

TR-A-0022

音の鋭さと振幅包絡

上田 和夫

赤木 正人

Kazuo UEDA

Masato AKAGI

1988. 3. 18

A T R 視聴覚機構研究所

## 目 次

1. まえがき	1
2. 心理実験 1	
2.1. 刺激音	2
2.2. 実験方法	3
2.3. 実験結果	3
2.3.1 類似性判断のMDS分析	
2.3.2 鋭さ判断のMDS分析	
2.3.3 明るさ判断のMDS分析	
2.4. 考察	6
3. 心理実験 2	
3.1. 実験方法	8
3.2. 実験結果と考察	8
4. 心理実験 3	
4.1. 実験方法	9
4.2. 実験結果と考察	9
5. むすび	10

Appendix 電子計算機による白色雑音合成法

# 音の鋭さと振幅包絡

Sharpness and amplitude envelopes of sound

上田 和夫  
Kazuo Ueda

赤木 正人  
Masato Akagi

A T R 視聴覚機構研究所  
ATR Auditory and Visual Perception Research Laboratories.

内容梗概： 200 ms 以下の広帯域雑音の振幅包絡の形、ピーク音圧レベル、持続時間の違いが、鋭さ、明るさに与える影響を調べることを目的として、一対比較法による心理実験を行い、多次元尺度法による分析を行った。さらに極限法、調整法により鋭さの主観的等価点を求める実験を行った。その結果、鋭さの判断に大きな個人差が存在すること、鋭さの判断の個人差は音の大きさと主観的長さの二つの心理次元の重みの違いとして説明できること、振幅包絡の形が音の鋭さに影響すること、実験で用いた刺激について明るさの判断を行うのは困難であること、がわかった。

## 1. まえがき

音声の個人性の認識、音韻認識において、非定常な音に対する聴覚の特性が大きな役割を果たしていると考えられる。さらに話者の感情、性別、年齢までを含めて音色の知覚という大きな枠組みの中でこれらの問題を扱うことができると考える。音色の知覚に影響を与える動的な物理パラメタとして振幅包絡の形の時間変化、スペクトル包絡の形の時間変化、持続時間等が考えられる。本年度はこれらの物理パラメタのうち、振幅包絡の形の時間変化、音圧、持続時間といった要因について心理実験を行い、物理量と心理的印象との対応関係を調べた。

さて、定常音の音色の属性については、音の鋭さがそのもっとも主要なものであるとされている(von Bismarck, 1974a)。そして Bismarck (1974b) は心理物理学的研究の結果、定常音の鋭さと物理パラメタは、大きさ-臨界帯域比パタンの重み付をした一次モーメントを用いて関係付けることができることを示した。これは定常音の鋭さがパワースペクトルに従属して決まるものであることを意味する。

一方、非定常音の鋭さについては難波、桑野、加藤(1974)の研究がある。彼らは、白色雑音に (a) 矩形波動的なもの(持続時間 5 ms から 500 ms)、(b) 持続時間と等しい立ち上がり時間を持つ(定常部を持たない)三角波動的なもの(持続時間 100 ms から 500 ms)、の二種類の振幅包絡を与えたものを刺激音として、semantic differential method (SD法)により、音の物理量と印象との関係について調べた。その結果、鋭さ(金属性)因子と大きさ(迫力)因子が得られ、実験に用いられた刺激音の印象がこの二つの因子で説明できることが示された。

また、刺激音のピーク音圧レベルが上昇する程、および持続時間が短縮される程、鋭さの印象が強まる、と結論している。ただし、この研究は本来、大きさの研究であるので、振幅包絡の形が鋭さにどの様に影響するかについては直接触れられていない。

ところで、我々が日常耳にする音声、楽器音等は言うまでもなく非定常なものである。また、楽器音の同定に際して、音の立ち上がり部分が知覚上重要であることが繰り返し指摘されている(Saldanha and Corso, 1964; Keeler, 1972; Grey, 1977; 山口, 1977; 安藤、山口, 1983)。楽器音の立ち上がり時間は通常数 ms ないし 200 ms 程度であるので、非定常音の音色知覚を調べるためには 200 ms 以下の立ち上がりを持つ音に対する聴覚の特性と音色の関係についての研究が必要である。

また、「鋭さ」の知覚メカニズムと「明るさ」(厨川、八尋、柏木、1978)のそれとは、音のスペクトル分布が影響する、という意味で共通性がある。しかし上田(1988)は、非定常音まで含めて考えた場合、音の「鋭さ」と「明るさ」には波形の振幅包絡の情報を必要とするかどうか、という点に知覚メカニズム上の違いがあるのではないかと考えている。即ち「鋭さ」の判断には波形の振幅包絡の情報を必要とするが、「明るさ」はそうではないのではないかと考えている。

そこで本研究では、振幅包絡の形、ピークの音圧レベル、持続時間が音の鋭さ、明るさに与える影響を調べることを目的として広帯域雑音を用いて一対比較法による心理実験を行い、多次元尺度法(MDS)による分析を行った。さらに定量的なデータを得るために、極限法、調整法による補足実験も行った。

## 2. 心理実験 1

### 2.1. 刺激音

Fig. 1 に示すような(A)矩形波状、(B)後方に頂点を持つ三角波状、(C)前方に頂点を持つ三角波状、の3種の振幅包絡を持つ広帯域雑音(200 Hz - 20 kHz)を刺激音として用いた。ピークの音圧レベル(正確には、刺激音の振幅包絡のピークが持続したときの音圧レベル)は 65 dB SPL, 85 dB SPL の2種類、持続時間は 5 ms, 10 ms, 20 ms, 40 ms, 70 ms, 100 ms, 200 ms の7種類とした。200 Hz 以下の成分を含まない広帯域雑音を用いたのは、持続時間の変化にともなう低域のスペクトルの変化を避けるためである。なお、B型とC型の刺激音については、同一音圧レベルの刺激音群内では、持続時間が短くなるほど立ち上がりあるいは立ち下がりの傾斜がきつくなる。

刺激音の総数は 42 個である。各刺激音を他の刺激音と組み合わせて対を作り、刺激対をランダムに並べて刺激系列を作成した。対の総数は 1722 対である。対にされた刺激音間には 500 ms、二つの対の間には約 2.5 sec の空白時間を設けた。刺激音は電子計算機(DEC/VAX 8600/ULTRIX, MASSCOMP/MC5600/UNIX System V)により、48 kHz サンプリング、16 bit で合成された。合成時には白色雑音として合成、そのまま録音し、被験者に提示する際に 200 Hz, 96 dB / Oct. のローカットフィルター(NF FV-665)を介して提示した。DEC/VAX 8600 上で合成された1秒分の白色雑音(具体的な合成法については Appendix を参照)のデータを MASSCOMP/MC5600 に転送し、振幅包絡の形の窓を掛けて白色雑音を切り出し、刺激音のデータを作成した。切り出し位置は刺激音毎にランダムに変えた。ただし、

後で切り出し開始点を固定して作った刺激音と聞き較べてみても差は判らなかつた。作成された刺激音のデータを DA 変換し、20 kHz, 96 dB / Oct. のローパスフィルター(NF FV-665)を通して刺激音とした。

## 2.2. 実験方法

刺激音は全て DAT (SONY DTC-1000 ES)に録音され、200 Hz, 96 dB / Oct. のローカットフィルター(NF FV-665)を通して、ヘッドフォン(STAX SR  $\Lambda$  Pro)により、被験者の両耳に提示された。刺激音の波形がヘッドフォンの出力として正確に再生されることは確認済みである。使用したヘッドフォンの周波数特性については上田(1987)に詳細に述べられている。音圧レベルは人工耳(B & K, 4153)にヘッドフォンを置き、精密騒音計(B & K, 2231)により測定した。

実験は(a)類似性、(b)鋭さ、(c)明るさ、を判断する3セクションに分かれ、被験者は類似性のみ7段階評定を行い、他は各対のうちどちらがより鋭いか、或は明るいかを判断する。被験者は20才台の大学生、大学院生13名である。鋭さ、明るさだけでなく、類似性の判断も含めたのは刺激音の全体的な印象がどの様に捉えられているかを知るためである。SD法による研究では測定時の形容詞対の選び方如何によっては重要な心理的因子を抽出できない可能性があるが、類似性判断のデータを多次元尺度法で分析することにより、このような危険を避けることができる。

## 2.3. 実験結果

実験結果は次のようにして整理した。まず、類似性の判断結果を被験者毎に42 × 42のマトリックスの形に整理し、対角線について対称な要素を平均した。これから、次式により定義されるプロフィール非類似性行列(柳井、高根、1977)を求めた。

$$d_{ij} = \left( \sum_{\substack{k \neq i \\ k \neq j}}^n (o_{ik} - o_{jk})^2 \right)^{1/2} \quad (1)$$

ここで  $d_{ij}$  は刺激  $i$  と  $j$  との心理的距離(非類似性)、 $o_{ij}$  は類似性行列の( $i, j$ )要素である。

鋭さと明るさの判断結果については、次のように整理した。まず各対のより鋭い、或は明るいとされた側の刺激に1点を与え、被験者毎に星取り表を作る。これは42 × 42のマトリックスの形をとる。得られたマトリックスについて(1)式によりプロフィール非類似性行列を求める。それぞれのプロフィール非類似性行列を入力行列として、MDS分析を行った。使用したプログラムはALSCALである。

### 2.3.1. 類似性判断のMDS分析

各被験者の類似性判断の結果をMDS分析したところ全ての布置にほぼ共通の構造がみられたので、全被験者の非類似性行列を合計したものを再び分析した。合計する際には各被験者の距離データの総和が同じになるように規準化した上で合計した。2次元解を採用した(ストレス、0.049)。Fig. 2に得られた布置を示す。この図でアルファベットはFig. 1の振幅包絡の型と対応し、アルファベ

ットの小文字は 65 dB SPL のピーク音圧レベルに、大文字は 85 dB SPL のピーク音圧レベルに、それぞれ対応している。数字は持続時間の短いものから順に割り当てられた番号である。図中の線は同じ振幅包絡の型、同じピーク音圧レベルの刺激を結んでいる。(以下、同様)。

この図から刺激音がピーク音圧レベルによって大きく二分され、各群内で持続時間の順に並んでいることがわかる。刺激音の振幅包絡による区別は—特に 65 dB SPL の刺激音群では—あまり明確ではない。

ここで刺激音の音圧を時間軸について平均した平均パワー ( $P_a$  と略す) 及びトータルパワー ( $P_t$  と略す) を次のようにして求めた。なお、ここで言うトータルパワーは難波ら(1974)の言う実効値音圧レベルと同じものである。

$$P_a = P + 10 \log_{10} S \quad (2)$$

$$P_t = P + 10 \log_{10} DS \quad (3)$$

ここで  $P$  はピーク音圧レベル、 $D$  は持続時間、 $S$  は A 型の刺激音については 1、B、C 型の刺激音については  $1/3$  の値を取る。

$P_a$ 、 $P_t$ 、 $D$  について Fig. 2 の布置に対する重回帰分析を行った。 $R^2$  は 0.8483 ~ 0.8728、得られた回帰係数は 0.0001 以下の水準で有意であった。 $P_a$ 、 $P_t$ 、 $D$  に対応するベクトルを図中に示す。これらのベクトルと 1 軸とのなす角度はそれぞれ  $85^\circ$ 、 $65^\circ$ 、 $38^\circ$  である。物理量としては  $P_a$  と  $D$  は独立したもので、 $P_t$  は  $P_a$  と  $D$  から得られる物理量である。しかしこの類似性の心理空間については  $P_t$  と  $D$  がほぼ直交している ( $103^\circ$ ) ので、この二つの次元でこの心理空間は説明できる。

### 2.3.2. 鋭さ判断の MDS 分析

各被験者について、個別に MDS 分析を行ったところ、13 名中 2 名の被験者については判断のばらつきのために適当な解が求められなかったため、この 2 名のデータは考慮外とした。各被験者の星取り表の各行の要素を横に合計して、各刺激が他の刺激よりも鋭いとされた割合を求めて、刺激布置の解釈をするときの参考にした。残り 11 名のデータの分析結果は次のようにまとめることができる。

1) 持続時間と振幅包絡の形が等しく、ピーク音圧レベルの異なる刺激については、ピーク音圧レベルの高い方をより鋭いとする傾向が一般にみられた。

2) 同じピーク音圧レベルの刺激群内では、持続時間の長いものほど鋭いとした被験者が大部分 (9 名) であった。例として、被験者 4 の刺激布置を Fig. 3 に示す (2 次元解、ストレス 0.076)。1 軸の正方向が鋭いとされた方向である。2 軸にははっきりした意味はない。この図でピーク音圧レベルが 85 dB の刺激音群について、同じ持続時間の B 型と C 型の刺激音を比べてどちらの型の刺激音がより鋭いとされているかを見ると、5 ms の刺激 (B1, C1) を除いて C 型の刺激音が B 型の刺激音より鋭いとされていることがわかる。65 dB の刺激音群については全体に小さく固まって布置され、B 型と C 型の差は明確ではない。

3) 同じピーク音圧レベルの刺激群内で、持続時間の短いものほど鋭いとした被験者は 2 名であった。例として、被験者 1 の刺激布置を Fig. 4 に示す (2 次元解、ストレス 0.099)。鋭いとされた方向は 1 軸の正方向である。これも 2 軸には意味はない。この図では B 型の刺激音と C 型の刺激音の鋭さの違いは明確では

ない。他の1名の被験者の刺激布置は、刺激音のピーク音圧レベルにより二つのクラスターが形成された。

各被験者の分析結果の1軸の座標値を  $P_t$  と  $D$  に重回帰させたところ、全ての被験者について回帰可能であった ( $R^2$ : 0.8009~0.9656)。大多数の被験者については 0.0001 以下の水準で有意な回帰係数が得られた。ただし、1名の被験者のデータについては  $D$  が0でないという帰無仮説が棄却できず (有意水準 0.6338)、また2名の被験者のデータについては  $D$  の回帰係数の有意水準が他の被験者よりも若干高かった (それぞれ 0.0541, 0.0036)。Fig. 3 については  $P_t$  の回帰係数が 0.1225、 $D$  の回帰係数が -0.0052 となった。Fig. 4 については  $P_t$  の回帰係数が 0.0724、 $D$  の回帰係数が -0.0203 となった。

ピーク音圧レベルと持続時間が等しいB型とC型 (Fig. 1) の刺激音を比べてどちらの型の刺激音がより鋭いとされたかを全被験者について調べたところ、7名の被験者に全体にC型の刺激音をB型の刺激音より鋭いとする傾向がみられた。残りの4名の被験者にははっきりとした傾向が見られなかった。

### 2.3.3. 明るさ判断のMDS分析

被験者別に個別の分析を試みたところ、5名のデータについては判断のばらつきのために適当な解を求めることができなかった。残りの8名のデータの分析結果から、85 dB SPL の刺激音については、その反応から被験者を大きく次の二つに分類することができる。

- a) 持続時間の長い刺激音ほど明るいとする被験者 (6名)。
- b) 持続時間の短い刺激音ほど明るいとする被験者 (2名)。

a) の被験者については、65 dB SPL の刺激についても同様に持続時間の長い刺激音ほど明るいとする反応を示した。さらに、a) の被験者のうち、65 dB SPL の刺激音と 85 dB SPL の刺激音を比べて、ピーク音圧レベルの高い刺激音ほど明るいとした被験者は5名で、あと1名の反応にはピーク音圧レベルによる差が見られなかった。

b) の被験者のうち、1名は 65 dB SPL の刺激音に対しては 85 dB SPL の刺激音とは逆に、持続時間の長い刺激音ほど明るいと反応し、他の1名は 65 dB SPL の刺激音については持続時間による反応の差が明確ではなかった。

a) の被験者の刺激布置の例として Fig. 5 を、b) の被験者の刺激布置の例として Fig. 6 を示す。いずれも1軸の正方向が明るいとされた方向で、2軸には特に意味はない。

各被験者の分析結果の1軸の座標値を  $P_t$  と  $D$  に重回帰させたところ、3名の被験者については  $R^2$  が小さく (それぞれ 0.6485, 0.6230, 0.2355)、重回帰が不可能であった。残りの5名については重回帰可能であった ( $R^2$ : 0.8624 ~ 0.9525)。ただし2名の被験者については  $D$  の回帰係数が0でないという帰無仮説が棄却できなかった (有意水準はそれぞれ 0.3225, 0.2666)。Fig. 5 については  $P_t$  の回帰係数が 0.1218、 $D$  の回帰係数が -0.0029 となり、ほぼ  $P_t$  のみで明るさの判断が説明可能である。Fig. 6 は重回帰不可能であった。

Fig. 6 の被験者は 65 dB SPL の刺激音についてはC型よりもB型の方が、85 dB SPL の刺激音についてはB型よりもC型の方が明るいとしている。他にどちらのピーク音圧レベルについてもC型の方がB型よりも明るいとした被験者が1名あった。その他の6名の被験者にはいずれもはっきりとした傾向が見いだせなかった。

## 2.4. 考察

類似性の判断については全ての被験者についてほぼ一致した傾向が見られたのに対し、鋭さの判断については大きな個人差がみられ、被験者によって相反するような結果が得られたことは注目に値する。明るさの判断については判断そのものが困難であったと考えられ、半数近くの被験者が一貫した基準のもとに判断を行うことができなかった。

類似性判断のデータによる刺激布置に対して、刺激音の平均パワー ( $P_a$ )、トータルパワー ( $P_t$ )、持続時間 ( $D$ ) を従属変数として重回帰分析を行った結果、 $P_t$  と  $D$  のベクトルはほぼ直交することがわかった ( $103^\circ$ )。  $P_t$  と  $D$  の組合せは  $P_a$  と  $D$  あるいは  $P_a$  と  $P_t$  の組合せよりも直交に近く、Fig. 2 の心理空間をよく説明できると考えられる。ここで  $P_t$  が音の大きさとよい対応を示すこと (難波ら、1974)、  $D$  のベクトルの方向は主観的な音の長さの判断とほぼ同じ方向であると考えられることから、この刺激布置は音の大きさと音の長さの二つの次元で表現できる、と言える。  $P_t$  と  $D$  は完全に直交しているわけではないが、これは  $D$  と主観的長さの次元との関係を次のように考えると、音の大きさの次元と音の主観的長さの次元とは  $P_t$  と  $D$  よりも直交に近いと考えられる。すなわち、Fastl (1988) によれば、急峻な立ち下がりを持つマスキング音の後ろ側に極めて非線形な減衰曲線を持つマスキング効果が約 100 ms の間みられる。そしてその減衰の様子はマスキング音の持続時間によって変化する。このマスキング効果は音響心理学的な興奮パターンを反映している。筆者はこの刺激がなくなっからの減衰部分を持つ興奮パターンが主観的長さの知覚に影響すると考える。つまり、急峻な立ち下がりを持つ刺激音 (本研究では A, C 型の刺激音) の主観的な長さは刺激音の物理的な長さよりも長く知覚されるものと考えられる。 Fig. 2 の類似性判断の平面で考えれば、主観的な長さの次元は  $D$  よりも大きさの次元に対してより直交に近くなると考える。

刺激音の振幅包絡の形の違いがあまり明瞭に見られなかったのは、類似性を判断する際、被験者にとって音の大きさと音の長さの次元の占めるウェイトが他の次元よりも相対的に大きかったためであると考えられる。

鋭さの判断のデータにより得られた各被験者の刺激布置について、その 1 軸の座標値に  $P_t$  と  $D$  を独立変数とする重回帰分析を行なった結果から次のようなことがいえる。いずれの被験者のデータも  $P_t$  と  $D$  による重回帰が可能であったことは、これらのデータが  $P_t$  と  $D$  に適当な重みを掛けて加え合わすことで表現できることを意味する。  $D$  の回帰係数が有意でなかった被験者のデータについては  $D$  に 0 の重みを掛けると考えればよい。  $P_t$  と  $D$  はそれぞれ音の大きさ、音の長さの判断方向に対応するので、大多数の被験者については、鋭さは大きさと音の長さが入り交じった形で判断されていたといえる。そして鋭さの判断の個人差は音の大きさと音の長さに掛けられる重み (0 を含めて) の違いとして説明することができる。

ここで、鋭さの判断において、 $D$  に重みがなく、 $P_t$  だけで回帰される反応を示した被験者について、この被験者は鋭さではなく大きさの判断をしていたのではないかという疑問が生じるかも知れない。確かにこの被験者がトータルパワーを手がかりとして判断したということはいえる。しかし、本実験で用いた刺激についてトータルパワーのベクトルの方向と鋭さの判断から得られた次元の方向が一



致したからといって、被験者が鋭さの判断でなしに大きさの判断をしたのだとは必ずしもいえない。なぜなら、 $P_t$  はあくまでも物理量であるので、この被験者の大きさの次元が  $P_t$  の方向とびったり一致していたという保証はない。言い替えればこの被験者の大きさの次元が  $P_t$  とはある（小さな）角度をなしていた可能性がある。また、たとえ  $P_t$  と大きさの次元の方向が一致していたとしても、今回用いた刺激よりも広い範囲の物理的な変化を含む刺激に対しては大きさの次元と鋭さの次元の方向が異なる可能性もある。

さて、類似性判断の平面も  $P_t$  と  $D$  の 2 次元で説明できるので、鋭さの次元を類似性の平面に重回帰することが可能である。各被験者の鋭さの次元（鋭さ判断の平面の 1 軸の座標値）を類似性の平面（Fig. 2）に重回帰したところ、鋭さのベクトルは類似性の平面の 1 軸に対して  $65^\circ$  から  $107^\circ$  の間に分布した。

Fig. 3 の鋭さのベクトルは類似性の平面の 1 軸に対して  $82^\circ$  の角度をなした。

Fig. 4 の鋭さのベクトルは類似性の平面の 1 軸に対して  $107^\circ$  の角度をなした。

$65^\circ$  の鋭さのベクトルは  $P_t$  と全く同じ方向であり（持続時間が長いほど鋭い）、 $107^\circ$  の鋭さのベクトルは 2 軸の正方向よりも少し左を向いている（持続時間が短いほど鋭い）。すなわち、大きさに重きを置いて判断すれば長い音が鋭くなり、長さに重きを置いて判断すれば短い音が鋭くなることがわかる。

本研究により、今回用いたような刺激に対しては、鋭さに大きな影響を与える物理量としてトータルパワーと持続時間をとることが適当であることが示された。また一般に、トータルパワーのみについては大きな方が鋭く聞こえ、持続時間のみについては短い方が鋭く聞こえることが示された。ただし、鋭さの判断には個人差があり（Fig. 3, 4）、大きさ（トータルパワー）に重きを置くか、長さ（持続時間）に重きを置くかによって、表面上、持続時間の短い刺激ほど鋭いとするか、持続時間の長い刺激ほど鋭いとするか、という個人差が生じる。そして、この個人差は大きさと長さへの重みの違いによって説明できることが明らかになった。

明るさの判断のデータによる刺激布置のうち、Fig. 5 については  $P_t$  のみでは説明がつく。つまりトータルパワーを主な手がかりとして明るさの判断が行なわれたといえる。一方、Fig. 6 は  $P_t$  と  $D$  の重回帰では説明できなかった。そして 65 dB SPL の刺激音と 85 dB SPL の刺激音とでは持続時間の増加による鋭さの変化方向が正反対になっており、ある音圧レベルを境にして持続時間の影響の仕方が反対になるのではないかと考えられる。

明るさの判断については、安定した判断を行うことが困難な被験者が約半数を占めた。この原因の一つとして、雑音の「明るさ」を判断する、ということが被験者にとって非日常的な経験であったために判断の枠組みが安定しなかった、ということが考えられる。しかし、多数の刺激対について判断したにもかかわらず判断の枠組みが安定しなかったとすると、これはむしろ、これらの被験者にとって、雑音について明るさを判断することに無理があったのではないかと考えた方が自然である。別の言い方をすれば、彼らは非定常音についても周波数軸上のエネルギー分布によって音の明るさを判断しようとしたのだと考える。

一方では明るさについても安定した判断を示した被験者もいた。これらの被験者について、明るさの判断結果と鋭さの判断結果を見比べてみたが、それらがほとんど同じ被験者は 2 名のみであった。つまり、大部分の被験者はなんらかの形で明るさと鋭さの区別はついていたものと考えられる。

B と C の型の振幅包絡を持つ刺激音については、ピーク音圧レベルと持続時間

が等しければトータルパワーも等しくなる。従って、もし同じピーク音圧レベルと持続時間を持つ B, C 型の刺激音の間で鋭さなり、明るさなりに違いがみられたとすれば、それは振幅包絡の形の違いに対する聴覚の非線形な応答の結果である可能性がある。

Fig. 3 では 85 dB SPL の刺激音群について C 型の刺激音が B 型の刺激音より鋭いとされる傾向が見られた。ただし、持続時間が短くなるに従って、両者の鋭さの差が小さくなり、5 ms の刺激音では B 型の刺激音の方が C 型の刺激音より鋭いとされている。これは持続時間が短くなるにしたがって B 型の刺激音の立ち上がりの傾斜がきつくなり、C 型の刺激音の立ち上がりとの差が小さくなっていったことが一つの原因として考えられる。5 ms の刺激音で鋭さが逆転していることは、C 型の刺激音の立ち上がり部分が後続部に対して順行マスクングの効果を持ち、音の大きさが減少した可能性が考えられる。65 dB SPL の刺激の布置が全体的に不明瞭であるのは、85 dB SPL の刺激との音圧差が大きいため (20 dB) この音圧差に着目して被験者が判断したためであると考えられる。

鋭さについては、7 名の被験者に C 型の刺激音を B 型の刺激音より鋭いとする傾向が見られた。残りの 4 名の被験者にははっきりした傾向が見られなかった。明るさについては、B 型と C 型の間で何らかの違いがみられたのは 2 名の被験者だけで、その他の被験者にはいずれもはっきりとした傾向が見いだせなかった。

そこで、鋭さに振幅包絡の形が影響するのかどうかを見極め、その影響をより定量的に調べるために極限法による実験を行った。

### 3. 心理実験 2

#### 3.1. 実験方法

この実験では極限法 (完全上下法) を用い、刺激音の持続時間は 70 ms 一定とした。ピーク音圧レベル 85 dB SPL の B 型、及び C 型の刺激音を標準刺激とし、比較刺激として A 型の刺激音の音圧レベルを 0.5 dB ステップで変化させたものを用いて鋭さの主観的等価点 (point of subjective equality, PSE と略す) を測定した。標準刺激と比較刺激の間隔は約 1 秒、試行間隔は約 4 秒とした。上昇系列、下降系列それぞれについて 12 回ずつ判断を行なった。被験者は 20 代から 30 代の男性 4 名である。実験装置は実験 1 と同様である。

#### 3.2. 実験結果と考察

被験者によって一つの刺激系列内で上限閾、下限閾の両方が求められなかった例が出たため、そのような被験者については該当する系列のデータを取り除いた。上昇系列と下降系列の数のバランスをとるため、取り除かれた系列に隣り合う変化方向が反対向きの系列のデータも除去した。こうしたところ最低の判断回数が上昇、下降それぞれ 10 回ずつとなったため、他の被験者のデータも最初の数試行分をカットしてデータ数のバランスをとった。上限閾と下限閾の平均をとって PSE とした。B 型の刺激音の PSE は平均 79.1 dB SPL、C 型の刺激音の PSE は平均 81.1 dB SPL であった。B 型と C 型の PSE の差について片側の t 検定を行ったところ、その差は 0.0004 以下の水準で有意であった。また、標準刺激とトータルパワーの等しい比較刺激の音圧レベルは 80.2 dB SPL である。この音圧レベルと PSE の差についても同様に検定を行なったところ、B 型との差は 0.003 以下の水準で有意、C 型との差は 0.02 以下の水準で有意であった。以上により、

少なくともこの実験に参加した被験者の判断結果によれば、振幅包絡の形の違いが鋭さの印象に影響し、持続時間とトータルパワーが等しければ、C型がもっとも鋭く、ついでA型、B型の順に鋭く聞こえることがわかった。

そこで振幅包絡のピークの位置が鋭さにどのように影響するのかをより詳しく調べるために、次の実験を行なった。

#### 4. 心理実験 3

##### 4.1. 実験方法

調整法を用いた。標準刺激として、Fig. 7 に示す 7 種類の振幅包絡を持つ広帯域雑音を用いた。ピークの音圧レベルはいずれも 85 dB SPL とした。持続時間はいずれも 70 ms である。比較刺激として、持続時間 70 ms の A 型の (Fig. 1) 振幅包絡を持つ広帯域雑音を用いた。刺激音の周波数帯域は実験 1, 2 と同様である。被験者は比較刺激の鋭さ (音圧) を標準刺激と等しい鋭さになるように調節する。刺激間の間隔は 500 ms とした。一回スイッチを押すことで比較刺激の音圧が 0.5 dB ステップで変化するようにした。比較刺激の音圧の初期値は 74.0 dB SPL, 76.7 dB SPL, 79.3 dB SPL, 82.0 dB SPL の 4 種類として標準刺激と組合せ、ランダムに提示した。各組合せについて 16 ~ 18 回判断を行い、後ろから 12 回分の判断結果をデータとして採用した。被験者は男女各 1 名である。実験の装置系はプログラマブルアッテネータ (多摩川電子、TRA-308A) を用いて被験者の手元のスイッチから音圧の調節、刺激対の提示開始を行えるようにし、DAT を使わずにリアルタイムで実験を行なった以外は実験 1, 2 と同様である。

##### 4.2. 実験結果と考察

得られた PSE の平均と標準偏差を被験者別に Fig. 8, Fig. 9 に示す。分散分析を行なったところ、標準刺激の種類、被験者の要因が共に有意であった。各被験者のデータについて多重比較法による分析を行なった結果を Table 1, 2 に示す。Table 1 が Fig. 8 のデータについて、Table 2 が Fig. 9 のデータについての分析結果である。これより、Fig. 8 についてはピークが先頭にある刺激音と、ピークが後ろにある刺激音との間に鋭さの有意な差があることがわかる。またピークが 20 ms の位置にある刺激の鋭さが落ち込んでいるのが特徴的である。Fig. 9 についてはピークの位置が 0 ms から 35 ms までと 50 ms から 70 ms までの二つに大きく分かれることがわかる。またピークが先頭にあるものと末尾にあるものとの鋭さに有意な違いがないことがわかる。実験 1 で B 型と C 型の刺激音に鋭さの違いが見られなかった被験者は Fig. 9 の被験者と同じタイプの判断様式を持っていた可能性もある。いずれにせよ、振幅包絡のピークの位置により、鋭さに変化が生じたことは事実である。また振幅包絡の形の違いが鋭さに及ぼす影響にも個人差が見られたことになる。

Fig. 8 と Fig. 9 を比べると PSE の変化の仕方の違いのほかに、PSE の値が全体に Fig. 9 の方が高いことがわかる。標準刺激と比較刺激は、比較刺激が 80.2 dB SPL の時にトータルパワーが等しくなるので、Fig. 8 は全体にそれよりも低め、Fig. 9 は全体にそれよりも高めになる。さらに実験 2 で得られたデータと比べてみても同様のことがいえる。これが何を意味するのかは今のところ不明である。

Fig. 8に見られる PSE の変化の様子は、もし被験者の鋭さの判断規準が標準刺激が変化しても変わらなかったとすれば、順行マスキングと逆行マスキングのトレードオフの結果であるかも知れない。一方 Fig. 9 では刺激の前半部にピークがあるものと刺激の後半部にピークがあるものとの間にカテゴリーカルな境界が存在する可能性がある。これら二人の被験者の結果の差が何を意味するのかは今後の課題として残されている。今のところ考えられる原因として、この実験では標準刺激と比較刺激の提示順序が常に一定（標準刺激が比較刺激に先行する）であったため、どちらか一方の被験者のデータに系列効果が現われた可能性がある。この点を確かめるためにはさらに実験を積み重ねることが必要である。

## 5. むすび

持続時間 200 ms 以下の広帯域雑音の振幅包絡の形、音圧、持続時間と、鋭さ、明るさとの関係を調べることを目的として、一対比較法、極限法、及び調整法による心理実験を行い、鋭さの判断に大きな個人差が存在すること、鋭さの判断の個人差が音の大きさと主観的長さの二つの心理次元の重みの違いとして説明できること、今回用いた刺激について明るさの判断を行なうことは困難であること、また、振幅包絡の形が音の鋭さに影響することが明らかになった。

謝辞 ご協力頂いた被験者の方々、御討論頂いた Dr. Hugo Fastl, Dr. Victor W. Zue, Dr. Stephanie Seneff, Mr. Jim Glass, Dr. Roy D. Patterson、ご支援頂いた平原主任研究員に対し、厚くお礼申し上げます。

## 文献

- 安藤繁雄、山口公典 (1983) 「三味線音の音響的性質について」日本音響学会誌、39, 433-443.
- von Bismarck, G. (1974a) "Timbre of steady sounds: a factorial investigation of its verbal attributes," *Acustica*, 30, 146-159.
- von Bismarck, G. (1974b) "Sharpness as an attribute of the timbre of steady sounds," *Acustica*, 30, 159-172.
- Fastl, H. (1988) "Noise measurement procedures simulating our hearing system," *J. Acoust. Soc. Jpn.(E)*, 9, 75-80.
- Grey, J.M. (1977) "Multidimensional perceptual scaling of musical timbres," *JASA*, 61, 1270-1277.
- Keeler, J.S. (1972) "The attack transients of some organ pipes," *IEEE, AU-20*, 378-391.
- 厨川守、八尋博司、柏木成豪 (1978) 「音の7属性の性格について」日本音響学会誌、34, 501-509.
- 難波精一郎、桑野園子、加藤徹 (1974) 「音の立ち上がり時間と大きさについて — エネルギー値との関係 —」日本音響学会誌、30, 144-150.
- Saldanha, E.L., Corso, J.F. (1964) "Timbre cues and the identification of musical instruments," *JASA*, 36, 2021-2026.

- 上田和夫 (1987) 「STAX SR  $\Lambda$  Pro の周波数特性--人工耳及びダミーヘッドによる測定」 ATR テクニカルレポート, TR-A-0016.
- 上田和夫 (1988) 「音色の表現語に階層構造は存在するか」日本音響学会誌, 44, 102-107.
- 山口公典、安藤繁雄 (1977) 「短時間スペクトル分析法の自然楽器音への適用」日本音響学会誌, 33, 291-300.
- 柳井晴夫、高根芳雄 (1985) 「新版 多変量解析法」(朝倉書店、東京) pp. 142-143.

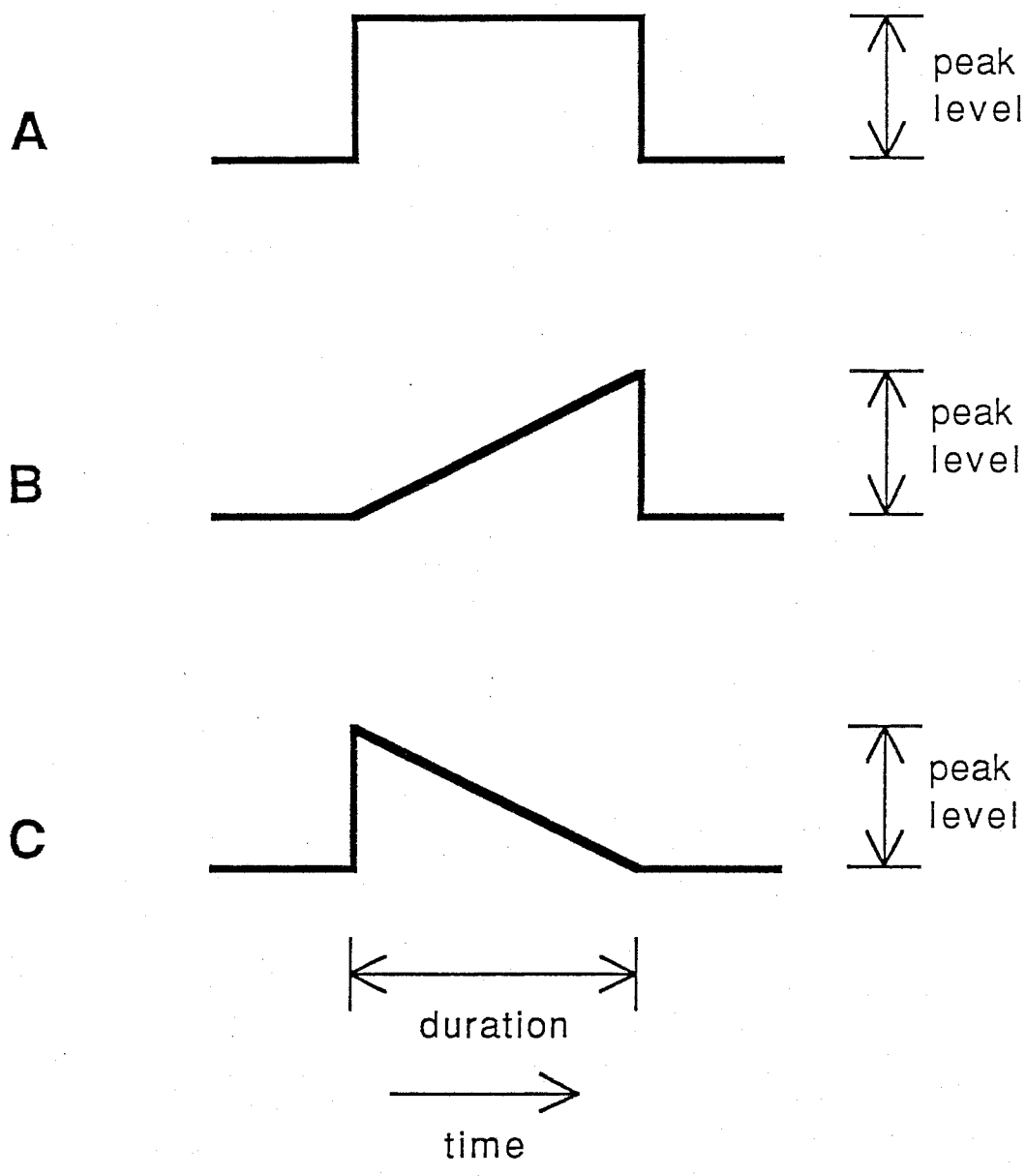


Fig. 1 Amplitude envelopes of the stimuli

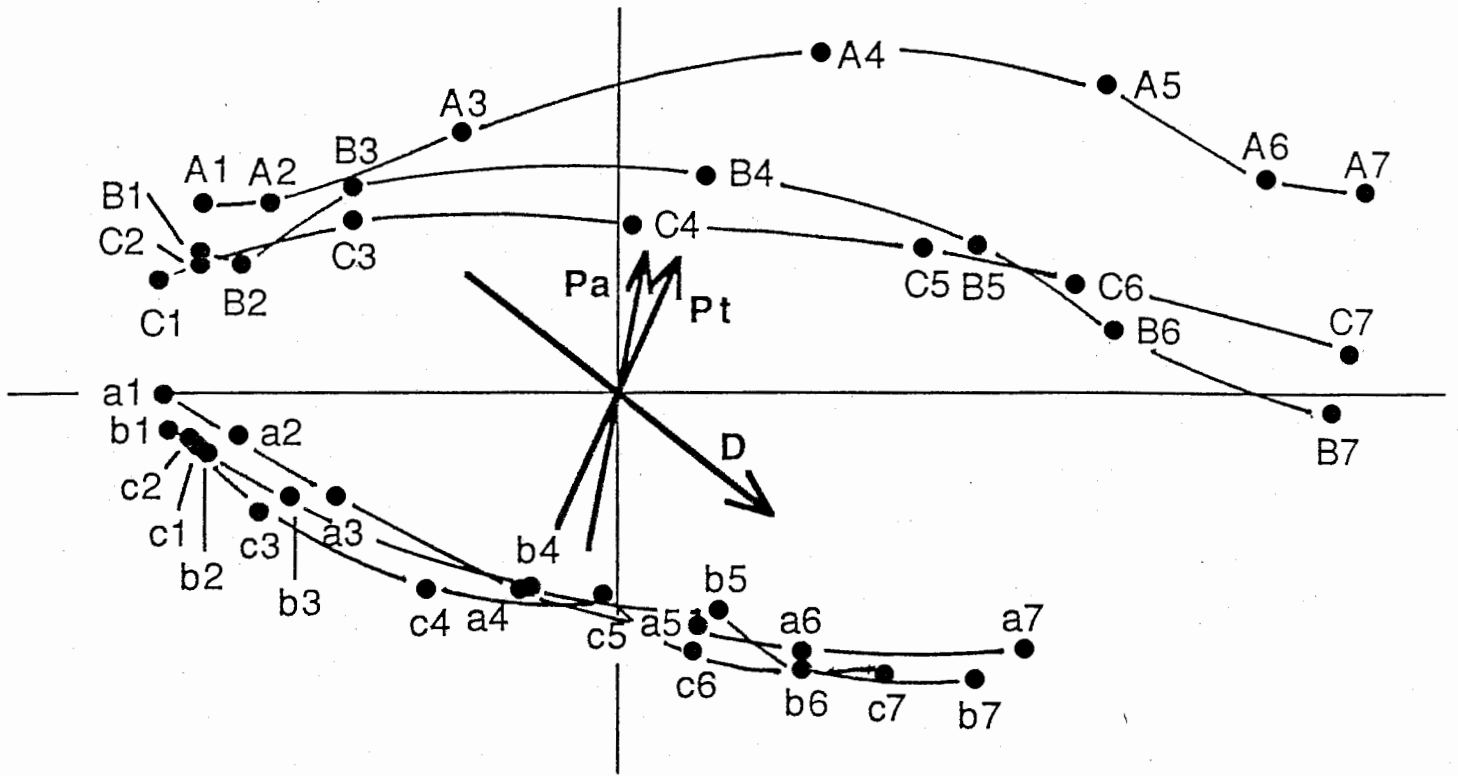


Fig. 2

Configuration derived in 2 dimensions from similarity data. Dimension 1 (horizontal) vs. dimension 2 (vertical). Stress = 0.049.

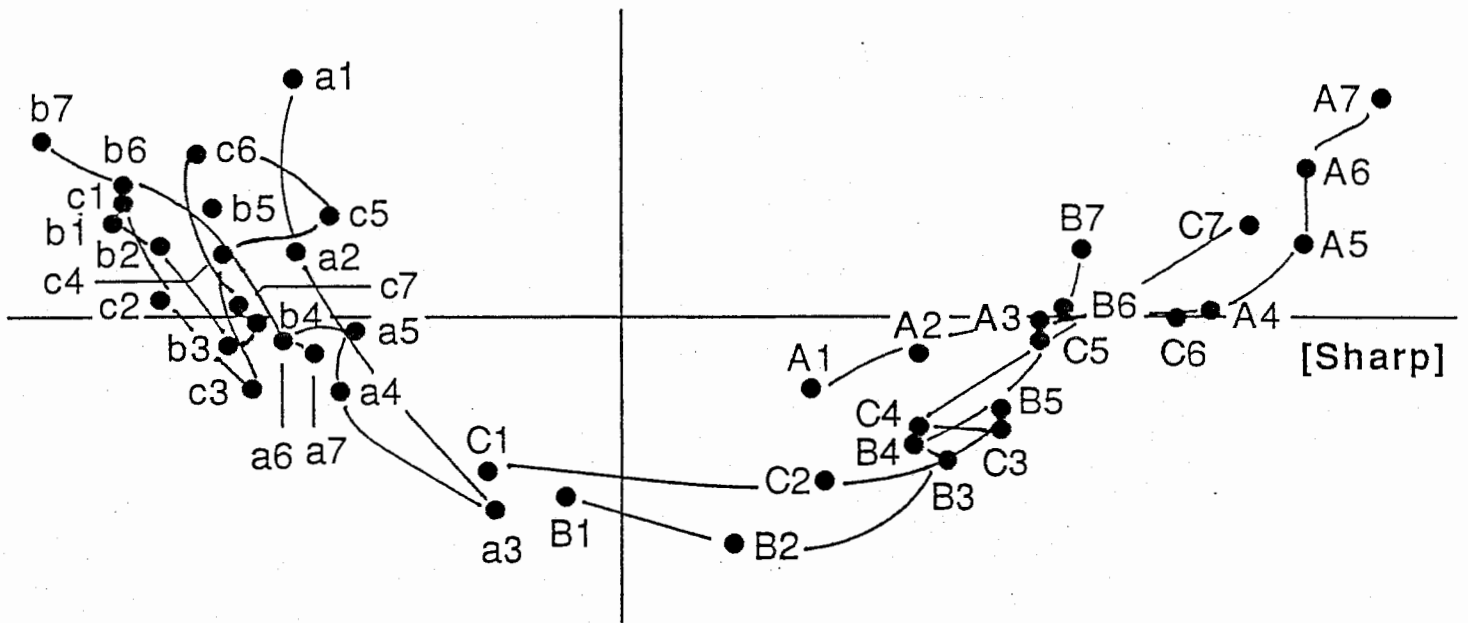


Fig. 3

Configuration derived in 2 dimensions. Sharpness judged by subject 4. Dimension 1 (horizontal) vs. dimension 2 (vertical). Stress = 0.076.



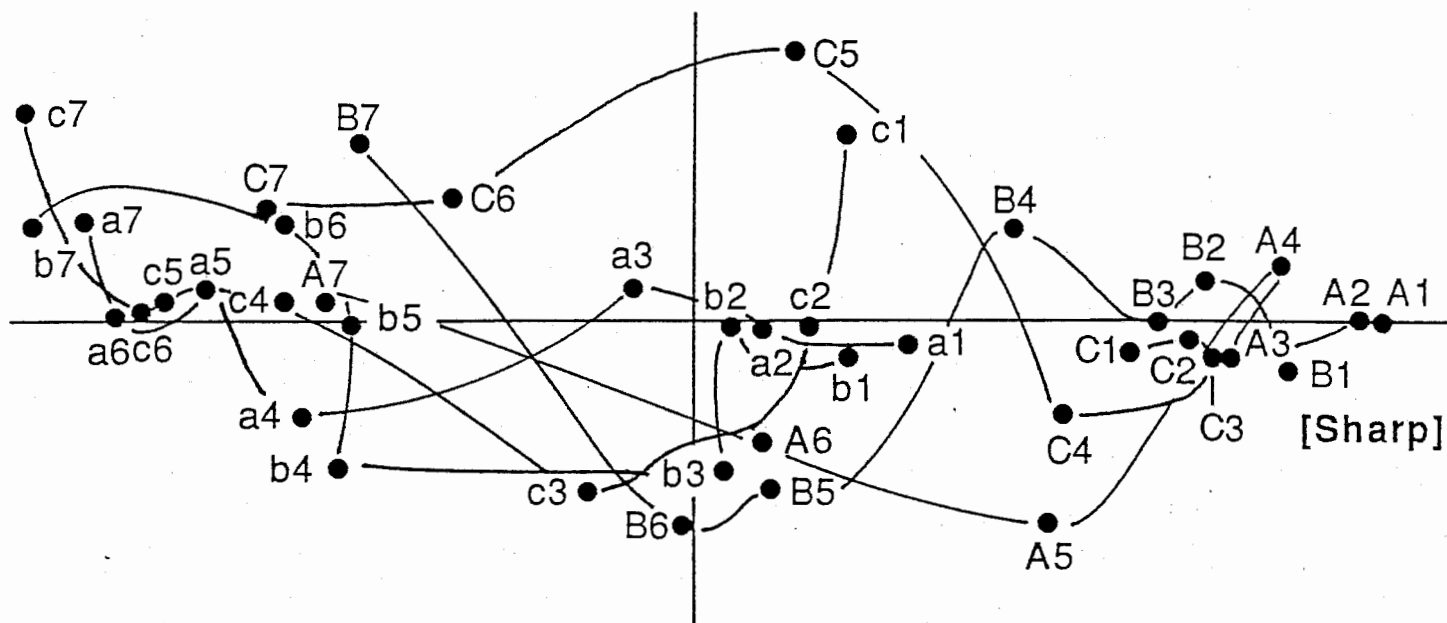


Fig. 4

Configuration derived in 2 dimensions. Sharpness judged by subject 1. Dimension 1 (horizontal) vs. dimension 2 (vertical). Stress = 0.093.

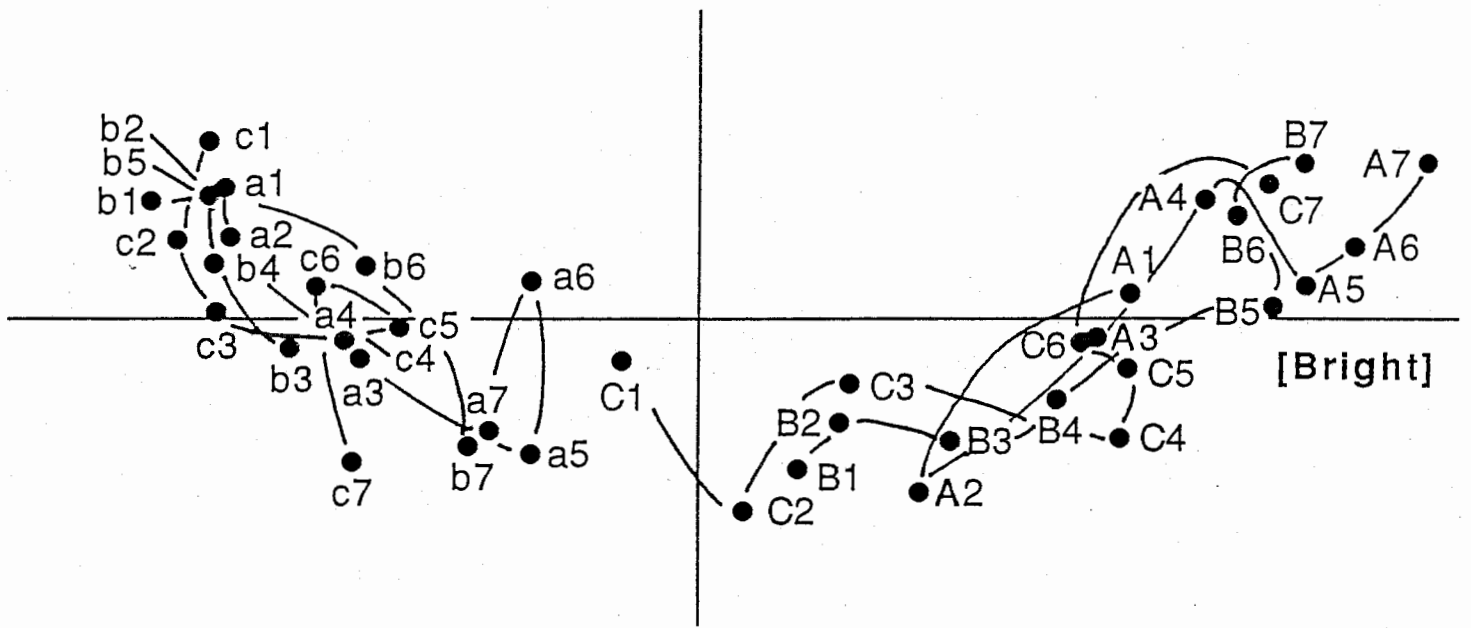


Fig. 5

Configuration derived in 2 dimensions. Brightness judged by subject 12. Dimension 1 (horizontal) vs. dimension 2 (vertical). Stress = 0.073.

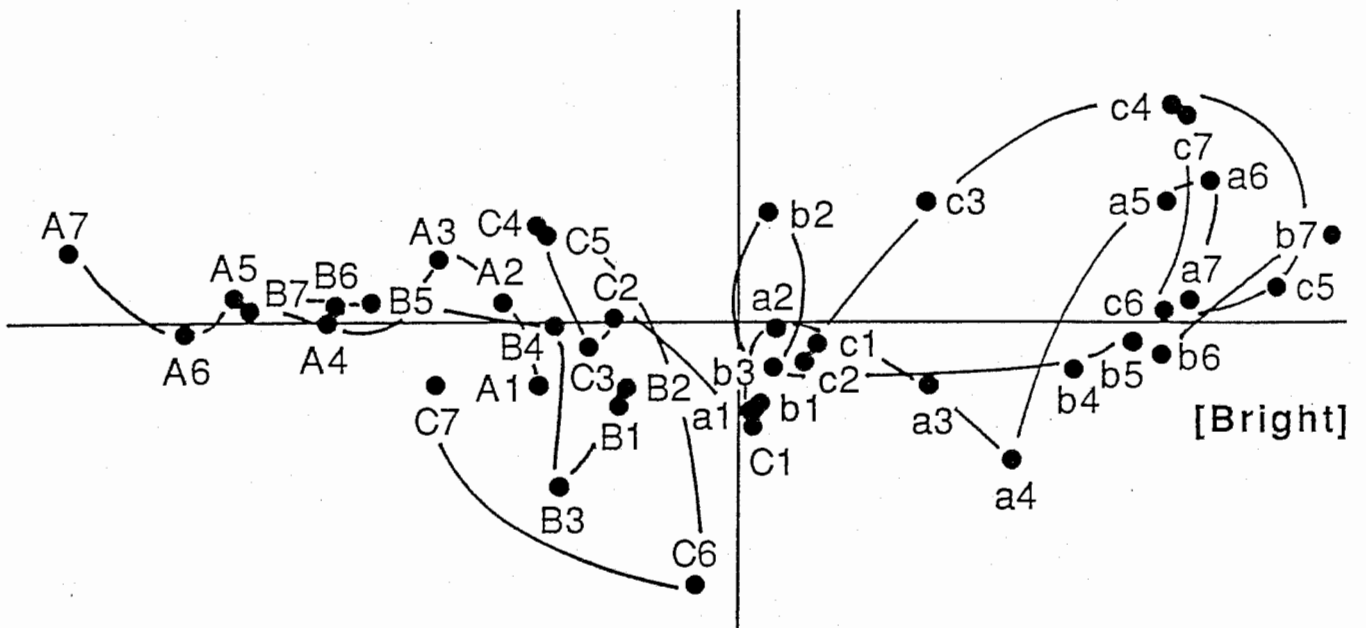


Fig. 6

Configuration derived in two dimensions. Dimension1 (horizontal) vs. dimension 2 (vertical). Brightness judged by subject 1. Stress = 0.099.

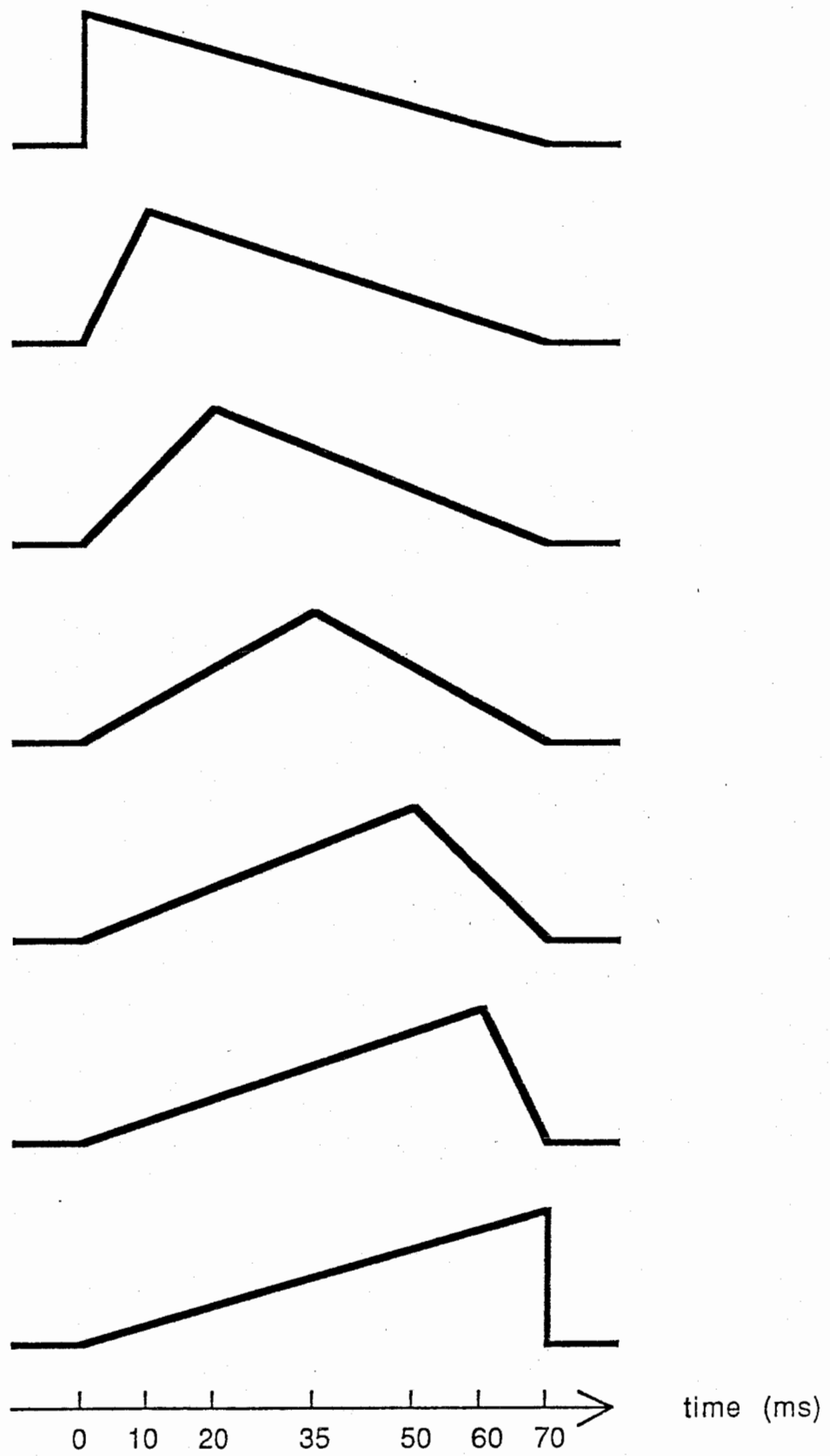


Fig. 7 Amplitude envelopes of the standard stimuli

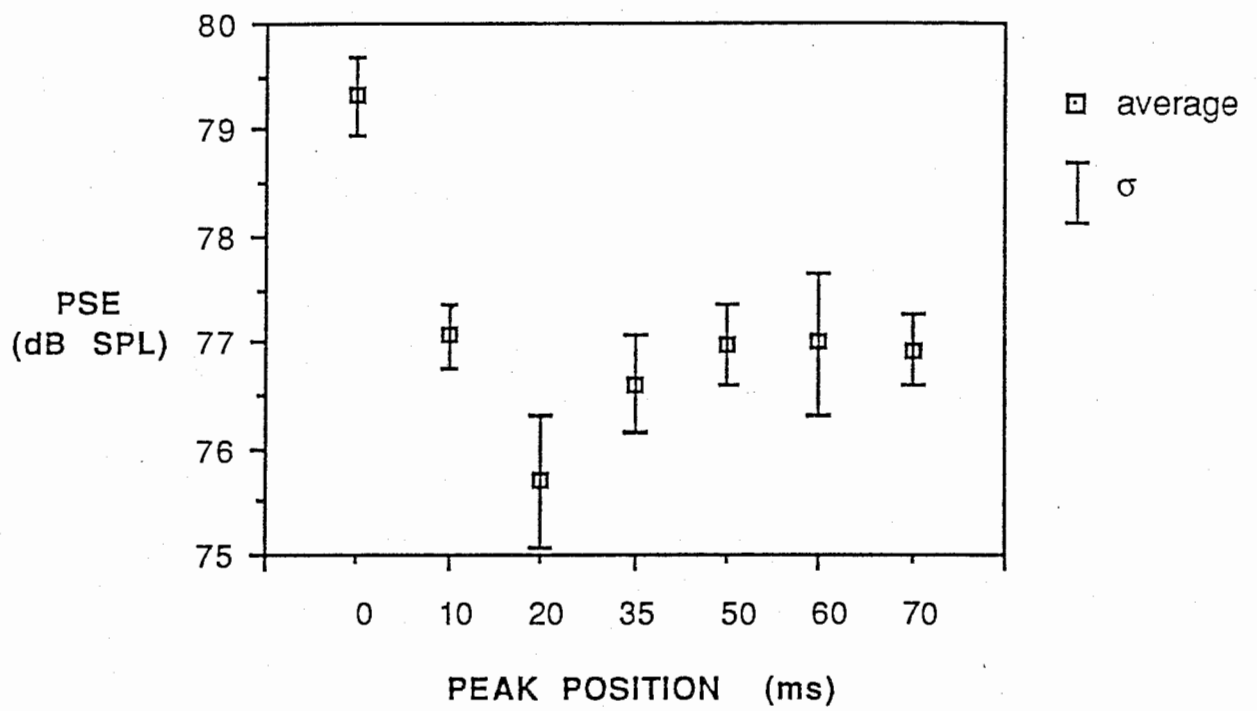


Fig. 8 PSE values for each standard stimuli.  
Subject A.

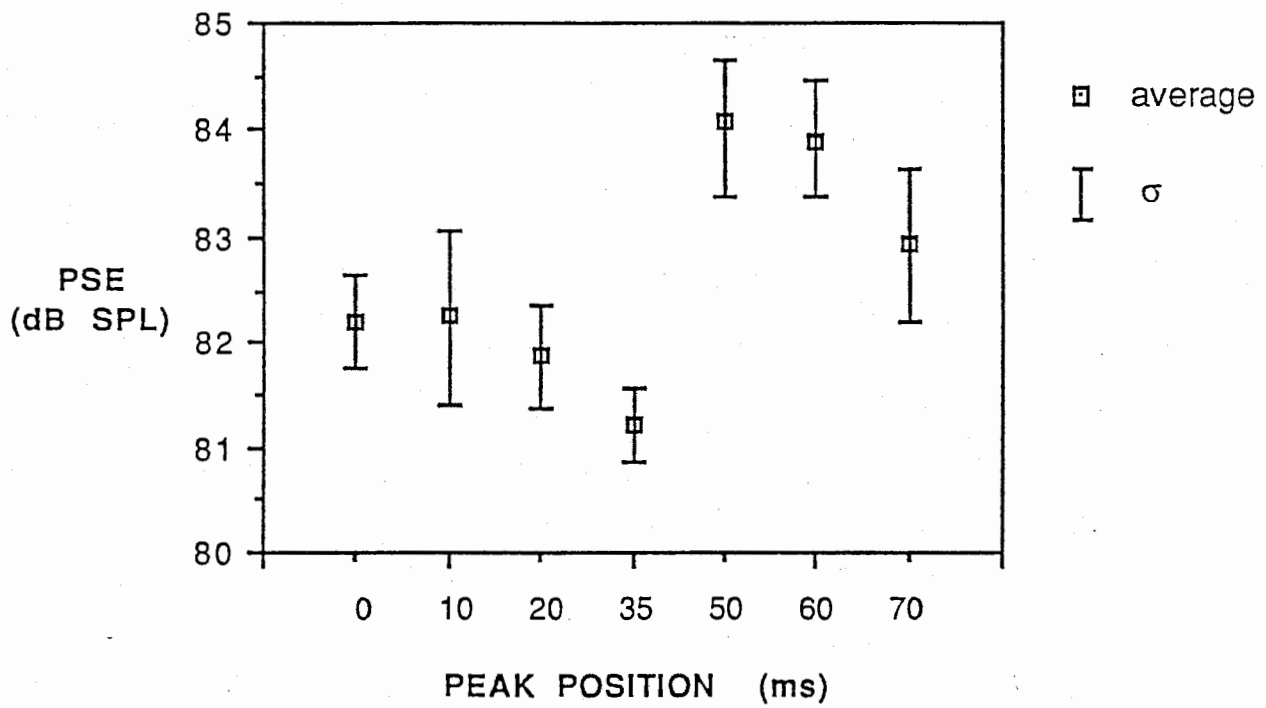


Fig. 9 PSE values for each standard stimuli.  
Subject B.

Table 1 (STNUM: 0~6 は C<sup>0</sup>-7 が矢印にあるものから順に7けたの刺激音であった)

ANALYSIS OF VARIANCE PROCEDURE

RYAN-EINOT-GABRIEL-WELSCH MULTIPLE RANGE TEST FOR VARIABLE: PSE  
 NOTE: THIS TEST CONTROLS THE TYPE I EXPERIMENTWISE ERROR RATE

ALPHA=0.05 DF=77 MSE=1.02019

NUMBER OF MEANS	2	3	4	5	6	7
CRITICAL RANGE	1.03084	1.12163	1.17097	1.2042	1.20517	1.24845

MEANS WITH THE SAME LETTER ARE NOT SIGNIFICANTLY DIFFERENT.

REGWQ	GROUPING	MEAN	N	STNUM	C <sup>0</sup> -7位置 (ms)
	A	79.3333	12	0	0
	B	77.0833	12	1	10
	B	77.0000	12	5	60
	B	76.9583	12	4	50
	B	76.9167	12	6	70
	B	76.5833	12	3	35
C	B	76.5833	12	3	35
C		75.7083	12	2	20
C		75.7083	12	2	20

Table 2

ANALYSIS OF VARIANCE PROCEDURE

RYAN-EINOT-GABRIEL-WELSCH MULTIPLE RANGE TEST FOR VARIABLE: PSE  
 NOTE: THIS TEST CONTROLS THE TYPE I EXPERIMENTWISE ERROR RATE

ALPHA=0.05 DF=77 MSE=1.711

NUMBER OF MEANS	2	3	4	5	6	7
CRITICAL RANGE	1.33498	1.45255	1.51646	1.55949	1.56074	1.61679

MEANS WITH THE SAME LETTER ARE NOT SIGNIFICANTLY DIFFERENT.

REGWQ	GROUPING	MEAN	N	STNUM	C <sup>0</sup> -7位置 (ms)
	A	84.0833	12	4	50
	A	83.8750	12	5	60
	A	82.9167	12	6	70
B	A	82.9167	12	6	70
B	C	82.2500	12	1	10
B	C	82.2083	12	0	0
B	C	81.8750	12	2	20
B	C	81.2083	12	3	35
	C	81.2083	12	3	35

目 次

1. まえがき	1
2. 白色雑音合成のアルゴリズム	1
3. プログラム例	2
4. スペクトルの測定結果	3
5. 応用	3

図 ( 図 1 - 図 4 )



## 1. まえがき

ここでは白色雑音を合成するプログラムを紹介し、合成された白色雑音のスペクトルの測定結果を示すこと、また合成された白色雑音の応用的な使い方について述べることを目的とする。まず始めに電子計算機で白色雑音を合成するための方法について述べ、次に筆者が作成したC言語で書かれたプログラムを紹介する。さらに、そのプログラムで合成された白色雑音のスペクトルの測定結果を紹介する。最後に応用的な側面について述べる。

以下に述べる方法により作成された白色雑音は極めて平坦なスペクトル分布を持つので、精密な聴覚心理実験を行う際には特に有用である。また、一度作成すれば、DA変換時のサンプリング周波数を落とすことにより、持続時間のより長い帯域雑音として用いることもできる。

## 2. 白色雑音合成のアルゴリズム

次に述べる方法により、電子計算機上で白色雑音を合成することができる。予め与えておかなければならないパラメタは次の通りである。

- ① サンプリング周波数 ( $s$  とする)。
- ② 作成する雑音の周波数範囲。最低周波数及び最高周波数。
- ③ 何 Hz ステップで周波数成分を足していくか ( $\Delta f$ )。
- ④ 作成する雑音の持続時間。
- ⑤ 振幅の最大値。

1 点は次式により計算される。

$$x_n = \sum_{i=m1}^{m2} [\sin(w_i n + \theta_i) + \cos(w_i n + \theta_i)]$$

$$w_i = 2\pi f_i / s$$

$$f_i = \Delta f \cdot i$$

ここで  $n$  はサンプル点、 $m1$  は  $\Delta f \cdot m1$  が最低周波数となるような値、 $m2$  は  $\Delta f \cdot m2$  が最高周波数となるような値、 $w_i$  は角速度、 $\theta_i$  は位相初期値、 $s$  はサンプリング周波数、 $f_i$  は周波数、 $\Delta f$  は周波数のステップである。 $\theta_i$  はランダムである。

与えられたパラメタに従い、次のアルゴリズムで雑音を合成する。

- ① 各周波数成分の初期位相を決定する。これは周波数によりランダムになるようにする。

②与えられたサンプリング周波数に従い、各点毎に各周波数についてサインとコサインを足し合わせて各周波数成分の振幅を求め、各周波数成分の振幅を加算していく。

③一旦作られたデータの振幅の最大値を求め、その最大値で各点の振幅を除し、正規化する。予め与えられた振幅の最大値を正規化された各点の振幅に乗ずる。

### 3. プログラム例

図1にC言語で書かれた白色雑音合成用のプログラム例を示す。このプログラムでは20 Hzから20 kHzの帯域を持つ1秒間の白色雑音が合成される。各周波数成分間の間隔は1 Hzステップとしている。このプログラムをDEC/VAX 8600(OS: ULTRIX 4.3 BSD)で実行したところ、CPU時間は約38時間であった。

### 4. スペクトルの測定結果

図2に作成された白色雑音のスペクトルを示す。図2の横軸はlog frequency (50 Hz - 20 kHz)、縦軸はパワー (1目盛り10 dB)である。この白色雑音は図1のプログラムで合成したバイナリデータをMASSCOMP/M5600により48 kHzサンプリングでDA変換し、20 kHz, 96 dB/Oct.のローパスフィルターを掛けたものである。分析は3 dBのアッテネータを介した後、スペクトルアナライザにより行った。分析窓長は20 msである。

図2においてスペクトルのピークとディップの差は最大で30 dB程度である。全体としては平坦なスペクトルであると言える。

図3に128回の時間加算を行った分析結果を示す。時間加算を行った以外は図2と同じ条件で分析を行った。この図からこの白色雑音は平均的には極めて平坦なスペクトルを持つことが分かる。

図4はDA変換時のサンプリング周波数を12 kHzに落とし、5 kHzのローパスフィルターを掛けたものの分析結果である。128回の時間加算を行っている。50 Hzから5 kHzまで、図3と同様に極めて平坦なスペクトルを持っている。

### 5. 応用

一度作成した白色雑音のデータに対して次の様な応用的な使い方が可能である。例えば、DA変換時のサンプリング周波数を下げることにより、最高周波数により低く、持続時間のより長い帯域雑音を作ることができる。図1のプログラムにより合成されたデータを例に取れば、DA時のサンプリング周波数を12 kHzに落とせば、10 Hzから5 kHzまでの持続時間4秒の帯域雑音を作ることができる。

或は、スペクトルが極めて密でかつ平坦であることを利用して、バンドパスフィルターを用いて任意の帯域を持つ帯域雑音を作ることができる。この場合 20 Hz から 20 kHz までのどの帯域から取り出してもスペクトルの平坦性は保証されている。

また図 1 のプログラムでは 1 Hz ステップで周波数成分を足し合わせて雑音を合成しているが、このステップをより粗くすることにより計算に要する時間を節約することができる。例えば 10 Hz ステップとすれば計算に要する時間は約 10 分の 1 になるはずである。どの程度までステップを粗くするかは雑音の使用目的による。

ここで紹介した方法により合成した白色雑音を使用して心理実験を行った（上田、赤木、1987）。この実験では刺激音のスペクトル包絡を一定に保つことが重要な要件であったので、この白色雑音は特に有用であった。

ここで述べた方法により合成された白色雑音はそのスペクトルの平坦性が合成時点から保証されているということがその利点である。また、計算機上に乗っているということを利用して任意の時点からの切り出しが自由にできるという利点もある。更に合成方法に僅かな変更を加えることにより、ピンクノイズの合成などにも応用可能である。

#### 文献

上田和夫、赤木正人(1987)「音の鋭さと振幅包絡」日本音響学会聴覚研究会資料、H-87-64。

```

/*
**      Noise Generating Program
**                               by K. Ueda, June 26, 1987
*/
#include <stdio.h>
#include <math.h>
#define PI 3.1415926536

double st[20001], r[20001];
10
main()
{
    double ti(), sin(), cos(), fabs(), maxi();
    double amp, theta, PI2, 1, maximum = 0.000001;
    int bottom_fr = 20;          /* lower limit of frequency in Hz */
    int upper_fr = 20000;       /* upper limit of frequency in Hz */
    int duration = 1000;       /* duration in ms */
    int sampling_fr = 48000;    /* sampling frequency in Hz */
    int step = 1;              /* adding step in Hz */
20
    int expand = 20000;         /* maximum amplitude (16bit: 32767) */
    int hz, numb_sample, n;
    short namp;
    FILE *output_temp, *input_temp, *output_file, *fopen();

    /***** open output file: temporary and permanent *****/
    output_temp = fopen("temp_dat", "w");
    output_file = fopen("band1_dat", "w");

    /***** computing number of sample *****/
30
    numb_sample = (sampling_fr / 1000) * duration;

    /***** printing parameters *****/
    printf("duration          = %7d ms\n", duration);
    printf("sampling frequency = %7d Hz\n", sampling_fr);
    printf("step width (fr)      = %7d Hz\n", step);
    printf("number of sample      = %7d\n", numb_sample);
    printf("bottom frequency     = %7d Hz\n", bottom_fr);
    printf("upper frequency       = %7d Hz\n", upper_fr);
    printf("expanding rate        = %7d\n", expand);
40

    upper_fr = upper_fr + 1;
    PI2 = 2.0 * PI;

    /*** computing sampling timing and initial phase for each fr. ***/
    hz = bottom_fr;
    while(hz < upper_fr)
    {
        st[hz] = ti(hz, sampling_fr);
        r[hz] = ((double)rand() / 32767.0) * PI2;
50
        hz = hz + step;
    }

    /**** computing amplitude at each point *****/
    for(n = 0; n < numb_sample; n++)
    {
        amp = 0.0;
        hz = bottom_fr;
        while(hz < upper_fr)
60
        {
            theta = (PI2 / st[hz]) * (double)n + r[hz];
            amp = sin(theta) + cos(theta) + amp;
            hz = hz + step;
        }
        maximum = maxi(maximum, amp); /* find out maximum amplitude */
        fwrite(&amp, 8, 1, output_temp); /* writing in temp. file */
    }
    fclose(output_temp);

    /***** adjusting maximum amplitude *****/
70
    input_temp = fopen("temp_dat", "r");
    for(n = 0; n < numb_sample; n++)
    {
        fread(&amp, 8, 1, input_temp);
        amp = amp / maximum;
        namp = expand * amp;
        printf("namp[%d] = %d\n", n, namp);
        fwrite(&namp, 2, 1, output_file);
    }
80
    fclose(input_temp);
    fclose(output_file);
}

/**** sampling timing ****/
double ti(helz, sampl)

```

```
int helz, sampl;
{
    double st;
    st = sampl / helz;
    return(st);
}
90

/**** find out maximum amplitude *****/
double maxi(m, l)
double m, l;
{
    double ab;
    ab = fabs(l);
    if(ab > m)
        m = ab;
    return(m);
}
100
```



## 1 白色雑音合成用プログラムの例

\*\* TR9403 DIGITAL SPECTRUM ANALYZER \*\*  
\*\*\*\* MFD BY ADVANTEST \*\*\*\*  
Pk 2 650.0Hz -17.4 dBV

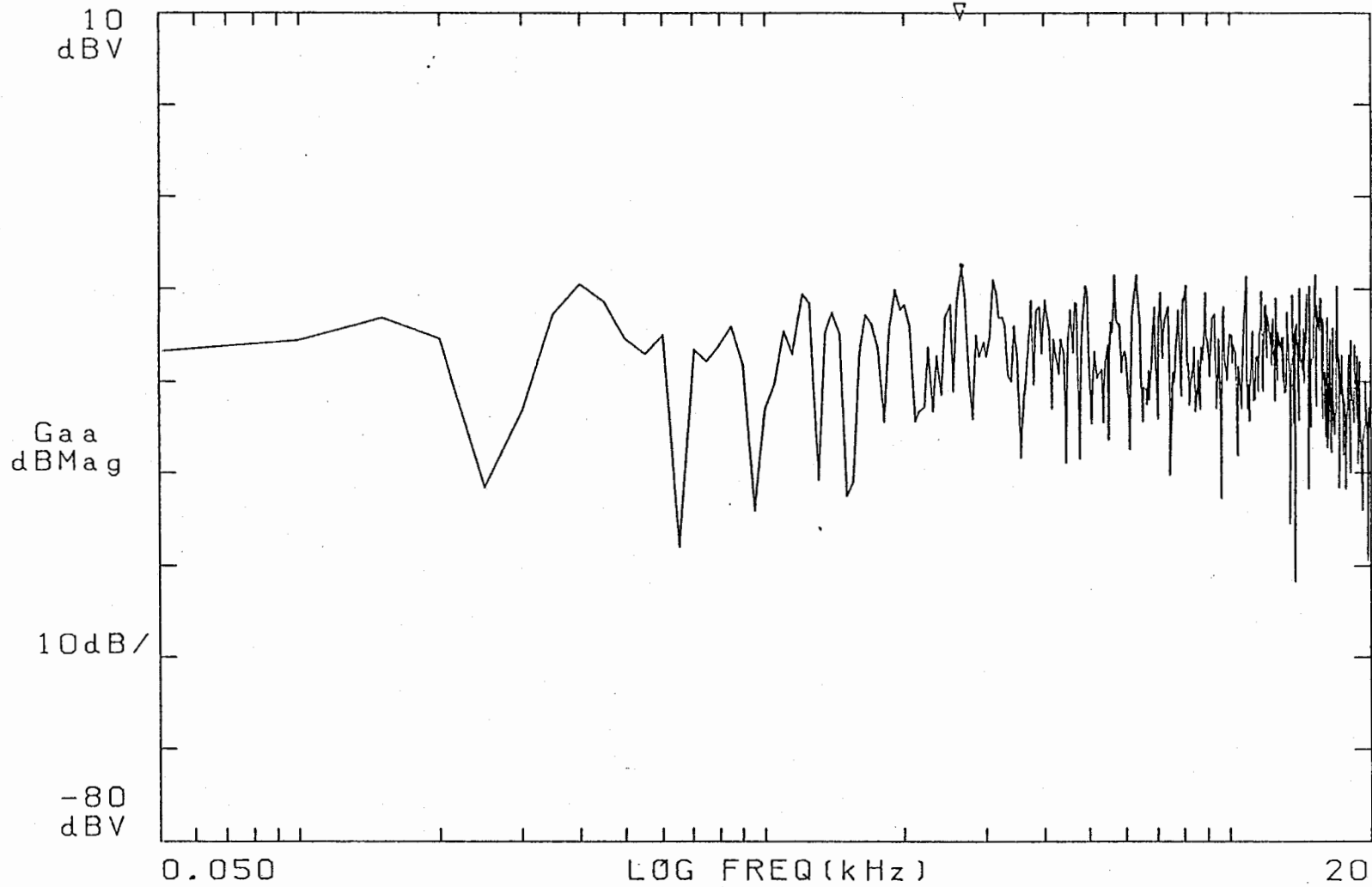


図 2 図1のプログラムで合成された白色雑音のスペクトル

\*\* TR9403 DIGITAL SPECTRUM ANALYZER \*\*  
\*\*\*\* MFD BY ADVANTEST \*\*\*\*  
Pk 3 100.0Hz -23.5 dBV

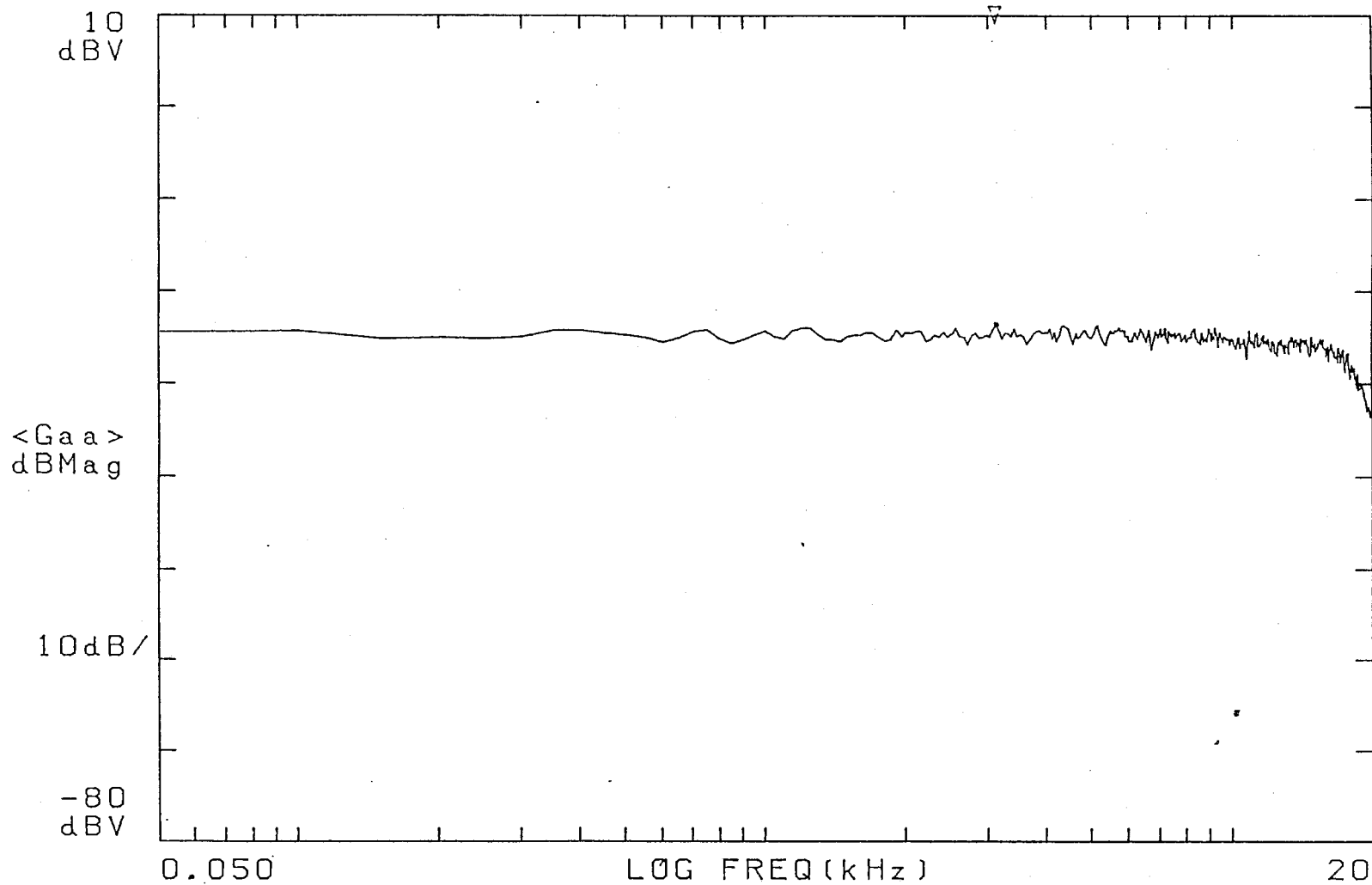


図 3 時間加算したスペクトル

\*\* TR9403 DIGITAL SPECTRUM ANALYZER \*\*  
\*\*\*\* MFD BY ADVANTEST \*\*\*\*  
Pk 1 750.0Hz -17.9dBV

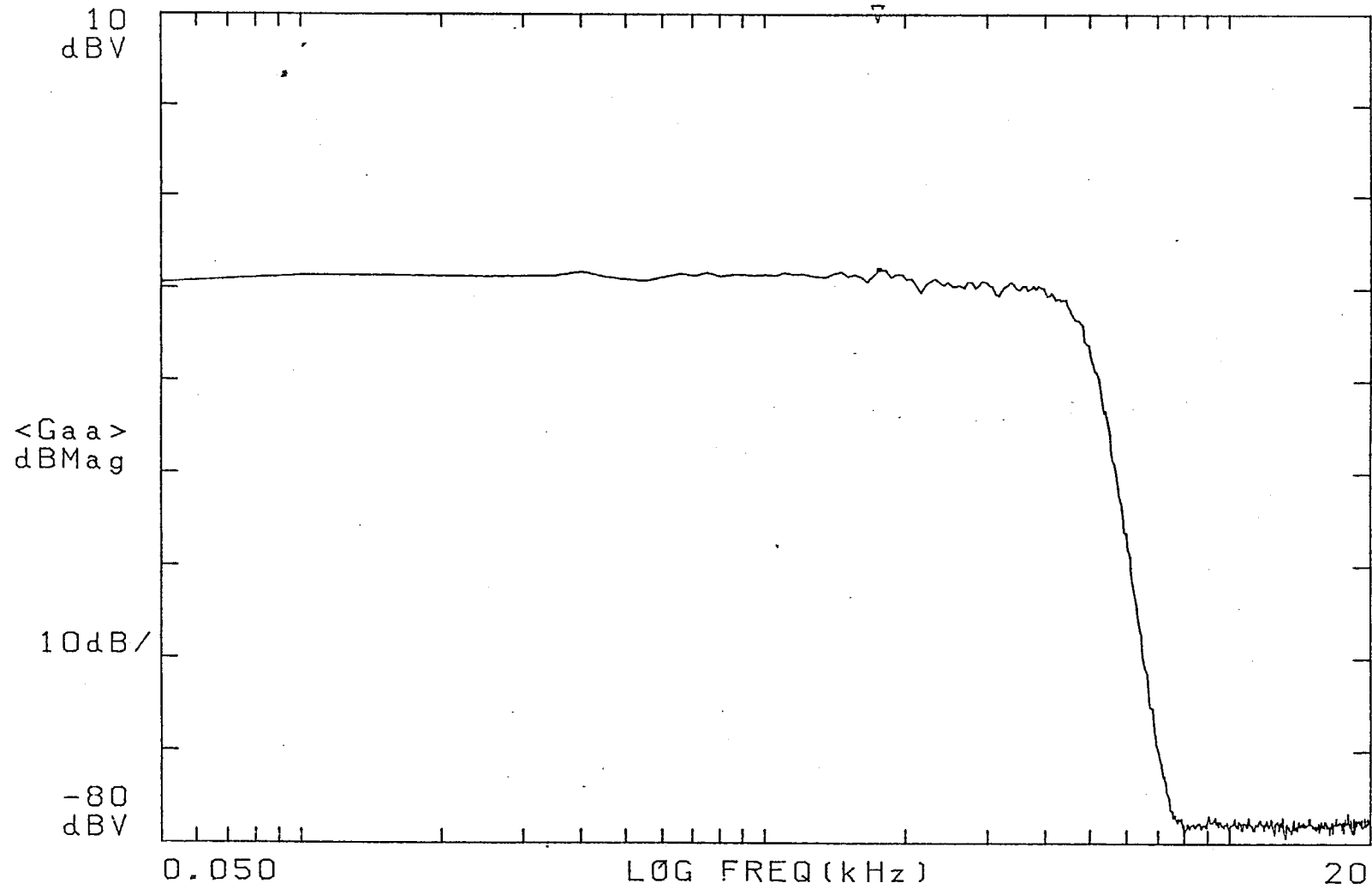


図 4 5 kHz 以下のローパスノイズのスペクトル (時間加算)