

博士学位論文
Ph.D. Dissertation

音の高さ，音色，音声の知覚と短期記憶に関する研究
Pitch, Timbre, and Speech: Perceptual and Short-term
Memory Interference

上田 和夫
Kazuo UEDA

1998
京都
Kyoto, Japan

博士学位論文

Ph.D. Dissertation

音の高さ，音色，音声の知覚と短期記憶に関する研究

Pitch, Timbre, and Speech: Perceptual and Short-term

Memory Interference

上田 和夫

Kazuo UEDA

1998

京都

Kyoto, Japan

内容梗概

音の高さ、音色、音声の知覚と短期記憶に関する研究

上田 和夫

本研究は、音の高さの知覚に与える音色の影響、非音声の高さの短期記憶に数詞音声系列が与える影響について調べたものである。音の高さと音色とは通常、それぞれ独立した属性であると考えられている。たとえば、音楽では音色は楽器の種類に対応し、同じ高さの旋律を違う楽器で演奏することができる。しかし、音の高さと音色との区別は必ずしも明確ではない。本研究では、無限音階複合音を変形した音の高さを調べた実験によって、音の高さには二つの心理的次元が含まれ、それらが螺旋形で表現できることを示した。二つの次元のうち一つは直線的、音色的な次元で、もう一方は循環的、音階的な次元である。また、音声は特殊で非音声とは別の貯蔵庫に記憶されるという考え方がある。しかし、非音声の高さと音声の系列が短期記憶で干渉し、両者の記憶が特殊化されていないことを、非音声（純音と調波複合音）と数詞音声（高さをコントロールした音声と自然な音声）を用いた三つの実験で確かめた。これらの結果から、音の高さの知覚に音色が、非音声の高さの短期記憶に音声の系列が影響することが明らかになった。このような聴覚系のふるまいを説明するためには、従来の心理的次元の階層的干渉モデルを拡張する必要があることを論じた。

Abstract

Pitch, timbre, and speech: Perceptual and short-term memory interference

Kazuo UEDA

This investigation focuses on perception and short-term memory of pitch. The first half of the paper is concerned with the effect of timbre brightness on pitch perception. Pitch and timbre are usually regarded as independent attributes. For example, instruments can be distinguished by timbre difference, and the same melody on the same key can be played with different instruments. The distinction between pitch and timbre, however, is not always obvious. An experiment that subjects judged relative pitch of specially generated complex tones revealed that pitch is consisted of two components: one is rectilinear and brightness-based dimension (tone height), and the other is circular and scale-based dimension (tone chroma). The second half of the paper is concerned with short-term memory interference between pitch of non-speech and series of spoken numbers. It had been suspected that these two might not interfered because of memory specialization for speech and for non-speech. However, the present investigation revealed that the interference does occur; both errors, of non-speech pitch recognition and of serial recall of spoken numbers, were significantly increased when the two tasks were simultaneously required. These results indicate that pitch perception is affected by timbre, and that short-term memory of non-speech pitch is affected by series of spoken nubmers. Some enhancements for a dimensional interaction model were proposed to account the results.

目次

第 1 章	まえがき	1
1.1	音の高さと音色：用語の定義	1
1.2	音の高さと物理量の関係	2
1.2.1	純音の高さと物理量	2
1.2.2	複合音の高さと物理量	3
1.2.3	雑音の高さと物理量	4
1.3	音色と物理量の関係	5
1.4	音の高さと音色の関係	6
1.5	音声と非音声：用語の定義	7
1.6	本研究の目的	8
1.7	本論文の構成	8
第 2 章	音の高さの二面性の空間的表現	10
2.1	音の高さの二面性に関する研究	10
2.1.1	音の高さの二面性	10
2.1.2	無限音階複合音の循環性	11
2.1.3	無限音階複合音と音の高さの二面性	13
2.1.4	音の高さの螺旋モデル	14
2.2	実験の目的	17
2.3	刺激の選択	18
2.4	刺激	20
2.5	実験手続き	23
2.6	実験結果	23
2.7	考察	29

2.7.1	刺激群 A, B による違い	30
2.7.2	二つの心理的次元：トーンハイトとトーンクロマ	30
2.7.3	判断の個人差と螺旋形	31
2.7.4	個人差と音楽経験との相関	32
2.7.5	過去のモデルとの比較検討	33
2.8	まとめ	34
第 3 章	非音声の高さの短期記憶と音声の干渉効果	36
3.1	短期記憶における音声と非音声の干渉効果に関する研究	36
3.1.1	音声と非音声の短期記憶の特殊化を支持する証拠	37
3.1.2	テスト音と挿入音の高さ関係の重要性を示す証拠	38
3.1.3	テスト音と挿入音の体制化の影響を示す証拠	39
3.2	短期記憶の特殊化を支持する証拠への疑問	40
3.3	高さをコントロールした音声と調波複合音の干渉効果	43
3.3.1	刺激	43
3.3.1.1	テスト音の合成	43
3.3.1.2	調波複合音の合成	44
3.3.1.3	数詞音声刺激の作成	44
3.3.2	実験条件	45
3.3.3	実験手続き	47
3.3.4	実験結果	47
3.3.5	考察	49
3.3.5.1	テスト音と挿入音の高さ関係の重要性	49
3.3.5.2	音声の高さの不明確さ	49
3.3.5.3	音声刺激の性質のあいまいさ	50
3.3.6	まとめ	50
3.4	自然な音声と純音の干渉効果	51
3.4.1	刺激	52
3.4.1.1	純音と音声	52
3.4.1.2	刺激呈示法と音圧の較正	52
3.4.1.3	音声の大きさの調整	53
3.4.1.4	音声の高さの測定	53
3.4.2	被験者の選抜	54
3.4.2.1	高さの再認能力による選抜	54

3.4.2.2 絶対音感テストによる選抜	55
3.4.3 実験条件	56
3.4.4 実験手続き	57
3.4.5 実験結果	58
3.4.6 考察	63
3.4.6.1 音声の大きさ調整の誤りについて	63
3.4.6.2 純音の高さの記憶と数詞音声系列の記憶は干渉する	63
3.4.6.3 再認, 再生誤り率と高さの不明確さ	64
3.4.6.4 Deutsch (1970) との相違の原因	65
3.4.6.5 音楽的被験者と非音楽的被験者	66
3.4.7 まとめ	67
3.5 自然な音声と調波複合音の干渉効果	67
3.5.1 刺激	68
3.5.2 実験条件	69
3.5.3 被験者	70
3.5.4 実験手続き	70
3.5.5 実験結果	71
3.5.6 考察	74
3.5.6.1 調波複合音の高さと数詞音声系列の干渉	74
3.5.6.2 高さの再認誤り率とテスト音の高さの明確さ	74
3.5.6.3 高さの再認誤り率と音声の高さの不明確さ	75
3.5.6.4 数詞の系列再生誤り率とテスト音の高さの明確さ	76
3.5.7 まとめ	76
3.6 総合考察	77
3.6.1 音声と非音声の短期記憶メカニズム：何が干渉を起こすのか	77
3.6.2 音声の有意義性	78
3.7 まとめ	79
第 4 章 音の高さと音色, 音声と非音声の関係	80
4.1 心理的次元の階層的干渉モデル	80
4.2 階層的干渉モデルの拡張	81
第 5 章 結論	83

謝辞	84
文献	85
本研究に関する研究業績	93
その他の研究業績	97

第 1 章 まえがき

われわれは日常、自明のことであるかのように音の高さと音色を区別し、音声と非音声とを区別する。しかし、このような二分法はいつでも成り立つのであろうか。本研究は、音の高さの知覚と短期記憶に他の聴覚属性が与える影響について調べたものである。特に、音の高さの知覚に関しては音の高さに音色が与える影響について、音の高さの短期記憶に関しては非音声の高さの短期記憶と数詞音声系列の短期記憶が与える影響について調べた。

本章では、本研究の対象となる音の高さと音色、音声と非音声について、研究の背景を説明し、本研究の目的と本論文の構成について述べる。

1.1 音の高さと音色：用語の定義

「音の高さ」や「音色（ねいろ）」という用語は、日本工業規格 (1986) では次のように定義されている。

定義：音の高さ (pitch)

聴覚にかかわる上の音の性質の一つで、高低で表現されるもの。

定義：音色 (timbre)

聴覚にかかわる上の音の性質の一つで、2 音の大きさ及び高さが共に等しくても、その 2 音が異なった感じを与えるとき、この相違に対応する性質。

これらの定義によれば、音の高さと音色はそれぞれ独立した、互いに重ならない二つの知覚的属性であり、各々を別々に操作できることになる。果たしてそうであろうか。

確かに、同じ高さの旋律を異なる楽器で演奏した場合には、「同じ高さの組

み合わせて構成された、同じ旋律の音色が変化した」ということができよう。また、旋律を楽器ではなく、人間の声で演奏する、すなわち「歌う」こともできる。これも楽器のときとは「音色」が変わったことになる。さらに声の場合には、同じ高さを違った母音で歌うこともできる。この母音の違いも「音色」の変化のなかに含まれる。ふだんの話し声では、音韻の区別は「音色」の変化、イントネーションは「高さ」の変化とみなすことができる。音声の質の違いによって話者を判断する際には、おもに声の基本周波数の違い（高さ）によって性別が、スペクトルの違い（音色）によって同性内での違いが判断されることがわかっている（これに関するレビューは Ueda, 1993 を参照）。

しかし、音の高さも音色も、複数の物理量が関係する多次元的な心理的属性である。そして、両者の判断には共通の物理量が影響する場合があり、音の高さと音色の区別にあいまいな部分があることをこれから説明する。まず、音の高さと物理量の関係、音色と物理量の関係について述べ、次に音の高さと音色の関係について述べる。

1.2 音の高さと物理量の関係

ここでは、単耳で聞き取れる音の高さについて、音の高さを決定する物理量が複数存在することを説明する。

1.2.1 純音の高さと物理量

純音の高さは周波数と密接な関係があり、周波数の高い音ほど高く聞こえる。その関係を分割法 (method of fractionation) により尺度化したものがメル尺度 (Stevens and Davis, 1938) である。また、純音の高さは周波数だけではなく、音の強さ、音圧とも関係する。すなわち、純音の周波数を一定に保って、音圧を変化させると、およそ 1000 Hz 以下では音圧をあげるほど低く聞こえ、およそ 3000 Hz 以上では音圧をあげるほど高く聞こえる (Stevens, 1935; Walliser, 1969)。したがって、音圧による高さの変化が比較的少ない、中程度の周波数帯域を除

いて、音の高さを一定に保ったまま音の大きさを変化させるためには、純音の周波数と音圧の両方を操作しなければならない。

ただし、音圧による高さの変化は個人差が大きく、また、その変化の仕方が非単調な被験者もいることを示すデータもある (Burns, 1982; Verschuure and van Meeteren, 1975). さらに、そのような変化の仕方は、同じ個人内でも左右のどちらの耳に刺激を呈示するかによって異なる場合もある (Burns, 1982). しかし、両耳に同じ刺激を呈示した場合には、左右の耳から入力された音の高さは中枢で融合され、単一の音の高さが知覚される。

1.2.2 複合音の高さと物理量

調波複合音の場合、基本周波数にはほぼ対応する高さが知覚される。しかし、フィルターによって基本周波数成分だけを取り除いても、複合音の高さは変わらない (Plomp, 1976; Demany, 1989; Moore, 1989; Hartmann, 1996). すなわち、物理的には存在しない、基本周波数成分の高さに相当する高さが知覚される。7-8カ月の乳児でも欠如した基本周波数成分に相当する高さを知覚する (Clarkson and Clifton, 1985). さらに、基本周波数成分を取り除いた残りの成分の周波数をシフトさせて調波関係からずらし、非調波関係にすると高さが不明確になる。大人でも 7-8 カ月の乳児でも、非調波の度合いが増すにしたがって知覚される高さが不明確になる (Clarkson and Clifton, 1995).

基本周波数成分のあるなしにかかわらず、音のスペクトル構造も高さに影響する。基本周波数が同じでも、母音（スペクトルの形状）が異なれば高さがずれて知覚され、そのずれ方は、スペクトルのエネルギー分布によって決まる (Chuang and Wang, 1978). 基本周波数が等しく、スペクトル構造の異なる複合音の高さを比較すると、高次倍音の豊富な、明るい音色の音はそうでない音よりも高く知覚されやすい (Hesse, 1982; Patterson, 1989; Singh and Hirsh, 1992). また、スペクトル構造の大きく異なる複合音どうしや複合音と純音との間で高さの弁別閾を測定すると、同じスペクトル構造の音どうしで測定した場合と比べて弁別閾が大きくなる。この原因は、二つの音の音色の違いであると考えられる

(Moore and Glasberg, 1990, 1991; Moore, Glasberg, and Proctor, 1992). すなわち、基本周波数が同じでも、スペクトル構造が異なり、音色が異なれば、高さやその明確さが異なる。

1.2.3 雑音の高さと物理量

純音や調波複合音には波形に周期性がみられ、周期性と高さとの間に強い関係があるが、一方で、波形にはっきりした周期を持たない雑音でも高さは知覚される。

35 dB/octave の遮断特性のフィルタを用いて作られた、高域通過雑音どうし、低域通過雑音どうしでオクターブマッチング（標準音よりも 1 オクターブ高く、あるいは 1 オクターブ低くなるように比較音の遮断周波数を調整する）を行った実験では、遮断周波数がオクターブになるように被験者は調整を行った (Small and Daniloff, 1966). すなわち、雑音の高さと遮断周波数に相関があることが示された。さらに、Fastl (1971), Zwicker and Fastl (1990) は、120 dB/octave という、きわめて鋭い遮断特性のフィルターを用いて、低域通過雑音、高域通過雑音、帯域通過雑音を作り、純音とのマッチングによって雑音の高さを調べた。ばらつきは大きいものの、被験者はおおむねスペクトルのエッジに純音の周波数を合わせた。この高さをエッジピッチ (edge pitch) と呼ぶ。雑音の帯域幅が 10 Hz 以下という狭帯域の場合には、調整値は雑音の中心周波数付近に集まる。エッジピッチの明確さはスペクトルのエッジの鋭さと正の相関がある (Fastl, 1980; Zwicker and Fastl, 1990).

しかし、標準音として 1/2 octave あるいはそれ以上の帯域幅の雑音を、比較音として 1/2 octave の帯域雑音を用い、比較音が標準音よりも低いかどうかという二肢の回答から被験者に強制選択させた実験では (Rainbolt and Schubert, 1968), 全体のデータとしてはエッジピッチは観測されなかった。むしろ、帯域雑音の中心周波数付近に相当する高さが知覚された。ただし、個人ごとのデータでは、帯域雑音の上限周波数、あるいは下限周波数のどちらか一方に相当する高さを知覚した被験者もいた。

以上をまとめると、波形にはっきりした周期を持たない雑音でも高さは知覚されるが、その高さは明確なものではない。雑音の高さには二種類あり、純音とのマッチングではエッジピッチが知覚されやすく、雑音どうしの比較では中心周波数に相当する高さが知覚されやすくなる。そして、エッジピッチはスペクトル構造の急激な変化が生じる部分を手がかりとして知覚され、中心周波数に相当する高さはスペクトル全体のある種の「平均」として知覚されると考えられる。

1.3 音色と物理量の関係

音色に影響する物理量は多種多様である。

音のスペクトルのエネルギー分布が変化すると、音色が変わって知覚される (Lichte, 1941; Berger, 1964; Plomp, Pols, and van de Geer, 1967; Slawson, 1968; von Bismarck, 1974a, b; Plomp, 1976; de Bruijn, 1978; Grey and Gordon, 1978; Benedini, 1980; 大串, 1980; Risset and Wessel, 1982; Preis, 1984; Bloothoof and Plomp, 1988). 特に、定常音の「鋭さ (sharpness)」は、臨界帯域ごとに重み付けされた、パワースペクトルの一次モーメントで予測できる (von Bismarck, 1974b). さらに、聴覚のマスキング特性を考慮して聴覚系内でのスペクトルを推定することにより、マスクトスペクトルの形状と音色とが密接に関係していることが示された (小澤, 1993). 基本周波数を変化させた調波複合音の音色についても、同様のことが確かめられている (森谷, 1995). そして、パワースペクトルだけではなく、成分の位相も、調波複合音の音色に影響する (舘, 磯部, 1973). さらに、楽器音は基本的に調波複合音であるが、成分音間の周波数関係には厳密な調波関係からのわずかなずれが存在し (安藤, 1971; Keeler, 1972), これが楽器音の自然さに寄与していることがわかっている (山口, 新実, 竹田, 1970).

音色に関する研究では、刺激の音圧や持続時間は一定にして実験が行われることが多いが、一方で、音圧や持続時間も音色の鋭さに影響する要因である (難波, 桑野, 加藤, 1974; Ueda and Akagi, 1990).

広帯域雑音の振幅包絡の形状も鋭さに影響する (Ueda and Akagi, 1990). 楽器音のような非定常な部分を含む調波的複合音では、全体的な振幅包絡の変化 (Miller and Carterette, 1975) だけでなく、成分ごとの過渡的な振幅の変化も音色に影響し、とくに音の始まりの部分の変化形状は楽器音の識別に大きく影響する (Berger, 1964; Saldanha and Corso, 1964; Beauchamp, 1966; Strong and Clark, 1967; Wedin and Goude, 1972; Grey, 1977, 1978; Grey and Moorer, 1977; Charbonneau, 1979; 安藤, 山口, 1983).

正弦振幅変調音の変調度は「粗さ (roughness)」と関係する (Terhardt, 1974). 特に、臨界帯域ごとの振幅包絡変動の大きさが全体の粗さを決定する.

同一の臨界帯域内に二つの成分が存在する場合、二成分の周波数が一致していれば澄んだ音色の、協和度の高い音に聞こえるが、周波数がわずかにずれるとうなりが生じ、次第に濁った音色の、不協和な音になる. さらに周波数が離れると、うなりが生じなくなり、濁りも減少していく (Plomp and Levelt, 1965). 二つの調波複合音が存在する場合は、このようなことが互いの成分音どうしの間で起こり、全体として二音の基本周波数が単純な比になるときほど協和度が高くなる. そして、二音の協和度は物理量から予測可能である (Kameoka and Kuriyagawa, 1969a, b).

1.4 音の高さと音色の関係

音色は多次元的な性質を持つ心理属性であり、このことは定常音でも (Lichte, 1941; Plomp, Pols, and van de Geer, 1967; von Bismarck, 1974a; Plomp, 1976; de Bruijn, 1978; 大串, 1980), 非定常音でも (Solomon, 1958; 曾根, 城戸, 二村, 1962; Kitamura, Namba, and Matsumoto, 1968; Miller and Carterette, 1975; Grey, 1977; Grey and Gordon, 1978; 厨川, 八尋, 柏木, 1978a, b; Bloothoof and Plomp, 1988), さらに、音色を表現することばの分析でも (上田, 1988) 確かめられている. そして、スペクトル構造は音色に大きな影響を与える物理的要因である.

一方で、音の高さの判断には音の周期だけでなく、スペクトル構造も影響する. 特に、雑音のようにはっきりした周期を持たない音の場合はもっぱらスペ

クトル構造によって高さが決まる．このように音の高さや音色と物理量との関係は，単純な一対一の関係にはなっておらず，多次元的，複合的で，それぞれ複数の物理量が関係し，なかにはスペクトル構造のように音の高さと音色の両方に影響する物理量もある．

1.5 音声と非音声：用語の定義

日本工業規格 (1986) では，「音声認識」などの複合的な用語の定義はあるものの，「音声」という用語それ自体は定義されていない．これは，「音声」という用語が自明のもので定義する必要がないというよりはむしろ，意味があいまいで，文脈によって意味が異なり，定義が困難であることを示すと考えられる．ここでは，本研究で使用する範囲での定義として，「音声」と「非音声」を次のように定義する．

定義：音声 (speech)

- (1) 人間の発話器官で言語として生成された音．
- (2) (1) と同等の物理的特性を持つ音．
- (3) (1), (2) の音として知覚される音，または言語として知覚される音．

定義：非音声 (non-speech)

音声以外の音の総称．

「音声」の定義として，もっとも基本的なものは (1) にあげたものであろう．(2) の定義は，録音・再生された音声や，計算機で規則合成された音声を含む．人間にとって，(1) と区別できない音は，やはり「音声」とみなすべきであろう．また，(3) はきわめてあいまいな定義であるが，一方で「音声」の本質的な一側面を表しているとも言える．それは，(1) の音とはかけ離れた物理的特性を持つ音でも，音声として知覚される場合があるからである．

たとえば、オウムや九官鳥の鳴き声は、人間の音声とは物理的特性が大きく異なり、音色も違うので人間の音声とは区別されるが、それでもやはり音声として（ことばを）知覚可能である。また、同じ音が前後関係や聴取者の構えによって音声として知覚されたり、非音声として知覚されたりする場合もある。たとえば、人間の発話した音声进行分析して、フォルマント（スペクトルでエネルギーの集中している部分）の周波数変化を調べ、三つのフォルマント周波数の変化パターンを正弦波でなぞる形で合成した、音声としてはきわめて不自然な合成音を被験者に聞かせた実験がある（Remez, Rubin, Pisoni, and Carrell, 1981）。それによると、刺激の性質について何も知らされなかった被験者では、この音を音声として自発的に知覚することはほとんどなかったが、「計算機合成された音声で語られた文章」を聞き取るように教示された被験者は、もとの文章をかなり正確に聞き取ることができた。

このように、音声として知覚されるための刺激手がかりが減少すると、「音声」として知覚されるか、「非音声」として知覚されるかは聴取者側の条件によって左右されやすくなる。

1.6 本研究の目的

本研究の目的は、音の高さの知覚と短期記憶に他の聴覚属性が与える影響について調べることである。まず、音の高さの知覚には音色の一部が影響し、両者の区別があいまいな例が存在することを示す。また、非音声の音の高さの短期記憶に数詞音声系列が影響するのかどうかを、短期記憶における干渉効果に基づいて明らかにする。これらの知見をもとに、音の高さと音色、非音声の高さの短期記憶と音声の関係するメカニズムについて考察する。

1.7 本論文の構成

本論文では、第 1 章において、音の高さと音色の用語の定義と、音の高さと

音色の物理的要因に関する過去の研究を紹介した。さらに、音声と非音声の用語を定義し、本研究の目的について述べた。第 2 章では、音の高さが単一の心理的次元ではなく、音色の一部を含んだ多次元的なものであることを明らかにし、被験者の音の高さの判断結果が螺旋形の図式で表現されることを確かめる。さらに、第 3 章で非音声の高さの短期記憶と数詞音声系列の短期記憶メカニズムが特殊化されていないことを、短期記憶における両者の干渉効果から明らかにする。そして、第 4 章で音の高さと音色、非音声の高さの短期記憶と音声系列の関係を説明するモデルについて考察する。第 5 章は結論である。

第 2 章 音の高さの二面性の空間的表現

本章では、音の高さには二つの側面があり、それらを螺旋形の図式で表示できることを被験者の高さ判断のデータに基づいて示す。

2.1 音の高さの二面性に関する研究

通常、「音の高さ (pitch)」は音の心理的属性の一つと考えられており、高さの感覚は独立した、単一の次元であると考えられている。しかし、高さの感覚は、少なくとも二つの要素から成っているとする考えも古くから示されてきた。

2.1.1 音の高さの二面性

Meyer (1904) は、音の高さの感覚は「高さ (pitch)」と「質 (quality)」の二つの要素に分けられると主張した。「高さ」はこの場合、音程の判断ができる心理的次元をさし、「質」は音色のことをさしていた。

Révész (1913) も音の高さには二つの要素、すなわち「高さ (Höhe)」と「質 (Qualität)」が含まれていると主張した。この用語は Meyer のものと似ているが、意味は反対であった。つまり、Meyer のいう pitch は Révész のいう Qualität であり、Meyer の quality は Révész の Höhe であった。Révész は、絶対音感を持つ人と持たない人の高さの記憶の違いを説明するためには、この区別が必要であると考えた。

Bachem (1950) は、(1) 音の高さはトーンハイト (tone height) とトーンクロマ (tone chroma) の二つの要素から成る、(2) トーンハイトは、音楽で用いられる音階の範囲内では——極めて低い（基本）周波数や、極めて高い（基本）周波数の音は音楽では用いられない——周波数の対数関数で近似できる、(3) トーンクロマは、周波数の対数（2 を底とする）の周期関数で表すことができると主張した。

本研究では、トーンハイトとトーンクロマという用語を次のような意味に定義する。

定義：トーンハイト (tone height)

音の高さの直線的な要素を表す用語。

定義：トーンクロマ (tone chroma)

音の高さの循環的な要素を表す用語。ここで「循環的」とは、オクターブ関係にある音の高さが類似している（同じ階名で表される）ことを指している。

2.1.2 無限音階複合音の循環性

音の高さの相対的な判断に循環性が見られる場合があることは、Shepard (1964) によって示された。彼は、計算機合成した複合音を用いて無限音階を作り出した。この複合音のスペクトルを Fig. 2.1 に示す。各複合音は 10 個の成分を含み、各成分の周波数間隔はオクターブであった。この音のスペクトル包絡は 5 番目の成分付近にピークがくる、ベルのような形をしており、すべての複合音について周波数軸上でピーク位置が固定されていた。したがって、複合音の各成分の周波数が上昇すると、低周波側の成分の振幅は次第に強くなるが、高周波側の成分の振幅は次第に弱くなり、ちょうどオクターブ分、各成分の周波数がシフトすると、もとの複合音と全く同じスペクトル構成に戻る。このような音を順に呈示すると、高さが無限に上昇し続けるような印象が得られる。この音階のもたらす印象は、Fig. 2.2 の無限階段の錯視になぞらえられている。

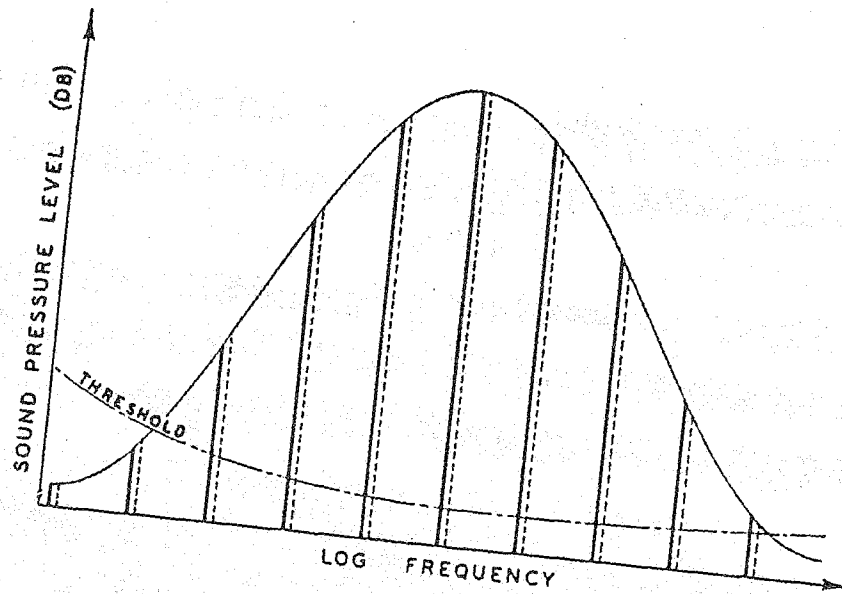


Fig. 2. 1 Spectrum of endless scale sound invented by Shepard. Components were spaced at octave intervals. The dotted lines correspond to an upward shift of the frequencies of all components. From Shepard (1964).

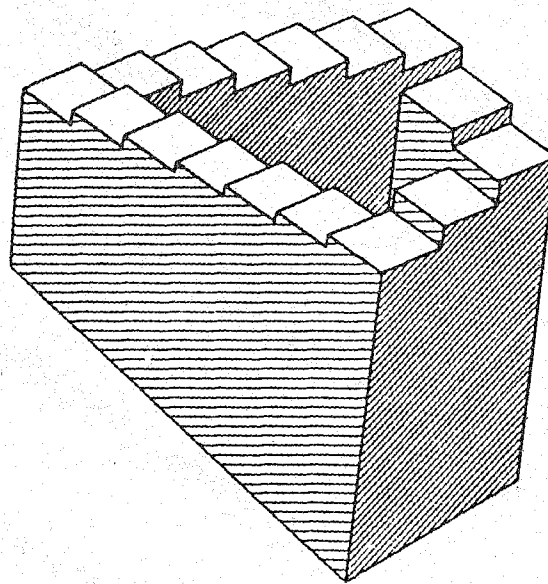


Fig. 2. 2 "Circular" staircase illusion. From Shepard (1964).

この現象は、複合音の成分間隔や、スペクトル包絡の形を変化させたより詳細な実験によってさらに確かめられた。それによれば、高さの判断に循環性が生じるためには、対数周波数軸上で等間隔の成分が、中音域にある程度の数存在すればよく (Pollack, 1978), その間隔はオクターブでなくても良い (Burns, 1981)。また、周波数成分の間隔がオクターブよりわずかに広い複合音と、その複合音全体の周波数を倍にした複合音をこの順序で呈示すると、大多数の人はあとの音の方が低く聞こえる (Risset, 1971, 1986; Schroeder, 1986)。

スペクトル包絡を台形にし、純音の長三和音を 3 オクターブにわたって繰り返したような成分周波数を持つ、9 成分からなる複合音で無限音階を作った試みもある (寺西, 津村, 1983; Teranishi, 1986; Nakajima, Tsumura, Matsuura, Minami, and Teranishi, 1988)。このような複合音でも、たいていの被験者で高さの判断に循環性が生じる。

2.1.3 無限音階複合音と音の高さの二面性

Shepard の複合音を変形した音を用いて、音の高さの概念のあいまいさが示された (Risset, 1969, 1971, 1978)。成分周波数を連続的に下降させ、同時にスペクトル包絡のピーク位置の周波数を連続的に上昇させて、両者の動く方向が逆向きのグリッサンドが作られた。このような刺激の高さが上昇するのか、下降するのかを判断させたところ、被験者によって判断が異なった。

Charbonneau and Risset (1973) は、無限音階複合音の高さの判断結果を循環的な図形で表現できることを示した。彼らは Shepard の複合音に類似した音を用いて、対にした刺激間の高さの違いを 0 から 10 の数字で被験者に答えさせ、その結果を多次元尺度法 (Multidimensional scaling: MDS) で分析した。すると、半音階に相当する刺激が平面で円を描くように順に並べた形に表示された。さらに Charbonneau and Risset (1975) は、成分周波数を固定し、スペクトル包絡のピークを移動させた刺激や、成分周波数を下げると同時に、スペクトル包絡のピークを高い周波数にシフトさせた刺激を用いて対比較実験を行い、被験者に高さの判断 (後の音が前の音よりもより高いか、低い、同じ高さか) を求め

た。判断結果から刺激間の心理的距離を求めて MDS で分析した結果、スペクトル包絡のみが移動する刺激では直線的な布置を、成分周波数とスペクトル包絡のピークが逆方向に移動する刺激では歪んだ螺旋状の布置を得た。

大串 (1981, 1984), Ohgushi (1985) は Shepard の複合音を変形し、成分周波数の上昇とともにスペクトル包絡の周波数も上昇する刺激を作成し、対にした刺激の高さの判断（どちらが高いか）を被験者に求めた。その結果を MDS によって分析し、刺激布置の 2 次元解を求めた。そして、被験者によって直線、一回転の円形、馬蹄形といった異なる刺激布置が得られることを示した。このことから、音の高さには直線的な次元と循環的な次元があり、両者へのウェイトの違いにより被験者の判断に個人差が生じると考えた。

上田 (1983) は、大串 (1981, 1984), Ohgushi (1985) と同じ刺激を使い、類似した実験を行った。この実験では、被験者は対呈示された刺激の高さの隔たりを数字で答えた。二つの実験結果は、大変よく似ていた。

2.1.4 音の高さの螺旋モデル

さて、音の高さを螺旋モデルで表示しようとする試みは、1846 年の Drobish にさかのぼる (Drobish, 1855; Ruckmick, 1929)。彼は、オクターブの関係にある音の高さは類似しているので、その類似性を螺旋の 1 回転をオクターブに対応させることで表現しようとした。Révész (1913) は、この考え方を支持し、音の高さの二要素が螺旋で同時に表現できると考えた。

Ruckmick (1929) は Drobish の考えを受け継ぎ、「音の釣り鐘 (the tonal bell)」というモデルを考え出した [Fig. 2.3 (a)]。このモデルでは、音の「幅 (breadth)」が螺旋の半径の大きさと、純音の高さの弁別閾の周波数による変化を螺旋のピッチで表している。縦線は、オクターブ関係にある音が同じ名前と呼ばれることを表す。

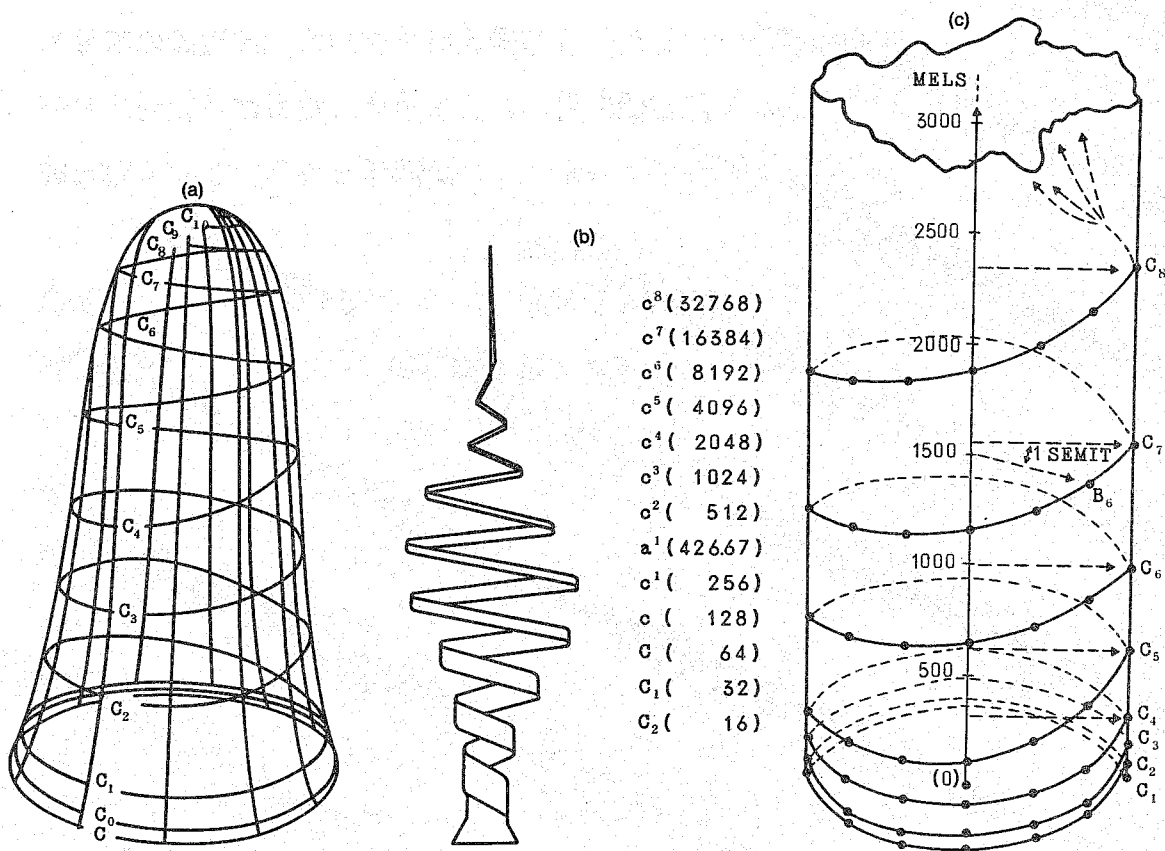


Fig. 2.3 Helicoid models of pitch. (a) The tonal bell (from Ruckmick, 1929); (b) *oto-no-horin* (from Yatabe, 1962); (c) pitch helix (from Ward and Burns, 1982).

音の「幅」は音色の属性の一部と考えられるが、Ruckmick は音の「幅」の周波数による変化を測定したわけではなく、螺旋の半径の大きさとの対応は見かけ上の効果をねらった、主観的なものといえる。そのうえ、音の「幅」とトーンクロマの次元における刺激間の心理的距離とが対応しているのかどうかは不明である。また、トーンハイトの次元は純音の周波数弁別閾で尺度化されているが、このモデルはトーンクロマの上限で終わっており、その上の周波数の、一般にトーンクロマ（オクターブ類似性）は感じられないが、周波数の弁別は可能な領域が表現されていない。したがって、このモデルは、刺激間の高さの心理的距離を表すものではなく、音楽に用いられる限られた周波数範囲での音色の属性と純音の周波数弁別閾、およびオクターブ関係を組み合わせて図式的に示したものというべきである。

矢田部 (1962) は、これによく似た「音の法輪」というモデルを提案した [Fig. 2.3 (b)]。このモデルでは、純音の周波数弁別閾は螺旋の半径の大きさと表されており、モデルの上限と下限は可聴閾と対応している。しかし、ここでも、モデルと刺激間の高さの心理的距離との対応は考慮されていない。

Ward and Burns (1982) も音の高さを螺旋で表現できると考えている [Fig. 2.3 (c)]。垂直軸に対する螺旋上のある点（ある音の高さ）の射影は、その点のトーンハイトをメル尺度上で表し、水平面に射影された中心角はトーンクロマを表す。トーンハイトが約 2500 mels (5000 Hz) 以上の音は、次第にトーンクロマが感じられなくなるので、このモデルの上側の境界はあいまいに作られている。メル尺度は、純音の周波数と高さの関係を分割法で尺度化したものであり、被験者は「音程では判断しないように」と教示された (Stevens and Davis, 1938)。このようなメル尺度とトーンハイトとの関係は実証されておらず、心理的距離との対応も明らかではない。

これらのモデルは、基本的に直観と心理物理学的なデータを組み合わせて螺旋形のモデルで表現したものと言える。これらとは違った観点からの研究として、刺激間の心理的距離を考え、音楽的文脈のなかでの高さを構造的に表現する試みがある (Shepard, 1982a, b)。これは、トーンハイトを表す直線、トーンク

ロマを表す円に加えて、5 度の円、長 3 度の関係を表す四角形など、全部で七つの構造成分を考え、これらを組み合わせて多次元的な構造を作るものである。たとえば、トーンハイト、トーンクロマ、5 度の円を組み合わせれば螺旋管のまわりを巻く二重螺旋が得られる。そして、構造成分のなかのどれを音楽上重要であると考え、より重視するかによってさまざまな組み合わせが可能であり、それによって得られる構造も変わってくる。

しかし、本研究で対象とする「音の高さ」は、音楽的文脈のなかでの高さではない。むしろ、そのような文脈抜きで、二つの音の高さを比べてどちらが高いかを判断するような場合に使われる「音の高さ」の概念を対象としている。したがって、Shepard のモデルについてはこれ以上触れないこととする。

2.2 実験の目的

音の高さの螺旋モデルに関して、これまでに提唱されたモデルはいずれも被験者が音の高さを判断した結果を直接反映するように作られたものではなく、これらのモデルが本当に音の高さの二面性を表しているのかどうかは不明である。

また、Risset (1971) は音の高さの概念にあいまいさがあり、トーンハイトとトーンクロマのどちらの次元に重きをおくかによって高さの判断に個人差が生じると述べているが、定性的な記述にとどまっている。

以下に述べる実験は、音の高さの二面性が螺旋で表現されるのかどうかを、被験者の音の高さの判断結果から多次元尺度法 (MDS) を用いて確かめ、また、音の高さの判断の個人差が、音の高さの二つの心理的次元 (トーンハイトとトーンクロマ) へのウェイトの違いで表現されることを MDS の INDSCAL モデルを用いて定量的に示すことを目的としている。

2.3 刺激の選択

日本工業規格 (1986) では、「音の高さ」を一次元的な（直線的な）心理的次元として定義している（1.1 参照）。これは、純音の高さを判断する際には、直線的な要素が循環的な要素よりも強いためであると考えられる。しかし、純音はどちらかといえば特殊な音であり、音声をはじめとしてわれわれが日常よく耳にする音は調波複合音であることが多い。そして、調波複合音の高さを判断する際には、循環的な要素がより重視されると考える。このように考えることの根拠は、純音と調波複合音に見られる協和性の違いにある (Plomp and Levelt, 1965; Kameoka and Kuriyagawa, 1969a, b)。

協和とは、二つ以上の音が同時に鳴らされたときに、よくとけあって聞こえることを指す。とけあって聞こえることは、複数の高さに分離しない、または分離しにくく、一つの高さに聞こえやすいこと、すなわち、それらの音の高さが類似していることを意味する。二つの純音の協和度は、二音が同じ周波数のときが最大で、そこから二音の周波数を少しずつ離していくと次第に協和度が低下し（不協和になり）、約半音ずれたときにもっとも協和度が低くなる。さらに周波数を離していくと今度は次第に協和度が高くなっていく。その間の協和度の変化の様子は連続的であり、特定の音程がその周囲の音程と比べて特に協和度が高いということはない。したがって、半音以上の音程を考えれば、純音の高さの変化は「カテゴリーカルな判断や学習による効果 (Allen, 1967) を除いて考えれば」単調であり、特定の音程で隔てられた二音の高さが特に類似しているということはないと考えられる。

一方、調波複合音は、調波関係にある複数の純音成分から構成される複合音であり、二つの調波複合音の不協和度は成分どうしの不協和度を重み付けして加算することによって求められる。このようにして求めた不協和度は心理実験の結果をよく表す (Fig. 2.4)。

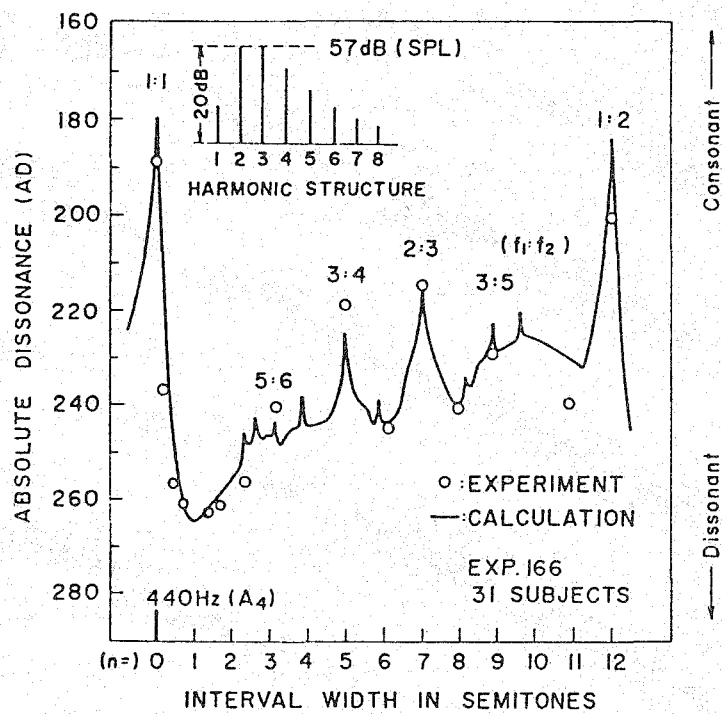


Fig. 2. 4 Comparison between calculated and psychophysically determined degrees of consonance. From Kameoka and Kuriyagawa (1969b).

さて、調波複合音の場合、オクターブずれた二音の協和度は、全く同じ高さの二音の協和度とほぼ等しい。すなわち、きわめて協和度が高い。そして、このような協和度の高い関係は、純音の場合とは違って、特定の周波数比の場合のみに見られる。また、調波複合音の成分構成を変化させた場合でもオクターブのみが一貫して協和度が高い (Kameoka and Kuriyagawa, 1969b)。したがって、オクターブの関係にある二つの調波複合音の高さの類似性は、飛び抜けて高いと考えられる。そして、この高さの類似性が高さの循環的な判断を引き起こすと考えられる。

そこで、被験者が音の高さの循環的な側面を知覚する可能性を高めるため、調波複合音を刺激として用いることを考えた。特に、オクターブにおける循環性を強調するため、Shepard の無限音階複合音を変形した音を電子計算機で合成し、刺激とすることにした。

2.4 刺激

刺激の合成方法は次の通りである。

$t_{\max}$ を成分周波数の組の総数とする。一つの複合音に $c_{\max}$ 個の成分が含まれるとき、 $(t+\alpha)$ 番目 ($t=1,2,\dots,t_{\max}; \alpha=0,t_{\max},2t_{\max}$) の複合音の c 番目の成分周波数は次のように定義される。

$$f(t+\alpha, c) = f_{\min} 2^{[(t-1)/t_{\max} + (c-1)]} \quad (2.1)$$

ここで、 $f_{\min}$ は $(1+\alpha)$ 番目の複合音の最低成分周波数である。各成分の音圧 (L) は次のように定義される。

$$L(t+\alpha, c) = L_{\min} + (L_{\max} - L_{\min}) \times [1 - \cos \theta(t+\alpha, c)] / 2 \quad (2.2)$$

ここで、 $L_{\min}$ と $L_{\max}$ はそれぞれ最小および最大音圧を表す。関数 θ は、

$$\alpha = 0 \text{ のとき, } \theta(t + \alpha, c) = 2\pi c / c_{\max} \quad (2.3)$$

$$\alpha = t_{\max} \text{ のとき, } \theta(t + \alpha, c) = 2\pi(c - 1) / c_{\max} \quad (2.4)$$

$$\alpha = 2t_{\max} \text{ のとき, } \theta(t + \alpha, c) = 0 \quad (c = 1) \quad (2.5)$$

$$\theta(t + \alpha, c) = 2\pi(c - 2) / c_{\max} \quad (c \geq 2) \quad (2.6)$$

によって定義される。

本研究では $t_{\max} = 6$, $c_{\max} = 10$, $f_{\min} = 4.863$, $L_{\min} = 22$, $L_{\max} = 56$ という値を用い、複合音を合成した。全部で 18 個の複合音が合成された。成分周波数が同じで、スペクトル包絡のピーク位置が異なる複合音は 3 個ずつある。複合音のスペクトルの例を Fig. 2.5 に示す。

各刺激は、他の刺激と可能なすべての組み合わせで対にされた。刺激対はランダムに並べられ、フルトラック、モノラル、9.5 cm/s でオープンテープに録音された。刺激対の総数は 153 対、各刺激の持続時間は 400 ms、各刺激対の刺激間間隔は 300 ms、試行間間隔は 2 s であった。

テープの再生スピードを 9.5 cm/s としたときの刺激を A 群、19 cm/s としたときの刺激を B 群と呼ぶ。B 群の刺激を用いたのは次のような理由からである。A 群の刺激では、刺激の最高周波数成分は 2.5-5.0 kHz 付近になるが、この周波数帯域では耳の感度が比較的高いため (Sivian and White, 1933)、最高周波数成分のみに着目して被験者が判断を行う可能性がある。B 群では、最高周波数成分は 5-10 kHz 付近となり、耳の感度が低下するので、このような問題が小さくなる。

テープレコーダーのイコライザは、テープスピードによって周波数特性が切り替わるように作られているが、本研究でテープの再生に使用したテープレコーダー (Nagra 4.2) の周波数特性は、テープスピードを変化させても刺激周波数の範囲内では 1 dB 以下の変化しか起こさなかった。一つの刺激の周波数成分の強さは 34 dB にわたって変化するので、テープスピードを上げたときのスペクトル形状の歪みは無視できる。

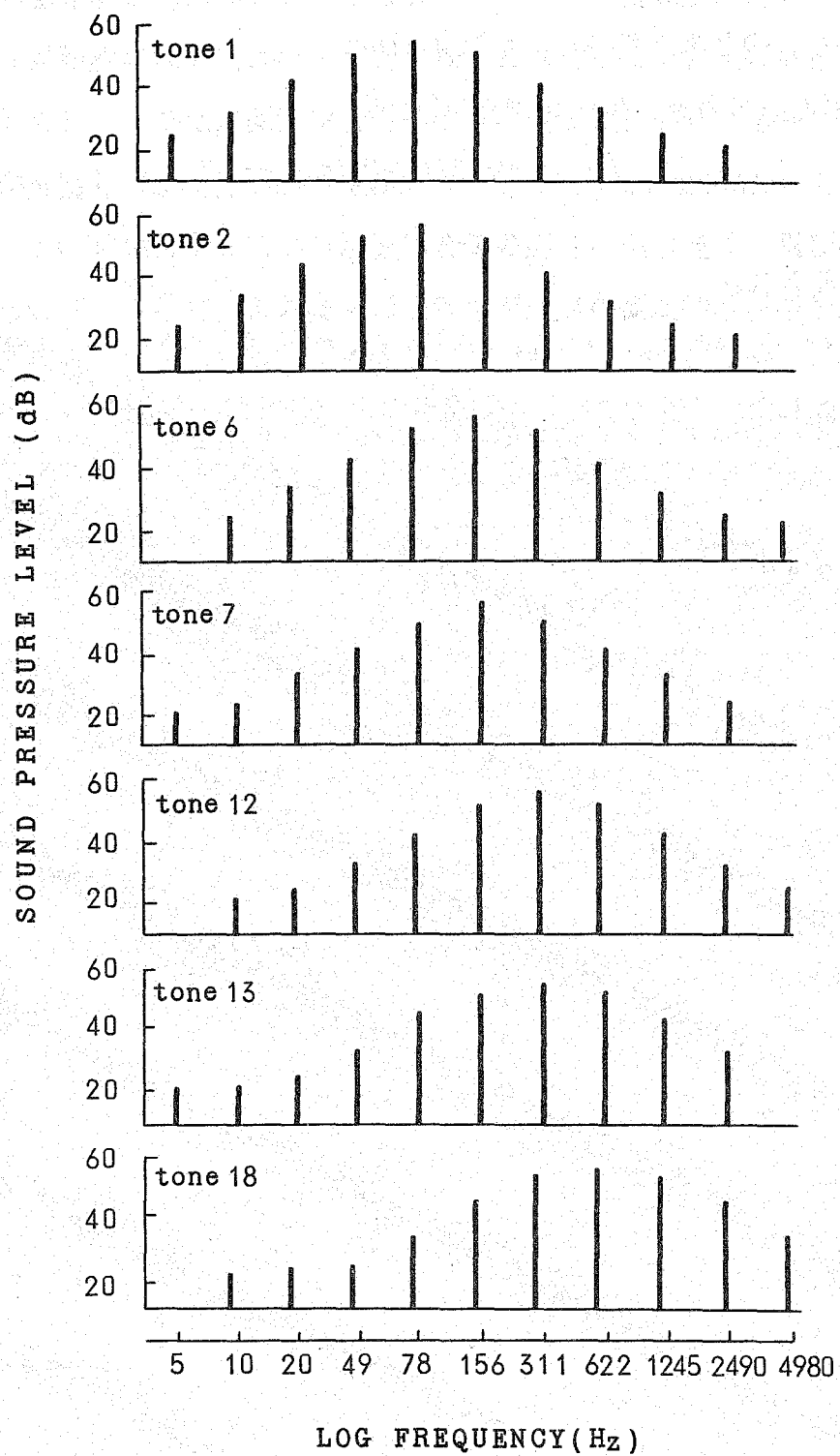


Fig. 2. 5 Spectra of tone stimuli.

2.5 実験手続き

11 名の被験者に防音室内でスピーカからテープを再生し、刺激を呈示した。呈示音圧はおよそ 70 dB SPL であった。すべての被験者に両方の刺激群を呈示した。テープの走行方向を反転させることで刺激の呈示順序の入れ替えを行い、各刺激群ごとに正転と逆転ともに一度ずつ呈示した。約半数の被験者には、両方の刺激群ともに逆転の系列から実験を開始した。どの被験者にも最初に刺激群 A を呈示し、休憩の後に刺激群 B を呈示した。これにより、より短い時間での反応が要求される刺激群 B の実験が支障なく行われることを期待した。

被験者は、試行間間隔中に反应用紙に反応を記入した。被験者には、「刺激対のどちらの刺激がより高く聞こえるかを答えてください。同じ高さに聞こえた場合には、そのように答えても構いません。」という教示を与えた。

被験者はすべて 20 代で、何らかの音楽経験のある者であった。楽器の演奏経験は平均 9.9 年、器楽合奏の経験は平均 1.0 年、合唱経験は平均 3.9 年であった。

2.6 実験結果

実験結果の分析に、非計量的多次元尺度法 (nonmetric MDS) を用いた。このタイプの MDS は、順位尺度以上の水準の入力データに適用され、入力データと刺激布置における刺激間の距離との間には順位尺度水準の関係だけを仮定する (Kruskal and Wish, 1978)。

被験者の判断結果から、次のようにして被験者ごとに 18×18 の行列を作り、得点表の形でデータを整理した。被験者が、ある刺激を他の刺激よりも高いと判断した場合には、その刺激に 2 点を与えた。もう一方の刺激は 0 点とした。被験者が、両方の刺激を同じ高さで判断した場合には、両方の刺激に 1 点を与えた。各対につき 2 回ずつ被験者は判断したので、その結果からこのような方

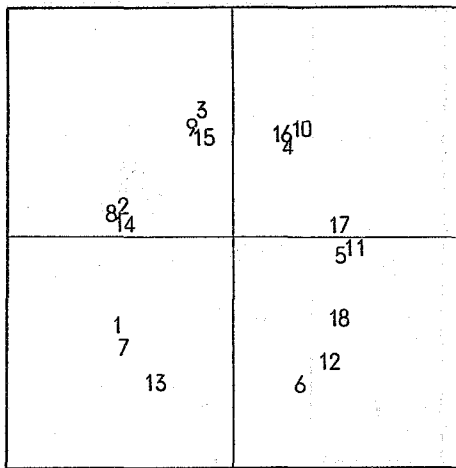
式に従って求めた点数を該当するセルに加算していった。行列の対角成分には、同じ高さで 2 回判断したことに相当する点数として 2 点を入れた。

MDS の入力データは、刺激間の心理的距離を表す距離行列（非類似性行列）である。得点表を距離行列に変換するため、得点表の各行を 18 次元空間の座標とみなし、各座標間のユークリッド距離を求めた。

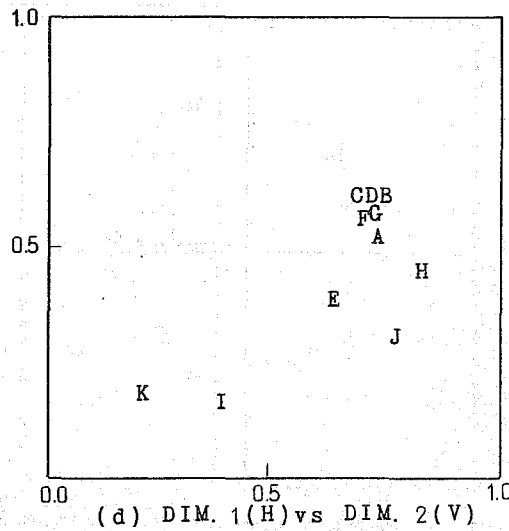
用いた MDS のプログラムは ALSCAL (Takane, Young, de Leeuw, 1977; 高根, 1980) である。まず、すべての被験者のデータをひとまとめにして INDSCAL モデルにより分析した。そして、数名の被験者のデータを個別に分析した。

ストレス値 0.1 以下で 3 次元解が求められた。一般に、ストレス値が 0.1 以下であることは、入力データと得られた刺激布置とのフィッティングが良好であることを示す (Kruskal, 1964)。Figs. 2. 6~2. 9 に、得られた刺激布置と被験者空間を示す。刺激布置 [Figs. 2. 6 (a)-(c), 2. 7 (a)-(c), 2. 8, and 2. 9] の番号は刺激番号を表す。曲線は刺激番号を順に結ぶように、著者が描いた。被験者空間 [Figs. 2. 6 (d)-(f) and 2. 7 (d)-(f)] のアルファベットは被験者を表す。

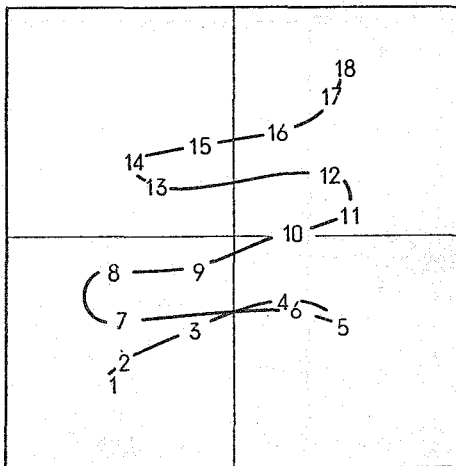
Fig. 2. 6 (a)-(c) および Fig. 2. 7 (a)-(c) の刺激は、螺旋曲線の線上に番号順に布置されている。螺旋のピッチは一定ではなく、刺激番号が増えるにしたがって増加するものとなっている。刺激番号 1-7-13, 2-8-14, 3-9-15 といった刺激の組み合わせは、同じ成分周波数の組み合わせを共有し、スペクトル包絡のピーク位置が異なるものの組み合わせとなっている。これらの刺激の組は、第 1, 2 次元の平面で互いに近く布置されている。この平面では、円周上を刺激が番号順に 3 回転する、循環的な布置が得られている。また、被験者は、図の第 3 次元で上の方向をより高いと判断した。このような特徴は、刺激群 A, B とともに共通している。



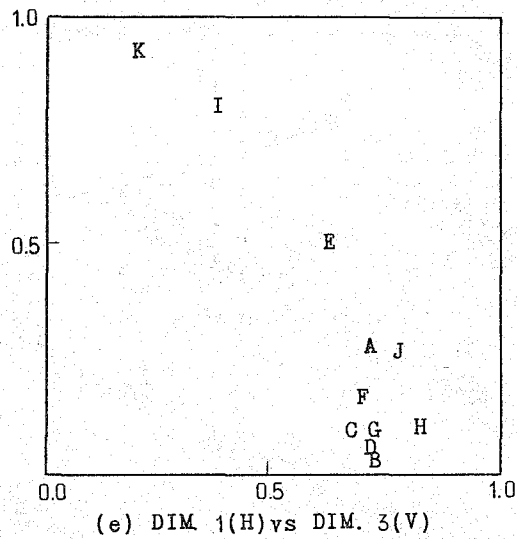
(a) DIM. 1(Horizontal) vs DIM. 2(Vertical)



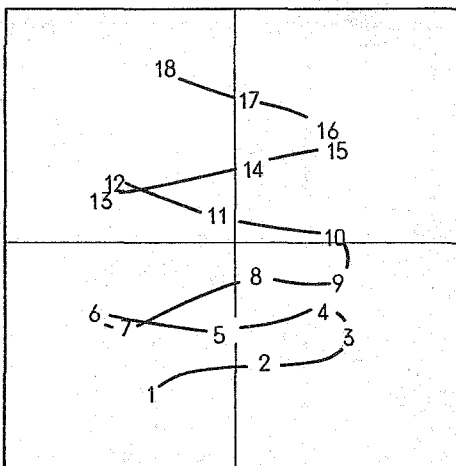
(d) DIM. 1(H) vs DIM. 2(V)



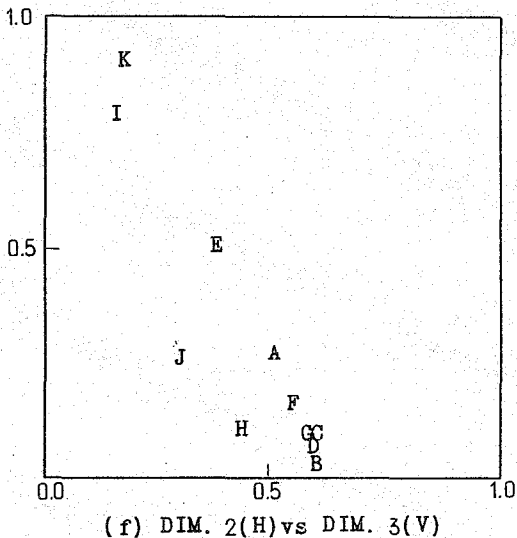
(b) DIM. 1(H) vs DIM. 3(V)



(e) DIM. 1(H) vs DIM. 3(V)

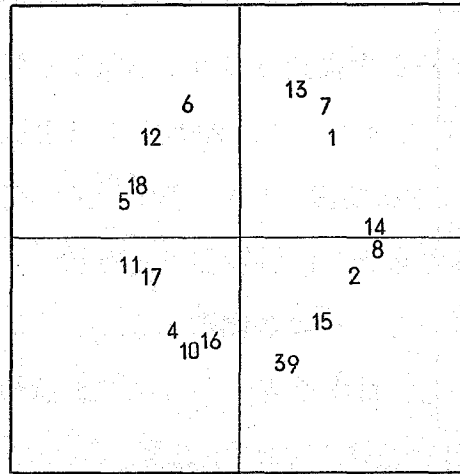


(c) DIM. 2(H) vs DIM. 3(V)

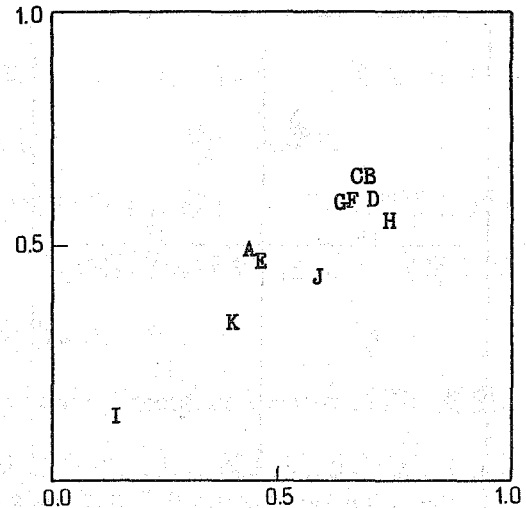


(f) DIM. 2(H) vs DIM. 3(V)

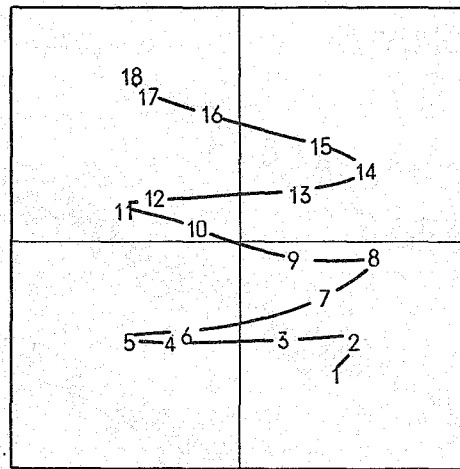
Fig. 2. 6 Configuration derived in three dimensions with the INDSCAL model. Stimulus set A; stress = 0.090. (a)-(c) Stimulus configuration; (d)-(f) subject weights.



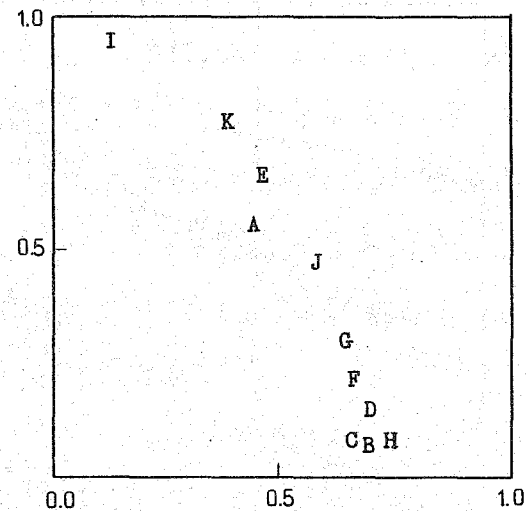
(a) DIM. 1(Horizontal) vs DIM. 2(Vertical)



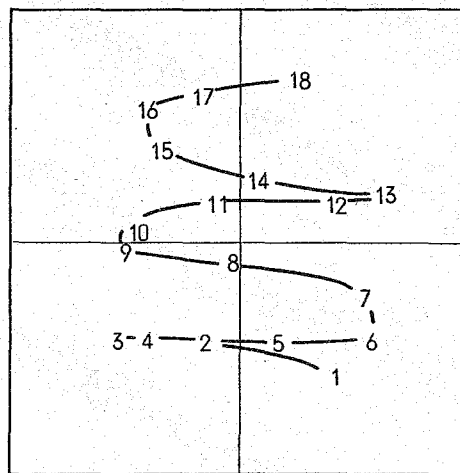
(d) DIM. 1(H) vs DIM. 2(V)



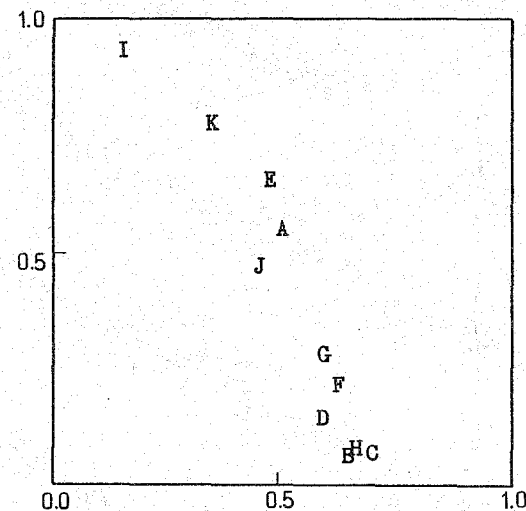
(b) DIM. 1(H) vs DIM. 3(V)



(e) DIM. 1(H) vs DIM. 3(V)

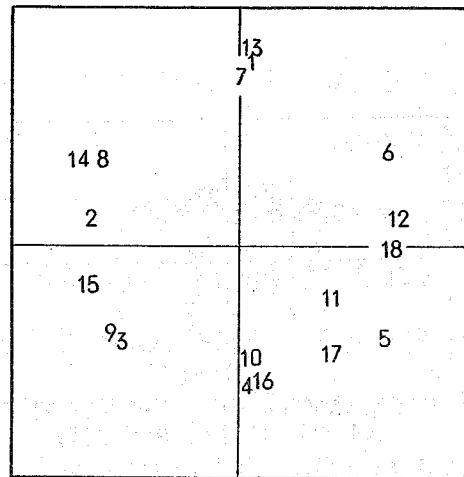


(c) DIM. 2(H) vs DIM. 3(V)

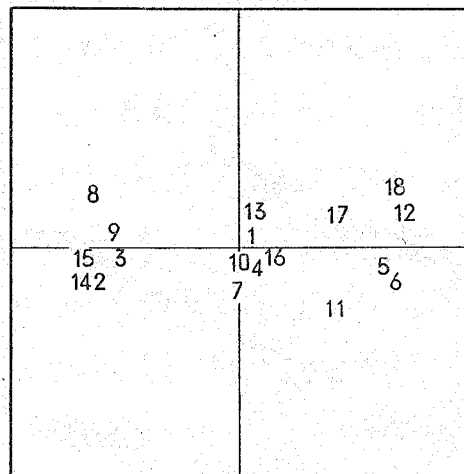


(f) DIM. 2(H) vs DIM. 3(V)

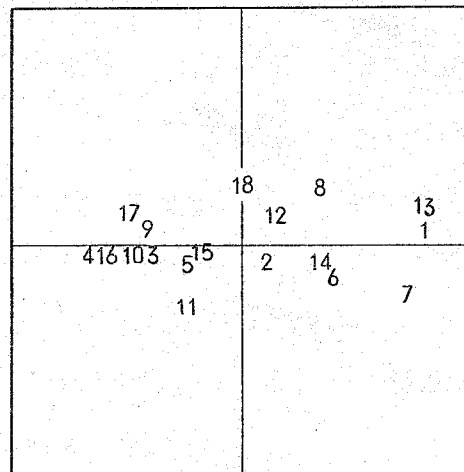
Fig. 2. 7 Configuration derived in three dimensions with the INDSCAL model. Stimulus set B; stress = 0.082. (a)-(c) Stimulus configuration; (d)-(f) subject weights.



(a) DIM. 1(H) vs DIM. 2(V)

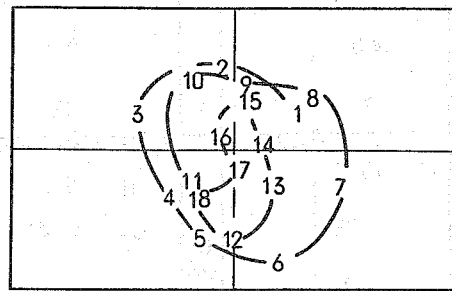


(b) DIM. 1(H) vs DIM. 3(V)

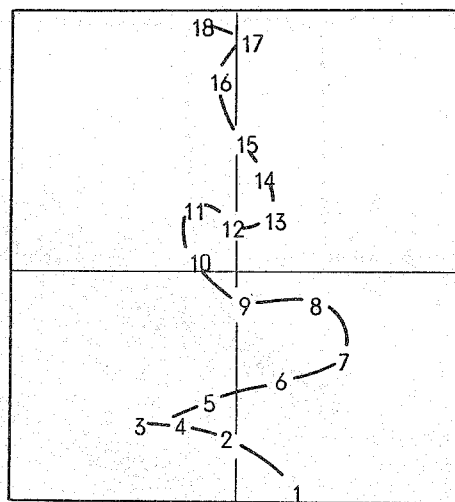


(c) DIM. 2(H) vs DIM. 3(V)

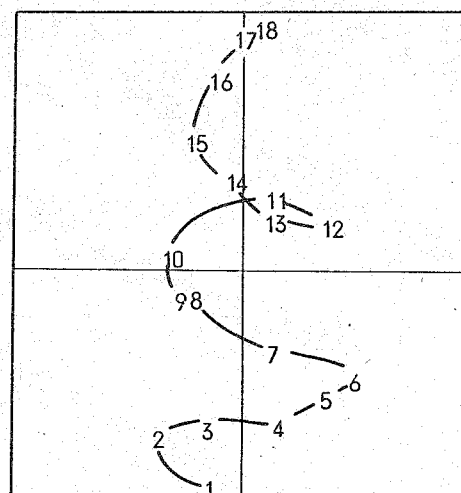
Fig. 2. 8 Stimulus configuration derived in three dimensions. Subject B; stress = 0.012.



(a) DIM. 1(H) vs DIM. 2(V)



(b) DIM. 1(H) vs DIM. 3(V)



(c) DIM. 2(H) vs DIM. 3(V)

Fig. 2. 9 Stimulus configuration derived in three dimensions. Subject I; stress = 0.016.

これらの刺激布置に対応する被験者空間では、被験者の個人差が各次元へのウェイトの違いとして表現されている。どちらの刺激群に対しても、第 1 次元と第 2 次元とのウェイトの関係には正の相関があり、第 1 次元と第 3 次元、第 2 次元と第 3 次元との間には負の相関が見られる。すなわち、被験者の個人差には、第 1, 2 次元の平面に大きなウェイトを置く被験者は第 3 次元へのウェイトが小さくなり、第 1, 2 次元の平面へのウェイトが小さい被験者は第 3 次元へのウェイトが大きくなるという関係が見られることになる。したがって、どれか一つの次元（たとえば、第 3 次元）のみに着目して個人差を考えることもできる。実際、Fig. 2. 7（刺激群 B）において第 3 次元へのウェイトがもっとも小さい被験者 B と、もっとも大きい被験者 I のデータを取りだし、個別に分析すると（Figs. 2. 8 and 2. 9），被験者 B ではほぼ第 1, 2 次元の平面（循環的な次元）だけの布置になり、被験者 I では第 3 次元（直線的な次元）が強調された布置になる。刺激群 A についても同様の結果が得られた。

第 3 次元へのウェイトと楽器等の経験年数との相関係数を調べたところ、楽器演奏、器楽合奏、合唱の経験年数との相関係数は、それぞれ刺激群 A で 0.310, 0.086, 0.376, 刺激群 B で 0.386, 0.091, 0.309 であり、いずれも 5 % 水準で統計的に有意ではなかった。

2.7 考察

実験結果を要約すると、次の 4 点になる。

- (1) 刺激群 A, B による違いは見られなかった。
- (2) 対呈示された刺激についての高さの判断を MDS で分析し、多次元的な、螺旋形の刺激布置が得られた。
- (3) 被験者による判断結果の違いは、第 1, 2 次元の平面と第 3 次元へのウェイトの違いとして表現された。
- (4) 高さの判断結果の個人差と被験者の音楽経験年数には、相関が見られなかった。

2.7.1 刺激群 A, B による違い

刺激群 B は刺激群 A と同じテープを倍速で再生して作ったものである。したがって、刺激群 B は刺激群 A と比べて、すべての成分周波数が 2 倍になると同時に、刺激の持続時間、刺激間隔、試行間隔は $1/2$ になっていた。また、すべての被験者に A, B の順序で実験を行ったため、順序効果も結果に影響する可能性があった。しかし、両者の結果には、刺激布置においても、被験者空間においても、顕著な違いは見られなかった。したがって、これらの条件の違いは被験者の判断の仕方には影響を与えなかったし、また、最高周波数成分のみに着目して被験者が高さの判断を行うこともなかったと考えられる。

2.7.2 二つの心理的次元：トーンハイトとトーンクロマ

被験者は刺激の「高さ」を判断した。通常、「高さ」は 1 次元的なものと考えられている。しかし、得られた刺激布置は、1 次元の、直線的なものではなく、3 次元的な螺旋形のものであった。このことは、音の「高さ」の概念が多次元的な性質のものであることを示している。では、ここで得られた心理的次元とはどのようなものであろうか。

Figs 2.6, 2.7 の第 1, 2 次元の平面に見られる循環的な布置は、この平面全体を一つの心理的次元と考えるべきであることを示している。なぜなら、円形の布置について、その平面上のどれか特定の方向に意味を考えることはできないからである。そして、この平面上で互いに近くまとまって布置されている刺激は、同じ階名の音に聞こえる——ここでは全音音階に相当する音が用いられているので、たとえば、ド、レ、ミ、ファ $\sharp$ 、ソ $\sharp$ 、ラ $\sharp$ の音が三つずつあるように聞こえるといってよい。したがって、この平面は音の高さの循環的な成分である、トーンクロマの次元と解釈できる。

また、第 3 次元は、これにそって刺激が番号順に並べられ、より高いと判断された刺激が図の上の方向に布置されているので、音の高さの直線的な成分、すなわちトーンハイトの次元と考えられる。本研究で操作された刺激の物理的

特性のなかで、このような一次元的な判断の手がかりになりうるものはスペクトル包絡のピーク位置の変化である。一般に、スペクトル包絡の形によって音色が変化し（1.3 参照）、特に高域にエネルギーが集中するほど明るさ (brightness) あるいは鋭さ (sharpness) が増す (Lichte, 1941; von Bismarck, 1974a, b)。したがって、スペクトルのピーク位置の周波数が上昇して明るさ、あるいは鋭さが増すと、「より高い」と判断されやすくなることになる。すなわち、通常は「音色」の一部と考えられている音の属性が、音の高さの判断に影響したことになる。

2.7.3 判断の個人差と螺旋形

螺旋は、回転成分と直線成分を組み合わせたもので、螺旋形の布置は、トーンハイトとトーンクロマの二つの心理的次元を同時に表現したものである。そして、螺旋の直線成分を伸縮させることで、二つの極端な形を考えることができる。一つは螺旋の直線成分の長さを 0 にして、円だけにした形である。トーンハイトの次元を使わずに、トーンクロマの次元だけで高さを判断した場合がこれにあたる。得られた刺激布置のなかでもっともこの形に近いものが、被験者 B の布置 (Fig. 2.8) である。一方、螺旋の高さ方向を伸ばし、底面の円を縮小していけば、一本の直線になる。トーンクロマの次元を使わずに、トーンハイトの次元だけで高さを判断した場合がこれにあたる。完全に直線的な布置で表されるような高さの判断をした被験者はいなかったが、これにもっとも近い形は被験者 I の布置 (Fig. 2.9) である。

さらに、円と直線の両極端の間に、様々な形の螺旋を考えることができ、高さの二つの心理的次元へのウェイトの違いとして高さの判断の個人差を考えることができる。実際、被験者空間 [Fig. 2.6 (d)-(f), Fig. 2.7 (d)-(f)] において、被験者 B と I の間には他の 9 名の被験者がほぼ連続的に存在しており、このような見方が成り立つことを示している。

2.7.4 個人差と音楽経験との相関

では、このような高さの判断の個人差は何によって生じているのであろうか。

音階や階名は音楽で用いられるものであるから、被験者の音楽経験によってトーンクロマにより重きをおいた判断をするようになるのであろうか。それとも、経験によって音色の微妙な差異に着目するようになり、トーンハイトにより重きをおくようになるのであろうか。

先行研究によれば、60名の被験者を使って調べたところ、オーディオマニアだが楽器演奏の経験に乏しい被験者はトーンハイトで高さを判断し、演奏経験の非常に豊富な被験者は、トーンクロマで高さを判断したという (Risset, 1971)。また、トーンクロマの基礎にはオクターブの類似性があるが、純音の類似性を7段階で評定させると、音楽教育を受けた被験者ではオクターブ関係にある音についてははっきりと類似していると判断するが、音楽教育を受けていない被験者ではそのような結果が見られない (Allen, 1967)。

しかし、本研究では、被験者の音楽経験年数と音の高さの心理的次元との間には、はっきりした関係が見いだされなかった。この原因として考えられる可能性は次の三つである。

- (1) 被験者集団の性質が異なったため。
- (2) 刺激と実験方法の違いによる。
- (3) 音楽経験の測定方法に問題がある。

(1) について、Risset (1971) によれば、トーンハイトの次元で高さを判断した被験者は、トーンクロマによる高さの判断が「できなかった」という。本研究の被験者は、すべて何らかの音楽経験があったので全員トーンクロマによる判断ができ、Risset の被験者のような極端な被験者が含まれていなかったため、相関が小さくなったと考えられる。この裏付けとして、トーンハイトの次元へのウェイトがもっとも大きかった被験者の刺激布置でさえ完全な直線ではなく、トーンクロマの次元が含まれていたことがあげられる (Fig. 2.9)。

(2) について、Risset (1971) の取った方法は、連続したグリッサンドを聞かせて、全体に高さが上がったか、下がったかを尋ねるやり方か、成分周波数とス

ペクトル包絡のピーク位置の両方が異なる刺激を対呈示し、高さの変化方向を調べるというやり方であった。このようなやり方では、トーンハイトのみによる高さの違いは結果に現れにくい。本研究では、成分周波数が同じでスペクトル包絡のピーク位置だけが異なる刺激対が含まれており、トーンハイトの違いがより判断しやすかったと考えられる。このような条件では、音楽経験により、むしろ音色の微妙な違いを聞くようになることも考えられる。

この考えを直接裏付ける証拠はないものの、次のようなことは言える。絶対音感を持っているような被験者は、幼少の頃から非常に音楽的な環境で育ったと考えられる (Ward and Burns, 1982)。少なくとも、他の被験者よりは音楽経験が豊富である可能性が高い。本研究の被験者で、被験者 K と J は絶対音感を持っていると申告し、それぞれ 15 年と 19 年の楽器演奏経験があった。そして、2 名ともどちらかといえばトーンクロマだけで判断せず、むしろトーンハイトにある程度ウェイトをおいた判断を示した [Fig. 2.6 (d)-(f), Fig. 2.7 (d)-(f)]。

(3) について、高さの判断に与える音楽経験の影響を調べるには、経験年数のような尺度では不十分であり、経験の質を調べられるような測定が必要だったのかもしれない。しかし実際には、そのような測定は非常に困難である。音楽経験や音楽的能力の意味するところは非常に広く、かつあいまいである。楽譜の読み取り能力や、聞き取りの能力は比較的簡単にテストでき、音楽大学等の入学試験でも用いられているが、ここで問題にするような、音の高さの概念に与える影響をどのようにして測定したらよいのかは今後の課題として残る。

2.7.5 過去のモデルとの比較検討

螺旋モデルと音の高さの心理的距離との関係には、次のようなことがいえる。螺旋モデルは、オクターブ関係にある二つの音の類似性を、螺旋の線上を伝っていく距離よりも二音間の直線距離の方が短いということで表現しているといえる (Shepard, 1982b)。オクターブの類似性がトーンクロマの基礎となるので、直線的な成分と循環的な成分を合わせ持つ螺旋モデルは、音の高さの二面性を表現できることになる。

2. 1. 4 で紹介した、音の高さを螺旋で表現しようとする過去のモデルは、音の高さの二面性を螺旋で表現しようとする点では共通している一方で、刺激間の高さの心理的距離については考慮されていない。そのかわり、各々の心理的次元を、純音の周波数という物理量と心理量との対応を考える、心理物理学の測定と結びつけてモデル化しようとしている。そして、どのような測定とどちらの心理的次元を結びつけるかという点が、提案者の恣意的な判断に任されている。それゆえ、いずれも螺旋形であるという意味では、本研究の結果と一致しているといえるが、それ以上の比較はできない。

本研究で取られたアプローチは、これらの過去のモデルで用いられたアプローチとはまったく異なる。すなわち、オクターブごとの高さの循環性を強調するために特別に合成された調波複合音を用い、刺激間の高さの心理的距離（非類似性）を測定し、MDS による分析を行った。したがって、音の高さの二面性を螺旋で表現することが、研究者の恣意的な判断に基づいたものではなく、被験者の高さの判断に基づいて得られるものであることが実証された。

2.8 まとめ

本章の主な結論は、以下のようにまとめられる。

- (1) Shepard の無限音階複合音を変形した刺激を用いて、刺激の高さの対比較実験を行い、被験者の判断結果を MDS で分析したところ、3 次元空間における螺旋形の刺激布置が得られた。すなわち、被験者の音の高さの判断結果が、螺旋で表現されることが示された。
- (2) 刺激空間は、直線的、音色的なトーンハイトと、循環的、音階的なトーンクロマの二つの心理的次元から成り立っていると解釈された。
- (3) 被験者の判断の個人差は、これら二つの次元へのウェイトの違いとして定量的に表現された。
- (4) 被験者の音楽経験と高さの判断の個人差との間には、はっきりした相関関係が見られず、両者の関係が先行研究で考えられていたほど単純ではない。

ことが示唆された。

第 3 章 非音声の高さの短期記憶と音声の干渉効果

第 2 章で論じた、音の高さの二面性に関する研究から、音の高さには音色の一部が関係していることが明らかになった。すなわち、音の高さと音色はまったく切り離されたものではなく、むしろ、影響し合う面があると考えられる。これは非音声についての研究であったが、第 1 章で述べたように、非音声でも音声でも、高さと音色の両方が知覚上重要な役割を果たしている。したがって、音声についても同様のことが考えられるし、両者の知覚メカニズムや記憶メカニズムには共通する部分があってもおかしくない。音声を知覚する際には、「高さ」や「音色」といった区別をすることなく、言語上の意味はもとより、情緒的な内容までが伝達されることを考えれば、高さと音色の統合的な処理は音声ではより重要であるともいえるだろう。

しかし、一方で、脳における音声の処理（特に音韻の処理）は特殊化されており、非音声処理するメカニズムとは別のメカニズムを必要とするという考え方があり（たとえば、Liberman and Mattingly, 1985）。果たして音声と非音声は切り離された関係にあるのだろうか。本章では、この問題について、特に非音声の高さの短期記憶と数詞音声系列の短期記憶の干渉効果を調べることで検討し、両者の間に干渉が見られること、すなわち両者が切り離されていないことを示す。

3.1 短期記憶における音声と非音声の干渉効果に関する研究

音声の処理は特殊化されているという主張を、もっとも極端な形でモデル化すれば、「音声は専用の特殊化されたメカニズムで処理され、純音などの非音声の処理メカニズムとは無関係である」と考えることができる。この考えを確かめる一つの方法として、聴覚短期記憶において音声と非音声とが干渉するかどうかを確かめる方法がある。すなわち、音声と非音声の記憶メカニズムがそ

それぞれ特殊化されていれば、干渉を起こさないであろうし、特殊化されていないのであれば、干渉を起こすと考えられる。

3.1.1 音声と非音声の短期記憶の特殊化を支持する証拠

Deutsch (1970) は、音声と非音声（純音）の高さは短期記憶で干渉しない、すなわち両者の記憶メカニズムが特殊化されていると考えられる実験データを提出した。

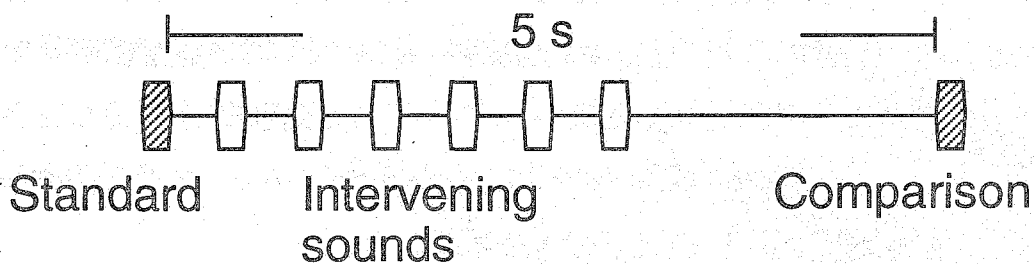


Fig. 3. 1 Schematic diagram showing the stimulus sequence in the Deutsch experiment paradigm (Deutsch, 1970). The intervening sounds were pure tones or spoken numbers.

この実験では、Fig. 3. 1 のような刺激系列を用いてテスト音（標準音と比較音）の間に 6 個の純音または数詞音声の挿入音を挿入し、テスト音の高さの再認識率と、数詞の系列再生誤り率が調べられた (Table 3. 1)。そして、純音を挿入した場合にはテスト音の高さの再認識が妨害されたが、数詞音声の挿入した場合には妨害が見られなかったこと、また、テスト音の高さの再認識課題と数詞の系列再生課題の両方を被験者に遂行させた場合でも、それぞれの課題を単独で遂行させた場合との成績に有意差が見られなかったことが、干渉のない証拠とされた。そのため Deutsch (1975c) は、純音の高さと音声はそれぞれ特殊化された短期記憶で保持されると主張した。

Table 3. 1 Percent errors in the different experimental conditions. From Deutsch (1970).

Condition	Task	
	Pitch recognition	Number recall
A. Pitch recognition with intervening tones ignored	32.3	
B. Pitch recognition with intervening numbers ignored	2.4	
C. Pitch recognition with intervening numbers recalled	5.6	25.3
D. Number recall with no pitch recognition required		27.4

Pechmann and Mohr (1992) は, Deutsch (1970) とよく似た条件を含む実験を行った。彼らは被験者を楽器を演奏するかどうかによって音楽的被験者と非音楽的被験者の 2 群に分けた。被験者は絶対音感を持っていなかったと報告されているが, フォーマルなテストをしたという記述はない。

テスト音に純音, 挿入音に男性話者の単音節語を用いてテスト音の高さの再認識率が測定された。音楽的被験者では音韻課題 (最後の二つの語が韻を踏んでいるかどうかを判断する) を挿入しても高さの再認は妨害されなかった (他の条件との有意差が生じなかった) が, 非音楽的被験者では妨害が生じた。この結果を彼らは, 「調音ループ (articulatory loop)」と「音ループ (tonal loop)」の二つのループを持つ作動記憶 (working memory) モデルで説明しようとし, 各々のループの動作に要する注意の負荷が被験者群によって異なるのではないかと考えた。

3.1.2 テスト音と挿入音の高さ関係の重要性を示す証拠

Semal and Demany (1991, 1993) は純音および調波複合音を挿入音として用いた。そして, テスト音と挿入音の高さの隔たりの程度が, テスト音の高さの再認識率を支配し, 両者の高さがある程度以上離れると誤り率が減少することを示した (Semal and Demany, 1991)。さらに, 彼らは挿入音のスペクトル構成や音圧, 振幅包絡の形状を変化させて実験を行った。スペクトル構成や振幅包絡

形状の違いは音色に影響し、音圧の違いは大きさに影響する。しかし、いずれの要因も、高さの明確さに影響しないかぎりテスト音の高さの再認識率にはほとんど影響しない (Semal and Demany, 1993)。たとえば、挿入音が純音であっても調波複合音であっても、高さが等しければ高さの再認識への干渉効果はほぼ同じである。また、テスト音と挿入音のスペクトルが重なるかどうかではなく、主観的な高さの範囲が重なるかどうか重要である。基本音の欠如した調波複合音を聞くと、基本音の高さに相当する高さが知覚される (1.2.2 参照)。このような音を挿入音として用いた場合、テスト音と挿入音との主観的な高さの関係で干渉効果が決定される。テスト音と挿入音の高さの関係が、テスト音の高さの再認識成績に影響することは Deutsch の一連の研究 (Deutsch, 1973a; Deutsch, 1973b; Deutsch, 1974; Deutsch, 1975a; Deutsch, 1975b; Deutsch and Feroe, 1975) によっても示されている。

3.1.3 テスト音と挿入音の体制化の影響を示す証拠

テスト音の高さの再認識率に影響するのは、テスト音と挿入音の高さ関係だけではない (Jones, Macken, and Harries, 1997)。たとえば、両者の時間的近接の度合いを操作することにより、テスト音と挿入音を一つの「流れ (stream)」として体制化させるか、二つの流れとして分離して体制化させるかによって、テスト音の高さの再認識率が影響される。すなわち、両者を分離して体制化できる条件では、誤り率が低下する。同様のことが、両者の空間的位置 (どちらの方向から聞こえてくるか) や、音色 (同じ音色か、別の音色か) による体制化によっても生じる。また、高さの近接による挿入音の体制化によっても誤り率は低下する (Deutsch, 1978a)。空間的位置による分離が誤り率の低下に寄与することは Deutsch (1978b) によっても示されている。これらの結果は、Bregman (1990) の言う、音の流れの分離 (stream segregation) が高さの再認識率に影響することを示す。

3.2 短期記憶の特殊化を支持する証拠への疑問

3.1.1 で述べた Deutsch (1970) の実験で、純音を挿入音とした場合のテスト音と挿入音の高さの変化範囲は完全に重なり合っていた (Fig. 3.2) . 3.1.2 で紹介したその後の研究結果から考えても、この条件での誤り率が高いという結果は妥当であろう。ところが、数詞音声の挿入音とした場合については、数詞音声の話者の性別 (男性と女性とでは声の高さが約 1 オクターブ異なる) や高さについての記述が Deutsch (1970) の論文には全くない。そして、テスト音の周波数は通常の音声の高さの範囲よりも高い (274~517 Hz, Fig. 3.3) . そのため、テスト音と挿入音の高さの関係の重要性、また、両者の群化の重要性を示した研究 (3.1.2, 3.1.3 参照) を考慮すると、得られた干渉効果の差が音声と非音声の記憶メカニズムの特殊化を反映したものなのか、高さの差によるものなのかは Deutsch (1970) の実験結果からは分らない。

同様の問題点は、Deutsch (1970) の実験の再現性を確かめたと主張する、Pechmann and Mohr (1982) についても指摘できる。彼らの実験では、テスト音の純音の周波数範囲は 262-494 Hz、純音の挿入音の周波数範囲は 185-698 Hz、音声は男性話者の単音節語であった。音声の高さについての記述はないが、通常の男性話者の音声の高さから考えると、テスト音も純音の挿入音も、音声より高いといえる。したがって、彼らの実験からも、音声と非音声の記憶メカニズムが特殊化されているかどうかは明らかではない。

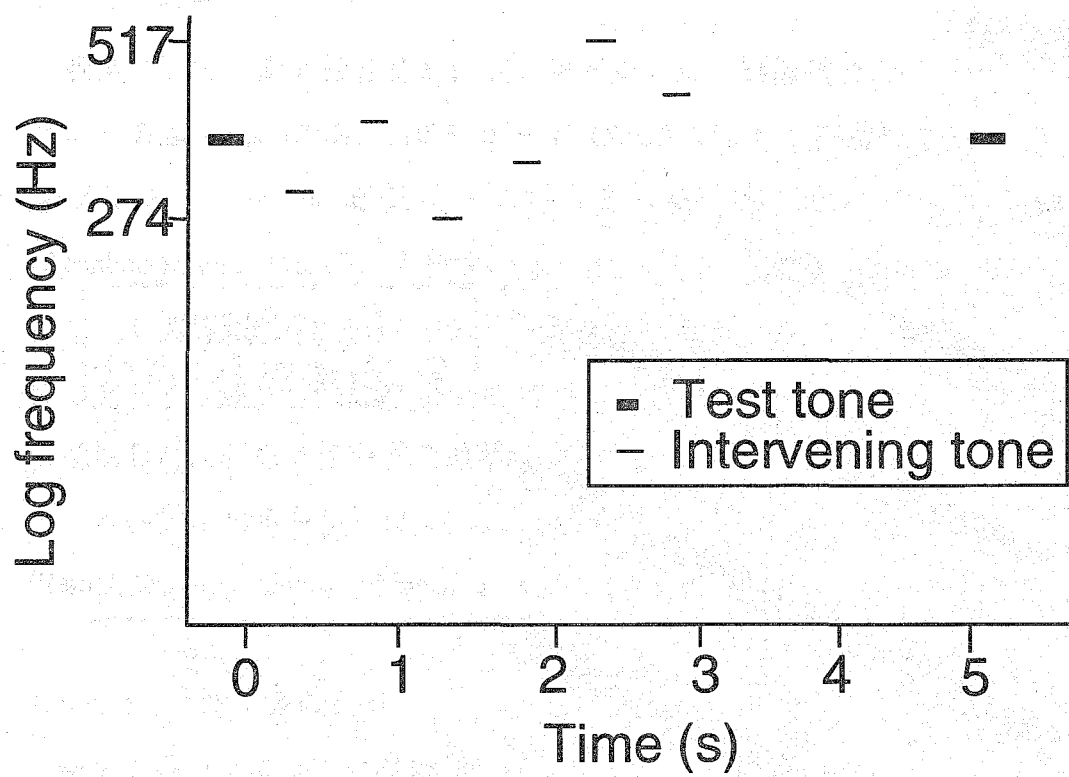


Fig. 3. 2 Schematic diagram showing a tone-intervening trial in the Deutsch experiment.

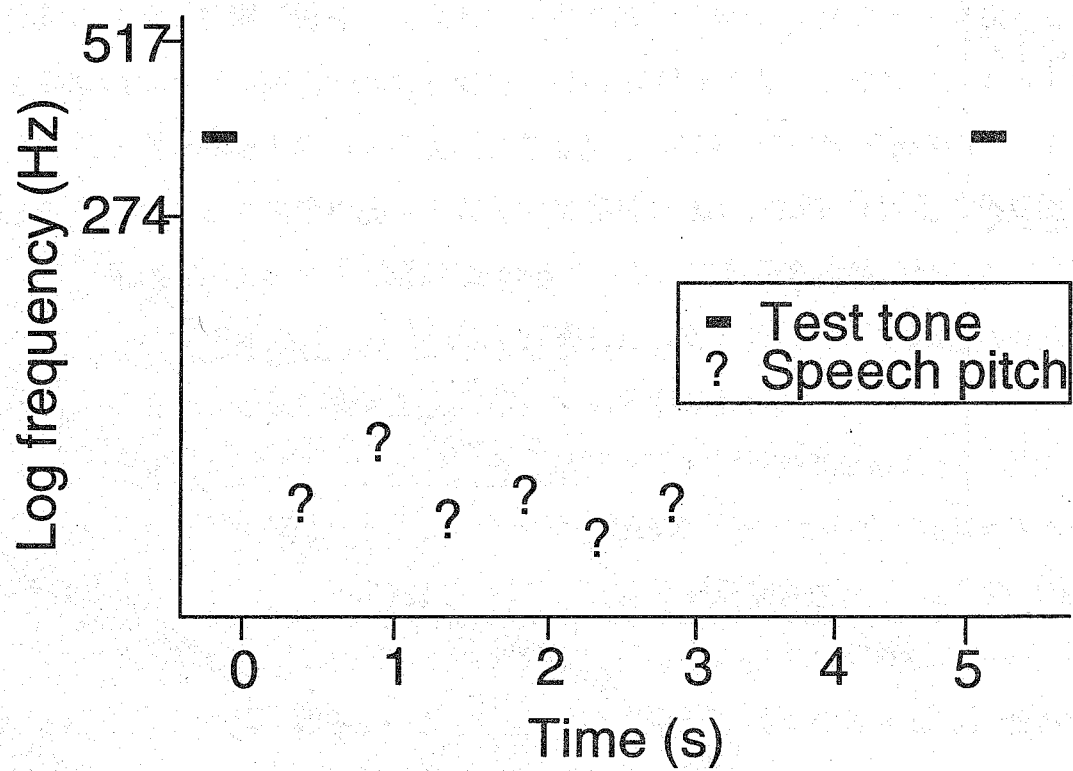


Fig. 3. 3 Schematic diagram showing a speech-intervening trial in the Deutsch experiment.

3.3 高さをコントロールした音声と調波複合音の干渉効果

Deutsch (1970) の実験で得られた干渉効果の差が、音声と非音声の記憶メカニズムの特殊化によるものなのか、それとも両者の高さの違いによるものなのかを調べるため、音声の高さと非音声の高さを等しくして実験を行った。

3.3.1 刺激

刺激として、調波複合音と数詞音声を用いた。調波複合音はパーソナルコンピュータ (Compaq 386/20, DSP ボード OROS AU22 を装着) により、20 kHz サンプリング、16 bit 直線量子化で合成した。数詞音声はマイクロフォン (Zennheiser, MD 416 U), ミキサー (Expelec, Mix 002SR) を用いてパーソナルコンピュータに同上のサンプリング条件で取り込んだ。刺激の DA 変換時にはカットオフ周波数 8 kHz, 80 dB/oct のローパスフィルタ (Frequency Devices 678) を使用した。刺激は被験者の両耳にヘッドフォン (TDH 39) で呈示された。すべての刺激の大きさがほぼ 65 phons で等しくなるように、振幅が調整された (Semal, Demany, Ueda, and Hallé, 1996)。

3.3.1.1 テスト音の合成

テスト音 (標準音と比較音) は、基本周波数成分から第 3 倍音までの等振幅の 3 成分からなる調波複合音である。各音の持続時間は 350 ms で、立ち上がり、減衰時間は 10 ms, その間の振幅包絡は直線的に変化させた。標準音の基本周波数は 110-220 Hz の範囲を 150 cents (100 cents は平均率の半音に相当する) 間隔で等分した周波数とした。全部で 9 個の標準音が用いられた。比較音は、標準音と同じものに加えて、標準音よりも 75 cents 高い基本周波数、および 75 cents 低い基本周波数のものが用意された。全部で 19 個の比較音が用いられた。

3.3.1.2 調波複合音の合成

挿入音のうち、調波複合音を次のようにして合成した。調波複合音の基本周波数は 103.9-138.6 Hz の範囲を 50 cents 間隔で等分した周波数とした。ただし、標準音で用いた基本周波数は除いた。用いた基本周波数は全部で 8 種類である。調波複合音のスペクトル構成は、各々の基本周波数について、基本周波数成分から第 6 倍音まで、第 9 倍音まで、第 13 倍音まで、第 20 倍音までの 4 種類が用意された。各刺激の持続時間は 600 ms である。

3.3.1.3 数詞音声刺激の作成

数詞音声刺激は次のようにして作成した。数詞音声として、フランス人男性の発話したフランス語の数詞 (sept, neuf, dix, quinze) を用いた。これらはいずれも CVC 音節である。基本周波数 110 および 130.9 Hz (130.9 Hz は 110 Hz の 300 cents 上の周波数である) の 3 成分調波複合音をガイドトーンとしてヘッドフォン (TDH-39) を介して呈示し、話者はこれらの高さに自分の声の高さを合わせ、防音室内で発話した。話者は、自分の声の高さがなるべく一定になるように発話した。

録音した音声サンプルの高さを調整法によって測定した。被験者は 2 名で、3 成分調波複合音を調整音とし、各サンプルについて 3 回ずつ測定を行った。そして、以下の基準すべてを満たすもののみを合格とし、基準に達しなかった場合は音声サンプルの収録をやり直した。

- (i) 各被験者内での標準偏差が 10 cents 以内であること。
- (ii) 被験者間の平均の差が ± 25 cents 以内であること。
- (iii) 全体の平均と意図した高さ (ガイド・トーンの基本周波数) とのずれが ± 25 cents 以内であること。

得られた音声サンプルをもとに、サンプリング周波数を操作することによって高さを操作し、調波複合音の挿入音と高さの一致する音声刺激を作成した。

3.3.2 実験条件

実験条件は次の三つからなる：(1) 予備テスト条件，(2) 実験条件 1，(3) 実験条件 2.

予備テスト条件では，テスト音のみが呈示された．テスト音として最初の 3 倍音までからなる調波複合音を用いた．9 個の標準音からランダムに一つを選択して呈示し，標準音の開始から 6 秒後に比較音を呈示した (Fig. 3.4)．標準音と同じ比較音が呈示される確率は 50%，75 cents 高い，あるいは低い比較音が呈示される確率はそれぞれ 25% である．

実験条件 1 ではテスト音の間に 4 個の調波複合音が挿入された (Fig. 3.4)．9 個の標準音は，基本周波数の低いものから順に 3 個ずつ，3 グループに分かれる．これらは，挿入音の基本周波数の中心周波数との差によって，Close, Middle, Remote と呼ばれる．Close では標準音の基本周波数は，挿入音の基本周波数の中心周波数と一致する．Middle では標準音の方が 450 cents 高く，Remote では 900 cents 高い．各試行のテスト音と挿入音は次のようにして決定した．標準音はランダムに一つを選び，予備テスト条件と同様にして比較音を組み合わせた．挿入音の基本周波数の中心周波数の -100, -50, +50, +100 cents に相当する 4 ステップの周波数の中からランダムに周波数を選んで組み合わせ，一連の挿入音とした．挿入音のスペクトル構成は 4 種類からランダムに選ばれた．

実験条件 2 では，数詞音声の挿入音として用いた．数詞音声の高さは実験条件 1 の挿入音の基本周波数と対応し，数詞音声の種類 (4 種類) は実験条件 1 の挿入音のスペクトル構成の種類 (4 種類) と対応している．実験条件 1 と同様にして，Close, Middle, Remote の 3 グループを作り，各試行のテスト音と挿入音を選択した．

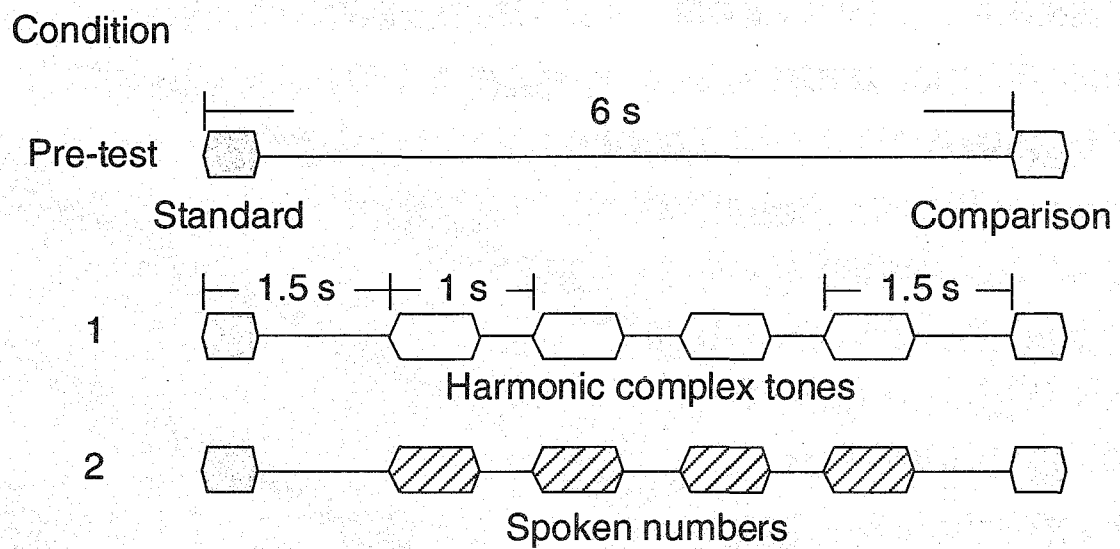


Fig. 3. 4 Schematic representation of the experimental conditions.

3.3.3 実験手続き

被験者は「比較音が標準音と同じか違うかを判断し、パーソナルコンピュータのキーボード上で反応する」ように教示された。挿入音の入った条件では、挿入音を無視するよう求められた。予備テスト条件で 27 試行中、誤答が 3 回以下だった被験者（34 名中 18 名）だけでそれ以降の実験を行った。実験条件 1, 2 はそれぞれ 4 ブロックずつに分けられ（各 1 ブロックあたり 27 試行）、両条件をブロックごとに交替させた。各実験ブロックの間で、被験者は休憩をとることを許された。両条件の実施順序は被験者間でカウンターバランスがとられている。実験は防音ブース内で行われた。

3.3.4 実験結果

誤り率の平均を Fig. 3.5 に示す。テスト音の高さの再認識誤り率（記憶への干渉）にもっとも大きく影響する要因は、テスト音と挿入音との高さの隔たりの程度である。テスト音と挿入音の高さが近い場合（Close）にはどちらの実験条件でも誤り率はかなり高く、両者の高さが離れていくにしたがって誤り率はゼロに近づく。Close と Middle との誤り率の差はいずれの実験条件においても有意である [Kruskal-Wallis 検定による。実験条件 1 で $H(1) = 136.58, p < 0.01$, 実験条件 2 で $H(1) = 113.49, p < 0.01$]。

挿入音の種類による誤り率の差はそれに比べてずっと小さい。ただし、Close 条件での誤り率の差は有意である [$H(1) = 19.774, p < 0.01$]。

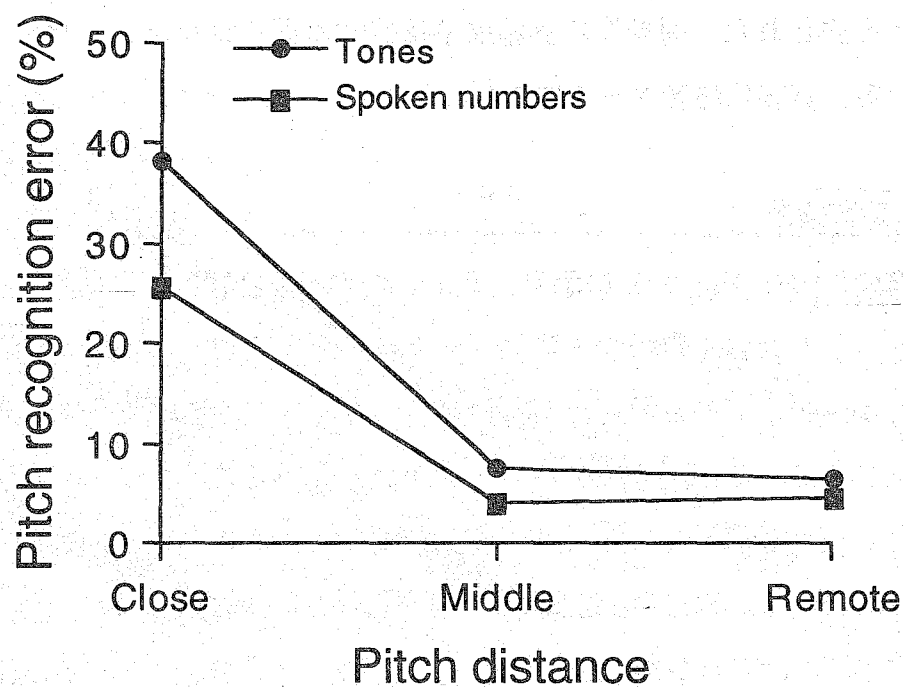


Fig. 3. 5 Percentage of pitch recognition error shown by 18 listeners. "Close," "Middle," and "Remote" depict pitch distances between test tones and intervening sounds.

3.3.5 考察

3.3.5.1 テスト音と挿入音の高さ関係の重要性

実験結果から、高さを一定にした音声を用いると、調波複合音を用いた場合と同様、テスト音と挿入音の高さの関係により、高さの再認識率が大きく左右されることが明らかとなった。この実験で操作された変数のなかで、高さの関係が従属変数にもっとも大きな影響を及ぼした。このように、テスト音と挿入音の間の高さの関係が高さの再認識率に大きく影響することは、音声以外の刺激を用いた先行研究の結果 (Semal and Demany, 1991; Deutsch, 1972) ともよく一致する。調波複合音 (テスト音) と、高さをコントロールした音声 (挿入音) の高さを近づけると、調波複合音の高さの再認識に誤りが生じるのだから、両者の間には高さに関して干渉が見られると言ってよい。

3.3.5.2 音声の高さの不明確さ

しかし、挿入音が音声の場合と調波複合音の場合とを比較すると、音声を挿入した場合は誤り率が低かった。特に、テスト音と挿入音の高さが近い条件では、その差は統計的に有意であった。この差は何を意味するのだろうか。

一つの可能性として、音声の高さが (わずかながら) 変動するため、調波複合音と比べて高さが不明確になり、テスト音に対する干渉が小さくなったことが考えられる。

そこで、音声と調波複合音の高さの明確さを比較するために、調整法による実験を行った。先の実験で用いた刺激の中からいくつかを選んで刺激として使用した。標準音として基本周波数が 113.3 および 134.7 Hz の調波複合音と、同じ高さの数詞音声を用い、調整音として 3 成分の調波複合音を用いた。2 名の被験者がそれぞれの標準音について 10 回ずつの調整を行った。刺激呈示の条件は 3.3.1 と同様である。調整値の標準偏差は、音声で 7.4 cents, 調波複合音で 4.4 cents となった。

刺激の種類による調整値の分散の違いについて F 検定を行ったところ、1 % 水準で有意差が認められた [$F(39, 39) = 2.83$]. この調整値のばらつきの違いは、

高さの明確さの違いを反映していると考えられる。したがって、このような高さの明確さの違いが、テスト音と挿入音の高さが接近して全体に誤り率が高くなった条件で、有意な誤り率の差として現れた可能性がある。

3.3.5.3 音声刺激の性質のあいまいさ

もう一つの考え方として、音声と非音声の記憶は特殊化されているが、実験で用いた音声刺激に音声と非音声の両方の性質が含まれており、音声の高さが非音声の高さの記憶に「紛れ込み」、非音声の高さの記憶に弱い干渉を起こした可能性もある。

この実験で用いた音声は、人間が発話したという意味では音声だが、高さをコントロールしようとしたために、自然に発話された音声と比べて高さの変化が少なく、平板な高さを持つものとなった。そのような音声は、音声としてよりはむしろ、非音声に近い音として知覚された可能性がある。しかし、この実験には、音声刺激が音声として知覚されたかどうかを調べる課題は含まれていないため、この点は不明である。

3.3.6 まとめ

この実験により、以下のことが明らかになった。

- (1) 調波複合音のテスト音と、高さを一定にした音声の挿入音の高さを近づけると、テスト音の高さの再認識率が上昇する。
- (2) 音声を挿入した場合の誤り率は、調波複合音を挿入した場合の誤り率よりも低い。このことは、音声の高さの不明確さが原因の一つと考えられる。音声と調波複合音の高さの明確さの違いを調べた実験の結果が、これを裏付ける。
- (3) もう一つの原因として、不自然な音声に非音声としての性質が含まれていたことが考えられる。

全体としてこの実験により、音声の高さの要因について検討されていなかった Deutsch (1970) や Pechmann and Mohr (1982) の実験結果を根拠として、音声

と非音声の短期記憶が特殊化されていると主張することに大きな疑問が生じるが、一方で、この実験にもいくつかの問題点が存在する。そこで、音声と非音声の短期記憶に干渉が見られるのかどうかについて、より明らかにするための実験を行った。

3.4 自然な音声と純音の干渉効果

次のような点が、先に述べた実験の問題点と考えられる。

- (1) 挿入音として使用した音声は、調波複合音の高さに合わせて高さをなるべく一定に保って発話されたものであること。
- (2) 音声刺激が音声として聞き取られていたかどうか不明であること。
- (3) 被験者が絶対音感を持っていたかどうか不明であること。

(1) は、音声としての自然性についての問題である。自然に発話された音声では、高さが大きく変動するのが普通である。一方、意図的に高さを定常に保って発話された音声は不自然で、調波複合音に近いものといえる。不自然な音声に対して被験者が通常の音声と同じような処理をしたという保証はない。これを解決するには、自然に発話された音声を用いて実験をしなければならない。

しかし、たとえ自然な音声を用いたとしても (2) の問題は残る。もし、被験者が音声刺激を音声として聞き取っていなかったとしたら、調波複合音との間で干渉効果に大きな差が見られなくても不思議ではない。したがって、音声刺激が音声として聞き取られていたかどうかをチェックしておく必要がある。そのためには、音声の系列再生課題を行うことが考えられる。

Deutsch (1970) の研究では、音声は単に無視されやすかっただけなのか、それとも音声をきちんと処理していても高さの再認の妨げにならなかったのかを調べるため、被験者に高さの再認をさせた上で、数詞の系列を再生させる条件を設けている。その場合でも、高さの再認のみの条件と比べて高さの再認の誤り率は 3.2% しか増加しなかった (Table 3.1)。しかし、この研究ではテスト音と音声の高さの関係が考慮されておらず、結果の再現性も依然として不明である。

(3. 2) . それゆえ, Deutsch の実験の再現性を検討できるように, Deutsch の実験条件を包括するような実験を行う必要がある.

(3) は音の高さの labeling の問題である. 絶対音感を持つ被験者は, 音の高さを瞬時に音名という言語的な label に変換し, 保持できると考えられている (Ward and Burns, 1982; Miyazaki, 1988; 1989). したがって, 絶対音感を持つ被験者は定常音の高さを言語的に記憶する可能性があり, このような被験者では「純音の高さの記憶と音声の記憶は特殊化されている」という Deutsch (1975c) の仮説の検証ができなくなる. それゆえ, 被験者が絶対音感を持っているかどうか調べておく必要がある.

以上のような観点から, 音声と非音声の短期記憶が特殊化されているのかどうかを明らかにすることが次に述べる実験の目的である.

3. 4. 1 刺激

3. 4. 1. 1 純音と音声

本研究で用いた刺激は, 次の 3 種類である: (1) 純音, (2) 女性話者 1 名の発話した日本語数詞音声, (3) 男性話者 1 名の発話した日本語数詞音声. 数詞音声はそれぞれ「0 (ゼロ)」から「9 (きゅう)」までの 10 個ずつである.

純音は UNIX ワークステーション (Silicon Graphics, Indy, R4600PC CPU) を用いて, 12 kHz サンプリング, 16 bit 直線量子化で合成された. 持続時間は 200 ms, 立ち上がり, 減衰時間は 10 ms とした. 実験で用いた純音の振幅は, 合成段階ですべて等しかった.

音声は ATR 人間情報通信研究所の音声データベース (匂坂, 浦谷, 1992) を使用した. サンプリング周波数, 量子化ビット数は合成純音と同じである. 音声ファイルをワークステーションに取り込んだ後, 各音声の前後の空白部分をカットした.

3. 4. 1. 2 刺激呈示法と音圧の校正

ワークステーションから出力された刺激は, DAT (Sony, DTC-95ES) に録音さ

れた。被験者に刺激を呈示する際には、DAT の再生出力をアダプタ (STAX, SRM-1/MK2 pp) を介してヘッドフォン (STAX, SR Λ Professional) に入力した。このヘッドフォンの周波数特性、インパルス応答特性は実耳装着時においても極めて優秀で (Ueda and Hirahara, 1991; 上田, 平原, 1990), 歪みも少ない (平原, 鈴木, 上田, 1990)。音圧は、1 kHz の純音が 75 dB SPL となるよう、IEC カプラー (Brüel & Kjaer, Type 4153), 精密サウンドレベルメータ (Brüel & Kjaer, Type 2231) を用いて ATR 人間情報通信研究所の無響室で較正を行った。

3.4.1.3 音声の大きさの調整

音声の大きさをそろえるため、1 kHz, 75 dB SPL, 200 ms の純音を標準音として上下法により各音声刺激の大きさの主観的等価点 (PSE) を求めた。順序効果を相殺するため、標準音—比較音 (音声) の順で呈示する組み合わせと、比較音 (音声)—標準音の順で呈示する組み合わせの両方で測定した。各音声について、一方の呈示順序につき、6 系列、一つの音声について 12 系列の測定を行った。比較音の音圧レベルは 2 dB ステップで変化させた。音圧の初期値は 0.5 dB 間隔の 3 段階のなかからランダムに選ばれた。

被験者は 2 名であった。聴力検査により、被験者の聴力が正常であることを確認した。ATR 人間情報通信研究所の防音室内で測定を行った。得られた PSE の被験者間の相関係数は、女性話者の音声で 0.72, 男性話者の音声で 0.97 であった。2 名の被験者の PSE の平均に基づいて、各音声刺激の振幅をワークステーション上で調整し、大きさをそろえた。

3.4.1.4 音声の高さの測定

話者ごとの音声の大まかな高さを調べるため、純音との対比較による実験を行った。純音の持続時間は 200 ms, 被験者は 1 名である。以下の純音についての記述では、周波数の代わりに音名を用いることとする。音名の基準周波数は A4 が 440 Hz, 音律は平均律である。比較した純音の範囲は、女性話者については C#3-C#5, 男性話者については C2-C4 で、4 半音 (400 cents) ステップ

で 7 個の純音をランダムに音声と組み合わせた。また、音声と純音の呈示順序の前後を入れ替えて組み合わせた。

被験者は純音の高さと音声の高さを比較し、どちらが高いかを答えた。ただし、純音の高さが音声の高さの変化幅に含まれている場合にはそのように答えた。判断回数は、各組み合わせにつき 20 回であった。静かな部屋で測定を行った。音声の方が高いとされた場合には 1 点、純音の高さに音声の高さが含まれているとされた場合には 0.5 点をあたえ、結果を刺激ごとの合計得点の形に整理した。男女各話者ごとに平均的な高さ（得点率が 50 %となる点にもっとも近い平均律の音名）をこの結果から求めたところ、女性話者では D#4 (311.13 Hz)、男性話者では A2 (110.00 Hz)となった。

3.4.2 被験者の選抜

3.4.2.1 高さの再認能力による選抜

主実験の被験者を選抜するため、高さの再認能力を調べるテストを行った。テスト音の範囲を広げたほかは、Deutsch (1970) の方法に従っている。

A. 実験手続き

音名で C#2 から C5 に相当する約 3 オクターブの範囲 (69.30-523.25 Hz) で、半音階を構成する純音 36 個を標準音とした。標準音はランダムに一度ずつ用いられたので、各被験者の試行数は 36 試行である。

各標準音には比較音として、(1) 同じ音、(2) 半音高い音、(3) 半音低い音のいずれかが組み合わせられた。組み合わせられる確率は (1) が 50 %、(2)、(3) がそれぞれ 25 % であった。被験者は、比較音が標準音と同じか違うかを判断する。したがって、チャンスレベルは 50 % であった。

刺激間隔は 5 秒であった。被験者は 3 秒の試行間隔中に、回答を用紙に記入した。被験者には、比較音が標準音と異なる場合には、その差は常に半音であることがあらかじめ伝えられた。被験者は健聴者 37 名である。

B. 結果と考察

Deutsch (1970) が用いた標準音の範囲は C#4-C5 であった。この範囲は、本研究で用いた標準音のもっとも高いオクターブの範囲に相当する (Deutsch の用いた音名は A4 が 435 Hz であったため、周波数にはわずかな差がある)。Deutsch の被験者の選抜基準は、誤り率がゼロであることであった。

本研究では、この範囲で誤り率がゼロの被験者は 25 名存在した。しかし、標準音が C#3-C4 の範囲ではそのうち 18 名が少なくとも一回誤答し、C#2-C4 の範囲では 24 名が少なくとも一回誤答した。したがって、Deutsch の被験者に同じ範囲のテスト音で実験を行ったとすれば、やはりテスト音の周波数範囲が下がるにつれて誤り率が増加した可能性が高い。

そこで、本研究では、以下のような基準に従って被験者を選抜した。

(1) 標準音が C#4 以上の範囲では、誤り率がゼロであること。

(2) 標準音が C4 以下の範囲では誤り率がチャンスレベルよりも低いこと。

この基準をクリアした被験者は 37 名中 20 名であった。このうち 18 名の被験者で 3.4.3 以降の実験を行ったが、実験終了後に一部の実験条件の手続きにミスがあったことが判明したため、18 名中 9 名の被験者で問題の部分だけ実験をやり直した。そのため、最終的に得られたデータは 9 名分となった。この 9 名の被験者の C#3-C4 の範囲での平均誤り率は 13.0 %、C#2-C4 の範囲では 29.6 % であった。

3.4.2.2 絶対音感テストによる選抜

高さの再認能力を調べるテストの直後に、全被験者について絶対音感テストを実施した。このテストは、絶対音感を持つ被験者を排除することが目的である。

A. 実験手続き

刺激となる純音は、C#4-C5 の半音ステップの 12 音で、これらが一つずつランダムに 4 秒間隔で呈示された。したがって、全試行数は 12 である。被験者

は、刺激の高さを C をドとする階名で回答用紙に記入した。絶対音感保持者の正答率について調べた研究 (Miyazaki, 1988) を参考にして、12 試行全部で正解した被験者を絶対音感保有者とみなした。

B. 結果

絶対音感保有者とみなす基準をクリアした被験者はゼロであった。9 名の被験者の正答率は、最高 33.3 %, 最低 0.0 %, 平均 17.5 % であった。この正答率の水準は Miyazaki (1988) のいう、非絶対音感保持者 (non-AP subjects) の水準と一致する。そのため、このテストの結果、排除された被験者はいなかった。

3.4.3 実験条件

実験で操作した要因は次の通り。

(1) 標準音の周波数範囲。

- a. Deutsch (1970) の用いた範囲：C#4-C5, 277.18-523.25 Hz.
- b. 女性話者の平均的な高さを中心とする範囲：A3-G#4, 220.00-415.30 Hz.
- c. 男性話者の平均的な高さを中心とする範囲：D#2-D3, 77.78-146.83 Hz.

(2) 挿入音の種類。

- a. 挿入音なし。
- b. 純音 I：比較音の周波数範囲の半音階から標準音とその上下半音の範囲の音を除き、それ以外の音から試行ごとにランダムに選ばれる。
- c. 純音 II：比較音の周波数範囲の半音階からランダムに選ばれる。
- d. 女性話者の数詞音声。
- e. 男性話者の数詞音声。

(3) 被験者の課題。

- a. 標準音の高さの再認。
- b. 数詞の系列再生。
- c. 標準音の高さの再認と数詞の系列再生の両方。

テスト音は、標準音を (1) の範囲の半音階からランダムに選び、3.4.2.1 A. と

同様にして比較音と組み合わせた。

(2) の純音 I, II は、以下のような理由で設定した。Deutsch (1970) は「挿入音はランダムに選んだ。ただしテスト音の高さは取り除いた」と述べているので、純音 I の条件は Deutsch の条件と近いと考えられる。挿入音が純音の場合は純音 I のような制約を課すことが可能であるが、音声で挿入音とした場合は、音声の高さが連続的に変化するため、そのような制約を課すことは不可能となる。そこで、制約条件がある場合（純音 I）と、ない場合（純音 II）とで、誤り率にどのような差が見られるのかを調べることにした。

数詞音声は、それぞれ 10 個のなかから試行ごとにランダムに 6 個が選ばれた。

(3) の課題のうち、標準音の高さの再認のみを求める条件では、挿入音がある場合はそれを無視するように被験者に教示がなされた。そして、標準音と比較音の高さが違う場合には、その差は常に半音であることがあらかじめ伝えられた。

数詞の系列再生のみを求める条件では、呈示される比較音は常に標準音と同じで、このことはあらかじめ被験者に伝えられた。被験者は反应用紙の高さの再認の欄に常に「同じ」と答えた後、数詞の系列を正しい順序で筆記により再生するよう求められた。

標準音の高さの再認と数詞の系列再生の両方を求める条件では、まず標準音の再認を行った後、挿入された数詞の系列を正しい順序で再生するよう求められた。

これらの要因を組み合わせる実験条件とした。音声で挿入される条件では、女性話者の数詞音声は Deutsch の用いた範囲および女性話者の平均的な高さを中心とする範囲のテスト音と、男性話者の数詞音声は Deutsch の用いた範囲および男性話者の平均的な高さを中心とする範囲のテスト音と組み合わせられた。

3.4.4 実験手続き

標準音と比較音の間には 5 秒の保持期間が設けられた。挿入音のある条件で

は、保持期間中に 6 個の挿入音が挿入された。比較音の直前を除いて、各刺激の始まりと始まりの時間間隔は 500 ms であった。

挿入音の種類によって、全体を次の 4 ブロックに分けた。

ブロック 1: 挿入音なし, 36 試行。

ブロック 2: 純音 I, II, 72 試行。

ブロック 3: 女性話者の数詞音声, 72 試行。

ブロック 4: 男性話者の数詞音声, 72 試行。

被験者ごとにブロックの順序、およびブロック内での実験条件の順序をランダム化して実験を行った。刺激呈示法は 3.4.1.2 に述べた。実験は静かな部屋で行われ、被験者は 10 秒間の試行間隔中に反応を用紙に記入した。

3.4.5 実験結果

数詞系列の再生結果については、一つの系列がすべて正しい順序で再生された場合のみを正解とし、他は誤りとして再生誤り率を求めた。実験結果を Fig. 3.6-3.8 に示す。いずれのグラフにも、Deutsch (1970) のデータと一緒にプロットされている。

誤り率を逆正弦変換 (Snedecor and Cochran, 1967) した後、SAS 統計パッケージの GLM プロシジャにより被験者の要因 (S) をブロック因子として分散分析を行った。

挿入音なし、純音 I、純音 II が挿入された条件の高さの再認識誤り率 (Fig. 3.6) について、2 要因 [標準音の周波数範囲 (R)、挿入音の種類 (I)] の分散分析を行った。S および R、I の主効果が有意水準 1 % で、交互作用効果 $R \times I$ が有意水準 5 % で、いずれも有意であった [それぞれ、 $F(8, 64) = 4.64, p = 0.0002$; $F(2, 64) = 10.44, p = 0.0001$; $F(2, 64) = 23.20, p = 0.0001$; $F(4, 64) = 3.07, p = 0.0225$] 。

Fig. 3.6 を見ると、純音が挿入された条件の誤り率は高く、ほぼ一定である。Deutsch が用いた範囲の標準音を用いた条件と女性話者の平均的な高さを中心とする範囲の標準音を用いた条件の誤り率は、あまり変わらない。しかし、これらとは周波数の離れた、男性話者の平均的な高さを中心とする範囲の標準音

を用いた条件では、挿入音のない条件の誤り率だけが急激に上昇する（交互作用）。純音の挿入音 I と II の差は小さい。

次に、挿入音なし、女性話者、男性話者の数詞音声を入れた条件の高さの再認識率（Fig. 3. 7）について、3 要因〔標準音の周波数範囲（R）、挿入音の種類（I）、課題（T）〕の分散分析を行った。S および R、T の主効果がいずれも有意水準 1% で有意であった〔それぞれ、 $F(8, 80) = 3.00, p = 0.0054$; $F(2, 80) = 67.48, p = 0.0001$; $F(1, 80) = 13.37, p = 0.0005$ 〕。それ以外の主効果、交互作用効果は有意水準 5% で有意ではなかった〔I: $F(2, 80) = 0.19, p = 0.8245$; $R \times I: F(2, 80) = 0.10, p = 0.9071$; $R \times T: F(2, 80) = 3.10, p = 0.0503$; $I \times T: F(1, 80) = 0.16, p = 0.6947$ 〕。

標準音の周波数範囲が低くなると高さの再認識率が全体に増加する（Fig. 3. 7）。また、高さの再認と数詞の系列再生の両方の課題が課せられた条件では、高さの再認のみの条件と比べて、高さの再認識率が増加する。

さらに、数詞の系列再生の誤り率（Fig. 3. 8）について 3 要因〔標準音の周波数範囲（R）、挿入音の種類（I）、課題（T）〕の分散分析を行った。有意水準 1% で S および R、T の主効果が有意であった〔それぞれ、 $F(8, 56) = 14.03, p = 0.0001$; $F(2, 56) = 7.72, p = 0.0011$; $F(1, 56) = 13.23, p = 0.0006$ 〕。それ以外の主効果、交互作用効果は有意水準 5% で有意ではなかった〔I: $F(1, 56) = 1.07, p = 0.3044$; $R \times T: F(2, 56) = 2.30, p = 0.1094$; $I \times T: F(1, 56) = 0.69, p = 0.4096$ 〕。

数詞の系列再生の誤り率は、標準音の周波数範囲が下がると減少する（Fig. 3. 8）。系列再生課題のみの条件より、高さの再認課題が加えられた条件の方が誤り率が高い。

全体に Deutsch のデータは、この実験結果の類似した条件下のデータとほぼ近い値を示している。ただし、数詞音声の系列再生誤り率で、Deutsch のデータは課題による効果が逆転している。すなわち、系列再生のみを行った条件での誤り率が、標準音の高さの再認を同時に遂行した条件よりも大きくなっている。

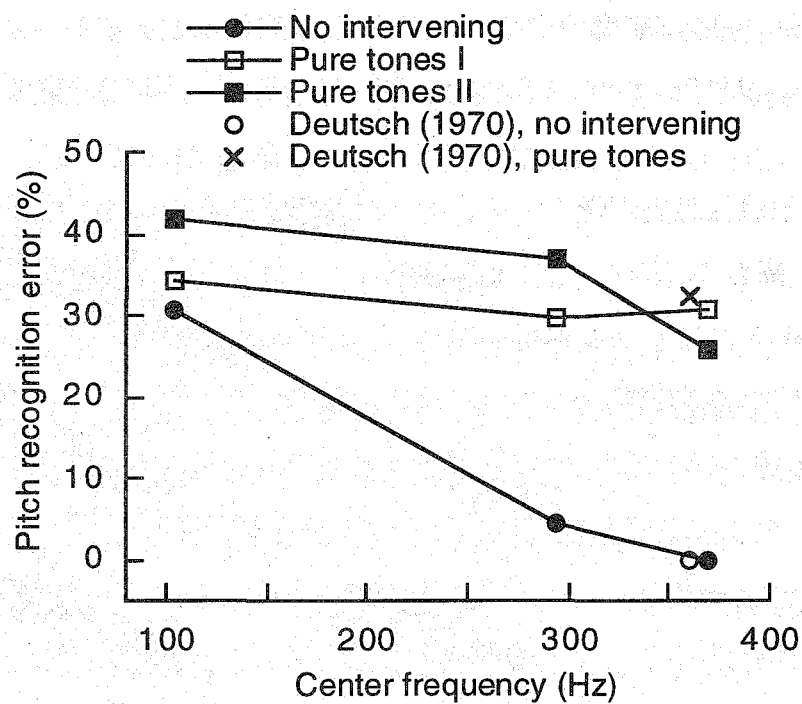


Fig. 3. 6 Percent of pitch recognition error plotted as a function of center frequency of the test tones. No-intervening and pure-tone intervening conditions.

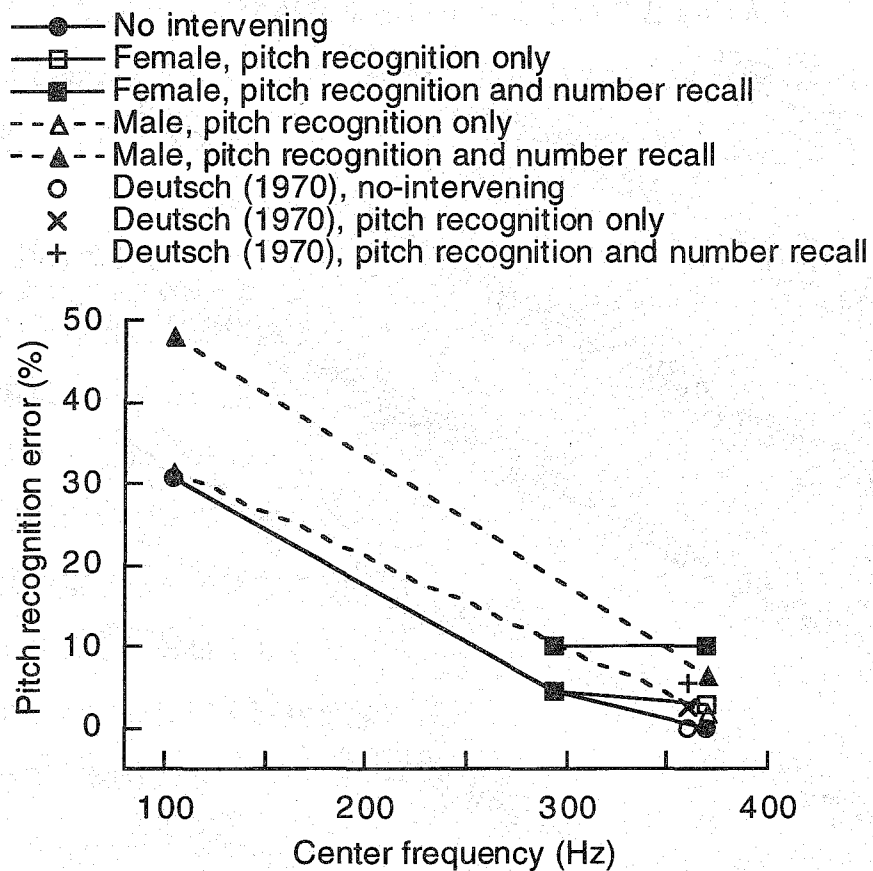


Fig. 3. 7 Percent of pitch recognition error plotted as a function of center frequency of the test tones. No-intervening and speech-intervening conditions.

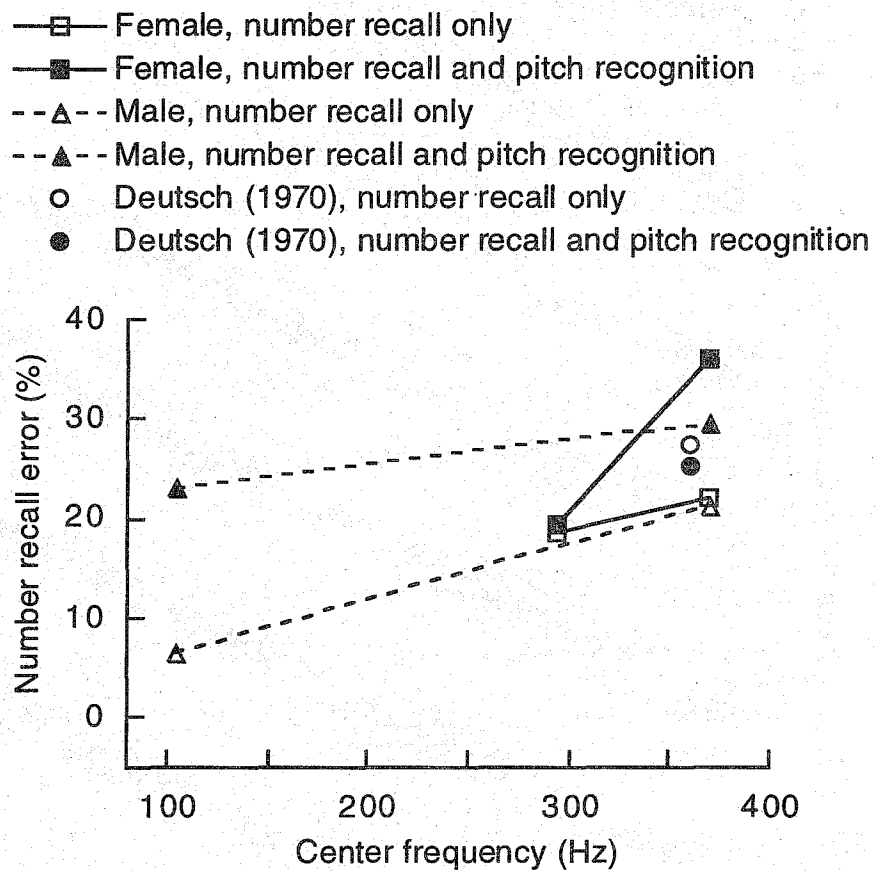


Fig. 3. 8 Percent of number recall error plotted as a function of center frequency of the test tones.

3.4.6 考察

3.4.6.1 音声の大きさ調整の誤りについて

実験終了後、男性話者の「9（きゅう）」の大きさ調整時に間違いがあったことが判明した。この音声の音圧は、意図した音圧よりも全体に 5.7 dB 低く設定されていた。しかし、挿入音（調波複合音）の音圧を 30 dB の範囲で変化させても高さの再認識率には有意な影響を与えないことが明らかにされており (Semal and Demany, 1993), また、この程度の音圧の変化は音声の明瞭性にもほとんど影響しないと考えられるので、実験結果への影響は無視できる。

3.4.6.2 純音の高さの記憶と数詞音声系列の記憶は干渉する

純音の高さの再認識課題に数詞音声の系列再生課題が加わると、高さの再認識率が有意に上昇する。数詞の系列再生誤り率も、高さの再認識課題が加わることで上昇する。また、挿入音を無視する条件で、純音が挿入された場合の高さの再認識率は統制条件（挿入音なしの条件）よりも高いが、数詞音声が入力された場合の高さの再認識率は統制条件と変わらない。それゆえ、音声の高さを無視することは純音よりも容易だが、純音の高さと数詞系列を同時に記憶しようとする両者に干渉が生じることになる。

したがって、純音の高さの記憶メカニズムと、音声の記憶メカニズムが特殊化されているとは言えない。逆に、両者の記憶には干渉を起こす場があると考えられる。

では、なぜ音声の高さは純音よりも無視しやすかったのだろうか。ここでは二つの可能性が考えられる。

- (1) 音声の高さの記憶は特殊化されているため。
- (2) 音声の高さが不明確なため。

(1) の考え方では、(a) 音声の高さと (b) 非音声の高さは、それぞれ特殊化された処理装置で処理され、特殊化された記憶貯蔵庫で保持されるが、音声が入力された結果、認識された (c) 数詞系列の記憶と (b) 非音声の高さの記憶は同じ記憶貯蔵庫で保持されると主張することになる。しかし、この考え方はあま

りにも不自然である。

(2) の考え方では、音声と非音声に特殊化された記憶貯蔵庫を必要とせず実験結果を説明できる。すなわち、記憶痕跡の強さの違いで干渉効果を説明する。もし、音声の高さが不明確で、その記憶痕跡も弱いものとすれば、より高さの明確な、純音の高さの記憶に対してはあまり干渉を起こさない。しかし、テスト音の純音と同程度に明確な高さを持つ、純音の挿入音は、テスト音の高さの再認に大きな干渉を及ぼすことになる。一方、数詞系列は音声の高さよりも明瞭な記憶痕跡を持つため、純音の高さの記憶と干渉することになる。

3.4.6.3 再認、再生誤り率と高さの不明確さ

この実験結果をさらに細かく説明するためには、低い周波数の純音および自然な音声の高さの不明確さを考える必要がある。

純音の周波数と弁別閾との比（相対弁別閾）は、1 kHz 以下では周波数が低下するほど増加する。十分に訓練された被験者でさえ、その大きさは 125 Hz, 200 ms の刺激で 0.47%（1 kHz での相対弁別閾の約 3 倍）に達する（Moore, 1973）。これは、低い周波数では純音の高さが不明確になることを意味する。同様のデータは Nordmark (1968) にも示されている。また、自然な音声は基本周波数が大きく変動するため、一定の高さは知覚されにくく、高さは不明確である。

高さの再認識り率 (Fig. 3. 6) について考える。テスト音の周波数範囲が低い条件で、挿入音なしの条件では高さの再認識り率が上昇する。これは低い周波数の純音の高さが不明確になることで説明できる。

純音 I と II を挿入音としたときの、テスト音の高さの再認識り率はほぼ同じである。つまり、挿入音を選ぶ際にテスト音と同じ音を除外するかどうかは、あまり結果に影響しなかった。ただし、挿入音がなくてもテスト音の周波数が低下するだけで、テスト音の高さの再認識り率が上昇するので、低い周波数では純音 I, II の差がはっきりと現れてこなかった可能性もある。

テスト音の高さの再認のみを行う条件では (Fig. 3. 9), 挿入音なしの条件と音声挿入した条件の間で高さの再認識り率に差がない。これは、音声の高さが

低い周波数の純音の高さよりも不明確であるためと考えられる。

挿入音からテスト音への干渉効果だけではなく、テスト音から挿入音への干渉効果もあると考えられる。数詞の系列再生誤り率が、テスト音の周波数範囲が下がると減少するのは (Fig. 3. 8) , テスト音の高さが不明確になるので数詞系列の記憶への干渉効果が小さくなるためと考えられる。

3. 4. 6. 4 Deutsch (1970) との相違の原因

このように、本研究では、純音の高さの記憶と音声の記憶が干渉することが示された。しかし、Deutsch (1970) は、このような干渉効果は見られないと結論した。この違いはなぜ生じたのであろうか。

まず、被験者集団の違いについて検討してみる。被験者の選抜基準に関して、比較可能な部分については、本研究と Deutsch の研究の選抜基準は同じである。また、Deutsch の条件 A に対応する本研究の実験条件は Deutsch の用いた範囲の標準音 × 純音 I 条件であるが、両者の高さの再認識率 (32.3% と 30.6%) にはほとんど差がないので、被験者集団の違いが原因である可能性は低い。

そこで、これ以外の要因について検討するため、Deutsch のデータの他の部分と本研究のデータの対応を調べる。まず、Deutsch のデータで条件 B (挿入音を無視して、テスト音の高さの再認識のみを行う) と C (テスト音の高さの再認識を行い、数詞の系列再生も行う) では、C の高さの再認識率は B より 3.2 % 増加している。しかし、Deutsch は、この差の有意性については述べていない。また、条件 C と D (テスト音の高さの再認識を行わずに、数詞の系列再生のみを行う) の数詞の系列再生誤り率の差 (D は C より 2.1 % 増加) は有意でなかったと述べている。しかし、本研究のデータでこれらに対応づけできる誤り率は、それぞれ 4.6 % から 9.6 % の増加、0.9 % から 16.6 % の減少が見られ、どちらの差も有意であった。

このような結果の相違を説明する要因として、次のようなことが考えられる。

- (1) 音声の高さの統制の有無。
- (2) 統計的検定法の違い。

(3) 音声の質（明瞭性）の違い。

(1) の音声の高さの統制について考える。Deutsch (1970) では、音声の高さの要因は研究の視野に入っていなかった。そのため、音声の高さとテスト音の高さが大きく離れていた可能性が高い。その場合、全体に誤り率が低くなり、条件間の誤り率の違いが小さくなる。本研究では、音声の高さにテスト音の高さを近づけた条件が含まれており、条件間の誤り率の差が増大した可能性がある。

(2) の統計的検定法の違いも影響した可能性がある。Deutsch は符号検定を用いた。この検定法では、実験条件全体を通しての被験者の要因が考慮されないため、被験者間のばらつきによって実験条件間の違いが有意とならなかった可能性がある。一方、本研究では分散分析を用いて、被験者の要因による分散を実験条件間の分散から分離したため、実験条件間の差が見いだされやすくなったと考えられる。実際、Deutsch (1970) の示したデータは、平均としては本研究の対応する条件におけるデータとかなり一致している。

(3) の音声の質は、Deutsch (1970) の実験については不明である。しかし、数詞の系列再生誤り率が全体に本研究結果よりも高めであることは、数詞が本研究よりも二つ多いことを考慮しても、Deutsch の用いた音声の質（明瞭性）に疑問を抱かせる。一方、本研究で用いた音声は、プロのアナウンサー・ナレーターが明瞭に発話し、放送スタジオで収録されたものであり、空白部分のカットなど編集や振幅の調整、保存などその後の操作も計算機上で行われたので、刺激系列の作成までのすべての過程で音質の劣化はきわめて少なく、音声の質はきわめて高い。このことも、実験条件間の差を明確にするのに寄与したと考えられる。

3.4.6.5 音楽的被験者と非音楽的被験者

Pechmann and Mohr (1992) は、音楽的被験者では音韻課題を挿入しても純音の高さの再認は妨害されないが、非音楽的被験者では妨害が生じると結論し、音韻課題を遂行する「調音ループ」と純音の高さの再認を遂行する「音ループ」への注意の負荷が、被験者群によって異なると考えた。しかし、彼らの研究で

は、統計的分析に際して被験者群内の個人差は考慮されておらず、個人ごとのばらつきによって条件間の差が不明瞭になり、有意差が検出されなかった可能性がある。本研究の被験者は、同様の高さのテスト音の範囲で、挿入音のない状態では高さの再認が完璧にできるものばかりであったので、Pechmann and Mohr の音楽的被験者に相当すると考えられる。このような被験者でも、個人差を考慮した分析を行えば、実験条件間に有意な差が見いだされた。

3.4.7 まとめ

この実験の結果明らかになったことをまとめると、次のようになる。

- (1) 純音と音声の短期記憶は干渉する。したがって、純音と音声の記憶メカニズムが特殊化されているという説は否定された。
- (2) 音声は純音よりも無視しやすいのは、音声の高さが不明確であるためと考えられる。
- (3) 低い周波数の純音の高さが不明確になったことが、結果に影響した可能性がある。
- (4) Deutsch (1970) の実験では、音声の高さという要因が考慮されていなかったこと、被験者間の分散を考慮した統計的分析が行われていなかったこと、音声の明瞭性の問題により、純音と数詞音声の間に記憶の干渉が起らなかったかのように見えた可能性がある。

3.5 自然な音声と調波複合音の干渉効果

先に述べた実験では、合成時の振幅が一定の純音をテスト音、挿入音として用いた。しかし、そのため、低い周波数では、大きさが減少し、高さも不明確になったと考えられる。このことが結果に与えた影響として、次の二つが考えられる。

- (1) 高さの再認誤り率が、テスト音が低い条件では全体に上昇した。このため、テスト音が低い条件では、挿入音の種類の違いによる誤り率の差が圧縮され、

明確に現れなかった可能性がある。

(2) 数詞の系列再生誤り率は、テスト音の高さが数詞音声の高さと近い条件の方が、高さが離れていた条件よりも低かった。高さが近い方が干渉を起こしやすいということからすると、矛盾する結果である。これは、低い周波数でのテスト音の大きさと高さの明確さが減少したため、高さが近いにもかかわらず、数詞系列の記憶に対する干渉効果が小さくなった結果かも知れない。

このような観点から、音声と非音声の短期記憶における干渉についてより明らかにするため、純音よりも高さが明確な調波複合音を用い、それらの大きさを一定にそろえ、男性話者の数詞音声と組み合わせて実験を行った。

3.5.1 刺激

テスト音は調波複合音、挿入音は調波複合音と 1 名の男性話者による数詞音声（「0（ゼロ）」から「9（きゅう）」まで）である。

調波複合音は基本音成分から順に 8 個の成分で構成されたもので、ワークステーション（Silicon Graphics, Indy, R4600PC CPU）を用いて 12 kHz サンプリング、16 bit で合成された。持続時間は 200 ms、立ち上がり、減衰時間は 10 ms とした。基本周波数を 73.42-155.56 および 261.63-554.37 Hz（音名では A4 を 440 Hz としたときの D2-D#3, C4-C#5 に相当）の範囲の半音ステップとし、計 28 個の複合音を合成した。

調波複合音の各成分の振幅（合成時）は次のようにして決定した。基本音成分の振幅を A_1 とするとき、第 i 倍音の振幅 A_i を、

$$A_i = 2^{-(i-1)} A_1 \quad (3.1)$$

となるよう合成した。

すべての調波複合音の大きさを等しくするため、1 kHz, 75 dB SPL, 200 ms の純音を標準音とする極限法（完全上下法）で各調波複合音の大きさの主観的等価点 (PSE) の測定を行った。それぞれの調波複合音について 2 dB ステップで

振幅を変化させ、上昇・下降系列を 2 系列ずつ、さらに標準音と比較音の呈示順序を入れ替えて計 8 系列の測定を行った。各系列における振幅の初期値は、各変化方向につき 1 dB ステップで 3 通りずつ用意し、それらをランダムに各系列に割り当てた。2 名の被験者の PSE を平均し、これをもとに各調波複合音の振幅を計算機上で調整して、大きさをそろえた。

男性話者による数詞音声は 3.4 で用いたものである。

3.5.2 実験条件

実験で操作された要因は次の通り。

(1) 標準音の基本周波数の範囲。

a. Deutsch (1970) の用いた範囲, C#4-C5: 277.18-523.25 Hz.

b. 数詞音声の平均的な高さを中心とする範囲, D#2-D3: 77.78-146.83 Hz.

(2) 挿入音の種類。

a. 挿入音なし。

b. 調波複合音 I: 数詞音声の平均的な高さを中心とした比較音の基本周波数の範囲, D2-D#3: 73.42-155.56 Hz から一定の制約条件（後述）のもとにランダムに選ばれたもの。

c. 調波複合音 II: 同上の範囲からランダムに選ばれたもの。

d. 数詞音声。

(3) 課題。

a. 標準音の高さの再認。

b. 数詞の系列再生。

c. 標準音の高さの再認と数詞の系列再生の両方

(1) の標準音は、所定の範囲内の半音階から 3.4.2.1.A と同様にして選ばれ、比較音と組み合わせられた。

(2) の調波複合音 I および II は、所定の範囲の半音階から試行ごとに 6 個がランダムに選ばれるが、調波複合音 I では標準音およびその上下半音と同じ音、あるいは同じトーンクロマの音が除かれている。純音を用いた 3.4 の実験

と同じ理由により、これらの条件を設けた。

数詞音声は 10 個のなかから試行ごとにランダムに 6 個が選ばれた。

(3) の課題は 3.4.3 と同様である。

これらの要因を組み合わせて実験条件とした。挿入音なしの条件は被験者選抜のための条件を含んでおり、先に実施した。

3.5.3 被験者

13 名の被験者全員にオーディオメータ (Dana, DA-301) による聴力検査を実施し、正常聴力であることを確認した。続いて、調波複合音の高さの再認テスト (挿入音なし条件, 24 試行) を行い、C#4-C5 の範囲で誤りがゼロであった 12 名を以後の実験の被験者とした。この被験者選抜の基準は、純音と調波複合音の違いをのぞいて Deutsch (1970) と同じである。

また、被験者全員が絶対音感を保有していないことを 12 試行のテストにより確認した。このテストでは C#4-C5 の半音階を構成する調波複合音がランダムに 4 秒間隔で呈示され、被験者はそれらの音名を同定するよう求められた。わからない場合は「わからない」という解答も認めた。絶対音感を持つ被験者と持たない被験者の音名判断に要する反応時間、および正答率について調べた過去の研究 (Miyazaki, 1988) をもとに、12 試行すべてで正解した被験者を絶対音感保有者とみなすこととした。しかし、この基準を通過した被験者はいなかった。正答率の最高は 33.3 %, 平均は 7.8 %, 最低は 0 % であった。この正答率の水準は Miyazaki のいう、非絶対音感保有者 (non-AP subjects) の水準と一致する。

3.5.4 実験手続き

刺激系列は DAT (Sony, DTC-95ES) に録音され、ヘッドフォン (STAX, SR A Pro.) で防音ブース (Dana, DA-0020) 内の被験者の両耳に呈示された。防音ブース内の暗騒音を普通騒音計 (小野測器, LA-210) で測定したところ、およそ 40 dB (A) であった。

Fig. 3.1 のような順序，保持期間で刺激を呈示した．各刺激の開始時間の間隔は 500 ms（比較音の直前を除く），試行間隔は 10 秒とした．各実験条件の試行数は 12 試行であった．被験者は試行間隔中に反応を用紙に記入した．実験の実施順序は被験者間でランダムに割り当てた．

3.5.5 実験結果

高さの再認識率を Fig. 3.9 に，数詞の系列再生誤り率を Fig. 3.10 に示す．

高さの再認識率は，テスト音の高さが低くなると全体に増加する傾向がある．特に，調波複合音が挿入された場合には，テスト音の高さが低い条件，すなわちテスト音と挿入音の高さが近い条件になると誤り率が大きく上昇した．

また，数詞が挿入された条件では，高さの再認のみを課題とした場合の再認識率は挿入音のない条件と差がないが，高さの再認と数詞の系列再生の両方を課題とした場合は高さの再認識率が増加した．

一方，数詞の系列再生誤り率は，テスト音の高さの変化に影響されない．しかし，やはり課題の違いによる影響は明らかで，数詞の系列再生のみを要求された場合に比べ，数詞の系列再生と高さの再認の両方を要求された場合の数詞の系列再生誤り率は常に高い．

個人ごとの誤り率を逆正弦変換 (Snedecor and Cochran, 1967) した後，SAS の GLM プロシジャにより被験者の要因をブロック因子とみなし，有意水準を 1% として分散分析を行った．

高さの再認識率 (Fig. 3.12) については，標準音の基本周波数の範囲 (R)，挿入音の種類 (I)，課題 (T)，被験者 (S) の主効果および交互作用効果 $R \times I$ が有意で [それぞれ $F(1, 99) = 110.56, p = 0.0001$; $F(3, 99) = 40.17, p = 0.0001$; $F(1, 99) = 12.69, p = 0.0006$; $F(11, 99) = 3.39, p = 0.0005$; $F(3, 99) = 12.38, p = 0.0001$]，交互作用効果 $R \times T$ は有意でなかった [$F(1, 99) = 1.07, p = 0.3029$]．

数詞の系列再生誤り率 (Fig. 3.13) については，主効果 T, S が有意で [$F(1, 33) = 8.45, p = 0.0065$; $F(11, 33) = 11.59, p = 0.0001$]，主効果 R，交互作用効果 $R \times T$ は有意でなかった [$F(1, 33) = 2.41, p = 0.1304$; $F(1, 33) = 0.26, p = 0.6145$]．

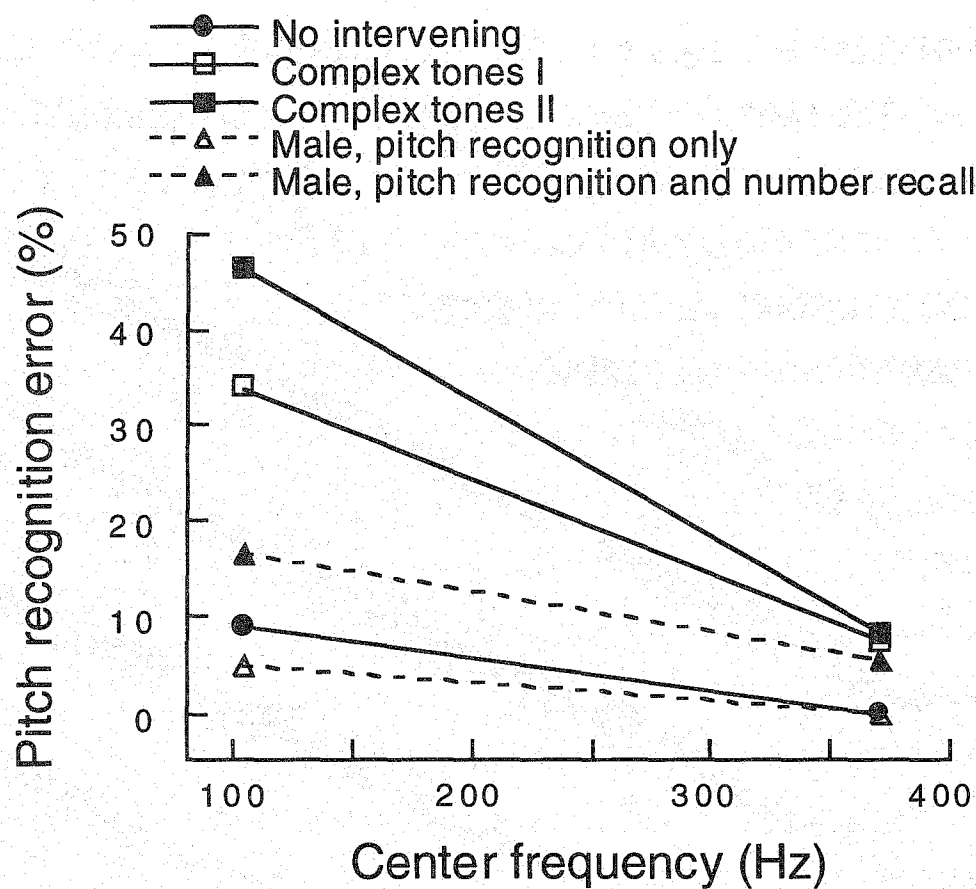


Fig. 3. 9 Percent of pitch recognition error plotted as a function of center frequency of the test-tone fundamentals.

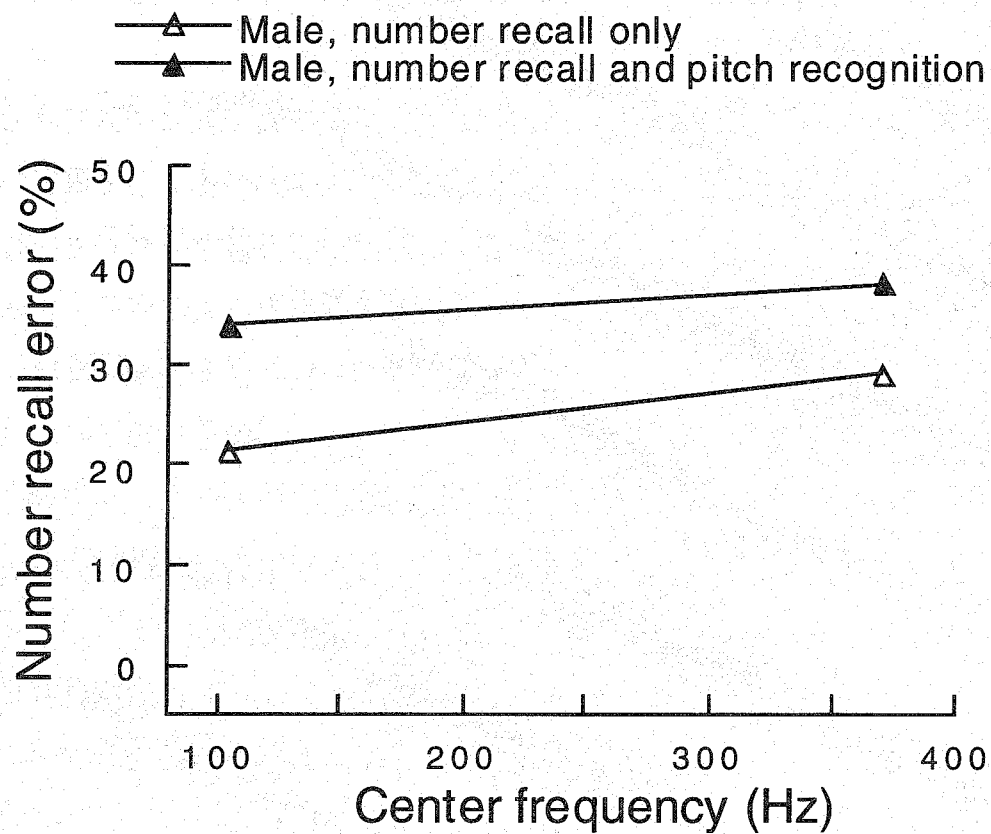


Fig. 3. 10 Percent of number recall error plotted as a function of center frequency of the test-tone fundamentals.

高さの再認識率について実験条件間の差の検定を行うため、テスト音の高さの範囲によってデータを 2 分割し、それぞれについて Tukey 法による多重比較を有意水準 5% で行った。

その結果、テスト音が低い条件では、調波複合音 I, II を挿入した条件とそれ以外の条件との間、および数詞を挿入して高さの再認のみを求めた条件と高さの再認と数詞の系列再生の両方を求めた条件との間で有意差が見られた。調波複合音 I と II の間には有意差が見られなかった。

テスト音が低い条件では、調波複合音 I, II を挿入した条件と挿入音なしの条件との間、および調波複合音 I, II を挿入した条件と数詞を挿入して高さの再認のみを求めた条件との間に有意差が見られた。調波複合音 I と II の間には、やはり有意差が見られなかった。

3.5.6 考察

3.5.6.1 調波複合音の高さと数詞音声系列の干渉

テスト音（調波複合音）の高さの再認と数詞音声の系列再生の両方を被験者に要求すると、それぞれの誤り率が有意に増大した。したがって、調波複合音の高さの記憶と数詞音声系列の記憶は互いに干渉しあい、特殊化されていない。すなわち、純音と数詞音声に関して 3.4 で得た結論が、調波複合音についてもあてはまることが明らかになった。

3.5.6.2 高さの再認識率とテスト音の高さの明確さ

まず、挿入音がない条件での高さの再認識率について考察する。

振幅の等しい純音を用いた研究 (3.4) では、テスト音が低い場合の誤り率が高かった (30.6%)。しかし本研究では、同様の条件での誤り率は 10% に達しなかった。これは明らかに、大きさの等しい調波複合音をテスト音として用いたことの効果である。すなわち、より明瞭な高さの記憶が残されるようになったと考えられる。また、挿入音がない条件での誤り率が低下したため、他の条件との差がより明確に見られるようになった。

次に挿入音のある条件について考える。

高さの再認識率にはテスト音と挿入音の高さの関係が大きく影響する (3. 1. 2). 本研究では、調波複合音を挿入した場合、テスト音と挿入音の高さが近い場合に大きな誤りが生じた。挿入音の高さもテスト音と同じ程度に明確であったため、大きな干渉が生じたと考えられる。これは、以前の研究と同様、高さの重要性を示す結果であるといえる。

3. 5. 6. 3 高さの再認識率と音声の高さの不明確さ

数詞音声が入挿された場合は調波複合音が入挿された場合に比べて、テスト音の高さの記憶は干渉されにくく、高さの再認のみを課題とした場合の誤り率は挿入音がない条件と有意差がなかった。また、数詞の系列再生課題が加えられると誤り率は有意に増加したが、それでも誤り率は 20 % 以下であった。この原因は何であろうか。

まず、音声の高さが連続的に変化するため、テスト音の高さと重なる場合も考えられるので、その効果について考察する。

調波複合音の挿入音として、テスト音と同じトーンクロマの音が出現しないようにした場合 (調波複合音 I) とそのような制約を設けない場合 (調波複合音 II) の両方で実験を行ったが、両者の間に有意差は見られなかった。

挿入音のなかに標準音と同じトーンクロマの音が出現した場合には、標準音の高さの記憶がリフレッシュされ、保持期間が実質的に短くなって成績が向上する (誤り率が低下する) 可能性も考えられるが、少なくとも今回の実験条件では、調波複合音に関してこの解釈を裏付けるような有意差は現れなかった。したがって、数詞音声を挿入した場合の高さの再認識率が調波複合音を挿入した場合よりも低いのは、音声の高さが連続的に変化するために標準音の高さを横切る場合が生じ、そのことによって標準音の高さの記憶が補強されるためであるとは考えられない。

これにかわる説明として、音声の高さの不明確さの影響が考えられる。すなわち、自然な音声の高さは連続的に大きく変化するため、はっきりした一定の

高さは知覚されにくく、テスト音の高さの記憶に対する干渉も調波複合音と比べて小さくなるためと考える (cf., 3.3.5.2) .

3.5.6.4 数詞の系列再生誤り率とテスト音の高さの明確さ

数詞の系列再生誤り率は、テスト音に純音を用いた研究 (3.4) と比べて全体に 10 % 程度増加し、その一方で、テスト音の高さの範囲に影響されなくなった。これはテスト音の高さの記憶が全体により強固になったため、テスト音の高さにかかわらず、挿入音に対してより大きな干渉効果が生じたためと考える。

3.5.7 まとめ

この実験で明らかになったことをまとめると、次のようになる。

- (1) 調波複合音の高さと数詞音声系列の記憶は干渉する。すなわち、両者の記憶は特殊化されていない。
- (2) テスト音の高さの再認識率には、テスト音と挿入音の高さの関係が大きく影響する。
- (3) テスト音の高さの明確さも、テスト音の高さの再認識率、数詞の系列再生誤り率に影響する。
- (4) テスト音の高さの再認に対する挿入音の干渉効果は、数詞音声の方が調波複合音よりも小さい。音声の高さの不明確さがこのことを説明する。

3.6 総合考察

3.6.1 音声と非音声の短期記憶メカニズム：何が干渉を起こすのか

本章では、音声と非音声の短期記憶が干渉しうることを示す実験事実を呈示してきた。すなわち、数詞音声の系列と非音声の高さの短期記憶は共有されているため、干渉が生じると考えられる。

Deutsch (1970, 1975c) や Pechmann and Mohr (1992) が主張するような、音声と非音声の記憶が特殊化されているという考え方の背後には、音声と非音声の区別が自動的になされ、その後それぞれ専用の処理装置によって処理が進み、記憶が保持されるという仮定がある。しかし、このような考え方では本研究で示されたデータ、すなわち、音声を単に無視して、非音声の高さの再認に専念すればよいという課題が、音声も記憶して系列再生し、同時に非音声の高さの再認も求められるという課題になると、高さの再認と系列再生の両方の誤り率が有意に増大するという実験事実がまったく説明できない。自動化され、特殊化された処理は、課題の影響を受けないはずだからである。また、Jones, Macken, and Harries (1997) の時間的、空間的な体制化による高さの再認成績への影響を明らかにした実験結果も、このような音声、非音声の特殊化という枠組みでは説明できない。

これに代わる考え方は、次のようなものであろう。

干渉効果の大きさに影響する要因として、次の四つを考える：(1) 刺激の心理的類似性、(2) 体制化、(3) 記憶痕跡の強さ、(4) 課題の負荷、である。

まず、心理的類似性の高い刺激の記憶は干渉を起こしやすいし、類似性の低い記憶は干渉を起こしにくいと考えられる。このことの裏付けは、本研究でテスト音と挿入音の高さが近いほど大きな干渉が見られるという実験結果が一貫して得られたことである。[ただし、類似性がきわめて高い場合には、逆に記憶を増強する場合もある (Deutsch, 1972).]

さらに、テスト音と挿入音が分離できるような何らかの体制化が行われれば、誤り率は低下する (Jones *et al.*, 1997)。

また、記憶痕跡の強弱も干渉効果の大小と関係すると考えられる。その裏付けとして本研究では、純音を用いた実験結果と調波複合音を用いた実験結果との比較により、高さが不明確な刺激の干渉効果はより小さいことが示された。音声についても同様のことが考えられる。

さらに、二つの課題を同時に課すことによって、負荷が増大すると成績が低下する。これは、限られた処理能力（記憶容量を含む）を二つの課題で奪い合うことになるためと考えられる。本研究では、高さの再認と数詞の系列再生の両方を課題とすると、どちらか単独の場合よりも両方の課題の誤り率が有意に増大することが示された。

これらの要因が組み合わさり、干渉効果の大きさを決定していると考えられる。

3.6.2 音声の有意味性

では、刺激が言語的な意味を持つものか、そうでないかが、短期記憶における干渉にどう影響するのかが次に疑問となるであろう。すなわち、音声と非音声が多様な心理的属性を共有するとしても、言語的な意味だけは共有できないと考えられるからである。

しかし、言語的な意味は基本的に長期記憶によって保持されているものであり、短期記憶での干渉効果には意味的側面は直接影響しないと考えられる。たとえば、数字の系列再生課題を視覚呈示し、その間に同時に音声を呈示すると、音声が多意味語であっても、無意味語であっても、被験者が理解できない外国語であっても系列再生の成績に干渉が生じる (Salamé and Baddeley, 1982, 1986)。

さらに、Jones (1993), Jones and Macken (1993), Beaman and Jones (in press) は一連の研究で同様の実験を行い、誤りを増大させるような聴覚刺激は、単調なものではなく、変化のあるものでなければならないことを示した。逆に、変化のあるものであれば、音声であっても非音声であっても、多意味音声であっても無意味音声であっても同様の効果を持ちうる。

これは二つのことを示している。すなわち、視覚呈示された系列が聴覚的に

保持されていること、また、干渉効果に関して、聴覚刺激の有意義性ないし意味的な関連性は本質的な要因ではないことである。したがって、Deutsch の実験パラダイムで考えると、言語的な意味を持たないテスト音も、音声の挿入音の系列記憶に干渉を及ぼすであろうと考えられる。このことは本研究により、テスト音の高さの明確さが、挿入された数詞音声の系列再生課題の成績に影響することで実際に示された。

3.7 まとめ

本章で得られた結論をまとめると、次のようになる。

- (1) 純音や調波複合音といった非音声の高さと、数詞音声の系列は短期記憶において干渉する。
- (2) テスト音と挿入音の高さの関係が、高さの再認識率に大きく影響する。
- (3) 刺激の高さの明確さが、高さの再認、数詞の系列再生の両方の成績に影響する。
- (4) 刺激の心理的類似性、体制化、記憶痕跡の強さ、課題の負荷が干渉の大きさに影響すると考えられる。
- (5) 音声の有意義性は、干渉の本質的な要因とは考えられない。

第 4 章 音の高さと音色，音声と非音声の関係

第 2 章で，音の高さと音色の間に両者の区別があいまいな部分が存在することを，また第 3 章で，音声の系列と非音声の高さとの間で干渉が生じ，両者の短期記憶メカニズムが特殊化されていないことを示した．これらの結果は，音の高さ，音色，音声，非音声といった用語から仮定される心理的次元が互いに独立ではなく，その境界があいまいで，そのような心理的次元が成立する過程では他の心理的次元と関係が深いと考えられている要素も影響を与えることを意味している．では，音の高さと音色，音声と非音声の処理メカニズムをどのようなものと考えればよいのだろうか．

4.1 心理的次元の階層的干渉モデル

心理的次元間の干渉，あるいは相互作用が生じるモデルとして，Melara and Marks (1990a-d) は次のようなものを考えた．このモデルでは，心理的次元ごとに処理チャンネルを考える．各チャンネルは階層的に構成され，各チャンネル内で下位の階層から順に処理が進んでいく．多次元的な刺激の分類課題を用いて，心理的次元間の干渉を調べた実験から，彼らは，階層内部で，また，異なるチャンネル間の同じレベルの階層どうし，および異なるチャンネル間の違うレベルの（上位と下位の）階層どうしで一方向または双方向の「漏洩 (crosstalk)」すなわち干渉作用が生じると考えた．

同様の考え方と実験方法のもとで，Krumhansl and Iverson (1992) は音の高さと音色の両方の次元を操作した実験を行った．その結果，音色は絶対的に判断され，高さは相対的に判断されることがわかった．しかし，音色の違いの度合いにも注意すべきであると彼らは指摘している．この実験では音色にはっきりした違いがあり，絶対的な判断がなされたが，音色の違いがより微妙だったなら，相対的な判断がなされた可能性がある．

さらに、この種の実験では、被験者の訓練の度合いも結果に影響することが明らかにされた (Pitt, 1994). 音楽的訓練の少ない被験者では、高さを判断するという課題とは無関係な音色の変化が高さの判断に影響する。しかし、音色を判断する課題で、課題と無関係な高さの変化が課題の成績に与える影響はより小さい。音楽的な被験者ではこのような無関係な次元が及ぼす影響は少なく、音色と高さの間の影響の仕方の非対称性もみられない。

もっとも、この実験で音楽的被験者の成績は ceiling performance に達しており、音楽的被験者にとっては非常にやさしい課題であったことがわかる。この種の実験では誤答がゼロに近づけば条件間の差をみることはできない。したがって、課題をより難しくすれば、音楽的被験者でも条件間の有意差がみられる可能性がある。訓練によって干渉が小さくなるということは、逆に、課題の難易度を調整すれば、どんな被験者でも干渉の存在を示すことができると考えられる。また、刺激の音色と高さは実験者が恣意的に設定したものであり、干渉の非対称性は刺激の設定によって変化する可能性がある。

4.2 階層的干渉モデルの拡張

本研究の立場からみて、この種のモデルの本質的な問題点は次のような点にある。それは、高さや音色はそれ自体が多次元であるのに、それらが基本的に単一の心理的次元として扱われていること、心理的次元の成立後の短期記憶過程がモデルに組み入れられていないことである。このモデルにおける心理的次元のチャネルは「被験者間で共通する、心理的に意味のある次元」(Melara and Marks, 1990d) と考えられているが、実験で操作された複数の次元は実験者が恣意的に設定した、大きく異なる物理的次元であり、それらの次元が被験者間で共有されているからといって、それぞれの次元が単一の次元であるとは限らない。

本論文で述べてきたような、音の高さや音色の多次元的な性質や個人差による次元間のウェイトの違いを説明するためには、さまざまな物理的特性からい

かにして特定の心理的次元（属性）が作り出され、選び出されて重み付けされ、さらにそれらが統合されてより大きな心理的次元が成立するのかを説明する必要がある。そのためには、心理的次元の階層的干渉モデルの枠組みで考えるならば、高さや音色といった単一の心理的次元を用語の上からアプリオリに仮定するのではなく、それらが単一の次元として成立する以前の段階で、さまざまな刺激の物理的特性を反映した神経情報の中から、徐々にそれらの一部が組み合わされて特定の心理的次元の成立に至るというようにモデルを拡張する必要がある。そして、このような情報の選択、組み合わせ、重みづけといった過程を考えることで、選択的注意、体制化、高さの概念の個人差の問題を扱えると考えられる。

さらに、そうやって成立した心理的次元が短期記憶に保持され、取り出される段階では、心理的類似性、記憶痕跡の強さ、記憶容量の限界を考えることで、非音声の高さと音声の系列の短期記憶における干渉を説明できる。また、課題の負荷の大きさは、心理的次元の成立、短期記憶での保持と取り出しのすべての段階で影響すると考えられる。

第 5 章 結論

本論文で明らかにされたことは、以下のように要約される。

聴覚におけるさまざまな心理的次元のうち、音の高さを中心として研究を行った。特に、音の高さの知覚に音色の及ぼす影響、非音声の高さの短期記憶に数詞音声系列の及ぼす影響を明らかにするために心理実験を行った。まず、音の高さと音色という用語の現在の定義と、音の高さおよび音色と物理量との関係について調べた過去の研究を紹介し、音の高さも音色も多次元的な性質を持ち、両者の間につながりが考えられることを説明した。また、音声と非音声という用語の定義を行い、本研究の目的を述べた（第 1 章）。

そして、無限音階複合音を変形した刺激と多次元尺度法を用いて、音の高さに二つの側面があり、それらを螺旋形の図式で表示できること、一方の側面は直線的、音色的なトーンハイトであり、他方の側面は循環的、音階的なトーンクロマであることを示した。被験者の判断の個人差は、これら二つの次元へのウェイトの違いとして表現された。また、被験者の音楽経験と個人差の関係は、先行研究で考えられていたほど単純ではないことが示唆された（第 2 章）。

さらに、従来は干渉を起こさないと考えられていた、非音声の高さと数詞音声の系列が短期記憶において干渉することを明らかにした。そして、刺激の心理的類似性、体制化、記憶痕跡の強さ、課題の負荷を考えることで実験結果が説明されることを示した（第 3 章）。

このような心理的次元間の相互作用を説明する既存のモデルとして、階層的干渉モデルを紹介した。しかし、このモデルでは、高さや音色の多次元性、短期記憶過程はまったく考慮されていないため、本研究の結果を説明できるようにモデルの拡張が必要であることを論じた（第 4 章）。

謝辞

本論文の執筆にあたってご指導いただいた、京都大学の清水御代明教授に深く感謝する。本論文のうち、第 1, 3 章で参照した、雑音の高さと音色に関する研究とヘッドフォンの特性に関する研究は、著者が ATR 視聴覚機構研究所に在職中に行われたものである。ご支援いただいた淀川英司社長（現工学院大学教授）、東倉洋一室長（現 NTT 基礎研究所所長）に感謝する。また、第 2 章の研究を進めるにあたって、NHK 放送科学基礎研究所の大串健吾先生（現京都市立芸術大学教授）にはひとかたならぬ援助をいただき、感謝のことばもない。大串教授なくしては私が聴覚心理学の道に進むことはなかったであろう。また、この第 2 章の実験は著者が京都大学文学部に学部学生として在学中に行われたものであり、当時ご指導いただいた本吉良治教授（現名誉教授）、故平野俊二教授に厚くお礼申し上げます。特に本吉教授には、早く博士論文を書くようにと折に触れて encourage していただいた。また、京都大学教養部（当時）の中島誠教授には防音室と実験装置の使用を許可していただき、感謝している。第 3 章のうち、3.3 の実験は、著者がフランス、ボルドー第 2 大学音響心理学研究所に客員研究員として滞在した折りに行われたものである。滞在費を支給してくれたフランス政府と、フランスに招いていただき滞在中も公私にわたり援助していただいた Catherine Semal と Laurent Demany に深く感謝する。3.4, 3.5 の実験には、元京都府立大学文学部学生の足立和佳美、永谷智美の両氏に実験、データ整理などで多大なご協力をいただき感謝している。京都大学の乾敏郎教授には著者が大学院在学中から研究会などを通じてお世話になり、また本論文執筆のきっかけを与えていただいた。ここに謝意を表する。本研究の一部は文部省科学研究費（課題番号 6710047, 8710052）、永井特別研究奨励金の補助を受けた。

文献

- Allen, D. (1967). "Octave discriminability of musical and non-musical subjects," *Psychon. Sci.*, 7, 421-422.
- 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.
- 安藤由典 (1971). *楽器の音響学* (音楽之友社, 東京) .
- Bachem, A. (1950). "Tone height and tone chroma as two different pitch qualities," *Acta Psychol.* 7, 80-88.
- Beauchamp, J. W. (1966). "Additive synthesis of harmonic musical tones," *J. Audio Eng. Soc.* 14, 332-342.
- Beaman, C. P., and Jones, D. M. (in press). "The role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis," *J. Exp. Psychol.: Learning, Memory, and Cognition*.
- Benedini, K. (1980). "Klangfarbenunterschiede zwischen tiefpassgefilterten harmonischen Klängen," *Acustica* 44, 129-134.
- Berger, K. W. (1964). "Some factors in the recognition of timbre," *J. Acoust. Soc. Am.* 36, 1888-1891.
- Bloothoof, G., and Plomp, R. (1988). "The timbre of sung vowels," *J. Acoust. Soc. Am.* 84, 847-860.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (Cambridge, MIT Press).
- Burns, E. M. (1981). "Circularity in relative pitch judgments for inharmonic complex tones: The Shepard demonstration revisited, again," *Percept. Psychophys.* 30, 467-472.
- Burns, E. M. (1982). "Pure-tone pitch anomalies. I. Pitch-intensity effects and diplacusis in normal ears," *J. Acoust. Soc. Am.* 72, 1394-1402.
- Charbonneau, G. (1979). "Effets perceptifs de la réduction des données dans la reconnaissance du timbre," *C. R. Acad. Sc. Paris, t. 289, Série B*, 147-149.
- Charbonneau, G., and Risset, J. C. (1973). "Circularité de jugements de hauteur sonore," *C. R. Acad. Sci. Paris, t. 277, Série B*, 623-626.
- Charbonneau, G., and Risset, J. C. (1975). "Jugements relatifs de hauteur: Schémas linéaires et hélicoïdaux," *C. R. Acad. Sc. Paris, t. 281, Série B*, 289-292.
- Chuang, C. K., and Wang, W. S. (1978). "Psychophysical pitch biases related to vowel quality, intensity difference, and sequential order," *J. Acoust. Soc. Am.* 64, 1004-1014.

- 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, 1521-1528.
- Clarkson, M. G., and Clifton, R. K. (1995). "Infants' pitch perception: Inharmonic tonal complexes," *J. Acoust. Soc. Am.* 98, 1372-1379.
- de Bruijn, A. (1978). "Timbre-classification of complex tones," *Acustica* 40, 108-114.
- Demany, L. (1989). "Perception de la hauteur tonale," in *Psychoacoustique et Perception Auditive*, by M. C. Botte, G. Canévet, L. Demany, and C. Sorin (INSERM, Paris). [上田和夫訳 (1989). 音の高さの知覚——「音響心理と聴知覚」より——, ATR テクニカルレポート, TR-A-0050.]
- Deutsch, D. (1970). "Tones and numbers: Specificity of interference in immediate memory," *Science*, 168, 1604-1605.
- Deutsch, D. (1972). "Mapping of interactions in the pitch memory store," *Science*, 175, 1020-1022.
- Deutsch, D. (1973a). "Interference in memory between tones adjacent in the musical scale," *J. Exp. Psychol.*, 100, 228-231.
- Deutsch, D. (1973b). "Octave generalization of specific interference effects in memory for tonal pitch," *Perception & Psychophysics*, 13, 271-175.
- Deutsch, D. (1974). "Generality of interference by tonal stimuli in recognition memory for pitch," *Quart. J. Exp. Psychol.*, 26, 229-234.
- Deutsch, D. (1975a). "Auditory memory," *Canadian Journal of Psychology*, 29, 87-105.
- Deutsch, D. (1975b). "Facilitation by repetition in recognition memory for tonal pitch," *Memory & Cognition*, 3, 263-266.
- Deutsch, D. (1975c). "The organization of short-term memory for a single acoustic attribute," in *Short-term Memory*, edited by D. Deutsch and S. A. Deutsch (Academic, New York).
- Deutsch, D. (1978a). "Delayed pitch comparisons and the principle of proximity," *Perception & Psychophysics*, 23, 227-230.
- Deutsch, D. (1978b). "Interference in pitch memory as a function of ear of input," *Quart. J. Exp. Psychol.*, 30, 283-287.
- Deutsch, D. and Feroe, J. (1975). "Disinhibition in pitch memory," *Perception & Psychophysics*, 17, 320-324.
- Drobisch, M. W. (1855). "Über musikalische Tonbestimmung und Temperatur," *Abhandl. Math. Phys. Kl. Königl. Sachs. Ges. Wiss.* 4, 1-120. [Cited in Pikler (1966). "Logarithmic frequency systems," *J. Acoust. Soc. Am.* 39, 1102-1110.]
- Fastl, H. (1971). "Über Ton höhenempfindungen bei Rauschen," *Acustica*, 25, 350-354.
- Fastl, H. (1980). "Pitch strength and masking patterns of low-pass noise," in

- Psychophysical, physiological and Behavioural Studies in Hearing*, edited by G. Van den Brink, and F. A. Bilsen (Delft University Press, Delft), pp. 334-340.
- Grey, J. M. (1977). "Multidimensional perceptual scaling of musical timbres," *J. Acoust. Soc. Am.* 61, 1270-1277.
- Grey, J. M. (1978). "Timbre discrimination in musical patterns," *J. Acoust. Soc. Am.* 64, 467-472.
- Grey, J. M., and Gordon, J. W. (1978). "Perceptual effects of spectral modifications on musical timbres," *J. Acoust. Soc. Am.* 63, 1493-1500.
- Gery, J. M. and Moorer, J. A. (1977). "Perceptual evaluations of synthesized musical instrument tones," *J. Acoust. Soc. Am.* 62, 454-462.
- Hartmann, W. M. (1996). "Pitch, periodicity, and auditory organization," *J. Acoust. Soc. Am.* 100, 3491-3502.
- Hesse, H. P. (1982). "The judgment of musical intervals," in *Music, Mind, and Brain*, edited by M. Clynes (Plenum, New York), Chap. 11.
- 平原達也, 鈴木真一, 上田和夫 (1990). "聴覚実験用ヘッドホンの歪率の測定," ATR テクニカルレポート, TR-A-0082.
- Jones, D. M. (1993). "Objects, streams, and threads of auditory attention," in *Attention: Selection, Awareness, and Control*, edited by A. Baddeley and L. Weiskrantz (Clarendon Press, Oxford).
- Jones, D. M., and Macken, W. J. (1993). "Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory," *J. Exp. Psychol.: Learning, Memory, and Cognition*, 19, 369-381.
- Jones, D. M., Macken, W. J., and Harries, C. (1997). "Disruption of short-term recognition memory for tones: Streaming or interference?" *Q. J. Exp. Psychol.* 50A, 337-357.
- Kameoka, A., and Kuriyagawa, M. (1969a). "Consonance theory Part I: Consonance of dyads," *J. Acoust. Soc. Am.* 45, 1451-1459.
- Kameoka, A., and Kuriyagawa, M. (1969b). "Consonance theory Part II: Consonance of complex tones and its calculation method," *J. Acoust. Soc. Am.* 45, 1460-1469.
- Keeler, J. S. (1972). "The attack transients of some organ pipes," *IEEE Trans. Audio. Electroacoust.* AU-20, 378-391.
- Kitamura, O., Namba, S., and Matsumoto, R. (1968). "Factor analytical research of tone colour," *Proc. 6th ICA*, Tokyo, paper A-5-11.
- Krumhansl, C. L., and Iverson, P. (1992). "Perceptual interactions between musical pitch and timbre," *J. Exp. Psychol.: Human Perception and Performance*, 18, 739-751.
- Kruskal, J. B. (1964). "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika* 29, 1-27.

- Kruskal, J. B., and Wish, M. (1978). *Multidimensional Scaling* (Sage, London). [高根 芳雄訳 (1980). 多次元尺度法 (朝倉書店, 東京)] .
- 廚川守, 八尋博司, 柏木成豪 (1978a) “音質評価のための 7 属性,” 日本音響学会誌 34, 493-500.
- 廚川守, 八尋博司, 柏木成豪 (1978b) “音の 7 属性の性格について,” 日本音響学会誌 34, 501-509.
- Liberman, A. M., and Mattingly, I. G. (1985). “The motor theory of speech perception revised,” *Cognition*, 21, 1-36.
- Lichte, W. H. (1941). “Attributes of complex tones,” *J. Exp. Psychol.* 28, 455-480.
- Melara, R. D., and Marks, L. E. (1990a). “Dimensional interactions in language processing: Investigating directions and levels of crosstalk,” *J. Exp. Psychol.: Learning, Memory, and Cognition*, 16, 539-554.
- Melara, R. D., and Marks, L. E. (1990b). “HARD and SOFT interacting dimensions: Differential effects of dual context on classification,” *Perception & Psychophysics*, 47, 307-325.
- Melara, R. D., and Marks, L. E. (1990c). “Interaction among auditory dimensions: Timbre, pitch, and loudness,” *Perception & Psychophysics*, 48, 169-178.
- Melara, R. D., and Marks, L. E. (1990d). “Perceptual primacy of dimensions: Support for a model of dimensional interaction,” *J. Exp. Psychol.: Human Perception and Performance*, 16, 398-414.
- Meyer, M. (1904). “On the attributes of the sensations,” *Psychol. Rev.* 11, 83-103.
- Miller, J. R., and Carterette, E. C. (1975). “Perceptual space for musical structures,” *J. Acoust. Soc. Am.* 58, 711-720.
- Miyazaki, K. (1988). “Musical pitch identification by absolute pitch possessors,” *Perception & Psychophysics*, 44, 501-512.
- Miyazaki, K. (1989). “Absolute pitch identification: effects of timbre and pitch region,” *Music Perception*, 7, 1-14.
- Moore, B. C. J. (1973). “Frequency difference limens for short-duration tones,” *J. Acoust. Soc. Am.*, 54, 610-619.
- Moore, B. C. J. (1989). *An Introduction to the Psychology of Hearing*, 3rd ed. (Academic Press, London). [大串健吾監訳, 上田和夫他訳 (1994). 聴覚心理学概論 (誠信書房, 東京)] .
- Moore, B. C. J., and Glasberg, B. R. (1990). “Frequency discrimination of complex tones with overlapping and non-overlapping harmonics,” *J. Acoust. Soc. Am.* 87, 2163-2177.
- Moore, B. C. J., and Glasberg, B. R. (1991). “Effects of signal-to-noise ratio on the frequency discrimination of complex tones with overlapping or non-overlapping

- harmonics," J. Acoust. Soc. Am. 89, 2858-2865.
- Moore, B. C. J., Glasberg, B. R., and Proctor, G. M. (1992). "Accuracy of pitch matching for pure tones and for complex tones with overlapping or nonoverlapping harmonics," J. Acoust. Soc. Am. 91, 3443-3450.
- 森谷彰 (1995). 高調波構造をもつ複合音の音色知覚過程に関する研究, 東北大学審査修士論文.
- Nakajima, Y., Tsumura, T., Matsuura, S., Minami, H., and Teranishi, R. (1988). "Dynamic pitch perception for complex tones derived from major triads," Music Perception, 6, 1-20.
- 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.
- 日本工業規格 (1986). JIS Z 8109 音響用語 (聴覚・音声・音楽) (日本工業規格協会, 東京) .
- Nordmark, J. O. (1968). "Mechanisms of frequency discrimination," J. Acoust. Soc. Am. 44, 1533-1540.
- 大串健吾 (1980). "複合音の音色を支配する物理的・心理的要因について," 日本音響学会学会誌 36, 253-239.
- 大串健吾 (1981). "無限音階複合音の高さの多次元分析," 日本音響学会聴覚研究会資料 H-81-36.
- 大串健吾 (1984). "複合音の高さの循環性とその応用," 電子通信学会論文誌, J67-A, 423-430.
- Ohgushi (1985). "Circularity of the pitch of complex tones and its application," Electron. Commun. Jpn. Part 1, 68, 1-10.
- 小澤賢司 (1993). 聴覚系における音色知覚過程に関する研究, 東北大学審査博士論文.
- Patterson, R. D. (1989). "The tone height of multi-harmonic sounds," in *Proceedings of the 1st International Conference on "Music Perception and Cognition,"* Kyoto, Japan, pp. 119-124.
- Pechmann, T., and Mohr, G. (1992). "Interference in memory for tonal pitch: Implications for a working-memory model," Memory & Cognition, 20, 314-320.
- Pitt, M. A. (1994). "Perception of pitch and timbre by musically trained and untrained listeners," J. Exp. Psychol.: Human Perception and Performance, 20, 976-986.
- Plomp, R. (1976). *Aspects of Tone Sensation* (Academic Press, London).
- Plomp, R., and Levelt, W. J. M. (1965). "Tonal consonance and critical bandwidth," J. Acoust. Soc. Am. 38, 548-560.
- Plomp, R., Pols, L. C. W., and van de Geer, J. P. (1967). "Dimensional analysis of vowel spectra," J. Acoust. Soc. Am. 41, 707-712.

- Pollack, I. (1978). "Decoupling of auditory pitch and stimulus frequency: The Shepard demonstration revisited," J. Acoust. Soc. Am. 63, 202-206.
- Preis, A. (1984). "An attempt to describe the parameter determining the timbre of steady-state harmonic complex tones," Acustica, 55, 1-13.
- Rainbolt, H. R., and Schubert, E. D. (1968). "Use of noise bands to establish noise pitch," J. Acoust. Soc. Am. 43, 316-323.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., and Carrell, T. D. (1981). "Speech perception without traditional speech cues," Science, 212, 947-950.
- Révész, G. (1913). *Zur Grundlegung der Tonpsychologie* (Vert & Comp, Leipzig). [高野瀏 訳 (1932). 音響心理學原論 (改造社, 東京)]
- Risset, J. C. (1969). "Pitch control and pitch paradoxes demonstrated with computer-synthesized sounds," J. Acoust. Soc. Am. 46, 88 (A).
- Risset, J. C. (1971). "Paradoxes de hauteur: Le concept de hauteur sonore n'est pas le meme pour tout le monde," Proc. 7th ICA Congr., Budapest, S10, 613-616.
- Risset, J. C. (1978). "Paradoxes de hauteur," IRCAM Rep. 10.
- Risset, J. C., and Wessel, D. L. (1982). "Exploration of timbre by analysis and synthesis," in *The Psychology of Music*, edited by D. Deutsch (Academic, New York), Chap. 2, pp. 25-58.
- Risset, J. C. (1986). "Pitch and rhythm paradoxes: Comments on 'Auditory paradox based on fractal waveform' [J. Acoust. Soc. Am. 79, 186-189 (1986)]," J. Acoust. Soc. Am. 80, 961-962.
- Ruckmick, C. A. (1929). "A new classification of tonal qualities," Psychol. Rev. 36, 172-180.
- 匂坂芳典, 浦谷則好 (1992). "ATR 音声・言語データベース," 日本音響学会誌, 48, 878-882.
- Salamé, P., and Baddeley, A. (1982). "Disruption of short-term memory by unattended speech: Implications for the structure of working memory," J. Verbal Learning and Verbal Behavior 21, 150-164.
- Salamé, P., and Baddeley, A. (1986). "Phonological factors in STM: Similarity and the unattended speech effect," Bulletin of the Psychonomic Society 24, 263-265.
- Saldanha, E. L., and Corso, J. F. (1964). "Timbre cues and the identification of musical instruments," J. Acoust. Soc. Am. 36, 2021-2026.
- Schroeder, M. R. (1986). "Auditory paradox based on fractal waveform," J. Acoust. Soc. Am. 79, 186-189.
- Semal, C., and Demany, L. (1991). "Dissociation of pitch from timbre in auditory short-term memory," J. Acoust. Soc. Am., 89, 2404-2410.
- Semal, C., and Demany, L. (1993). "Further evidence for an autonomous processing of

- pitch in auditory short-term memory," J. Acoust. Soc. Am., 94, 1315-1322.
- Semal, C., Demany, L., Ueda, K., and Hallé, P. A. (1996). "Speech versus nonspeech in pitch memory," J. Acoust. Soc. Am. 100, 1132-1140.
- Shepard, R. N. (1964). "Circularity in judgments of relative pitch," J. Acoust. Soc. Am. 36, 2346-2353.
- Shepard, R. N. (1982a). "Geometrical approximations to the structure of musical pitch," Psychol. Rev. 89, 305-333.
- Shepard, R. N. (1982b). "Structural representations of musical pitch," in *The Psychology of Music*, edited by D. Deutsch (Academic, New York), Chap. 11, pp. 343-390.
- Singh, P. G., and Hirsh, I. J. (1992). "Influence of spectral locus and F_0 changes on the pitch and timbre of complex tones," J. Acoust. Soc. Am. 92, 2650-2661.
- Sivian, L. J., and White, S. D. (1933). "On minimum audible sound fields," J. Acoust. Soc. Am. 4, 288-321.
- Slawson, A. W. (1968). "Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency," J. Acoust. Soc. Am. 43, 87-101.
- Small, A. M. Jr., and Daniloff, R. G. (1966). "Pitch of noise bands," J. Acoust. Soc. Am. 41, 506-512.
- Snedecor, G. W., and Cochran, W. G. (1967). *Statistical Methods*, 6th edition (Iowa State University Press, Ames). [畑村又好, 奥野忠一, 津村善郎共訳 (1972). 統計的方法: 原書第6版 (岩波書店, 東京)].
- Solomon, L. N. (1958). "Semantic approach to the perception of complex sounds," J. Acoust. Soc. Am. 30, 412-425.
- 曾根敏夫, 城戸健一, 二村忠元 (1962). "音の評価に使われることばの分析," 日本音響学会誌 18, 320-326.
- Stevens, S. S. (1935). "The relation of pitch to intensity," J. Acoust. Soc. Am. 6, 150-154.
- Stevens, S. S., and Davis, H. (1938). *Hearing: Its psychology and physiology* (American Institute of Physics, New York).
- Strong, W., and Clark, M., Jr. (1967). "Perturbations of synthetic orchestral wind-instrument tones," J. Acoust. Soc. Am. 41, 277-285.
- 舘田章, 磯部孝 (1973). "調和音の音色に及ぼす部分音の位相の影響," 医用電子と生体工学, 11, 108-116.
- Takane, Y., Young, F. W., and de Leeuw, J. (1977). "Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features," Psychometrika 42, 7-67.
- 高根芳雄 (1980). 多次元尺度法 (東京大学出版会, 東京).

- Teranishi, R. (1986). "Endlessly rising or falling chordal tones which can be played on the piano; another variations of the Shepard demonstration," 12th ICA, Toronto, K2-3.
- 寺西立年, 津村尚志 (1983). "多次元尺度解析を用いた相対音感の研究," 聴覚研究会資料, H-83-70.
- Terhardt, E. (1974). "On the perception of periodic sound fluctuations (roughness)," *Acustica* 30, 201-213.
- 上田和夫 (1983). "無限音階複合音の高さの多次元分析——直接判断データによる——," 京都大学文学部哲学科心理学専攻, 演習 II 発表資料 (未出版).
- 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.
- Ueda, K., and Akagi, M. (1990). "Sharpness and amplitude envelopes of broadband noise," *J. Acoust. Soc. Am.* 87, 814-819.
- Ueda, K. (1993). "Perceptual aspects of voice individuality," *Scientific Reports of Kyoto Prefectural University; Humanities*, 45, 159-171.
- 上田和夫, 平原達也 (1990). "聴覚実験用ヘッドフォンの検討," 日本音響学会聴覚研究会, 資料番号 H-90-5.
- Ueda, K., and Hirahara, T. (1991). "Frequency response of headphones measured in free field and diffuse field by loudness comparison," *J. Acoust. Soc. Jpn. (E)*, 12, 131-138.
- Verschuure, J., and van Meeteren, A. A. (1975). "The effect of intensity on pitch," *Acustica*, 32, 33-44.
- 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.
- Walliser, K. (1969). "Über die Abhängigkeiten der Tonhöhenempfindung von Sinustönen vom Schallpegel, von überlagertem drosselndem Störschall und von der Darbietungsdauer," *Acustica*, 21, 211-221.
- Ward, W. D., and Burns, E. M. (1982). "Absolute pitch," in *The Psychology of Music*, edited by D. Deutsch (Academic, New York), Chap. 14, pp. 432-434.
- Wedin, L., and Goude, G. (1972). "Dimension analysis of the perception of instrumental timbre," *Scand. J. Psychol.* 13, 228-240.
- 山口公典, 新実尚久, 竹田俊夫 (1970). "倍音の非調和性と楽音の自然性との関係," 日本音響学会誌, 26, 512-524.
- 矢田部達郎 (監修) (1962). 心理学初歩, 三訂版 (培風館, 東京).
- Zwicker, E., and Fastl, H. (1990). *Psychoacoustics: Facts and Models* (Springer-Verlag, Berlin), Chap. 5., pp. 120-130.

本研究に関する研究業績

論文

- 上田和夫, 大串健吾 (1984). “多次元尺度法による音の高さの二面性の空間的表現,” 日本音響学会誌 40, 798-804.
- Ueda, K., and Ohgushi, K. (1987). “Perceptual components of pitch: spatial representation using a multidimensional scaling technique,” J. Acoust. Soc. Am. 82, 1193-1200.
- 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.
- Ueda, K., and Akagi, M. (1990). “Sharpness and amplitude envelopes of broadband noise,” J. Acoust. Soc. Am. 87, 814-819.
- Ueda, K., and Hirahara, T. (1991). “Frequency response of headphones measured in free field and diffuse field by loudness comparison,” J. Acoust. Soc. Jpn. (E), 12, 131-138.
- Semal, C., Demany, L., Ueda, K., and Hallé, P. A. (1996). “Speech versus nonspeech in pitch memory,” J. Acoust. Soc. Am. 100, 1132-1140.

招待講演

- Ueda, K. (1996). “A hierarchical structure for adjectives describing timbre,” Third joint meeting of Acoust. Soc. Am. and Acoust. Soc. Jpn., J. Acoust. Soc. Am. vol. 100. no. 4, Pt. 2, 2751.

国際会議

- Ueda, K., and Hirahara, T. (1990). “Frequency response of headphones measured in free field and diffuse field by loudness comparison,” 119th meeting of Acoust. Soc. Am., J. Acoust. Soc. Am. Suppl. 1, vol. 87, S142.
- Hirahara, T., and Ueda, K. (1990). “Investigation of headphones suitable for psychoacoustical experiments,” 119th meeting of Acoust. Soc. Am., J. Acoust. Soc. Am. Suppl. 1, vol. 87, S142.
- Semal, C., Demany, L., Ueda, K., and Hallé, P. (1995). “Short-term pitch memory: Speech versus nonspeech,” Proceedings of the 11th meeting of the International Society for Psychophysics (Cassis, France, October 21-25, 1995), edited by C.A. Possamai, pp.115-120.
- Ueda, K. (1996). “Auditory short-term memory interference: Deutsch's demonstration revisited,” Third joint meeting of Acoust. Soc. Am. and Acoust. Soc. Jpn., J. Acoust. Soc. Am. vol. 100. no. 4, Pt. 2, 2628-2629.

シンポジウム

上田和夫 (1988). “Shepard の無限音階とその発展,” 日本心理学会第 52 回大会
自主シンポジウム, “音楽的高さの知覚・認知をめぐって.”

研究会

上田和夫, 大串健吾 (1984). “多次元尺度法による音の高さの 2 面性の空間的表現,” 日本音響学会聴覚研究会, 資料番号 H-84-16 (講演は英語で行なわれた).

上田和夫 (1987). “音色の表現語に階層構造は存在するか,” 日本音響学会聴覚研究会, 資料番号 H-87-18.

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, 資料番号 H-87-64.

上田和夫, 平原達也 (1989). “ラウドネス比較法によるヘッドフォン周波数特性の測定,” 日本音響学会聴覚研究会, 資料番号 H-89-38 (共催: 電気音響研究会, 資料番号 EA89-60).

上田和夫, 平原達也 (1990). “聴覚実験用ヘッドフォンの検討,” 日本音響学会聴覚研究会, 資料番号 H-90-5.

上田和夫 (1991). “話者の性の同定における音の高さの知覚,” 第 24 回知覚コロキウム (佐渡).

上田和夫 (1991). “フォルマント構造を持つ雑音の高さの相対判断,” 日本音響学会聴覚研究会, 資料番号 H-91-41.

Ueda, K., Semal, C., and Demany, L. (1992). “Tones and numbers: Little specificity of interference in immediate memory,” SFA - Groupe Audition, Séminaire Jeunes Chercheurs (Gradignan).

上田和夫, キャトリーヌ・スマル, ローラン・ドゥマニ (1993). “短期記憶への干渉効果,” 日本音響学会聴覚研究会, 資料番号 H-93-77 (共催: 音楽音響研究会, 資料番号 MA-93-30).

上田和夫, 足立和佳美 (1996). “聴覚短期記憶における干渉効果: 純音と数詞音声,” 日本音響学会聴覚研究会, 資料番号 H-96-4.

上田和夫 (1996). “聴覚短期記憶の干渉効果: 純音と数詞音声,” 第 29 回知覚コロキウム (神戸).

上田和夫, 永谷智美 (1997). “聴覚短期記憶の干渉効果: 調波複合音と数詞音声,” 日本音響学会聴覚研究会, 資料番号 H-97-13.

上田和夫 (1997). “聴覚短期記憶の干渉効果: 調波複合音と数詞音声,” 第 30 回知覚コロキウム (千葉).

一般講演

- 上田和夫, 大串健吾 (1984). "多次元尺度法による音の高さの二面性の空間的表現," 日本音響学会春季研究発表会 1-4-7, 講演論文集 pp. 227-228.
- 上田和夫 (1987). "音色の表現語に階層構造は存在するか," 日本音響学会春季研究発表会 2-4-4, 講演論文集 pp. 265-266.
- 上田和夫, 赤木正人 (1988). "音の鋭さと振幅包絡," 日本音響学会春季研究発表会 2-4-1, 講演論文集 pp. 311-312.
- 上田和夫, 赤木正人 (1988). "振幅包絡のピーク位置と鋭さ," 日本音響学会秋季研究発表会 1-1-6, 講演論文集 pp. 373-374.
- 平原達也, 上田和夫 (1989). "聴覚実験用ヘッドホンの検討," 日本音響学会秋季研究発表会 2-2-4, 講演論文集 pp. 427-438.
- 上田和夫, 平原達也 (1989). "STAX SR Λ Pro 実耳装着時の周波数特性," 日本音響学会秋季研究発表会 2-2-5, 講演論文集 pp. 429-430.
- 上田和夫, 平原達也 (1990). "ラウドネス比較法によるヘッドフォン周波数特性の測定," 日本音響学会春季研究発表会 1-5-5, 講演論文集 pp. 353-354.
- 上田和夫 (1992). "フォルマント構造を持つ雑音の高さの相対判断," 日本音響学会春季研究発表会 1-6-10, 講演論文集 pp. 381-382.
- 上田和夫, キャトリヌ・スマル, ローラン・ドゥマニ (1993). "短期記憶への干渉効果: 定常音と数詞の差は小さい," 日本音響学会秋季研究発表会 3-9-5, 講演論文集 pp. 433-434.
- 上田和夫, 足立和佳美 (1996). "聴覚短期記憶の干渉効果: 純音と数詞音声," 日本音響学会春季研究発表会 1-3-23, 講演論文集 pp. 441-442.
- 上田和夫 (1996). "聴覚短期記憶の干渉効果: 純音と数詞音声," 日本心理学会第60回大会, 発表論文集, p. 792.
- 上田和夫, 永谷智美 (1997). "聴覚短期記憶の干渉効果: 調波複合音と数詞音声," 日本音響学会春季研究発表会 2-8-11, 講演論文集 pp. 367-368.
- 上田和夫 (1997). "聴覚短期記憶の干渉効果: 調波複合音と数詞音声," 日本心理学会第61回大会, 発表論文集, p. 800.

その他

- 上田和夫 (1983). "無限音階複合音の高さの多次元分析——直接判断データによる——," 京都大学文学部哲学科心理学専攻, 演習 II 発表資料 (未出版).
- 上田和夫 (1987). "STAX SR Λ 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, 赤木正人 (1988). "音の鋭さと振幅包絡," ATR テクニカルレポート, TR-A-0022.
- Ueda, K., and Akagi, M. (1989). "Sharpness and amplitude envelopes of broadband

noise," ATR Technical Report, TR-A-0051.

平原達也, 鈴木真一, 上田和夫 (1990). "聴覚実験用ヘッドホンの歪率の測定," ATR テクニカルレポート, TR-A-0082.

Ueda, K., and Hirahara, T. (1990). "Frequency response of headphones measured in free field and diffuse field by loudness comparison," ATR Technical Report, TR-A-0089.

Ueda, K. (1992). "Perceptual aspects of voice individuality," ATR Technical Report, TR-A-0132.

Ueda, K. (1992). "Relative pitch judgments for formant structured broadband noise," ATR Technical Report, TR-A-0138. (非公開)

Ueda, K. (1993). "Perceptual aspects of voice individuality," Scientific Reports of Kyoto Prefectural University; Humanities, 45, 159-171.

学会賞

日本音響学会 栗屋 潔 学術奨励賞, 1988 年 3 月 24 日.

受賞対象: "音色の表現語に階層構造は存在するか," 昭和 62 年度春季研究発表会での講演発表.

その他の研究業績

論文

- Ueda, K., and Ohtsuki, M. (1996). "The effect of sound pressure level difference on filled duration extension," *Journal of the Acoustical Society of Japan (E)*, vol. 17, no. 3, 159-161.

国際会議

- Ueda, K., and Seo, N. (1998). "Short-term auditory memory interference: The effect of speech pitch salience," *Proceedings of the 16th International congress on Acoustics and the 135th meeting of Acoustical Society of America, Seattle*, 2355-2356.

シンポジウム

- 上田和夫 (1997). "科学と疑似科学, 心理学と疑似心理学, 治療と疑似治療: 音楽療法の評価," *日本音楽知覚認知学会, 公開シンポジウム資料* pp. 13-19.

研究会

- 上田和夫, 大串健吾, 藤原義章 (1986). "テンポのゆらぎと快さ," 第 19 回知覚コロキウム (仙台).
- 上田和夫, 岡田斉, 東倉洋一 (1988). "単音節のダイコティックリスニング: 耳の優位性における個人差," *日本音響学会聴覚研究会, 資料番号 H-88-67*.
- 上田和夫, 岡田斉, 東倉洋一 (1989). "明瞭性のバランスされた単音節のダイコティック・リスニング," 第 22 回知覚コロキウム (伊豆).
- 上田和夫, 大舘称 (1995). "時間伸長錯覚における先行音の音圧レベルの効果," *日本音響学会聴覚研究会, 資料番号 H-95-36*.
- 上田和夫, 瀬尾直子 (1998). "聴覚短期記憶の干渉効果: 音声の高さの明確さによる変化," *日本音響学会聴覚研究会, 資料番号 H-98-9*.

一般講演

- 上田和夫, 大串健吾, 藤原義章 (1986). "演奏における拍の長さの時間変化と快さ," *日本心理学会第 50 回大会, 感覚・知覚 13-1, 発表論文集* pp. 190.
- 上田和夫, 岡田斉, 東倉洋一 (1989). "明瞭性のバランスされた単音節のダイコティック・リスニング," *日本音響学会春季研究発表会 2-8-8, 講演論文集* pp. 323-324.
- 上田和夫, 東倉洋一 (1989). "単音節のダイコティック・リスニング," *日本心理学会第 53 回大会, 認知 4-10, 発表論文集* pp. 694.

- 上田和夫, 大拙称 (1995). “時間伸長錯覚における先行音の音圧レベルの効果,” 日本音響学会春季研究発表会 2-9-11, 講演論文集 pp. 467-468.
- 田原俊之, 上田和夫 (1996). “時間伸長錯覚における先行音の周波数の影響,” 日本音響学会春季研究発表会 1-3-21, 講演論文集 pp. 437-438.
- 上田和夫, 瀬尾直子 (1998). “聴覚短期記憶の干渉効果: 音声の高さの明確さによる変化,” 日本音響学会春季研究発表会 2-8-19, 講演論文集 pp. 443-444.

翻訳

- 上田和夫 訳 (1989). 音の高さの知覚——”音響心理と聴知覚” より——, ATR テクニカルレポート, TR-A-0050. Laurent Demany (1989). "Perception de la hauteur tonale," in *Psychoacoustique et perception auditive*, by M. C. Botte, G. Canévet, L. Demany, and C. Sorin (INSERM).
- 赤木正人, 上田和夫, 平原達也 訳 (1990). “音とは何か,” ATR テクニカルレポート, TR-A-0067 (翻訳担当箇所: 7, 8 章, pp. 77-103.) "Sound interpretation," in *Natural Computation*, edited by W. Richard (MIT Press, 1988).
- 大串健吾 監訳, 上田和夫, 岡田斉, 倉片憲治, 山田真司 訳 (1994). 聴覚心理学概論 (誠信書房, 東京) [翻訳担当箇所: 3 章 § 6-12, 5 章, 8 章, 用語集]. *An Introduction to the Psychology of Hearing*, by B. C. J. Moore (Academic Press, 1989), 3rd ed.
- 平原達也 監訳, 上田和夫, 赤木正人 訳 (1994). ナチュラル・コンピュテーション 2, [翻訳担当箇所: ”6 壊れたのかバウンドしたのか? 音で事象を判断することの心理物理学,” ”7 旋律の知覚”] (パーソナルメディア社, 東京). "Sound interpretation," in *Natural Computation*, edited by W. Richards (MIT Press, 1988).
- 上田和夫 訳 (1994). “2nd ICMPC (UCLA) 会議報告 (訳),” 日本音楽知覚認知研究会会報, 7 巻 1 号, 2-8. Justin London (1992). "The second international conference on music perception and cognition: University of California, Los Angeles, 22-26 February 1992," *The Journal of Musicology*, X, 537-545.
- 大串健吾 監訳, 上田和夫他訳 (1998). 音楽の認知心理学 (誠信書房, 東京) [翻訳担当箇所: 1 章, ”音楽における情動と意味,” 10 章, ”熟練した音楽家による時間比率の知覚、生成、模倣”]. Rita Aiello ed. (1994). *Musical Perceptions* (Oxford, New York).

その他

- 上田和夫 (1988). “VMS 版 SAS の使い方,” ATR テクニカルレポート, TR-A-0029.
- 上田和夫, 岡田斉, 東倉洋一 (1989). “単音節のダイコティック・リスニング: 耳

の優位性における個人差, 刺激の優位性,” ATR テクニカルレポート, TR-A-0044.

上田和夫 (1990). “UNIX 版 SAS の使い方,” ATR テクニカルレポート, TR-A-0081.

p