bucklein, 1962; Green, 1988; Moore *et al.*, 1989)。したがって、周波数応答特性が $\pm 1\text{ dB}$  以内に平坦であれば、完全に平坦な応答特性との違いを検知することはできない。

周波数応答特性で着目すべき重要な特徴をまとめると、次のようになる。理想的には、30 Hz から 16000 Hz の範囲で応答が $\pm 1$  dB 以内に平坦であるべきである。40 Hz から 14000 Hz の範囲で $\pm 2$  dB 以内に平坦であってもまだ大変良い特性であるといえるであろう。

ほとんどのコンポーネントは多かれ少なかれ信号をひずませる。通常、ひずみは入力になかった周波数成分が、出力に現れる比率として表示される。残念ながら、ひずみがどの程度聞こえるのかは、必ずしも簡単にわかるように表示されるわけではない。最も一般的には、高調波ひずみ率を示す方法が取られる。ある特定の周波数の純音が、コンポーネントへの入力として用いられる。もしそのコンポーネントが信号をひずませれば、出力は正確な正弦波ではなくなるが、入力と同じ繰り返し周期をもつ規則的な波形になるはずである。この規則的な波形は、(入力された正弦波と同じ周波数の)基本周波数成分と一連の高調波成分として表すことができる。基本周波数の振幅に対する、第2次高調波以上の高次の高調波の振幅の総和の比率(%)が「高調波ひずみ率」と呼ばれるものである。

高調波ひずみ率から高調波ひずみがどの程度聞こえるかは、単純に予測できるわけではない。なぜならば、ひずみの可聴度は、高調波成分間のエネルギーの分布に関わるからである。もし、ひずみ成分のエネルギーが、おもに第2次と第3次高調波にあれば、それらの成分は基本周波数によってある程度マスキングされてしまう。さらに、もし信号が音楽であれば、いずれにせよそのひずみ成分は、入力信号の中に通常含まれる高調波によって、マスキングされるであろう。このような場合、2~3%のひずみは、気づかれないことが普通であろう。逆に、もしひずみ成分が高次の高調波であれば、この成分はずっと簡単に知覚され、そして主観的にはより不快な感じになるであろう。このような場合、ひずみ率は約0.1%以下にすべきである。

ひずみを測定する方法として、コンポーネントに周波数の異なる二つの音を入力して出力を調べ、入力に含まれない周波数成分の振幅を測定する方法もある。高調波ひずみと同様に、ひずみ成分の振幅は、もともと入力された音に対する比率として表されることが普通である。各音はそれぞれ高調波ひずみ成分を生じるが、それに加えて、二つの音の非線形的な相互作用によって生じる周波数成分も見られる。もし、入力した音の周波数が $f_1$ と $f_2$ であれば、「結合」成分は $f_1 - f_2$ ,  $f_1 + f_2$ ,  $2f_1 - f_2$ ,  $2f_1 + f_2$ のような周波数について生じるであろう。このタイプのひずみは「相互変調ひずみ (intermodulation distortion)」と呼ば

れる(聴覚機構も同様のひずみ成分,特に $2f_1-f_2$ を生じることに注意,第1章,第3章,第5章を参照)。さらにここでも高調波ひずみと同様,記載された相互変調ひずみの比率から,通常の聴取条件において,それがどの程度聞こえるのか予測することは難しい。メーカーによって表示されている図は,測定のためにテストのために使われた周波数と相対レベルによって変わってくるからである。一般に,相互変調ひずみは高調波ひずみよりも簡単に聞こえるが,0.5%以下のひずみでは検知されにくいと考えられる。

コンポーネントの多くは,信号にある程度望ましくない雑音を加える。この雑音は普通,低い周波数成分のハム(交流電源の周波数に関連する)と高い周波数のヒスとなって現れる。この点についてのコンポーネントの性能は,比較的高いレベルの信号を入力したときの出力電力と,ハムと雑音だけのときの出力電力の比として表される。この比率は,普通,デシベルで表示される。ときには,ハムと雑音は,それぞれ異なった周波数に対する耳の感度を反映するように「重みづけられ」ることがある。これは第2章で論じたような騒音計で用いられる重みづけにやや似ている。そして,これは普通より高い信号対雑音比(S/N比)を示すことになる。重みづけなしで70dB,有れば80dBの比率があればほとんどの目的には十分であるが,音源のダイナミックレンジが非常に大きい場合(CDに見られるように)もう少し高い方が望ましい。

## B アンプ

たとえ安価なハイファイ用のアンプであっても,その基本的な性能は大変良く,したがって技術的に性能を改善してもほとんど聞き分けられるような差は生じない。例えば,あまり高級でないアンプであっても周波数応答特性は20から20000Hz $\pm$ 1dBにおよび,ひずみは0.1%以下,SN比は80dB以上である。これらの値は,聴覚によって要求される限界よりも良い。しかし,検討に値する性能の側面はいくつかある。

アンプにとって必要な出力は,用いられるスピーカとの関係のみによって決定される。電気的なエネルギーを音響的なエネルギーに変換する能率は,スピーカによって大変異なる。「ホーン」スピーカは電気的なエネルギーを音に変換するのに30%あるいはそれ以上の能率を示すのに対して,「transmission line」スピーカは1%以下の能率である(後述)。したがって,あるメーカーがいうように,あるアンプが「44ワットのピュアサウンド」を再生するというのはナンセンスである。実際,普通の大きさの部屋では,1音響ワットですら極端

に大きな音に相当する。ごく普通の電氣的な入力、例えば1Wでも、高い音圧レベル(90dB SPL以上)を発生するスピーカはたくさんある。言い換えれば、スピーカの能率が適当であれば、高出力のアンプにたくさんお金をかける必要はまったくないのである。加えて、2倍の出力(これはおそらく2倍の価格を意味する)は音響出力をたった3dB上げるだけであり、これはやっと聞き分けられる程度にしかならないのである(第2章参照)。

### C スピーカ

能率はスピーカの性能の重要な側面である。前節で触れたように、電氣的なエネルギーを音響的なエネルギーに変換する能率は、スピーカによって非常に異なる。残念ながら、メーカーは能率の表示法を今のところ標準化していないが、スピーカから1mの距離で90dB SPLの音圧レベルを生じるのに要した電氣的な入力(W)を示すことが多い。他に、1Wを入力したときに1mの距離に生じる音圧を表示する方法もある。このレベルは、最も低いもので77dB SPLから100dB SPLを越えるものもありいろいろである。1Wのアンプを用いてこの音圧が100dB SPLを越えるスピーカで再生する場合と同じ音圧を、先にあげた能率の低いスピーカで出そうとすると、200Wのアンプが必要となる。

ほとんどのスピーカにおいて周波数応答特性は、アンプのそれに比べるとずっと起伏がある。あまりに特性が悪く、メーカーが「周波数帯域」として音圧範囲を表示しないスピーカもたくさんある。デシベル表示のない周波数応答特性は無意味である。

スピーカの周波数応答特性は、無響室において測定されることが普通なので、結果に影響を与えるような壁、床、ドアからの反射音はない。一般家庭においては反射音は常に存在し、知覚される音質にある程度影響を与える。高い周波数は、通常低い周波数に比べて吸収されやすいので、低い周波数のレベルが相対的に強調される。加えて、低周波数領域では周波数によって、音楽に不快な「色づけ」をする部屋の共振が起こることもある。最後に、高周波数領域では複雑な相互作用のパターンが生じるので、耳に入ってくる音は頭の微妙な位置の違いに強く影響される。

もし、室内の音響特性がスピーカの「実質的な」周波数応答特性を制限する効果があるとすると、無響室で測定されたスピーカの応答は、そのスピーカがどんな音で再生するかとはあまり関係がないと思われるかもしれない。しかし、実際はそうではない。なぜならば「無響室での」周波数応答特性のピークやデ

リップは、室内の音響によって生み出されたピークやディップと結合して、さらに不規則な総合的応答特性を生じるからである。この「不規則性」を最小限にするためには、「無響室での」周波数応答は、できるかぎり平坦にすべきである。加えて、音源の性質を判断する場合、聴取者は部屋の特徴を補正するようである。これは、親しい人の声の質は、その人のいる部屋が変わってもそう大きくは変化しないという日常経験によってもわかる。この効果は、知覚的恒常性の一例である。知覚的恒常性とは、観測者によって知覚される対象（この場合は声）の性質は、観察条件が大きく変化してもほとんど変化しないということである。

先行音効果 (precedence effect) (第6章7節) と類似したなにかが音の定位の場合と同様に、音質の判断にも働いているのかもしれない。もし、そうであれば、音質の知覚は反射された音よりは、むしろスピーカから直接くる、まず最初に耳に到達する音によって決まることになるだろう。したがって、スピーカの欠陥によって生じる信号の変化は、室内の反射によって生じる変化よりもずっと気づかれやすくなるであろう。これは、室内音響が影響を与えることを否定しているわけではないが、この影響は物理的な特性から予測されるものよりは、より検知されにくいであろうということである。スピーカに関するかぎり、周波数応答特性が  $50\sim 15000\text{ Hz}\pm 3\text{ dB}$  あれば非常に良いといえる。そして、もしそのスピーカのほかの性能も同じように良ければ、このような特性のスピーカによって再生された音にはそのスピーカらしさはほとんどないであろう。これは、正確な変換器 (transducer) である。

最近までほとんど無視されてきたスピーカの性能の一つの側面に、位相応答がある。波形を正確に再生するためには、すべての周波数成分の相対振幅を正確に再生するばかりではなく、成分間の相対位相も保持すべきである。これは、聴取者の耳のところですべての周波数成分の時間遅れが同じになるべきである といってもよい。各成分の相対位相が変化すれば、波形は大幅に変化することがある。スピーカのメーカーは、位相応答についてはあまり関心を払ってこなかった。それは、耳は相対位相の変化には鈍感であると一般に考えられてきたからである。耳は「位相聾」であるという説は、基本周波数を含む八つの連続した高調波からなる調波複合音で実験を行った Helmholtz (1863) に起因するといわれている。実のところ、Helmholtz は6~8次以上の高次倍音の位相変化が、検知できる可能性があることを否定しているわけではない。そして、より最近の研究は、位相が変化すると、音色 (Plomp and Steeneken, 1969;

Patterson, 1987a) と音の高さの明瞭性 (Bilsen, 1973; Moore, 1977) の両方が変化することを示している。しかし、定常音については、位相変化の効果は比較的小さい。これは特に驚くには足りない。つまり、定常音については室内の反射によって、成分間の相対位相は顕著に変わるが、先に論じたように室内音響が再生された音の知覚的な質に与える影響は比較的小さいからである。

音楽の中で、当然、普通に見られる過渡的な音や短い音の場合、室内音響が音質に与える影響はかなり異なる。多くの実験結果によれば、音の持続時間がたとえ 2~3ms であっても、われわれは成分の相対位相の違いだけで(各成分の振幅ではなく)音の識別ができることを示されている (Patterson and Green, 1970; また第 4 章を参照)。もし、聴取者の頭が部屋のどの表面からも 60~90 cm 離れており、スピーカが壁にあまり近くなければ、そのように短い音がスピーカから再生された場合、それがどんな音であっても反射音 (echo) が耳に届く前に終わってしまう。したがって、成分間の相対位相が異なる短い音の違いを識別するわれわれの能力に、室内の反射はほとんど影響を与えない。これは Hansen と Madsen (1974) によって確認されている。

これらの結果は、スピーカの位相応答は、再生音の聴感上の品質を決定する上で重要な要因となる可能性があることを意味する。各周波数成分の相対位相の変化は、各周波数成分が広い周波数範囲に分布しているときよりも近接して分布しているときの方がはるかに検知されやすいので(第 3 章 2 E 節)、周波数軸上で不連続に起こる位相変化は、広い周波数範囲にわたってゆるやかに起こる変化に比べて、主観的にはより大きな影響を与える。あいにく、前者の方がより普通の状況である。最新の「高忠実度な」スピーカは、異なった周波数範囲を取り扱う二つあるいはそれ以上のユニットから構成されている。つまり、「ウーファ」が低い周波数領域に、「トゥイータ」が高い周波数領域に用いられる。スピーカへの電気的な入力、ネットワークという電気的なフィルタによって高い周波数領域と低い周波数領域に分けられる。そして、高い周波数領域と低い周波数領域の間の移り変わりの周波数が、クロスオーバー周波数と呼ばれる。このクロスオーバー周波数こそが、急激な位相変化がよく起こる領域なのである。したがって、このネットワークこそが位相ひずみの原因である。加えて、ある種のキャビネット、折り曲げ型エンクロージャーや反射型キャビネットは特定の周波数領域で著しい位相変化をひき起こす。周波数軸上で急激に位相が変化しないようにつくられているスピーカ、例えば Quad 社の静電型は位相変化が急激であったり不規則であったりするスピーカに比べると、より明

瞭で現実感のある音を再生すると判断される傾向にあることは注目に値する。残念ながら、スピーカの間で見られる位相応答の違いを他の性能の違いから分離することは難しい。このため、あるスピーカが他のスピーカより位相特性が優れているために、より良い音で再生するとは、はっきりとはいえない。

Blauert と Laws (1978) は、語音、音楽、雑音、短いパルスなどを含む多くの複合音を用い、周波数成分の一部だけを残りの成分に対して遅延させ、位相ひずみの検知閾について検討した。彼らは、遅延する周波数が1~4kHzの範囲にある場合、遅延時間の検知閾は約400 $\mu$ sであることを見いだした。また、彼らは、いくつかのスピーカやヘッドホンについて実際に生じる遅延時間を測定した。ヘッドホンについては、最悪のもので遅れ時間が500 $\mu$ sに達した。この値は検知閾よりもわずかながら大きくなっている。スピーカについては1500 $\mu$ sに達するものさえあった。この値は、検知閾よりはずっと大きく、スピーカから再生される音の知覚的な品質に重大な影響を与えるものと思われる。位相応答は知覚的には重要な側面であり、将来より明確になることが期待される。

スピーカの周波数応答特性は、普通、スピーカの正面に直接マイクを置く、いわゆる「軸上 (on axis)」と呼ばれる位置で測定される。もし、このマイクが、距離は同じであるが、少し角度がずれた「軸からはずれた」ところに置かれると、全般に測定された出力は低くなる。これはスピーカが、音をすべての方向に均一に放射するのではなく、前面に音を「集中する (beamed)」からである。放射の幅が特定の周波数で狭くなることは、望ましくない。再生音の特質が聴取位置に強く影響されるからである。逆に、「無指向性」スピーカに見られるように、放射幅が非常に広い場合にも問題がある。そのようなスピーカを使った場合、知覚される音質は、あまり聴取位置とは関係がない。しかし、聴取者の耳もとでの、直接音に対する壁や天井からの反射音の比率は、普通のスピーカよりも高くなる。この比率があまりに高いと先行音効果(第6章7節)が抑制されるので、音の定位は大変難しくなる。さらに、先行音効果が通常生じるときとは違って、室内での反射は知覚的に抑制されないという結果もある。したがって、室内の反射によって生じる振幅や位相応答の変化は、主観的にはよりわかりやすくなり、再生音の明瞭性は損なわれることになるのである。

この分野の研究は数少ないが、ほとんどの聴取者は、ある程度は集中度のある指向特性を好むことが普通のようなものである。ただし、軸のどちらの端についても約45°の範囲については、かなり均一に応答することが必要である。普通のスピーカの場合、指向性の幅は、再生される音の波長と変換器 (transducer)



の大きさ（振動面の面積）との関係に依存する。変換器が小さくなればなるほど、また周波数が低くなればなるほど指向性の幅は広くなる。したがって、一つの変換器についていえば、周波数が高くなるほど指向性の幅は狭くなる。多くのメーカーは、再生する周波数帯域内については指向性が十分に広い、二つ以上の変換器（先に述べたウーファとトゥイータ）を用いることでこの問題を解決しようとしてきた。しかし、各変換器の指向角（angle of disparison）は周波数によって異なるので、軸から外れたときの応答は、軸上の応答に比べるとかなり不規則になる傾向がある。指向性についての性能は、十分表示されていないことが多いので、メーカーがどのようにしてこの問題を取り扱ってきたのか確かめることは難しい。さまざまな周波数について応答を角度の関数として示す指向性特性図（polar diagram）を入手できることがときにはある。どの周波数についてもこの図ができるかぎり類似することが理想的である。

Quad ESL-63 スピーカは、広い周波数範囲にわたって指向性を均一にするという問題を巧妙に解決している。このスピーカは、平坦な放射表面をもち、この平面の裏側の中心に点音源があるかのように働くようにできている。そのような点音源であれば、波面は表面の端よりも先に平坦な表面の中心に到達するはずである。したがって、中心から放射する音と端から放射する音の伝播に時間遅れをつくりだすことによって、平坦な表面は仮想的に点音源とすることができるのである。遅延時間の長さは、平面から「仮想的な」音源までの距離に、つまりは指向角に影響する。そして、指向角を最適化するように遅延時間を選ぶことができ、この角度は広い周波数範囲にわたって維持することができる。

最後に、スピーカの性能について葬り去らなければならない迷信がある。ときに、技術雑誌やハイファイ雑誌で「ドップラひずみ」という言葉が使われることがある。このひずみは、スピーカのコーン紙が同時に二つ以上の周波数を再生しなければならないために起こる。低い周波数の音と高い周波数の音が同時に再生される場合を考えてみよう。低い周波数成分の位相が、コーン紙が聴取者に向かって動いている状態にある場合、高い音の周波数は見かけ上高くなる。逆に、コーン紙が聴取者から離れる方向に動いている状態に低い周波数成分の位相がある場合、高い音の見かけ上の周波数は低くなる。このドップラシフトは、駅で目の前を通過する列車の警笛の音の高さのずれとよく似ている。このため、原理的には低い周波数の音は高い周波数の音をわずかながら周波数変調する。これがひずみとして聞こえる可能性があるのである。しかし、実際には、Villchur と Allison (1980) は、聴取実験からも理論的な分析からもドッ

プラひずみはコーン紙の速度にかかわらずまったく聞こえないことを示した。

#### D カートリッジ (Pickups)

われわれは、スピーカの性質について、かなりの多く論じてきた。それは、スピーカが通常、一連の再生系における「最も弱い部分」であるからであり、システムがどのような音質で再生するかについて、最も大きな影響を与えるからである。しかし、この連鎖の始まりとなる信号源（例えば、レコードプレーヤやカセットデッキ）もまた、再生音に大きな影響を与える可能性がある。一般に、最高級のカートリッジの性能は周波数応答、過渡応答、位相応答については、ほとんどのスピーカをしのぐほど優秀であるが、廉価なカートリッジについては必ずしもこれはあてはまるわけではない。しかし、たとえ高級なカートリッジであっても、生じる高調波ひずみと相互変調ひずみの程度は無視できない。後者は、2~5%に達するのが普通である。高いレベルのひずみがあると音の透明さは失われ、各楽器の解像度は悪くなるので、相互変調ひずみのレベルが、カートリッジの再生音の質全体にかなり影響することはもっともである。

カートリッジに固有な性能の側面のうちで重要なものがいくつかある。これらのうちの一つであるトラッキング (tracking) は、深く変調(すなわち高い振幅の)された録音の一節で、再生針がレコードの溝と接触を維持する能力を指す。最近の録音においては、再生針は重力の1000倍の加速度を受けることもあるから、1gもしくはそれ以下の針圧で、溝とうまく接触を維持することができるのは奇跡のようなものである。溝との接触が完全でない状態になれば（これはミストラッキングといわれるが）再生音は、ぼやけたりひずんだりあるいはバチバチという音を伴ったりするであろう。この問題については、メーカーの指定する範囲内で、針圧を上げることで対処できる。カートリッジのトラッキング能力には大きな違いがあるが、この側面について正確に特性を表示しているメーカーはほとんどない(Shure社は除く)。したがって、良い商品を選ぶとするなら、最大限できることはテストレポートに頼ることである。カートリッジのトラッキング能力は、それに見合った品質のアームとともに用いられなければ、実現されないことはいうまでもない。1gあるいはそれ以下でトラッキングできる品質の良いカートリッジを使う場合には、軸受けの摩擦が非常に少なく、重量が非常に軽いアームが必要となる。

ステレオすなわち2チャンネルの再生が要求される場合、カートリッジにはレコードの溝のそれぞれたった一つの壁の動きから得られる、二つの電気信号

を抽出できなければならない。ここで二つの要因がこの問題に関連する。チャンネルセパレーションとチャンネルバランスである。チャンネルセパレーションは、完璧に一つのチャンネルだけから再生されるように録音した信号を用いて測定することが普通である。信号が録音されているチャンネルの出力に対して、信号が録音されていないチャンネルの出力を測定し、その比率をデシベルで表した値をとる。それが、そのカートリッジが二つのチャンネルを分離する程度についての尺度となる。15~20 dB の値があれば一般に適切であると考えられており、多くのカートリッジは楽にこれ以上の性能を示す。しかし、このセパレーションは、広い周波数範囲にわたって維持されることが重要である。

チャンネルバランスは、両方のチャンネルが同じ強度で再生できるように録音した信号によって測定される。両チャンネルの出力の比率をデシベルで表した値が、このバランスの尺度として用いられる。信号の周波数に依存しない両チャンネル間のレベルの違いは問題とはならない。アンプのバランスコントロールで修正できるからである。しかし、周波数によってレベルの差が変化する場合には問題が生じる。ステレオ録音においては、どんな楽器の主観的な定位も、二つのスピーカによって再生される相対レベルによって規定される。聴取者が二つのスピーカから同じ距離にいるときに、ある音が二つのスピーカから同じレベルで再生されれば、その音は二つのスピーカの中央に定位して聞こえるであろう。もし、左のスピーカからくる音のレベルが、右からくる音のレベルより少し高ければ、音像は少し左にかたよる。もし、左のスピーカから再生された音が、右から再生された音のレベルに比べて10~15 dB 強ければ、音は完全に左のスピーカからしか聞こえない。チャンネルバランスが、各楽器の主観的な定位を決定するのに重要であることは明らかである。もし、バランスが周波数とともに変化するなら、同じ楽器で演奏された音であっても、異なる周波数成分は主観的に異なった場所に定位する。したがって、はっきりとした「ステレオ音像」は生じないことになるであろう。そして、演奏されたどんな楽器についても、その「音像」はぼけてくることになり、そして各楽器の分離は悪くなることになるだろう。これと同様のことが、ステレオの二つのスピーカについてもあてはまる。すなわち、周波数および位相応答特性についてはできるかぎり揃える必要がある。そうしなければ「ステレオ音像」は悪くなるであろう。

## E ターンテーブル (Turntables)

ターンテーブルの速度に変動があれば、それに応じて再生音の周波数に変動

する。ワウ (wow) と呼ばれる遅いゆらぎは、音の高さの変化をひき起こす。これはピアノの音の定常部分を聞けば分かりやすい。これに対してフラッター (flutter) と呼ばれる速いゆらぎは、再生音を「粗く」する。ワウとフラッターを表示する規格は二つある。一つはヨーロッパの DIN 規格であり、これはおもにワウとフラッターのピーク値に基づく。もう一方はアメリカの NAB 規格であり、これはワウとフラッターの平均値に基づく。ワウとフラッターは、約 0.2% (DIN) または 0.1% (NAB) 以下であれば通常検知できない。最近の良いターンテーブルでならば、この値より悪いことは少ない。

ターンテーブルそのものの機械的な振動が、カートリッジを通してランブル音 (rumbling sound) として再生されることがある。ランブル音は、特定の基準レベルで録音された音 (溝での速度が 10 cm/秒のとき (水平方向で) の 1 kHz の純音を用いられることが多いが) との相対レベルを測定することによって表示される。品質の良いターンテーブルにおける SN 比は、重みづけなしで 50 dB もしくは重みづけがあれば 70 dB である。ランブルが聞こえる程度は、もちろんスピーカの低い周波数の再生能力に依存するであろう。したがって、低域まで周波数特性が延びた、品質の高いスピーカを使うのなら、ランブル音が少ないターンテーブルと組み合わせるとよい。

## F カセットプレーヤとレコーダ (Cassete players and recorders)

もともと小型のカセットテープは、高忠実度で録音するための媒体を意図したものではなかった。しかし、テープやレコーダ/プレーヤ (「テープデッキ」) は、この媒体の限界を越えるべく非常に進歩してきた。カセット固有の SN 比は約 55 dB にすぎない。高忠実度で再生するためには、この値はきわめて低い。そこで、カセット録音に固有の「ヒス」の影響を低減する目的で、多くの「ノイズリダクション」システムが開発されてきた。最も広く用いられているシステムはドルビー B (Dolby B) とドルビー C である (ドルビー A もあるが、おもにプロのレコーディングスタジオで用いられる、より複雑なシステムである)。これらのドルビーシステムは、録音の前に入力信号の高い周波数領域を強調する。再生するときには、これらの高い周波数成分を低減し、もとの高い周波数領域と低い周波数領域のバランスを復元しながら、テープから生じるヒスだけを減衰させる。しかし、困ったことに、入力信号が強い高周波成分を含んでいると、さらに強調するとテープは過負荷になり、かなりのひずみを発生させることになるであろう。このひずみを避けるためには、入力レベルが低いと

きに限って高い周波数領域の強調を行うようにしなければならない。したがって、高い周波数領域のレベルを再生時に下げる程度もまた、全体的なレベルに対応しなければならない。テープに録音される信号のレベルによって、どの程度高い周波数領域のレベルを下げるべきかを規定されるのが普通である。

ドルビー B とドルビー C の動作の概要は、以上のとおりである。ドルビー B は SN 比を約 65dB に改善するのに対して、ドルビー C は約 73dB に改善することが普通である。これらのシステムは、録音時に行った高い周波数領域の強調が、再生時に行われる高い周波数領域の減衰と正確に一致するように、慎重に回路を設計しなければ適切に機能しない。困ったことに、テープの「感度」は、ブランドによって異なる。ここでいう「感度」とは、あるレベルで録音した音について得られた再生レベルである。ドルビー処理はレベルに依存する。このため、テープの感度に違いがあるということは、テープデッキは特定のブランドのテープに合わせて調整する必要がある、この調整によってそのブランド、あるいは基本的に同じ感度をもつほかのブランドのテープに限ってうまく働くということを意味する。ドルビー C はドルビー B に比べてノイズリダクションの効果が大きいのであるが、テープデッキの調整が不正確であったりテープの感度が違ふと、より影響を受けやすい。

カセットには、録音レベルが高い場合、テープが「飽和」してしまうという別の問題点がある。その結果、高い周波数領域は減衰し、ひずみは増加する。テープデッキの周波数特性は、一般的には最大値から 20dB 下がったレベルで録音した特性として表示される。このレベルでは、30 から 16000 Hz で  $\pm 3$ dB の応答特性があれば良いデッキといえる。しかし、もし、録音レベルが上がれば、周波数特性は極端に悪化するのが普通である。すなわち、録音レベルメータが「0」を指すときには高い周波数領域の応答特性は 8000 Hz あるいはそれ以下に狭まるであろう。高レベルの周波数応答を改善する方法に、ドルビー HX (Dolby HX) [上限拡大 (headroom extention)] がある。これは、高レベルでの録音をひずみの増加や高い周波数領域での減衰なしに行わせるものである(これがいかに行われるかについての技術的な詳細についてはこの本の範囲外である)。

最後に、廉価なテープデッキの中には、ワウとフラッターが聞こえるものもあることにも注意すべきである。良い製品を買おうとするなら、ワウとフラッターが 0.2% (DIN) または 0.1% (NAB) 以下のデッキを探すべきであろう。

## G コンパクトディスクプレーヤとデジタルオーディオ テープデッキ (DAT)

CDプレーヤとDATは用いる録音媒体が大きく異なるが、その性能は本質的には同じなのでここでは一緒に扱う。CDとDATは、ほかのステレオシステムのコンポーネント、特にカセットデッキやスピーカなどに比べて一般には格段に性能が良い。原理的にこれらの出力信号は、レコーディングスタジオで制作されたマスターテープからの出力と弁別はできないであろう（スタジオレコーディングもいまやデジタル録音が普通である）。したがって、CDやDATがその規格どおりに動いているかぎり、検知できるような音質の劣化をひき起こすことはない。これゆえ、すべてのCDとDATは同じように再生できるわけである。

この主張にはCDとDATのエラー修正能力にやや違いがあるという制限をつけ加えておくことが必要である。したがって、損傷を受けたり欠陥のあるディスクやテープを再生できるかどうかについて、再生装置の間で生じるわずかな違いを見いだすことが、ときには可能であろう。しかし、ほとんどのCDやDATは、一般に起こり得るエラーに対しては大変効果的に対処できる。ただし、CDプレーヤやDATが、聞いてわかるような効果を生じる欠陥をもつこともあり得るのである。しかし、次のようにはいえる。最も安いCDやDATを買いなさい、それでもあなたの要求を満たすことはできる。しかし、評判の良いメーカーから買いなさい、と。

## H スピーカケーブル (Loudspeaker cable)

アンプとスピーカの間特別なケーブルを使うことによって、再生音の音質を良くすることができるという主張するメーカーがいくつもある。ケーブルの抵抗を低くすることによって、最良の状態で働くようになるスピーカがあるということは事実である。しかしこれは、特別なケーブルを必要とするということの意味するわけではない。もし、長いケーブルを使うなら、全体の抵抗が約 $0.5\Omega$ を超えない程度に太くするべきであるということだけを意味するだけなのである。特別なケーブルが「豊かな低音」だの「輝かしい高音」などを再現できるというような宣伝は無視しなさい。そのような主張には何の根拠もないからである。

## I 結 論

ハイファイシステムのコンポーネントを選択する場合、各コンポーネントは個別的に選択すべきではない、という点が重要である。むしろ、各コンポーネントは他のコンポーネントと同じ品質になるように選ぶとよい。もし、予算が限られているなら、能率の高いスピーカを購入すれば、音は最も良くなるであろう。このようなスピーカは、適度な出力のアンプで駆動すれば十分大きな音で鳴るからである。しかし、高品質のスピーカの多くは、比較的能率が悪い。もし、このようなスピーカを使うなら高出力のアンプがおそらく必要となろう。低能率のスピーカは高能率のスピーカが出すのと同じ音圧を出すためには100倍もの電力を要する場合がある。すなわち、前者に100Wのアンプ必要なら後者には1Wで十分ということである。レコードプレーヤ(ターンテーブル、アーム、カートリッジ)とテープデッキは、スピーカの後に選択すべきであろう。スピーカの質が良くなるほど、プレーヤやデッキの欠点が明らかになる傾向があるからである。ドルビーCやドルビーHXのついたテープデッキはヒスが少なく、高レベル録音時に周波数特性が良い。細かいデータ訂正能力の違いを考慮しなければ、CDプレーヤの音質はどれをとっても本質的に同じである。スピーカは、ハイファイシステムの中で最も重要なコンポーネントである。しかし、音質に影響を与える要因の多くは、メーカーによって表示されていないか、もしくは表示はされていても主観的印象とは単純に結びつけられないか、のいずれかである。したがって、メーカーの表示した特性からそのスピーカが、どのような音質で再生するかを知ることは難しい。スピーカを選ぶ最良の方法は、もしできるなら慣れ親しんだ録音の良いCDを使って自分の家で聞くことであろう。

## 4 コンサートホールの音響効果の評価

このトピックは最も複雑なものの一つであり、その詳細について網羅することは、この本の範囲を越えている。その詳細について、さらに興味のある読者はBeranek (1962) と Ando (1985) を参照されたい。コンサートホールの音響効果を比較するのに固有の問題を解決するための、いくつかの技法を短くまとめるだけにとどめておきたい。

コンサートホールの音響の質 (acoustic quality) の判断や、さまざまなコンサートホールの比較は、二つの要因があるために難しい。まず第1に、ある演奏で知覚される音の質は、コンサートホールそのものの特性に加えて、音楽家の演奏の仕方や演奏者の席の配列によっても影響を受ける。第2に、音についての長期記憶はあまり良くないので、音質の細かい部分の多くは、聴取者があるホールから別のホールへ移動するうちに思い出せなくなる。結局、聴取者の大多数の意見が一致することで、あるホールが「良い」か「悪い」かが決定される。しかし、ホールの評価の判断に影響を与える決定的な要因を特定することや、新しいホールを設計するときに応用できるような一般的な法則性を見つけることは簡単ではない。実際、新しいホールの多くは、良いと判断されてきた他のホールを単純に手本としてつくられるものである。この方法にはある種の妥当性があるが、その反面柔軟性がない。

Schroeder ら (1974) は、同一の音楽素材を用い、実際的な自由音場の条件のもとで、異なったホールの音の質を直接的に比較することを可能にする録音再生技法を開発した。まず第1段階では、オーケストラを無響室で録音する。このため、録音室の音響効果は評価されるホールの音響効果に影響を与えない。これは、音楽素材が常に同一であることを保証する。この録音は、評価されるホールのステージで再生される。Schroeder ら (1974) は2チャンネル録音を用い、ホールのステージでオーケストラの空間的な配列をありのまま再現するように、そのステージの上に二つのスピーカを置いて再生した。しかし、この技法が、生のオーケストラの演奏をより正確に再現できるようなマルチチャンネル録音に発展しなかった理由についてはわからない。

第2段階では、ダミーヘッドの「鼓膜」のところに置いたマイクで信号を録音する。ダミーヘッドとは、実物によく似た耳介と外耳道があり、「平均的な」頭の音響特性にできるかぎり一致するようにつくった、人間の頭の模型である。各ホールについて、普通聴取者の頭があると思われる位置に対応する、異なった2カ所で録音された。

第3段階では、評価のためにダミーヘッド録音を聴取者に呈示する。これを行うために、ステレオヘッドホンで再生する方法がよく用いられるが、この技法には限界がある。Schroeder と彼の共同研究者たちが指摘している一つの問題点は、次のとおりである。『ヘッドホン受聴では、知覚された音響「空間」または、われわれが客観的な方法で記述したい重要な特性の一つである、音に包み込まれているような感覚(囲繞感)が正確に再現されることはないのである』。



この問題を解決する目的で、無響室内で選択された二つのスピーカから再生するとき、録音信号が聴取者の鼓膜の位置で再生されるように録音が電氣的に処理された。Schroeder ら (1974) はこの技法について以下のように説明している。

各スピーカから放射された音は、「両方」の耳に到達するのであって、期待されるように聴取者の「近い側の」耳だけに届くわけではない。換言すると「遠い側の」耳に対しても、各スピーカからの「クロストーク」があるということである。しかし、フィルタを通した補正信号を適切に混合してスピーカから放射すると、この望ましくない「クロストーク」は相殺できるのである。

この補正方式に適切なフィルタの応答特性は、短い電氣的なインパルスを二つのスピーカのうちの一つに入力し、スピーカの正面からある程度の離れたところに置いたダミーヘッドの両耳に挿入した二つのマイクから得られた信号を録音した結果から算出できる。詳細については、原著 (1974) と Schroeder と Atal (1963) を参照されたい。

この実験結果は、Schroeder ら (1974) によると次のようであった。もとのダミーヘッド録音が再生されたときには、

……ダミーヘッドの右耳からの信号は聴取者の右耳だけに、ダミーヘッドの左耳からの信号は聴取者の左耳だけに呈示されることになる。それは、ちょうどヘッドホンで聞いているときのように、しかし、外耳道が適切に自由音場と結合しており、聴取者の頭が垂直軸の回りを回転しているときも、われわれの求める音響空間の知覚的不変性が得られるのである。

Schroeder ら (1974) は、ほぼ $\pm 10^\circ$ までの頭の動きに対しては、外在化された音のイメージは、リスニングルームに対して不変であり、ヘッドホンで受聴するときのように頭とともに回転することはない、と報告している。

Schroeder とその共同研究者たちは、この方法を用いて、異なったコンサートホールを一对にして呈示し、どちらを好むかを答えるよう被験者に求めた。それぞれの対において、被験者には二つのホールでの録音を好きなだけ切り替えて聴くことが許された。好まれたホールの得点は1であり、そうでないものは-1とした。もし、好みに違いがなければ、両方のホールに0が与えられた。以上の手続きが、すべての被験者について、またすべてのホールの対について繰り返された。そして、その結果は各ホールが各被験者に何回好まれたかを示す、嗜好マトリクスにまとめられた。

この結果は、データの分散を説明する有意な因子を抽出する因子分析 (Slater, 1960) により分析された。有意水準を越えた因子は一つであった。この因子は全分散の 50% を説明するもので、嗜好合意因子 (consensus preference factor) と命名された。すべての被験者は、この因子に正の因子負荷量を示した。この結果は、もしあるホール (例えば X) の因子負荷量が別のホール (例えば Y) より高ければ、ホール X はすべての聴取者によって、ホール Y より好まれているということを意味するのである。分析によって抽出された別の因子は、個人の好みの違いを表す因子であると思われる。もとの録音は、人のいないホールで行われた。人間の有無は、現代のホールについては音響にほとんど影響を与えないが、堅い木の椅子のある古いホールについては、人がいなければ残響時間が過剰に長くなることがある。残響時間が 2.2 秒以下のホールに分析を限定した場合、嗜好合意因子は全分散の 88% を説明するほど重要になるが、逆に好みの個人差の因子はより重要ではなくなった。

Schroeder ら (1974) はまた、嗜好合意因子とホールの幾何学的、音響学的なパラメータとの相関を探る目的で、ホールの種々の物理特性を測定した。彼らの考察した測度のすべてについては各々列記しないが、最も重要と思われるものについて触れることにする。残響時間が 2 秒以下のホールについては、残響時間は、嗜好性と高い相関をもつ。すなわち残響時間が長くなるほど、より好まれるようになる。しかし、残響時間が 2 秒以上のホールについては、残響時間は嗜好合意因子とやや負の相関を示すが、個人差の因子とかなり大きな相関を示すようになる。したがって、残響時間が 2 秒を越える範囲では、長い残響時間を好む聴取者がいる反面、短い残響時間を好む聴取者もいるということになる。

ホールの残響時間が長くても短くても、好みと強い負の相関を示した要因は、両耳間相関であった。両耳間相関とは二つの耳に到達する信号の相関である。聴取者は両耳の相関が低くなる、すなわち、両耳に到達する信号は相対的には互いに独立であるホールを好むのである。Schroeder ら (1974) は次のように示唆している。『おそらく、この効果は両耳の相関が少ない信号によって引き起こされる、よりはっきりした感覚——音に浸されているような——によって生じるのであろう』。両耳間の結合は、コンサートホールの壁や天井に音の拡散装置をつけることによって制御することができる。

最後に、残響時間が長いホール、これはまた大きいホールが多いが、ホールの容積は嗜好合意と強い負の相関を示す。したがって、いったんホールがある

程度の大きさに達すれば、さらに大きくすると知覚される音の質は悪くなるのである。

## 5 要約とあとがき

この節では、聴覚研究を実用的に応用した研究について、いくつか取り上げてきた。

音響心理学的なテストは、難聴の種類診断と、障害をもつ人が日常生活での程度困難を感じるかを予測することに役立つであろう。周波数選択性の測定は、この点で大変有効であると思われる。障害のある耳に入力する以前に信号に前処理を施す補聴器は、障害によって生じた知覚異常のいくつかを、部分的に補償することができる。例えば、AGCのある補聴器は音の大きさの補充現象を補償する上で役立つ。完全に聾の患者の場合、人工内耳は、限界はあるものの聴覚を復元する上で役に立つ手段を提供することができる。

音、特に音楽の再生系においては、再生システムにおける変換器、例えばスピーカ、カセットプレーヤ、カートリッジなどが全体的な音質に大きな影響を与える。残念ながら、システムが再生する音質に関わる性能の側面の多くは、あまり明示されていない。アンプ、CDプレーヤ、DATの性能は十分に良いので音質にはほとんど、あるいはまったく影響を与えない。

ダミーヘッドを用いた録音は、いろいろなホールの音響的な違いを比べる一つの方法となる。嗜好性に最も強く相関する因子は、残響時間と両耳間相関である。

音の知覚の機構に関する知識は、聴覚の科学的な研究がまだその創世期であった頃の Helmholtz (1863) の先駆的研究以来、飛躍的に増大してきた。生理学的なレベルでは、今や聴神経や蝸牛神経核での音の符号化について多くの知識が蓄積されている。しかし、このような神経系の基本的な情報が、聴覚系の高次のレベルでどのように処理されるのかについては、より低次の処理に比べるとあまり解明されているとはいえない。知覚的なレベルでは、純音や帯域雑音といった単純な刺激の変化を検知する能力については、多くの知識が蓄積されているが、音声や音楽といった複雑な聴覚パターンが、どのように知覚されるかについて理解するまでにはまだ遠い道のりがある。あいにく、われわれの理解が初歩的なレベルであるからといって、そこに含まれる概念が単純である

というわけではない。多くの学生たちが、聴知覚や聴覚系の神経生理や解剖学の科学論文に現れる専門用語のおかげで、思いとどまってきたのである。物理学の予備的知識のない学生にとってみれば、聴覚刺激の性質や、このような刺激のフーリエ分析ですら概念的には大変難しいものであろう。はじめの章から徐々に進んで来て、この本の最後まで読み終えた読者ならば、これらの最初の困難さを克服することを、そしてわれわれの音の知覚の背後にある魅力的で複雑な過程について大いに興味が啓発されることを、私は望んでやまないのである。