

音に関する物理的特性と心理的特性

鈴木 寛*

(平成 13 年 10 月 31 日受理)

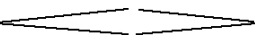
強弱の概念

ピアノという楽器の名前の由来ははピアノやフォルテを表現できる機能を備えたのでそれまでのチェンバロやハープシコードと区別して「ピアノフォルテ」と命名されたことからというのは周知の事実である。

楽譜に p や f の記号が出現したのは後期バロック音楽の頃からであり、グレゴリアンなどの楽譜やネウマによる楽譜にはその記述が見あたらない。しかし、チェンバロやパイプオルガンのような複数の鍵盤を有する楽器ではバロック音楽固有の特徴である「テラス状のダイナミック」つまり強い音の出る鍵盤か弱い音の出る鍵盤かを使い分けてリトルネロ形式やコンチェルト・グロッソ等の音楽様式が一般的なダイナミック表現に用いられてきたのである。

パイプオルガンには、SWELLER と呼ばれるパイプ群そのものを大きな箱の中で鳴らし、その箱に取り付けられたベネシアン・ブラインドのような複葉扉を開閉することで強弱というより明暗や遠近というニュアンスに近い音量制御を行う装置と、CRESCENDO 又は ROLL SWELLER と呼ぶストップの増減を行う装置で音源の数や音色の違うものを追加する装置がある。

同様にハープシコードやチェンバロにおいても隠れた位置にある弦や 4 フィート等の倍音を重ねることで明暗をコントロールをしたり、弦そのものの太さの違うものを区別することで音量を調節したのである。

本来自由に音量がコントロールできる声楽や弦楽器ですら当時は  のような表現は極力避けるように演奏されたのである。

従って当時の音楽はピアノフォルテの概念は今日的な概念とは異なるものであったと推察できる。

現代では学生や現場の教師に「ピアノ」の意味を尋ねると「弱く」と答える。又「フォルテ」の意味を尋ねると「強く」と答える。音楽辞典（標準音楽辞典―音楽之友社）に拠れば、ピアノとは「Piano (伊) <やわらかく>、<低く><弱く>」で、フォルテは「forte (伊) <強く><大きく>」とある。

所が小中学校の教科書や楽典の記述は全てピアノ＝弱く、フォルテ＝強くとなっているのである。因みにアメリカの教科書では PIANO=softly FORTE=loudly となっている。元来ピアノは「柔らかく」というのが第 1 義的に用いられるべきであったのだが、「強弱

という漢語にあわせて「弱く」が一般的になってしまったようである。このピアノ＝弱くという概念が日本の音楽教育や演奏家教育に大きな弊害をもたらしていることは以外と知られていない。例えば合唱指導などで、指導者が「ほら楽譜にピアノと書いてあるでしょ、だからもっと弱く歌って頂戴。」という言葉で指導するのと、「ほら楽譜にピアノと書いてあるでしょ、だからもっと柔らかに歌って頂戴。」と言うのでは結果が全然違ってくるのである。

弱く歌うと「貧弱」で「痩せた」演奏になってしまい、柔らかに歌うと「優しく」「暖かい」演奏になる。ピアノや打楽器のようにタッチだけで p f を表現する音楽の場合はこの差は無い。しかし、持続音を発する弦楽器や管楽器或いは人声の場合このタッチに相当するものと、ボリュームの時系列変化では意味が違ふ。

いずれにせよデジタルであろうとアナログであろうと音のデータは瞬間の振幅のみで記録再生されるのである。この「振幅」のことを英語で Amplitude と呼び、ボリューム変化等では相似形の波形のままズームで拡大縮小する場合にもこの Amplitude (増幅の度合い) ということばを使う。このように音のエネルギーに関する言葉は日本語でも「大小」「強弱」などの表現があるように英語でもこの Amplitude については Volume や Loudness、Intensity や Velocity、などのそれぞれニュアンスの異なる言葉がある。このニュアンスについては後述するとして音に関する迷信がいくつかある。

音の迷信

音の迷信その 1 は音の「大小」と「強弱」は同じであるとする考えである。「大小」は連続変化やズーム的变化で、「強弱」はニュアンスの違いのことなのである。音量を変えて聴いても強弱のニュアンスは変わらない。

前に述べた「ニュアンス」と言う言葉をもう少し掘り下げて考えて見よう。「はい」と言う平かなの台詞を何通りのニュアンスで言えるであろう。肯定の「ハイ」もあれば渋々の「はい」もあり、元気いっぱい「はい！」もある。これらの台詞を喋っているドラマの音量を変えて聞いてもニュアンスは同じである。

これらのニュアンスの因子は「アタック」「ディケイ」「サスティン」「リリース」等のパラメータで構成され

* 兵庫教育大学 学校教育学部附属実技教育研究指導センター (音楽教育分野)

るいわゆるエンベロープ的なものと、音域やピッチ変化等によるイントネーション的なもの、音質や音色に関するもの、スタッカートやレガート等のゲートタイムに関するもの、ビブラートやトレモロのようなモジュレーションに関するものなどがある。これらのどの要素もCD再生装置等のボリュームの位置とは関係がない。(勿論極端に小さいとか大きい場合は別であるが)

ここで述べた音の「大小」と「強弱」はこの中でも特にボリュームとエンベロープ的なものの区別についてのことであった。ボリュームは「絶対量」のことであるが、エンベロープやベロシティ、エクスプレッションなどは「相対量」である。

ボリューム等の「絶対量」はいわば固定状態で保持され、ニュアンス等のために細かく変化させない。にもかかわらず、かなりのDTMシーケンス・ソフトではチャンネル・ボリューム或いはトラック・ボリュームしか触れないようになっていて。本来これらのコントロールはパート間の音量バランスを調整するためのものであり、大きな段落や楽章の変更の時に調整されることはあっても演奏中にコチョコチョ触るものではない。本来の目的のためにチャンネル・ボリュームを設定してもあつと言う間にそれが動いてしまつて何のために・・と言うような経験もよくある。そのような場合に備えてGM規格ではコントロール番号8にバランスコントロールというのが用意されているが、使われているケースは極めて希である。

この音の絶対量についてやはり迷信がある。それは大きな楽器からは大きな音が出るとか、歌手の人数が2倍になれば音量も2倍になるというようなものである。

視覚的に大きいものは大きな音が出るというのはおかしい。もしそうならばパイプオルガンのいちばん短い小さなパイプの音は一番太くて長いパイプにかき消されてしまうはずであり、それを防ぐためにパイプオルガンでは短いパイプほど数を増やさなければならないはずであるが、実際にはそんなことはしない。

Sの教育(本紀要シリーズSMLの音楽科教育に説明済み)ではこの「視覚的に」大きい或いは多いものがすべて音量も大きいと信じさせてしまう傾向に注意しなければならない。

音響物理学では振幅＝音量かも知れないが、音楽の世界では「振幅×振動数」が音量なのである。従って高音楽器や高音域では振幅が小さくても振動数が多いので音のエネルギーは強いのである。つまり、小さなジャブも繰り返されると大きなパンチに匹敵するダメージを与えるのと同じである。初期のデジタルシンセはこのことが考慮されていなかったため、高音域はどうるさく、低音が不足する感じがしたのもそれが原因だったと思われる。Moog等のシンセではアンチログ

と呼ぶ回路でこれを補正し、あの豊かで芯のある低音を出していた。これをVU (Volume Unit) メータと呼ぶテープレコーダーなどに付いているレベルメーターで見ると、基本的には振幅計であるため当然低音はブンブンとメーターが振れ、高音では振れない。優秀な録音技師はこのことを知っているので故意に高音のレベルを上げる(或いは低音のレベルを下げる)ことはしない。TBSのこども音楽コンクールのテープ審査等ではリコーダーから大プラスまで全部同じメーターの振れにしていまい、何でも同じ出力で再生するので実物大の臨場感や厚みのない審査をしなければならないことをいつも不満に思っている。

あらゆる音楽ソースは実物大(等身大)の絶対音量を持っており何でも同じ音量で聴くというスタイルは住宅事情もあるだろうが自然な姿ではない。

しかし、一方では絶対音量がズームアウトされて縮小されても我々は臨場感以外は損なわれることは無いため支障はないという考え方も間違いではない。つまり、音源から一番遠いところの聴衆とかぶり付きでは違う音楽を聴いているわけではないとも言えるからである。

パイプオルガンでは一番短いパイプ(一番高い音)でも一番長い(低い)パイプでも本数は同じである。ピアノでは高域は弦が3本であるが、低域では1本である。この音源の数と音量の関係にもひどい迷信がある。つまり、音源が2倍になると音量も2倍になると言う考え方である。うちのクラブは人数が少ないのでボリュームがない等の言い訳によく使われる。

実際には我々が受容できる音量は0デシベルから130デシベルくらいまでで、もし甲子園に5万人の観衆が入って全員が80デシベルの絶叫をしたならばその日の観客は全員聴覚障害者になってしまうはずである。そうならないのは、音源の数が2倍になってもたった3デシベルしか絶対音量は増えないからに他ならない。従って誰もピアノの弦が増えたり減ったりする場所に気づかないし、音階が斑模様に聞こえることはないのである。

視覚的な楽器の大きさと音の大きさは基本的には関係がないことは冷静に考えれば理解できるが、子どもたちにはなかなか理解しにくい事であろう。それでもこのことはS即ち音知覚の教育ではとても大切なことなのである。音に限らず外見だけではわからないものがこの世の中にはたくさんあることを学習すべき場面の一つだと思う。それでなくても多数決が支配する今の世の中で少数派の中にも優れたものがあることを知ると言うような経験は大切である。例えばあの大オーケストラといえどもピッコロは1〜2本しか使われていないことなどからもそれは学習されると思う。音量もさることながら、どんな場面でどんな風に演奏されるから

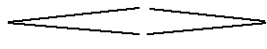
ピッコロの見かけ上の大きさと関係がない聞こえ方があることを知るような学習の素地が大切なのである。

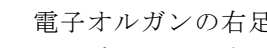
昨今の学校におけるDTMの指導者の内何人がこの言葉の分類や区別ができるであろうか。あるいは一体どのメーカーの学校用ソフトがこれらのニュアンスにこたえうるのであろうか。

例えばチェロを弾くとき、左手はピッチをコントロールし、右手は弓の圧力や擦る速度を変化させてVolume

Loudness、IntensityやVelocityのコントロールをしている。

MIDIデータではチャンネル・メッセージやコントロール・チェンジでこれらの表現をするのであるが、「強弱」に該当するVelocityはノート・オンのデータであるチャンネル・メッセージにチャンネル番号とキー番号にくっつけて処理される最低限度のAmplitude情報である。チェロ等のボウイングにUpやDownがあるがこのようなAmplitude情報は音源そのものがデータとして持っているのが普通である。

のような連続的な変化の場合はコントロール・チェンジの7番や11番を小刻みに変化させる。ピアノやティンパニーのような場合は

のような連続的な変化は前後の音のヴェロシティを変化させながら行うので、連続的な変化量でありながらMIDIデータは音符の数しかなく、弦楽器や管楽器の演奏データに比べて極端に少ない。

電子オルガンの右足で操作するエクスプレッション・ペダルはアンプのボリューム操作のようなものである。連続変化には向いているが、アクセントや個々の音のニュアンスにはタッチセンスが必要である。このタッチセンスとはベロシティを検出する装置のことである。

ピアノ練習の問題点

ベロシティの辞書的意味は「速度」である。ピアノにおいては「打鍵速度」であり、管楽器では「息の速度」であり、弦楽器では「弓の運動速度と圧力」で表現される。いわゆる電子オルガンや電子キーボードの多くはこのベロシティを検出する装置を持たないのでベロシティでしか表現できない音楽表現は制御できないのである。このことは簡易キーボードや足踏みオルガンがピアノの代用にならない最も大きな理由である。ベロシティは音のエネルギーを価として有するパラメータである。ボクシングの「パンチ力」のことである。あの小さなゴルフクラブのヘッドスピードがグランドピアノを持ち上げる瞬発力を生み出しているのもベロシティである。そこで結論としてピアノ・フォルテは実はベロシティのことでは無かったかと考えるに至った。

経験的に分かっていたことであるが、ピアノの学習者が陥る典型的な事例がそれを裏付ける。

1, 同じフレーズをpで弾かせるとfの時より遅く弾く。

2, 連続的に打鍵しながらpからfへクレッシェンドさせると次第に早くなる。

などのケースがそれにあたる。

初心者がppp等の特に弱い音を弾くときは時には「鳴らない」ことがある。これは鍵盤をそっと押さえれば弱い音が出ると信じる錯覚がそうさせているのであって、いかなる場合でもピアノはハンマーが弦を打たなければ音を発しないことをまだよく学習していないのである。

そこでゆっくり打鍵することでその効果を得ることが練習の結果得られる。しかし、ゆっくり打鍵することで曲のテンポそのものがゆっくりになってしまうのである。オルガンのような楽器ではこのコントロールは学習されない。音楽表現をベロシティで、言い換えればp fで表現するピアノと言う楽器の練習の難しさがここにある。

減衰音の心理的効果

パイプオルガンなどでは演奏される楽器（場所）によって同じ演奏者でも異なるレジストレーションで演奏する。世界に一台と同じ楽器が存在しないパイプオルガンでは当たり前のことであるが、何百個というストップを持つものから4個程度しかないものまでパイプオルガンでは絶対出せない音がある。（シアターオルガンは除く）

それはピアノやギターのような減衰音と呼ばれる音である。どんなに多くのストップを備えていても大きく分けてフリュー・パイプと呼ぶ笛の原理で鳴るものと、リード・パイプと呼ぶもの2種類しかなく、後は管の太さや材質、形状や位置等を変えてバリエーションを創っているに過ぎない。ハモンド・オルガンもドローバーの組み合わせで倍音構成（スペクトラム）を変えてあたかも多くのストップを持つかのような機能からスタートしたが、パーカッシブスイッチの採用によりコンデンサーの蓄電・放電の原理で減衰感のある音が出せるようになりパイプオルガンと決別した。

そして、アナログシンセサイザーではエンベロープジェネレータと呼ぶ概念でこの「持続音」「減衰音」のデザインを自由にできるようにしたのである。

オルガンは典型的な持続音による楽器である。それに対してピアノは典型的な減衰音の楽器である。明治や大正のハイカラさんの家庭にはピアノやオルガンがあった。ついこのあいだまで田舎の学校にはピアノの代用として「足踏みオルガン」（正確にはアメリカ・オルガンと呼び空気を吸い込んで鳴らし、空気を吹き出して鳴らすハーモニウムと区別する）が使われてい

た。

はたしてオルガンはピアノの代用になるのでしあろうか。ラベルのピアノ曲をオルガンで表現できるのであらうか。多分何かが失われることに気づくであらう。

この失われる何かとは「自己生成するイメージ」なのである。つまり、「打つ・はじく」などによるピアノやギターは聞き手にピッチとベロシティとタイミングだけを提供するので、聞き手はその減衰して消えつつある「実音」に代わって自分の頭の中で「虚音」を自己生成して新しい音イメージを創るのである。そのことが創造的な音楽的行為なのである。

それに対して「吹く」や「擦る」などの動作から生まれる持続音では聞き手に対して「強制的に追従する」ことを要求し聞き手の自由なイマジネーションによる自己生産的な行為を妨げるのである。雑念を許さない教会の典礼音楽にオルガンが用いられるのもその理由による。持続音はついてゆけるかついてゆけないかのどちらかの音楽でもある。

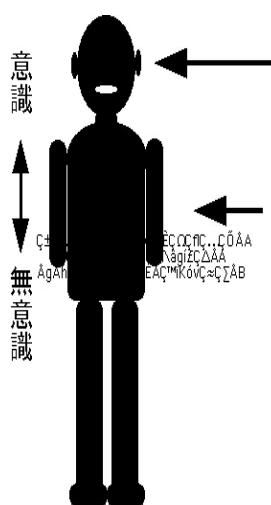
従って、このことからオルガンはピアノの代用にはならないし、白黒鍵盤の姿が似ているからという理由で同じような音楽表現ができると考えるのは大きな誤りである。

「ボン・ボン」という表現で代表されるピアノ音は典型的な減衰音である。時間軸に沿って次々と鳴らされるピアノの音は音域にもよるが長くても7〜8秒で聞こえないレベルにまで減衰する。しかし、我々の意識の中では鳴り続けているのである。この鳴り続ける意識の音のことを「内的聴覚」と呼び、音を音楽に変換するメカニズム（メタ認知とも言う）では不可欠なものである。ピアノはこの「内的聴覚」を誘導するあるいは誘発するという原理で音を音楽に変換しているのである。

ジャズ・ピアノなどではアクセントやシンコペーシ

ョンを表現するのに直接的な方法でリズム感を刺激する。この直接的刺激をインパクトと呼んだりする。MIDIデータではベロシティやデュレーションのデータだけでこれらの表現が可能であるが、受け取る側がそれを意識して捉えるのとはそうでないのではかなり違ってくる。

この直接的な刺激だけで音や音楽を聴く場合はいわば「意識の要らない」聞き方で、内的聴覚を使って意識しながら聴く場合とは



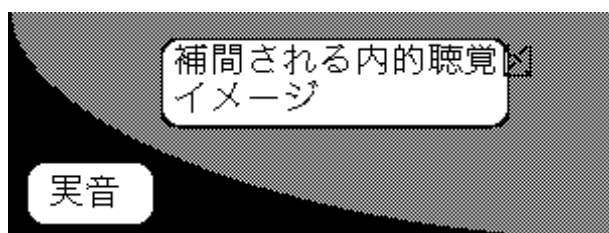
音や音楽に対する活動がかなり違うのである。

別の言い方をすれば「首から上」で聴く音楽と、「首から下」で聴く音楽があるとも言える。ピアノ曲などでは右手の旋律は「首から上」で認知されるが、左手の伴奏は「首から下」で感じられる。(図1)

一概には言えないが、伴奏は旋律のバックグラウンドであるからあまり意識を刺激しすぎても旋律刺激が弱くなる。

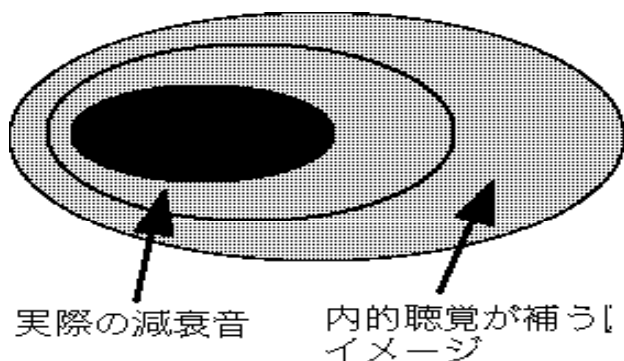
優れた作曲家や演奏家たちはこの「何を意識に」「何を無意識に」という配分をうまくやってきた。そもそも減衰系の音はインパクト即ち衝撃的刺激であるから鼓膜や聴覚を占有している時間は短い。従ってインパクトの密度や強さで意識の部分と無意識の部分の両方の世界に存在できるのである。

この直接的刺激の最たる物は「打楽器」であらう。持続音の出る打楽器はないので減衰時間に若干の差はあれすべては消えゆく音である。これはすべての楽器音の音色の違いの殆どはアタックの数ミリ秒の高調波成分で決定されるという事実に基づいているので、打楽器ほど叩いた時の音が多様な楽器も少ないのである。カナで云うと「ボン」とか「カン」とか麻雀用語のようであるが、語尾が「ン」になるという特徴がある。一般に「あ・え・い・お・う・」のような母音的なスペクトラムを持つ性質を「フォルマント」と呼び、例えばトランペットは「あ」、オーボエは「え」とかに近いが、打楽器ではそのフォルマントが識別できるほどの時間がなく、高調波が少なくなり減衰し始めた余韻の部分が「ン」と聞こえるのである。



この「ン」の部分で内的聴覚が働くのだが、図2の網掛けの部分が「ン」の部分である。勿論この図は概念図であるから要は「実音」の外側に「虚音」が存在するような聞き方もあるということを示しているに過ぎない。

インパクトとしての音でもお寺の鐘のように「ゴーン」と「ン」までの減衰時間が長い場合は余韻の中に様々な思いやイメージを展開するであらうが、鼓の音のように「ボン」と短いような場合は無音の沈黙を「間(ま)」と呼ぶイメージで補う。(図3)



このようにイメージで補う能力が音楽的能力と密接な関係があるので、音楽的能力の低い人たちにとってピアノの音楽はかなり難解なものと思われる。また、ロックやポップスの多くはパーカッションの助けで成り立っているが、音楽を難しいと感じさせないインパクトのあるリズムセクションに合わせて身体を動かすだけでその音楽の世界に入り込めるので、漫画世代の若者にはうってつけなわけである。現代音楽の多くがこの減衰音の「イメージ自己生成機能」に頼っていることは経験的に理解できる。

要するに内的聴覚は高等な精神活動であり、特に旋律線などの点を結んだ線としての認知は今日の発達したコンピュータでも不可能な世界なのである。

コンピュータで音を出す技術はシンセサイザーを頂点として随分発達してきた。元々発振（発音）の原理はアコースティックの世界では「打つ」「はじく」「吹く」「擦る」の四つしかないが、電子技術の場合無限に近いほどある。電子音の場合の発音動作はそのどれでもない「トリガー」とよぶパルススイッチとして開始される。そして、「ゲート・タイム」と呼ぶ時間の間ずっとこのトリガーはオンのまま保持される。つまりトリガーは持続音と同じ概念なのである。

1955年ハリー・オルソンによって発明されたシンセサイザーはこのトリガーによるゲート・タイムの間に、音量（ボリューム）や音色（スペクトラム）、様々な変調（モジュレーション）の連続的変化を行えるように設計されていた。このことがそれまでの電子オルガンとは決定的に違うことになった。従って、今日の電子オルガンはもはや電子オルガンではなくコンピュータ付きのシンセサイザーであると言った方が正しい。初期のシンセサイザーのコントローラは白黒鍵盤やフット・ボリューム、リボン・コントローラ、ジョイスティックのようなものが主流で、それらが発するアナログの直流電圧変化を音楽的表現に利用していたことはもうかなり忘れられてきている。アナログ信号そのものである直流電圧は極めてデリケートで部屋の温度にさえ左右されるというやっかいなものであった。

それに対してコンピュータのインターフェイスに用いられるデジタル信号は極めて安定がよく、確実であったため電子音源の制御に用いられるようになり、初期のアナログシンセサイザーを電圧制御型シンセサイザー（Voltage Controlled Synthesizer = VCS）と呼ぶのに対して、現在のシンセサイザーはデジタル制御型シンセサイザー（Digital Controlled Synthesizer = DCS）と呼ぶわけである。

もともとシンセサイザーには合成方式、抽出方式、変調方式、PCM方式、ハイブリッド方式、演算方式等があるが、最近ではDSP即ちデジタル・シグナル・プロセッサなるものが音源になる傾向が次第に強くなってきたのと、コントロール系にスライダーや回転つまみのようなアナログ系のものが使われるようになってきているので目や耳にはアナログ回帰のように映る。

このことは持続音から減衰音への連続的変化をスムーズにできるようになったことを意味する。つまり、現在のシンセサイザーではエンベロープデザインが固定されたものではなく極めて自由でフレキシブルになっているということを認識しなければならない。

本論では持続音と減衰音の間に「境界線」を引いてきたわけであるが、今日的な「打たない」「はじかない」「吹かない」「擦らない」電子音源ではこれらの動作が持つメッセージを伝えることができない。

「ドーン」と鳴る太鼓にはバチを振り下ろす動作がイメージとして付随するものであり、単に音響イメージだけが人間に伝わるものではない。作曲家や指揮者が頭に描く音のイメージは音楽からくるものか、音のイメージから音楽ができるのかは分からないが、バッハの最後の作品である「音楽の捧げもの」に楽器の指定がないことから考えれば初めに音楽的アイデアがある場合もあれば、ホルストの「惑星」のように一切の楽器変更が認められないような「音のイメージ」が絶対的な音楽イメージを支配している場合もある。バレイ等の舞踊にも音楽的イメージだけで躍るものと、音のイメージで躍る現代舞踊があるように、物理的というか生理的というか鼓膜をどう刺激したかが即イメージとなるものと、その音のスキーマとなるメタ認知を必要とする心理的聴覚による間接的イメージによるものがある。

シンセサイザーのエンベロープの概念では音の鳴り始めを「アタック (Attack)」と呼び、鳴り終わりの部分（正確にはトリガー信号がオフとなった瞬間或いはゲートタイムの終わった後）を「リリース (Release)」と呼ぶが、その (Attack) と (Release) に挟まれた部分を持続部分と考えると、その部分が減衰する方向へ変化するのが減衰音で、その部分が減衰しないのが持続音と定義することができる。

当たり前のようであるがこの定義は非常に大切で、シンセサイザーのような「打たない」「はじかない」「吹かない」「擦らない」電子音源ではあたかも「打った」ように聞こえる音や、そうでない音の間に存在するイメージを無段階連続的にデザインするパラメータがここに隠れているからである。それ故「音のイメージ」に関する限り「動作のイメージ」までも考慮していないのが普通で、純粋な「音のイメージ」だけによる音楽を考えているのが「電子音源」の基本的な姿勢であろう。しかもバッハのように純粋に音楽を捉える立場には意外なことにシンセサイザーなどの電子音源がピッタリなことにも気づく。

電子音源は粘土細工のように自由に編集できるので「プラスチック・サウンド」であるともいえる。あたかも何かの楽器の音のような形（エンベロープ）をつくることもそれほど難しいことではない。それでもリアルなサウンドにならないことが多い。その一番大きな理由はその楽器音の「顔」とでもいう特徴がうまく再現できないことがあげられる。

すべての音の典型的な特徴はその「鳴り始め（アタック）」の瞬間に識別される。50音の母音以外はすべて発音される直前に摩擦音や破裂音や鼻腔音や舌の振動などの母音以外の音が付加される。例えば「Sー」という摩擦音の後に「ア」という母音が来ると「さ」と聞こえるように、50音を発音するときには必ず1音声ごとに子音の口形とそれに続く母音の口形の二つの組み合わせが瞬間的に行われていることがわかる。母音は口腔や鼻腔の形状や容積を変化させて共鳴のレゾナンス（強調される周波数帯）を特徴づけることにより識別されるが、子音は母音の前に発せられる子音形成のための準備音を正確にキャッチしないと誤認識されてしまう。

さて、フルートの音を録音したテープのアタックとリリースの部分を削除したものと、ピアノの音を同様にカットしたものを聴き比べるとどちらがピアノでどちらがフルートなのかが分からない。ピアノとフルートの「母音成分（フォルマント）」は極めて似ているので、鳴り始めと終わりをカットした中間部分だけでは殆ど識別できないのである。

何故我々はピアノとフルートが識別出来るのかはこのカットされた部分と関係がある。

よく似たエンベロープと音域を持つホルンとユーフォニウムですら我々は耳だけで識別できる。強いて言うならホルンのアタック部分の発音は「P」か「PH」に近い「プォ」とか「フォ」に近いのに対してユーフォニウムの場合は「B」とか「W」に近い発音に聞こえることや、アタックの時間がわずかにホルンの方が長いことなどがその識別の拠り所となるが、さらに持続部分のフォルマントがホルンでは「O」や「U」

に近いのに対してユーフォニウムでは「A」や「U」に近いことも最終的な判断の根拠となるであろう。

このように色々な楽器音を特徴づけるいわば楽器音の「顔」はアタックの部分にかなり集中している。先行研究によれば人間の識別力はアタックの1000分の1秒程度まで可能だそうである。

例えば声楽の専門家はあるピッチから離れたピッチまで実際には「ポルタメント」のような動作で求めるピッチを探るのそうだが、探し求めるピッチを捕まえるまでの時間が2000分の1秒以内なので聞き手には一発でその音に到達したように聞こえるのだそうだ。

初期のムーグ・シンセサイザーのエンベロープ・ジェネレータでは一番早いアタックが500分の1秒程度であったがそれでも十分特徴あるアタックが表現できたものである。高い音域の音ほどこの識別に要する時間は短く、言い換えれば高い音ほど識別し易い傾向はあるが、およそ4千ヘルツをピークとする人間の耳の感受特性の関係でピアノの一番右端の鍵盤より高い周波数成分を含む部分は極めて感度が落ちるので、結果的にはいわゆる中音域が一番特徴を捉えやすいと言える。

シンセの音作りで一番注意が必要なのはこの中音域の仕上がりである。パイプオルガンのストップで一番バリエーションが多いのが8フィート系であるのも頷ける。

500ヘルツ以下の低音域では1000分の1秒という短い時間ではピッチ認知が不可能であるのでピッチを感じるまで少し時間（と云ってもほんの僅かであるが）がかかる。そのために低音楽器は発音が遅いと感じられることもあるし、低音域の振動は質量の大きなものが振動した結果であるから、高音域がスポーツカーなら重低音はダンブカーのように重いので振動の加速に時間がかかるのは当然であろう。

バイオリンのボウイングにアップやダウンさらにスピカート、ピッチカートなどがありそれらの特徴は殆どアタックの部分に集中していることがわかる。

管楽器においてもアンブッシャーやタンギングが演奏の決め手になっていることが多い。

ピアノに至ってはアタックがすべてで、一度ある強さで鳴ってしまった音は後から修正出来ない。

SMLの教育理論ではこのSの部分のアタックに注目する。鳴らす前にどんなアタックで鳴らすかをイメージさせることと、そのイメージ通りの音を出す運動の訓練と、結果的に出た音をフィードバックしてイメージと照合させることが必要であるからである。