

学 位 論 文

言語音と非言語音の知覚的差異をもたらす内的過程

広島大学大学院
教育学研究科
教育人間科学専攻

高山智行

目 次

第 1 章 言語音と非言語音の知覚的差異をもたらす内的過程に 関する研究の現状と本研究の目的	
第 1 節 言語音知覚の基本的問題	
1. 言語音とは (1)
2. 言語音の特異性 (2)
第 2 節 言語音と非言語音の知覚的差異をもたらす内的過程 に関する研究の現状	
1. 言語音知覚様式と非言語音知覚様式 (22)
2. 言語音知覚の理論 (39)
第 3 節 本研究の目的 (42)
1. 正弦波アナログ音の知覚 (43)
2. 本研究の目的 (44)
第 2 章 言語音と非言語音の知覚的差異をもたらす内的過程に 関する実験的検討	
第 1 節 破裂音－母音アナログ音の識別学習（実験Ⅰ）	
1. 目的 (47)
2. 方法 (49)
3. 結果 (51)
4. 考察 (58)
第 2 節 母音アナログ音の識別学習（実験Ⅱ）	
1. 目的 (60)
2. 方法 (63)
3. 結果 (66)
4. 考察 (73)
第 3 節 正弦波アナログ音に対するプライミング効果（実験Ⅲ）	
1. 目的 (75)
2. 方法 (78)
3. 結果 (83)
4. 考察 (90)

第3章 総合的考察と将来の課題

第1節 本研究の成果と意義

- 1. 実験結果の概要 (93)
- 2. 正弦波アナログ音を用いた音韻研究における本研究の
位置づけと成果 (97)
- 3. 運動理論および直接知覚理論による解釈と問題点 (104)

第2節 将来の課題

- 1. 音韻記憶表象の内容と形式 (109)
- 2. 音韻記憶表象の活性化 (114)

第3節 まとめ

..... (117)

引用文献

..... (119)

第1章 言語音と非言語音の知覚的差異をもたらす内的過程に関する研究の現状と本研究の目的

第1節 言語音知覚の基本的問題

1. 言語音とは

音声は、原則として、人間の発声器官により発せられた音^{注1)}であり、人間のコミュニケーションにおいて重要な役割を担っている。古井(1994)によれば、音声にはさまざまな情報が含まれており、大きくは言語的情報(音韻性情報)と声質情報に分けることができる。言語的情報は、伝えたい意味内容に関連しており、音声が担う最も重要な情報であると言える。他方、声質情報には、誰が話しているかという個人性情報、話者の感情を表現する情緒性情報、声の良し悪しに関する情報などが含まれる。現実のコミュニケーションにおいては、これらの情報は相互に複雑に絡み合っただけでなく、またそれを担う音声信号内の音響変数も相互に独立というわけではない。

言語音とは音韻性(言語的)情報の側面に着目したときの音声の呼び方であり、これにも、言語学的には、発声の結果としての具体的な音響的現れに着目した音の単位である単音と、その言語体系の中で意味的区別に寄与する音の機能的単位である音素(あるいは狭義の音韻)とが区別される^{注2)}。個人の発声に関して言えば、音素は単音よりは抽象度の

注1) 2.1.2 項でも述べるように、人間の音声を機械的に模擬することも可能である。以後、記述上必要であれば、機械による音声を人工音声(あるいは合成音声)、人間の音声を自然音声として区別する。

注2) 言語音の表記には、音声学的表記と音韻論的表記があり、例えば、「アキ」という語は、前者では[]と単音記号を用いて[aki]と表記し、後者では / / と音素(音韻)記号を用いて /aki/ と表す

高い言語音の記述であると言うことができるが、同じ音素に分類される音でも注意深く聞けば違いに気づくこともある。すなわち、言語音の知覚には、単音レベルと音素（音韻）レベルの両方の情報が含まれている。津崎（1994）はこのような単音と音素の両面性をもった知覚内容に言及する概念として「（広義の）音韻」という語を用いている。

本研究においても、音韻性情報に関する知覚に言及する場合に音韻知覚という用語を用いる。また、音声・非音声に限らず、音韻性情報に関わる知覚とそれ以外の情報に関わる知覚の区別に言及する場合には、言語音知覚と非言語音知覚という用語を用いる。

2. 言語音の特異性

言語音は、先に述べたように、原則的には、人間が自らの発声器官を使って作り出す音である。それでは、言語音とそれ以外の音（すなわち非言語音）、例えば楽器の音や騒音など環境の中で発生するさまざまな音、とは、単にその発生源が異なるだけなのであろうか。言語音の特異性について、Studdert-Kennedy（1976）は、刺激としての言語音信号と、それによってもたらされる知覚内容との両方において、言語音と非言語音とは異なっている、と述べている。

2.1 刺激としての言語音信号の特異性

刺激としての言語音の特異性は、その生成と機能という2つの側面からとらえることができる。

2.1.1 音声学的・音韻論的特異性

まず、その生成面から見ると、言語音は人間の発声メカニズムが生成することのできる音の下位セットである。ある言語文化において、発声可能な音すべてが言語的情報の伝達に用いられているのではなく、そのある一部分が意味の区別に寄与する音として用いられる。意味の区別に寄与しない音声のうち、それに類似した音が意味的区別にかかわる音(すなわち、音韻)にあれば、たとえ物理的にも知覚的にも違いを指摘することができたとしても、その音韻に同化して知覚されがちである。

例えば、「パパ [p^hapa]」の最初の「パ [p^ha]」の音のように、語頭の無声破裂音は息の音をとめない、「パパ [p^hapa]」の2番目の「パ [pa]」の音のように、語中の無声破裂音は息の音が弱いか、息の音がない場合もある。しかし、日本語では両者を同じ /pa/ の音であるとして扱い、異なる音としては扱わない。このように、(同じ話者の発話において、)知覚的には(聞こえの上では)違いを指摘できるが、意味の区別に関しては等価なものとして扱う言語音同士を異音と呼ぶ。すなわち、異音は、同じ音素(音韻)の単音レベルでのさまざまな表れである。

あるいは、そのような物理的な違いすら気がつかれないこともある。英語の /l/ と /r/ の音は、図 1.1 に示すように音響的には明確に異なってい

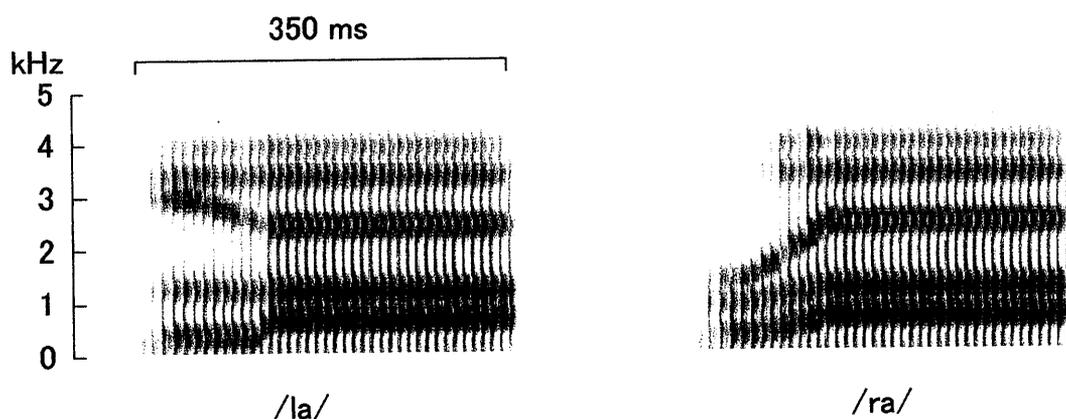


図 1.1 英語の /l/ と /r/ (合成音) のスペクトログラム

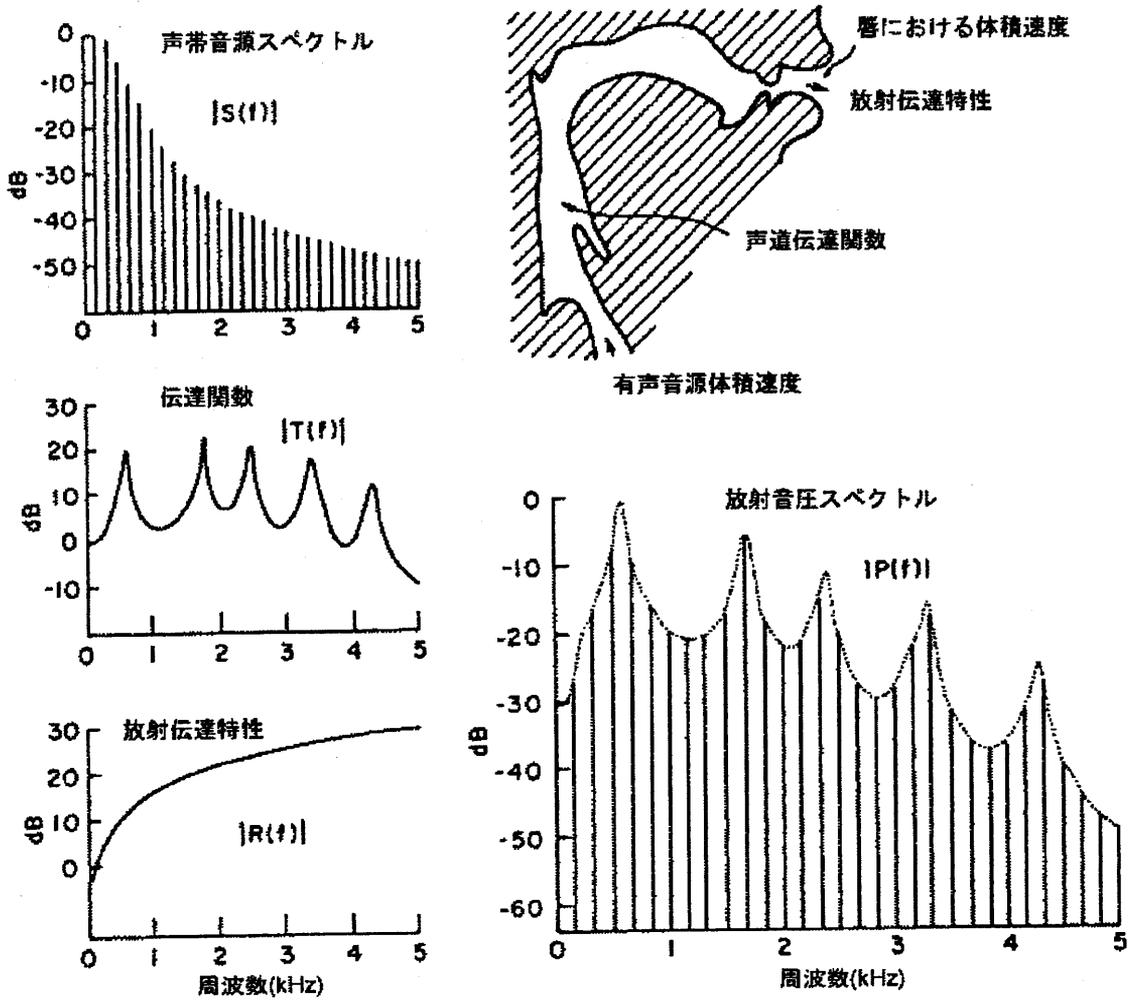
る（音響的特徴の記述については 2.4.1 項で論じる）が、日本語では意味の区別に寄与しないので、多くの聴者はそれらを区別できず（Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975）、ともに日本語のラ行の音として聞く（e.g., Best, 1995）。しかし、訓練することにより、英語圏の人々と同等に聞き分けたり、発音したりできるようになることも示されている（e.g., Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999）。

2.1.2 言語音のシミュレーション

刺激としての言語音の特異性をその機能面から見るならば、音響信号は音韻内容を伝達可能な言語音か、そうでない非言語音かのいずれしかなく、中間的な音は存在しない、ということが指摘される。

一方で、言語音は、独立した音源により励振されるフィルター群からの出力として近似可能である（Fant, 1960）。したがって、この原理をもとに、コンピュータ等で発声をシミュレートして、物理的には人間の言語音に無限に近似する音声信号を合成することも可能である。図 1.2 に言語音信号を模擬する場合の原理を示す。

他方、このような信号の知覚に関しては、物理的近似の程度に応じて、非言語音から言語音までの連続体を形成するわけではなく、言語音として知覚されるか知覚されないかのいずれかでしかない。House, Stevens, Sandel, & Arnold (1962) は、音響構成の上で明らかな非言語音から言語音までの音響的連続体を構成するいくつかの複合音群を合成し、それらの複合音群それぞれに対する識別学習に要する時間を比較した。その結果、彼らは、識別学習時間は言語音で最も短かったが、近似の程度に応じて徐々に時間が減少するのではなく、聴者が信号を言語音としてとらえるか否かにより学習時間が規定されることを見いだした。Stevens & House



$$P(f) = S(f) \times T(f) \times R(f)$$

図 1.2 言語音信号生成の原理 (Pisoni, 1976 より一部改変)

(1972) は、「...音が次第次第に音声に近くなっていくような音響的連続体を想像することは可能であるが、音の知覚に関する限り、そのような連続体は存在しない——知覚は二分法的なものである。音は言語的なものと知覚されるか、非言語的なものと知覚されるかのどちらかしかない」(p. 11)と述べている。ある音響信号が言語音として知覚されるやいなや、その信号は音韻内容を伝達する。

2.2 知覚内容の特異性

言語音の特異性の第2の重要な点は、その知覚内容が言語音自体の音韻カテゴリーの名前を表すということである。言語音を表記するために用いられるアルファベットや仮名文字においては、知覚された形とそれに与えられている音との関係は任意である。しかし、言語音の音韻名は言語音信号の知覚内容に一致し、両者は「任意と言うよりも、必然的な関係を担っている」(Studdert-Kennedy, 1976, p. 244)。

本節の冒頭で述べたように、言語学における言語音についての記述には、発声の具体的な現れである単音と、当該言語体系の中で意味的区別に寄与する音素(狭義の音韻)という2つのレベルがあり、この両者がともに音韻性情報に関する知覚に関わっている(津崎, 1994)。すなわち、知覚される音韻内容は、言語音信号の音響的現れを反映する知覚的カテゴリーであるとともに、異音変動やその他の音響的変動が捨象された抽象的カテゴリーであるという二面性をもつ。音韻性情報の取得という機能面からは、後者が優先される。

2.3 言語音の発声と聴取の過程

言語音の特異性は、その発声と聴取の過程に由来するものと考えられる。古井(1994)は、言語音の発声過程を図1.3のように3つの段階で記述している。すなわち、それらの段階は、「(1) 相手に伝えたい内容を表現する言葉を選択して正しく並べる言語的段階、(2) 脳から発声器官へ運動指令を発し、それに従う筋肉によって発声器官が活動する生理学的段階、(3) 音波の発生と伝搬が行われる物理学的段階」(古井, 1994, p. 1118)であり、「言語学的段階と生理学的段階の一部は離散的な性質を有するが、生理学的段階のほとんどと物理学的段階は連続的な

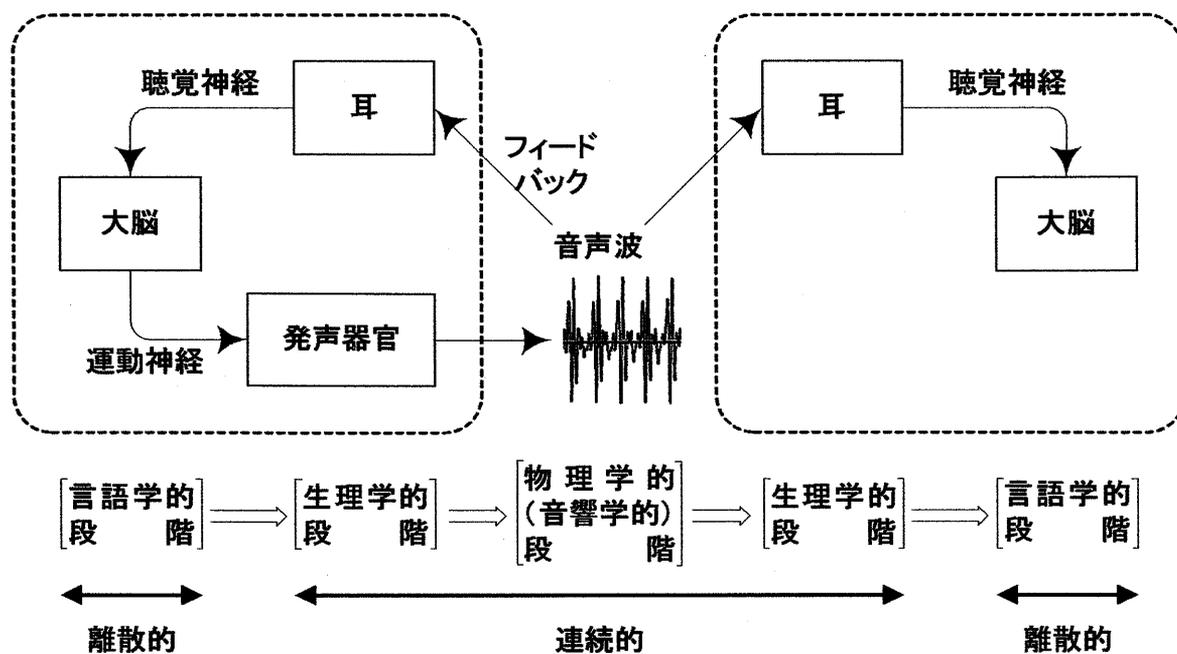


図 1.3 言語音声の発声と聞き取りのプロセス (古井, 1994 より一部改変)

性質を有する」(古井, 1994, p.1118), としている。また, 話者は相手が聴取することを前提として発声するのであるが, それと同時に自らの声を聴取しながら発声を調整しており, 発声と聴取は極めて密接に関係している。このことは, 自らの発声を 4 分の 1 秒程度遅らせてフィードバックされる状態では, 発話が大きく乱される聴覚遅延フィードバック (delayed auditory feedback, DAF と略す, e.g., Lee, 1950; 江副, 2000; 府川, 1980; Fairbanks, 1955) で確認することができる^{注3)}。

言語学的段階での離散的な表象に対して, 生理学的段階では, 声帯の準周期的な振動により生じる声帯音源, 舌や口唇による声道の狭めによって生じる摩擦音源, あるいは声道の閉鎖とその急速な解放による破裂

注3) ただし, 河原(2003)によると, DAF 効果の生起にはフィードバックされる信号に話者音声の音響的パラメータが保持される必要はないこと, そして DAF による妨害効果は音声に特化した現象でもないことが指摘されており, 従来の聴覚フィードバックによる単純な制御モデルは見直されている。

音源に対して、舌、口唇、顎などを動かして声道の形を変形することによって、音韻性情報を付加し音声信号として実現する。これを調音といい、母音、半母音、摩擦音、破裂音、破擦音、鼻音という種類がある。さらに、音節や単語、文の調音に関しては、「発声器官がきわめて複雑な動きをし、隣接する音の影響を受けて、一つひとつの音を発声するときとは異なった、平滑化された動きを示す」（古井，1994，p. 1120）ことになる。これを調音結合と言い、結果的に物理学的段階では、音声信号は一つひとつの音韻の順序づけられた系列としてではなく、それら音韻間の境界が不明確な連続的信号として実現されることになる。したがって、言語音の知覚過程は、物理学的段階での連続的な性質をもった信号に対して、最終的に、離散的な性質の言語内容（音韻内容）を回復する過程と言える（図 1.3）。

2.4 言語音知覚過程の問題

日本語を聞き話すわれわれの日常的な経験は、言語音は離散的な音韻や音節の連続であるかのように聞こえる、というものである。しかし、先述したように、発声時の調音動作が連続しているので、言語音信号自体も連続的である。ここに、言語音知覚における分節化・離散化の問題がある。

加えて、発話の仕方（例えば、囁き声、アクセントやイントネーション）や発話者自体の性別や方言の違い、発話者の内的状態（例えば、興奮や沈静）、あるいは発話が行われた場面（例えば、まわりの騒音）などによりもたらされる声の高さや大きさ、発話速度、音質など、物理的音響的には言語音信号は大きく変動しているはずである。しかし、われわれは、そのような多様な音響的現れの中から、話者の意図したメッセ

ージの音韻内容をほぼ恒常的に知覚することができる。同じ言語環境にいる限り、「あ」という言語音は、誰がどのように発音しようと、話者が「あ」を意図して発音していれば、少なくとも良好な聴取環境のもとでは、聴者には「あ」と知覚されるのである。これは、言語音知覚における正規化（標準化）の問題である。

音韻知覚の主要な問題は、これら分節化・離散化の問題と正規化の問題に要約することができるであろう。これらの問題をめぐってまずとられたアプローチは、いくつかの音響信号に対して同じ音韻が知覚される時、それらの音響信号には何らかの共通する音響的特性（あるいは音響手がかり）が見いだされるはずだ、として音響変動に左右されない不変的な手がかりを見つけ出そうとすることであった。

2.4.1 音韻知覚の音響的手がかり

音声は、さまざまな構成要素の周波数と強さが時間軸上で変化する音響パターンであり、そのような変化を同時に示す分析ツールがスペクトログラムである。この表示は、一定周波数帯域に含まれるエネルギー量が時間の関数としてプロットされる。時間は横軸に、周波数は縦軸にとられ、強さはプロットの濃淡で表現される。図 1.4 に成人男性話者による「青いズボンを買った」という発話のスペクトログラムを示す。この図の上側は音声波形、中段は広帯域スペクトログラム、下側は狭帯域スペクトログラムである。分析にあたって、スペクトル分布の時間的変動を検討するためには広帯域スペクトログラムが、周波数スペクトルの構成を調べるためには狭帯域スペクトログラムが用いられる。

広帯域スペクトログラムにおいて見られる顕著な特徴は、ほぼ一定周期で繰り返す縦縞と、濃淡で表されるエネルギー分布が全周波数帯域に

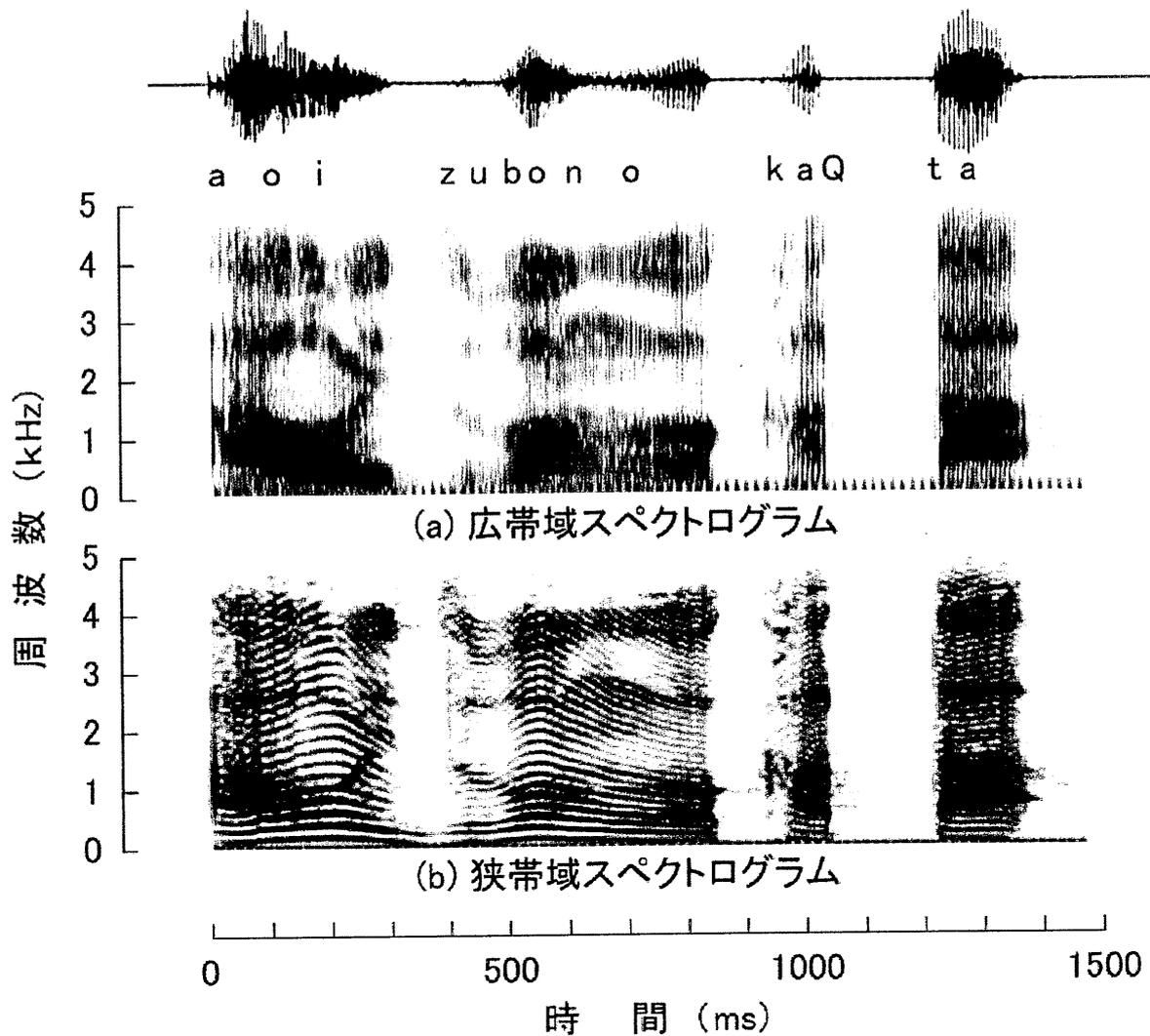


図1.4 男性話者による「青いズボンを買った」のスペクトログラム
 ローマ字表記は、それぞれの音声セグメント（単音あるいは音素）を表してお
 り、それらが開始されるおおよその時点に配置してある。「Q」は促音（「っ」）
 を表している。

わたって一様ではないという点である。縦縞は、声帯の振動（すなわち発話のピッチの逆数）を表している。また、周波数-時間軸に対する不均質なエネルギー分布の中での濃い帯は、フォルマントと呼ばれ、通常、母音の種類を区別する音響的手がかりとなる。例えば、最初の「青い」の部分では、4つのフォルマントが認められ、母音の変化に応じて下側の3つのフォルマントの周波数領域が特徴的に変化している。フォルマ

ントは下側から第 1 フォルマント (F1), 第 2 フォルマント (F2), 第 3 フォルマント (F3), 第 4 フォルマント (F4), 等々と名づけられており, 母音の識別には F1, F2 が重要であることが知られている。ここでは, 「アオイ/aoi/」と母音が変化するにつれて, F1 が次第に下降しているのに対して, F2 が「オ/o/」の部分で下降し, 「イ/i/」で上昇している。F3 も特徴的な動きを示し, 「アオ/a/」ではほぼ一定で, 「イ/i/」で下降している。ここで重要なことは, 先にも述べたように, 3 つの母音を連続して発音するために生じた調音器官の連続した動き (すなわち, 調音結合) の結果, 母音の音響的手がかりも連続的に変化している, ということである。

また, 同じ母音ではほぼ同じ周波数領域にフォルマントがあることが見て取れる。例えば, 「青い」の「ア/a/」と, 「買った」の「カ/ka/」と「タ/ta/」の母音部, あるいは「青い」の「オ/o/」と「ズボン」の「ボ/bo/」と「ヲ/o/」の母音部に関して, 時間的長さは異なるが, ほぼ同じ周波数領域にフォルマントが認められる。

子音に関して見ると, 「ズボン」の「ズ/zu/」は比較的薄い帯で示されており, 「買った」の「カ/ka/」の開始時にも比較的帯域の広い薄い帯が見られる。また, 「ズボン」の「ボ/bo/」ではあまりはっきりしないが, 「買った」の「カ/ka/」と「タ/ta/」では, 「ア/a/」のフォルマントに到達する直前に短時間で急速に変化するフォルマント部分がある。

狭帯域スペクトログラムは周波数解析が主要な目的である。周波数方向に見られる横縞は, 周期的成分である基本周波数成分とその高調波成分を表している。それが揺らいでいるのは, 発話の中で声の高さが変わっていることを示す。また, 「ズボン」の「ズ/zu/」ではフォルマント構造が見られるものの, 高い周波数領域では非周期的成分が混じっている

ようであり、「買った」の「カ/ka/」の開始部は広帯域の非周期的成分からなっている様子が見て取れる。

多くの言語音信号に対して、このようなスペクトログラム分析を実施することで音韻知覚に関連する可能性のある音響特性が選り出され、さらにそれらの音響特性を用いて合成された言語音に関する聴取実験を行うことによって、特定の音韻に共通する音響特性のリストが作られた。そこには、特定の音韻が不変的な音響特性を持って物理的に実現されているのであれば、その逆を辿って音響特性から一意的に音韻内容を特定することが可能であろう、との期待があったと考えられる。

2.4.2 不変性の欠如と符号化

しかし、そのような研究から見いだされたことは、例えば後続母音が異なる場合のように異なる文脈の中で、一方では、同じ音響手がかりから異なる知覚内容が生じ、他方では、異なる音響手がかりから同じ知覚内容が生じる、という事実であった（このような事実は、不変性の欠如と呼ばれている、Lieberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967）。前者の例として、Lieberman, Delattre, & Cooper (1952) は、さまざまな中心周波数の狭帯域バーストノイズ（破裂音ノイズ）をさまざまな英語母音の直前に提示し、その聞こえを調べたところ、同じ破裂音ノイズ、例えば 1400 Hz の中心周波数をもつノイズが /i/ の前に置かれると、/pi/ という音節の知覚が生じ、/a/ の前に置かれると /ka/ が、そして /u/ の前では /pu/ が知覚される、ということを見いだした（図 1.5）。すなわち、同じ破裂音ノイズが異なる母音文脈の前に置かれるとき、異なる破裂音の知覚の手がかりとなる。また、図 1.4 の例では、「買った」の「タ/ta/」の部分だけを切り出して単独で聞くと、「ダ/da/」に聞こえる。これは、

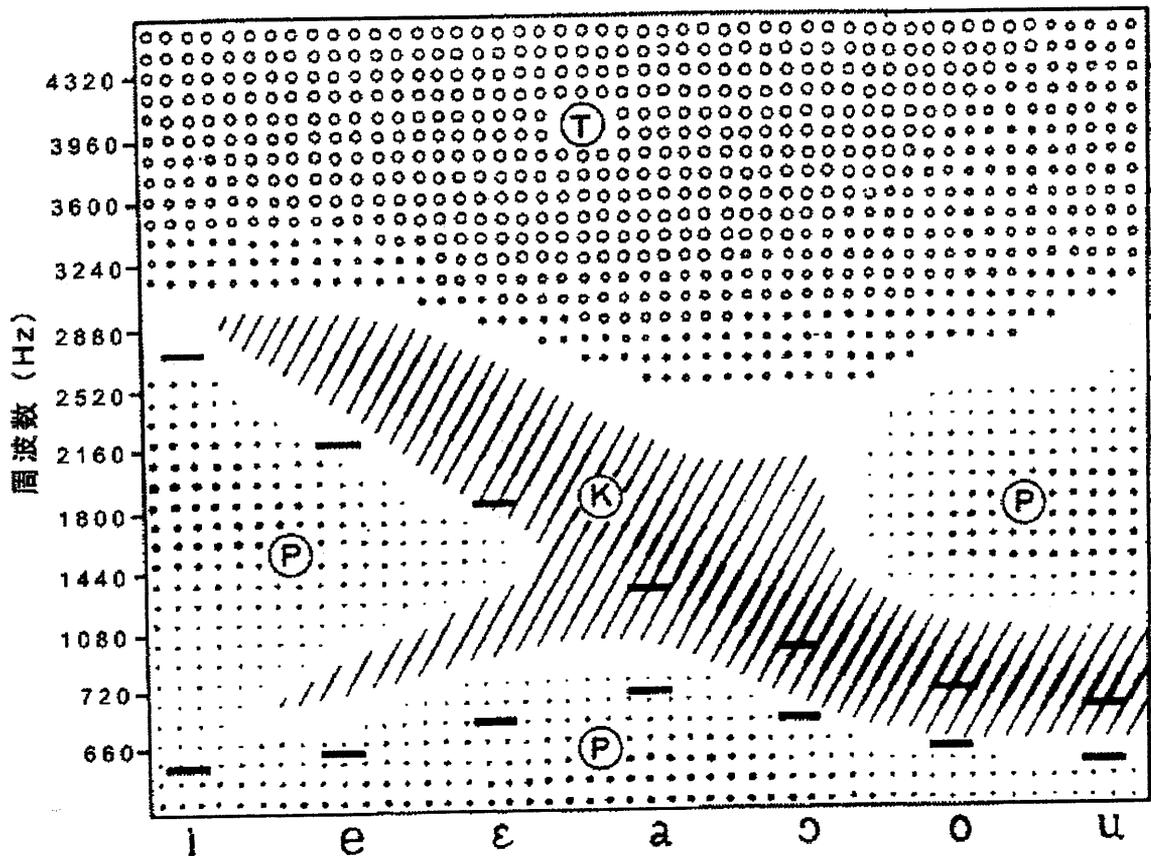


図1.5 破裂音ノイズの中心周波数と母音フォルマント周波数に応じた /p/, /t/, /k/ 反応の分布 (Liberman et al., 1952 より一部改変)

同じ音響パターンが、語頭と語中という異なる文脈では異なって知覚されるという例である。

また、後者の例として、Liberman, Delattre, Cooper, & Gerstman (1954) は、子音の知覚におけるフォルマント遷移の効果を調べ、破裂子音 /d/ が知覚されるためには、母音 /i/ の前では F2 は上昇し、/u/ の前では下降することが必要であることを示した (図 1.6)。すなわち、同じ音韻 /d/ の知覚が生じるために、異なる母音文脈では異なる音響手がかりが用いられていた。

これらの結果を受け、Liberman et al. (1967) は、破裂音のように音響的手がかりが前後の音韻環境により大きく変化する、すなわち結果とし

て得られる音声信号の音響パターンが大きく再構成されるような音韻を符号化 (encoded) された音韻と呼び、母音のようにあまり再構成されない音韻を符号化されていない (unencoded) 音韻

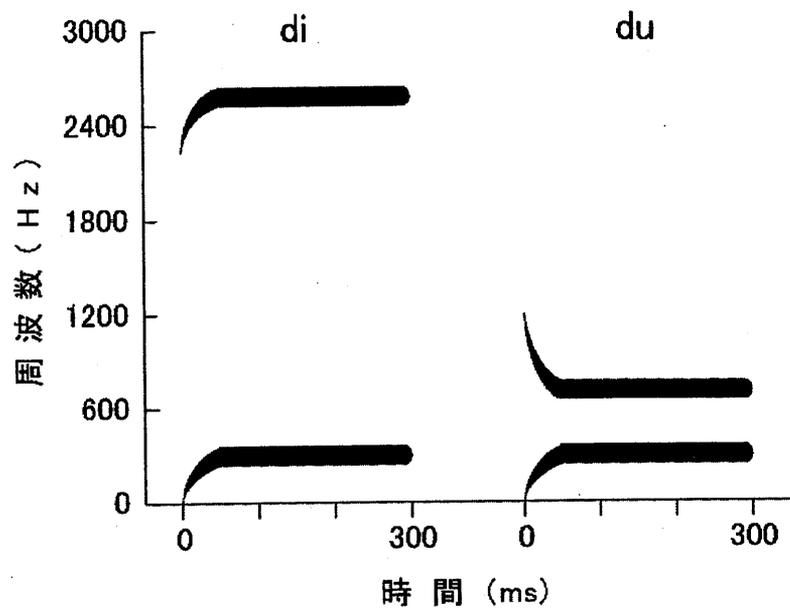


図1.6 母音 /i/ と/u/ の前における /d/ の手がかり
(Liberman et al., 1967 より一部改変)

と呼んだ。そして、このように、言語音信号の音響的模式と音韻との関係は複雑であるため、Liberman et al. (1967) は、そのような言語音の符号解読に特殊化したメカニズムが存在するに違いない、と主張した。

2.4.3 不変性への新たなアプローチ

音韻知覚と音響信号との対応関係を見いだそうとした以前の試みでは、スペクトログラムに見られる顕著な特徴が不変的の手がかりの候補として取り上げられていた。そこには、「(スペクトログラムにおいて) 視覚的に目立つ物理的特徴は、重要な知覚的特徴に対応している、という暗黙的な仮定」(Jusczyk, 1986, p.27-30) があったものと考えられる。先に述べたように、Liberman et al. (1967) は、そのような試みが成功しなかったため、音韻知覚における不変性の基盤を生物学的に特殊化した仮説的メカニズムに求めた。

他方、「人間の聴覚システムにより遂行される分析はスペクトログラ

ムで見いだされるものとは異なる出力を生じるかもしれない」(Jusczyk, 1986, p.27-30) ので、言語音に対する人間の聴覚神経系の応答を考慮して、言語音信号を適切に分析するならば、不変的特性が見いだされるはずである、との主張もある。以下に述べる 3 つのモデルでは、いずれもこのような聴覚系で得られると想定される分析結果を利用している。

① Stevens & Blumstein (1978) のモデル

Stevens & Blumstein (1978; Blumstein & Stevens, 1979, 1980) は、Stevens (1975) や Fant (1960) の研究を受けて、示差的な音声的特徴^{注4)}に直接対応する不変的な統合的音響特性が言語音信号内にあると仮定し、破裂子音に関する不変的特性として、破裂音開放時のスペクトル形状に基づくスペクトルの鋳型 (template) を提案している。すなわち、語頭の破裂音の手がかりであるフォルマント遷移や開放バーストノイズなどの要素は、それぞれ単独では不変的な手がかりとは言えなくても、全体として統合されたパターンとしては不変性を示すであろう、と考えられた。Stevens & Blumstein (1978) によれば、両唇破裂音 (/b/, /p/) は拡散的で下降あるいは平坦なスペクトルによって、歯茎破裂音 (/d/, /t/) は拡散的で上昇するスペクトルによって、そして軟口蓋破裂音 (/g/, /k/) は集約的なスペクトルによって特徴づけられる (図 1.7)。

注4) 示差的特徴とは、通常、調音動作に基づいてある音声分節 (phonetic segment, 単音) あるいは音素を他の音声分節あるいは音素と区別する二項的特徴あるいはその特徴次元のことをさす。たとえば、破裂子音 [b] と [p] は、調音の仕方 (閉鎖破裂) やその調音位置 (両唇) は共通するが、発音の際に声帯振動を伴うか否か、すなわち有声か無声かで異なっている。したがって、両者は有声性という示差的特徴で区別され、[b] は [+有声性]、[p] は [-有声性] とされる (ここで、+はその弁別素性のあることを、-はないことを示す)。このような弁別素性としては 12 ないし 13 のものが用いられており、個々の音声分節あるいは音素は、これら示差的特徴の値により一義的に決められる。

このモデルに関して多くのテストが行われているが、例えば、Blumstein & Stevens (1979) は、6名の話者による5つの母音文脈での語頭と語尾の破裂音の発話について、それらのスペクトル分析の結果がスペクトルの鑄型にどの程度うまく適合するかを検討した。その結果、語頭の破裂音については83% とかなり良い成績が得られることを示した。

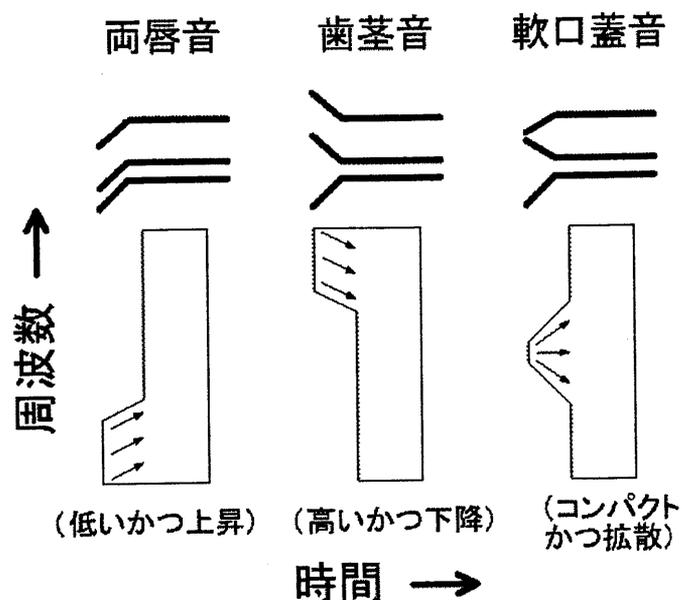


図1.7 語頭破裂子音の調音場所による仮説的な不変的スペクトルパターン

(Stevens & Blumstein, 1978 のモデルの前身である Stevens, 1975 のモデルについて引用した Pisoni, 1976 の図を一部改変した。)

しかし、Blumstein, Isaacs, & Mertus (1982) と、Walley & Carrell (1983) は、合成言語音に関する人間の聴者の識別と、スペクトルの鑄型モデルからの予測とを比較し、人間の聴者による識別は一貫して、破裂音開放時の不変的なスペクトル形状ではなく、文脈的に変動するフォルマント遷移に基づく、ということを示している。

② Searle, Jacobson, & Rayment (1979) のモデル

Stevens & Blumstein のモデルは静的なスペクトル形状に不変性を求めていたが、Searle, Jacobson, & Rayment (1979; Searle, Jacobson, & Kimberley, 1980) は、動的なスペクトル変化に着目した。彼らは、人間の末梢聴覚システムに関する心理物理学的データと生理学的データを整理して、ほぼ 1/3

オクターブの周波数解析力と、1.6 ms の時間解析力で言語音信号を分析する方法を選択した。1/3 オクターブ帯域幅を選択することで、低い周波数でのスペクトル解析力により、声開始時間（voice onset time, VOT と略す）のような比較的ゆっくりした事象を追いながら、第 1, 第 2 フォルマントを分離して追跡することが可能となり、他方、高い周波数での時間解析力により、時間的に短い破裂音バーストについての正確なタイミング情報を提供できる。1.6 ms 毎の動的スペクトルを用いて得られた特徴をもとにして、声の有無についてはほぼ完全に予測でき、調音場所に関する予測についても比較的精度が良いという結果が得られている。

③ Kewley-Port (1983) のモデル

Kewley-Port (1983; Kewley-Port & Luce, 1984; Kewly-Port, Pisoni, & Studdert-Kennedy, 1983) も、語頭の破裂音の調音場所と相関のある不変的な音響事象を特定するために、Searle et al. (1979, 1980) と同様に、動的なスペクトルに着目している。しかし、彼女のシステムでは、人間の周波数解析力を表す臨界帯域幅をモデル化する多くの研究で約 1/4 オクターブ帯域幅のフィルターが用いられていることを参考にして、結果的に 1/6 オクターブ帯域幅が用いられた。また、分析フレームも、Searle et al. (1979, 1980) が用いた 1.6 ms では詳細すぎるので、5 ms が用いられている。

モデルの予測性を検討するために、例えば、Kewly-Port & Luce (1984) は、6つの破裂音 /b, d, g, p, t, k/ と 5つの英語母音 [i, æ, a, ɔ, u] の組み合わせそれぞれについて 5回繰り返して発せられた 4名の話者（うち 2名については、速く話す場合とゆっくり話す場合も含められた）の発話を 5ms 毎に分析した結果を、(1) 高周波数域におけるエネルギーの急激な増加（急激な子音の開放を表す）、(2) 低周波数域での狭いエネルギーピ

ークの開始（声の開始を表す），（3）開始時のバーストあるいは無声スペクトル周波数の相対的傾き（調音場所を表す），（4）中位の周波数域での時間とともに広がるスペクトルピーク（軟口蓋音を表す），という4つの特徴を用いて分類した。分類の結果は，調音場所に関して86%の正確さであり，（異なる速さで発せられた）異なる話者の有声と無声の破裂音の識別に対して，それらの動的特徴が比較的有効であることが示された。

これらのアプローチは，言語音信号の文脈変動に対して，スペクトログラムに現れる音響特性に見いだせなかった不変性を，周波数×時間的広がりの中で音響特性を統合することによって見いだそうとしている。しかし，語中や語尾の音韻について，あるいは破裂音以外の音韻，例えば，鼻音，について，十分正確に特定できないなど，問題点もあった。また，Stevens & Blumstein (1978) のモデルについては，Searle et al. (1979, 1980) や Kewley-Port (1983; Kewley-Port & Luce, 1984; Kewly-Port et al., 1983) が指摘する動的特性を取り入れながら，その適用範囲が徐々に拡大されているが，その他のモデルについては，その後あまり進展していないようである。

2.4.4 分節化

言語音知覚において解決すべきもう一つの問題は，分節化の問題である。文字言語においては，音節，単語，文を一つひとつの表音文字の連なりとして表現できるが，音声言語においては，調音結合のために，一つひとつの音韻の発音の連なりとして表現することはできない。例えば，図1.4のスペクトログラムにおいて，「青い」という3つの母音間の境界，あるいは「ズボン」における「ボ」と「ン」と「ヲ」の境界を，

明確に位置づけることはできない。図 1.8 は、各音韻が、調音結合によって、言語音信号の中でどのように実現されているかを模式的に表したものである。この図に示されているように、ある時点での言語音信号はその時点の音韻と前後の音韻の情報を同時並行的に伝達している (Lieberman et al., 1967)。

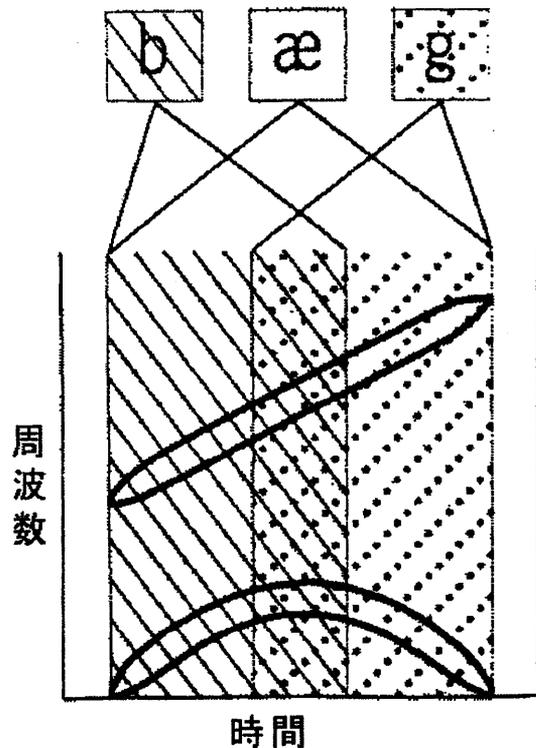


図1.8 言語音フォルマントパターンと音韻情報の配置

(Lieberman, 1970 より一部改変)

また、言語音信号の言語学的分節は、物理的分節とは必ずしも一致しないことも明らかである。図 1.4 において、「青い」と

「ズボン」との間、「ズボンを」と「買った」の間に見られるほとんどエネルギーがない部分は、文法上の句の区切りを表している。しかし、同じようにほとんどエネルギーのない部分が、「買った」の促音部にも現れている。すなわち、スペクトログラムを見る限り、あるいは他の音響分析の結果では、発話における文法上の区切りと促音部との区別はできないと考えられる。ムーア (1994) は、この問題の難しさを指摘し、「例えば、通常の会話のような調子で話された『recognize speech (音声を認識する)』という発話と『wreck a nice beach (素敵な浜を台なしにする)』という発話を区別するのは、たいていの音声認識機械にとって非常に困難であろう。」(Pp. 285) と述べている。

他方、調音結合による音韻性情報の連続性は、音韻性情報の急速な伝

達にとって重要な役割を演じている。Liberman et al. (1967) によると、通常の会話で毎秒 8 ~ 10 音韻、速い発話では 30 音韻が伝達されるが、後者では非言語音で測定された聴覚系の時間分解能を超えている (Miller & Taylor, 1948)。Warren, Obusek, Farmer, & Warren (1969) は、4 種の音が次々に休みなく交替する系列において、正しく順序を報告するためには各音の持続時間が 200 ~ 700 ms 必要であることを示している。すなわち、言語音が一つひとつの独立した単音の連なりであるとする、せいぜい毎秒 5 単音、音節に換算すると 2 ~ 3 音節程度でしか正しい順序で識別できないことになる。われわれがそれ以上の速さの発話を聞き取ることができるのは、音韻性情報が言語音信号の中にある程度同時並行的に伝達されているからである。

2.4.5 正規化 (標準化)

調音結合による音響-音韻不変性の欠如の問題に加えて、言語音信号の正規化 (標準化) に関わる 2 つの問題、すなわち話者の標準化の問題と時間と速さの標準化の問題、がある。

話者の標準化問題は、声道の形状や長さ、性別、年齢、方言などの個人間の差や、強勢やイントネーションなどの個人内での発話差により多様に変動する言語音信号から、どのようにして恒常的な音韻を知覚するのか、ということに関わっている。例えば、母音のフォルマントの周波数が上方へ (ほぼ比例的に) 変移すると、より小さい声道で (例えば、女性や子供) その母音が発せられたことを意味する (Repp, 1982)。Ladefoged & Broadbent (1957) は、先行する「Please say what this word is . . . ?」という文中の母音の第 1 および第 2 フォルマントの周波数を操作することによって、後続の単語 /bVt/ (V は母音を表す、例えば bat, bit など)

に組み込まれたあいまいな母音の知覚を調べた結果、先行文のフォルマント周波数の高さに応じて、物理的には同一の母音異なる母音として知覚されることを見いだした。すなわち、先行文で形成された同一話者の発話に対する判断基準が、後続の母音の知覚を左右したと言える。

また、時間と速さの標準化問題は、発話速度により持続時間が変化する言語音セグメントに対して、どのようにして安定した音韻を知覚するのか、ということに関わっている。例えば、物理的には全く同一の言語音信号が、ゆっくり話す文脈に置かれたときと急いで話す文脈に置かれたときとは異なる音韻として知覚される。例えば、一般に、発話速度が速くなれば、個々の音声セグメントの長さも短くなる。Miller & Liberman (1979) は、語頭のフォルマント遷移の持続時間が手がかりとなる破裂音-半母音 ([ba]-[wa]) 刺激連続体において、刺激全体の持続時間が増加するほど、カテゴリー境界の位置が、発話速度を考慮しているかのように、長い遷移持続時間の方向に移動することを示した。

いずれにしても、これらの正規化は、音韻知覚が、言語音信号の絶対的特性を手がかりにしているのではなく、何らかの関係的特性に基づいている、ということを示している。次節の文脈効果の項でも述べるように、このような現象が可能となる知覚メカニズムについて、言語音生成メカニズムと共通性を前提とした説明もなされているが (Repp, 1982)、非言語音刺激を用いた研究においても類似の現象が示されており、正規化がどのようなメカニズムによって実現されているのか、という問題について、十分な説明はなされていないように思える。

第2節 言語音と非言語音の知覚的差異をもたらす内的過程に関する 研究の現状

言語音は、物理的にも、知覚的にも非言語音とは異なる性質を持っていることが多くの研究で示されている。Liberman & Mattingly (1985; Liberman, 1982; Liberman et al., 1967) のように、この言語音の特異性に着目し、言語音の知覚に関して生物学的に特殊化したメカニズムが存在し、人が言語音を聞くときには知覚の言語音様式 (speech mode of perception) という特殊な知覚様式が自動的に働く、と主張する研究者もある。この節では、言語音知覚と非言語音知覚のメカニズムの違いを示すと考えられる現象とその問題点を紹介し、それらの現象を説明する理論についても概説する。

1. 言語音知覚様式と非言語音知覚様式

言語音と非言語音との知覚的差異を示すと考えられる現象が、いくつかの研究により列挙されている (Cleary & Pisoni, 2001; Liberman et al., 1967; Repp, 1982; ムーア, 1994; 津崎, 1994)。そのような現象には、次のようなものがある。

1.1 カテゴリー的知覚

非言語音信号に関して、Weber-Fechner の法則や Stevens の法則として知られているように、ある物理的次元を連続的に変化させると、それに対応する感覚次元は物理量の対数あるいはべき乗に比例して変化する。そ

して、これらの法則からは、一定間隔で段階的に物理量が増加するある範囲内の刺激連続体に対して、対応する感覚次元での2値的なカテゴリー判断を求めると、通常、特定の判断の比率は物理量の増加（あるいは減少）とともに徐々に増加（あるいは減少）すること、また、連続体上で隣り合った刺激間の弁別の精度はその物理量に反比例し、物理量の増加にともない次第に減少することが予想される。図1.9の右側はこれらの様子を模式的に表したものである。

他方、言語音に関して、その音響的手がかりを同様に一次的に変化

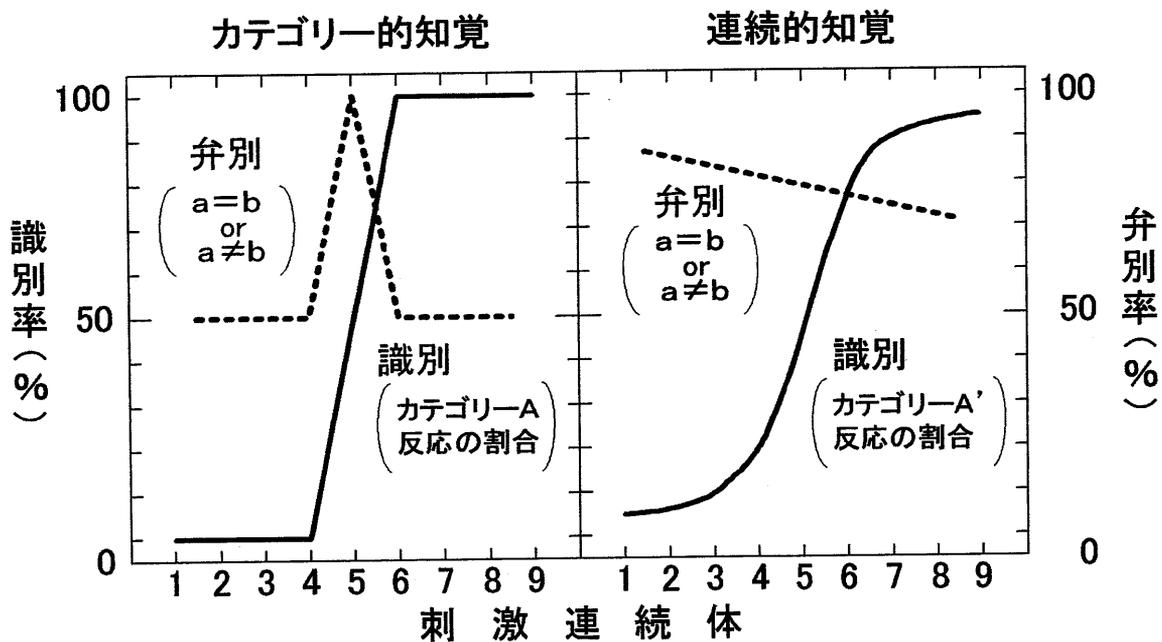


図1.9 カテゴリー的知覚と連続的知覚

この図において、刺激連続体は物理的に等間隔に構成されているものとする。識別率とは、連続体上の各刺激が特定のカテゴリーの属すると判断される割合であり、弁別率とは、連続体上の隣り合った（あるいは特定間隔の）刺激の同異が正しく判断される割合のことである。理想的なカテゴリー的知覚の特徴は、①急峻なカテゴリー境界、②低いカテゴリー内弁別率と高いカテゴリー間弁別率（この図においては、50%の弁別率は偶然レベルである）、③識別率から予測された弁別率と実測された弁別率の一致（必要条件ではないが）の3点である。他方、連続的知覚においては、カテゴリー反応の移行は緩やかで、弁別率も刺激の物理量に規定されて、Weberの法則に従うとすると、物理量が増加するにつれて、隣り合う刺激間の弁別率は低下する。

させても、ある音韻から別の音韻へ徐々に変化していくという知覚的印象は生じない。例えば、第2フォルマントの起点周波数を段階的に変化させることで /ba/ から /da/ までの音響的連続体を構成するとき、その連続体上の刺激に対する識別判断は、非言語音の連続体のようであるカテゴリーから他のカテゴリーへ徐々に変化するのではなく、連続体上の狭い範囲で急速に一方のカテゴリーから他方のカテゴリーへ変化する。しかも、連続体上の隣り合った刺激の弁別精度は、非言語音では心理物理学的限界により影響されるが、言語音連続体では、それぞれの刺激の音韻的同一性に影響され、同じカテゴリー内の刺激に対しては弁別精度は低下し、カテゴリー境界で弁別精度が上昇する。理想的には、刺激間の弁別がそれぞれの刺激に対する識別確率から予測されることになる。すなわち、言語音の知覚は、その弁別が識別内容に支配されるという意味でカテゴリー的である。図 1.9 の左側はこの様子を表したものである。このような特徴を持つ知覚は、非言語音のように弁別が刺激の物理的性質と感覚システムの閾特性により支配される連続的知覚と区別し、カテゴリー的知覚と呼ばれている (Liberman et al., 1967; 高山, 1986)。

カテゴリー的知覚は、破裂子音や鼻音のような音響的に過渡的特性を持つ言語音に対して現れるが、母音のような比較的定常的な言語音に対しては現れない、あるいは現れにくいということが知られている。しかし、課題要請や信号の減衰によって聴覚的な記憶への負荷を操作することで、このような音響信号に対してもカテゴリー的知覚が生じることも見いだされ (Fujisaki & Kawashima, 1969, 1970; Macmillan, Goldberg, & Braid, 1988; Pisoni, 1971, 1973), 課題における記憶要因の重要性が指摘されている (Schouten, 1987)。これはさらに、音韻知覚において、内的標準あるいはプロトタイプへの参照が生じるという主張に発展している (Kuhl, 1992;

Massaro, 1987; Miller, 1994)。

また、カテゴリー的知覚が、言語音知覚と非言語音知覚の違いを示す現象とは言えないという指摘もある。Pastore, Li, & Layer (1990) は、破裂子音の調音場所の手がかりとなる第2フォルマント遷移部(チャープ音)並びにそれに母音の第2フォルマント部を続けたブリート音の連続体に関して、聴者を適切に訓練することで、言語音の場合と類似したカテゴリー的知覚を生じさせることができる、ということを示した。同種の刺激は、Mattingly, Liberman, Syrdal, & Halwes (1971) によって、言語音刺激と非言語音刺激の知覚が異なる知覚様式あるいはモジュールによってもたらされることを支持する証拠とされていたものである。また、Sawusch & Gagnon (1995) も、後に述べる正弦波アナログ音を用いて、以前の研究 (Best, Studdert-Kennedy, Manuel, & Rubin-Spitz, 1989) において非言語音様式では連続的に知覚されることが示されていた同じ刺激に関して、非言語音としての識別訓練を繰り返し、カテゴリー的知覚が生じることを示している。彼らは、言語音、非言語音に共通して、信号のカテゴリー化を支える抽象的な聴覚表象の存在を主張している。

1.2 音声トレーニング

カテゴリー的知覚実験においては、典型的には、言語音信号内の単一の関連音響特性を一次元的に変化させることにより刺激連続体を構成する。しかし、自然音声の音韻知覚に関する研究においては、通常、信号内の複数の音響特性が特定の音韻の手がかりとして機能している、ということを示している (Bailey & Summerfield, 1980)。したがって、複数の音響的手がかり間で、その物理値のある範囲内において、ある手がかりに対して通常は音韻知覚の変化に導くような変化を与えても、それらを打

ち消すように別の手がかりを変化させることによって、もとの音韻知覚を保つことができる (Best, Morrongiello, & Robson, 1981; Fitch, Halwes, Erikson, & Liberman, 1980; 高山, 1989b, 1990)。このような効果を、手がかりトレーディング (cue trading) あるいは音声トレーディング (phonetic trading) と呼ぶ。図 1.10 に、摩擦音-破裂音-母音音節と摩擦音-母音音節の区別に関する 2 つの音響手がかり、摩擦音と有声音部までの無音区間の間隔と、有声音部の F1 遷移起点周波数、の間のトレーディング関係を示す。摩擦音に続いて破裂音を知覚するためには、有声音部の F1 起点周波数が高いとき、無音区間を十分長くとらなければならないが、F1 起点周波数が低いときには、無音区間はあまり長くする必要はない。

手がかりトレーディングは、言語音信号あるいはその起源についての知識を利用していると仮定しなければ説明できないと主張され、言語音知覚の特異性を示す重要な現象とされている (e.g., Repp, 1982)。Best et al. (1981) は、広帯域ノイズと正弦波アナログ音により摩擦音-破裂音-母音音節を模擬した音響刺激を用いて、それらが言語音様に知覚されるときには音声トレーディングが生じるが、非言語音として知覚されるときにはそのような関係は生じないことを示した。

他方、Parker, Diehl, & Kluender (1986) は、母音-破裂子音-母音音節を模擬した非言語音刺激について、言語音において破裂閉鎖区間に現れる声道パルスと閉鎖持続時間との間で成立するのと同種のトレーディング関係が認められることを示したが、その際、刺激が言語音として知覚されていない場合であっても、言語音ラベルを用いて言語音として知覚させた場合よりは効果の大きさは小さいものの、それと同方向のトレーディング関係が生じるということを見いだしている。

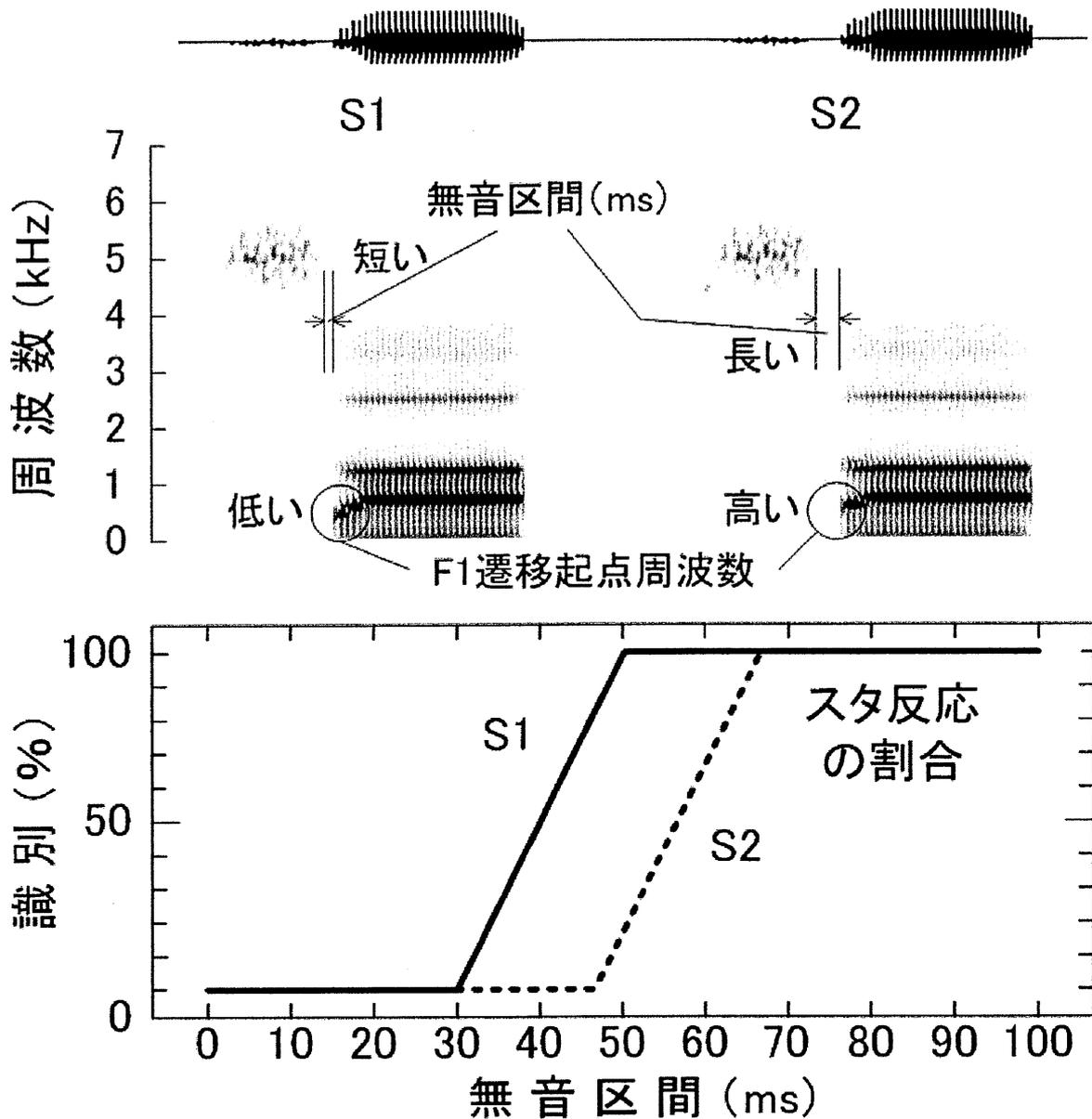


図1.10 摩擦音－破裂音－母音と摩擦音－母音の知覚における手がかり 트레이ディング

摩擦音に続いて様々な時間間隔で有声破裂音母音音節「ダ /da/」を提示するとき、音節の F1 起点周波数が低ければ、無音間隔が比較的短くても、「スタ /sta/」の知覚が生じるが、F1 起点周波数が高いと、等価な知覚を得るためには、より長い無音区間が必要になる。

1.3 文脈効果

音韻知覚における文脈効果とは、当該の音韻が置かれた言語的文脈の

影響で、音韻知覚内容が変化することを言う。発声における調音結合により言語音の音響的再構成が不変性の欠如をもたらすことは前述したが、その知覚的影響が音韻知覚の文脈効果である。文脈効果には、隣接音素によるもの、時間的に離れて先行・後続する音素によるもの、あるいは話者情報によるものなどさまざまなものがある。例えば、隣接音素による文脈効果として、破裂音ノイズを手がかりとする破裂子音の知覚が後続の母音により影響されるということを報告した前述の Liberman et al. (1952) の研究はその例である。また、話者情報による文脈効果としては、先行文中の母音フォルマントの周波数位置が後続のあいまいな母音の識別に影響することを示した前述の Ladefoged & Broadbent (1957) の研究がある。さらに、語頭のフォルマント遷移の持続時間が手がかりとなる破裂音—半母音連続体において、刺激全体の持続時間とともにカテゴリー境界の位置が移動することを示した Miller & Liberman (1979) の研究は、発話速度という文脈の影響を示している。

時間的に離れて先行する言語音文脈が後続の音韻知覚に影響する代表的な現象としては、選択的順応効果がある (Eimas, Cooper, & Corbit, 1973; Eimas & Corbit, 1973)。この効果は、当初、言語音知覚システムにおける示差的特徴検出器の存在を示すものと考えられていたが、多くの研究の結果、先行刺激による後続刺激の知覚に及ぼす文脈効果の一種、それも言語的処理というよりも聴覚的処理にかかわる効果であると考えられるようになった。例えば、Sawusch & Jusczyk (1981) は、順応刺激の知覚的同一性が、VOT を操作した有声—無声破裂音連続体 (/ba-/pa/) の知覚に及ぼす効果を検討した。順応刺激として連続体の両端の刺激、/ba/ と /pa/、並びに摩擦音 /s/ の後、無音区間に続いて /ba/ を提示する刺激 (/spa/ と知覚される) を用いたときの順応効果を比較した結果、/ba/ と /spa/ で

の順応効果は等価であった。すなわち、選択的順応において後続刺激の識別に影響するのは、先行する言語音の知覚的（音韻的）同一性ではなく、物理的同一性である、ということが示された。

他方、佐久間・高山（1986）は、順応刺激として言語音とその識別の手がかりとなる第2フォルマント部（ブリート音）とを用いて、言語音と非言語音による順応効果の大きさを比較する実験を行い、言語音刺激を順応刺激とするとき、非言語音刺激を順応刺激とする場合よりも順応効果が大きく、対側耳への順応の転移効果も大きいという結果を得た。この結果は、選択的順応効果が聴覚的処理と言語的処理の両過程に影響することを示唆している。

Repp（1982）によると、音声トレーディングと文脈効果は、言語音の生成メカニズムを考慮しなければ説明できないとされる。しかし、例えば、Pisoni, Carrell, & Gans（1983）は、Miller & Liberman（1979）と同様にフォルマント遷移の持続時間と刺激全体の持続時間を操作した実験を繰り返して、文脈効果の存在を確認した後、それら刺激のスペクトル構造を周波数変調音で模擬した正弦波アナログ音（後述）を用いて、非言語音としての識別に関しても同様の文脈効果が生じることを見いだした。彼らは、この効果が発話速度の標準化ではなく、言語音、非言語音にかかわらず、すべての音響信号のカテゴリー化と弁別に影響する一般的な心理物理的法則を反映しているにすぎないのかもしれない、と述べている。

1.4 耳の有利性

言語音知覚が特殊であることを主張する証拠として、言語音の知覚と非言語音の知覚が、大脳の異なる半球の処理に関わっている、というこ

とを示す研究が行われた（レビューとしては，Bryden, 1982, 1988; 利島・富永・高山, 1980）。Kimura (1961, 1964, 1967) は，一方の耳から対側の半球へ向かう交差神経路が，同側半球へ向かう非交差神経路よりも優勢であることを前提に，左右の耳に同時に異なる音響刺激を提示（両耳分離聴提示という）したとき，それらが言語刺激であれば（左言語半球と対側の）右耳での認知が優れ（右耳有利性 right ear advantage, 以後 REA と略す），非言語音刺激であれば左耳での認知が優れる（左耳有利性 left ear advantage, 以後 LEA と略す）か左右差が生じない（以後 noEA と略す），ということを示した。図 1.11 に，健常者における両耳分離聴に関する Kimura のモデルを示す。

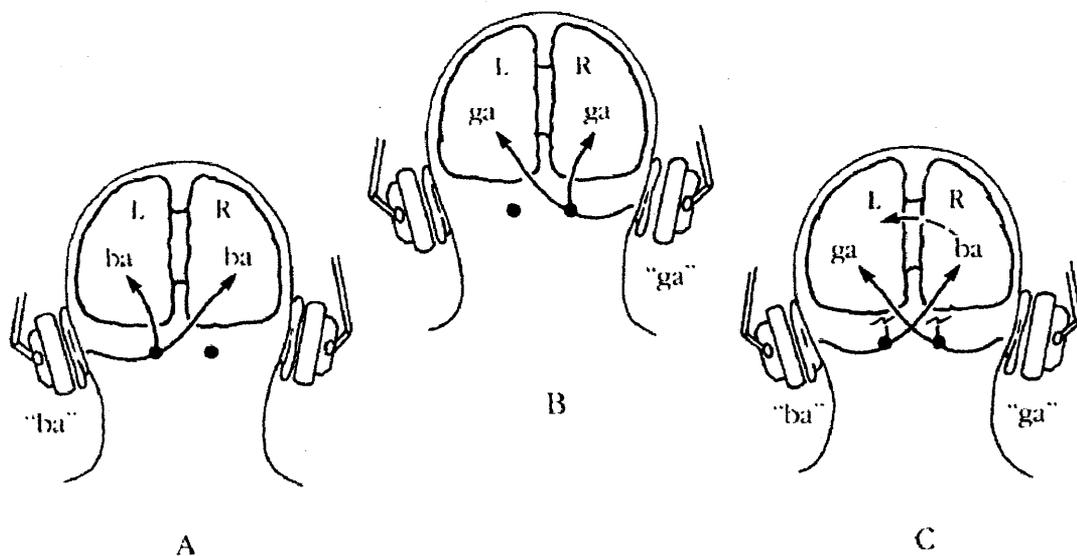


図 1.11 健常者の両耳分離聴についてのKimuraのモデル

(Springer & Deutsch, 1989 より)

単耳提示された言語音信号は，交差神経路と非交差神経路により両側の半球の言語音知覚中枢に伝達されるが（A, B），両耳分離提示された信号は，非交差経路が抑制されるために，対側半球にのみ伝達される（C）。その後の研究では，REA が，右半球から脳梁を経て左半球に伝達される左耳信号の減衰や遅れに由来すると考えられた（Berlin, Lowe-Bell, Cullen, Thompson, & Loovis, 1973; Sparks & Geschwind, 1968）。

音韻知覚に関しては、Studdert-Kennedy & Shankweiler (1970; Shankweiler & Studdert-Kennedy, 1967) が、自然の破裂子音－母音－破裂子音音節を用いて、語頭と語尾の破裂音に対する REA、中間の母音に対するわずかな REA (あるいは noEA) を示して以来、多くの研究でさまざまな言語音を用いて繰り返し検討されてきた (e.g., 高山・佐久間, 1990)。これまでの研究では、概ね、符号化の程度の強い破裂音や摩擦音では REA が認められるが、符号化の程度の弱い母音では耳の有利性は明確でない、ということが示されている。

他方、非言語音であってもその処理のされ方によ REA が示されることがあり (Bever & Chiarello, 1974; Halperin, Nachson, & Carmon, 1973; Natale, 1977; Papçum, Krashen, Terbeek, Remington, & Harshman, 1974)、また母音であっても記憶負荷などの課題要求を高めることで明確な REA が得られている (Godfrey, 1974; Weiss & House, 1973)。これらの研究は、耳の有利性が音韻知覚に関する特殊化したメカニズムの存在を示すとしても、そこに処理様式や注意など知覚的処理における主体の能動性が関与する余地があることを示すものであろう。

両耳分離聴で得られる耳の有利性と大脳半球の機能的非対称との関係については、十分に確立された事実とは言えないとして疑問視する主張もある (例えば, Bryden, 1988; 利島他, 1980; 高山・佐久間, 1990)。他方、その研究手段として用いられてきた両耳分離聴法は、音韻知覚の機能的単位としての音声的 (調音的) 特徴の心理学的実在性を検討する手段として有効であることが多くの研究で認められ、用いられている (Blumstein, 1974; Blumstein, Tartter, Michel, Hirsch, & Leiter, 1977; Hayden, Kirstein, & Singh, 1979; 高山, 1996)。

1.5 二重知覚

ある言語音の知覚の手掛りとなる音響的構成要素（チャープ）が一方の耳に提示され、それと同時に、その要素が取り除かれた残りの部分（基部）が他方の耳に提示されるとき、後者の側でもとの言語音が知覚される（図 1.12, e.g., Mann & Liberman, 1983; Rand, 1974）。基部自体、単独で提示されると言語音として知覚されるが、ここで知覚されるのは、要素単独と、その要素が取り除かれる前の言語音の 2 つであって、基部を含めた 3 つの音ではない。その意味も含めて、二重知覚と呼ばれている。

通常、ある音響的要素は同時に 2 つ以上の音源に割り当てられることはなく、このことを分離配置の原理（ムーア、1994）と言う。したがって、同じ音響信号が、非言語音として知覚されると同時に、言語音の一部として音韻知覚の成立に寄与するという事実は、両者に対してそれぞれ独立した知覚モジュールあるいは知覚様式が存在することを示す証拠

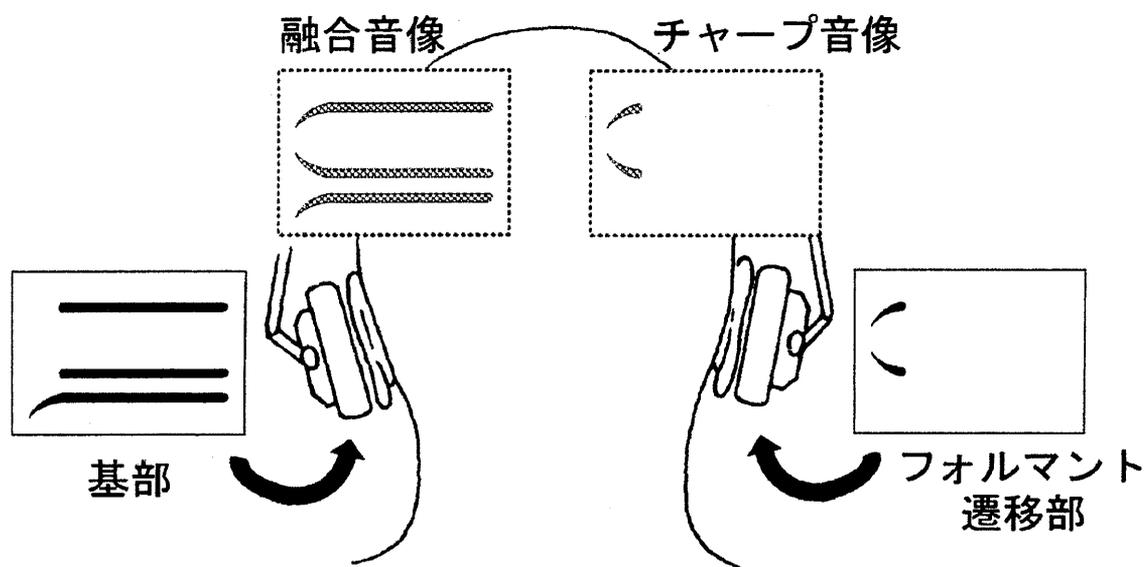


図 1.12 二重知覚

融合音像（言語音）が基部側に、チャープ音像がフォルマント遷移部側に同時に知覚される。

とされた (Lieberman & Mattingly, 1985)。二重知覚におけるそれぞれの知覚内容について検討した結果も、非言語音側では連続的な知覚が生じ、言語音側ではカテゴリー的な知覚が生じることを示しており (Mann & Liberman, 1983)、さらに文脈効果 (Mann & Liberman, 1983) や音声的トレーディング (高山, 1990) といった言語音知覚に特有とされる現象も確認されている。

他方、言語音から分離して提示される音響的手がかり部 (チャープ) が単独では非言語音として知覚され、基部との両耳間統合により音韻性情報の伝達が可能となる、という解釈に関して、異論も出されている。Nusbaum, Schwab, & Sawusch (1983) は、破裂子音の調音場所の手がかりとなる第 2 フォルマント (F2) 遷移について、遷移部単独であっても、通常の言語音刺激に匹敵する程度に識別が可能であることを示した。Repp, Milburn, & Ashkenas (1983) は、破裂子音の調音場所の手がかりとなる第 3 フォルマント (F3) 遷移について、遷移部単独での識別を検討し、Nusbaum et al. (1983) の主張を否定しているが、高山 (1989a) は破裂子音の構成要素に対する音韻識別の程度を比較した結果から、聴者は F2 遷移と F3 遷移それぞれから偶然よりも高い確率で音韻情報を取り出すことができること、その際 F3 遷移よりも F2 遷移のほうをより良く識別できることを示した。また、高山 (1990) は、摩擦音-破裂音-母音音節における破裂音の知覚に関する無音区間の長さとの第 1 フォルマント (F1) 遷移起点周波数との音声トレーディングについて、通常の単耳提示条件と、破裂音の音響的手がかりであるフォルマント遷移部を分離し、残りの部分と両耳分離提示する条件とで比較し、破裂音の知覚に関して、分離された遷移部からの情報が独立に評価されている可能性を示唆している。

1.6 正弦波アナログ音

House et al. (1962) が示したように、言語音をさまざまな程度に模擬した非言語音信号の知覚においては、非言語音として知覚されるか言語音(様)の知覚が生じるか、のいずれかである。Stevens & House (1972) は、「聴者が信号を聞くのに先だって、言語音に対する構えをもっている必要はない。適切な音響特性をもった信号が存在することが引き金となって、聴者の構えが生じるからである。」(p. 13) と述べている。しかし、正弦波アナログ音のように、言語音の音韻関連特性を保持しながら、聴者の構えを自動的に引き出せない刺激もある。

Remez, Rubin, Pisoni, & Carrell (1981) は、男性話者が発声した文のフォルマント構造を分析し、第1～第3フォルマントの中心周波数と振幅の時間変動を正弦波で模擬した複合音を作成した。この音は、言語音のような調波構造も、フォルマントを中心とした広い周波数帯域へのエネルギー分布ももたないので、言語音とは非常に異なって聞こえる。事実、この正弦波アナログの聴取実験において、音の性質について何の教示も与えず、自由に印象を報告するよう求めた場合、ほとんどの被験者は非言語音としての知覚内容を報告した。他方、それがコンピュータによって合成された文であり、できるだけ忠実に識別するように求められた被験者の多くは、聞こえ自体は不自然に感じるものの、言語音として聞くことができ、文全体あるいは文中の多くの音節を正しく識別することができた。

図 1.13 に、先に示した「青いズボンを買った」という発話の広帯域スペクトログラムとその正弦波アナログ音のスペクトログラムを示す。ローマ字表記は、各音韻のおおよその発話の時点を示している。

正弦波アナログ音は、聴者に与える教示や聴者の構え・注意によって、

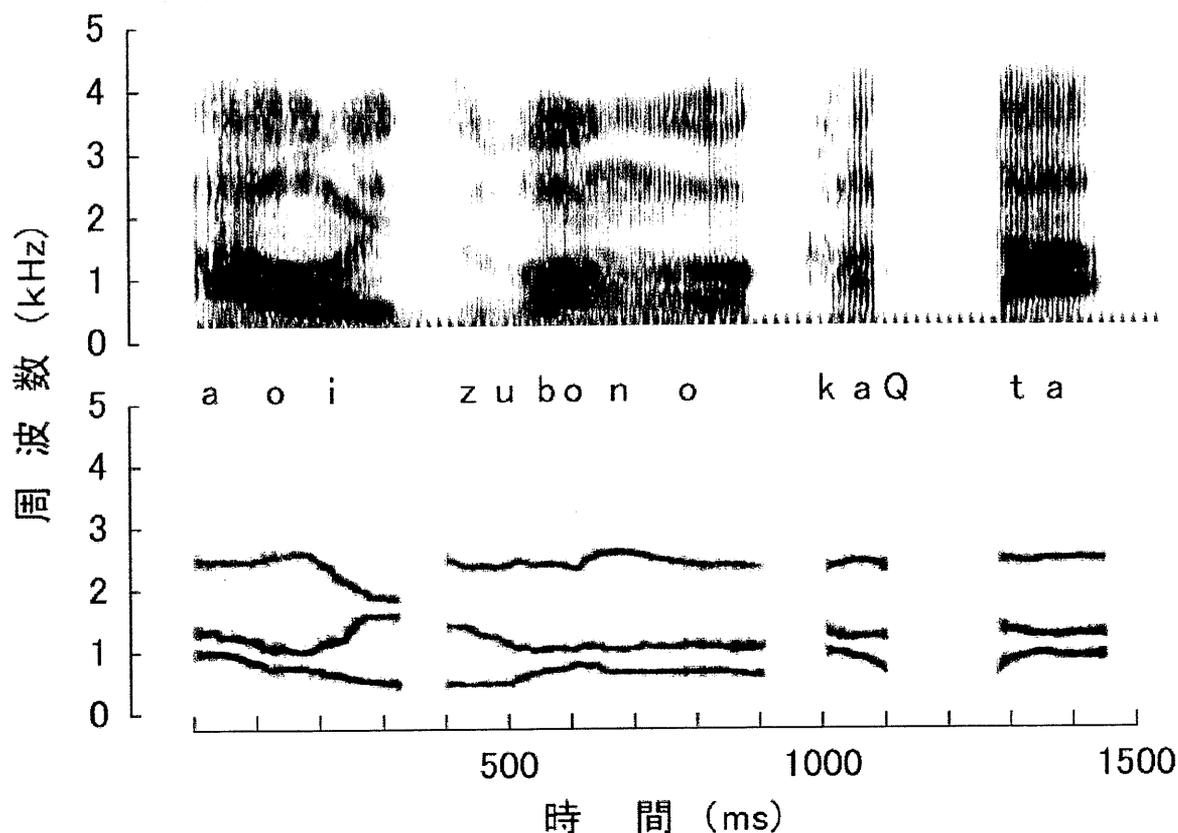


図 1.13 言語音「青いズボンを買った」のスペクトログラム（上）とそれを模擬した正弦波アナログ音のスペクトログラム（下）

同一の音響刺激に対する知覚内容を操作することができることから、音韻知覚の特異性あるいは言語音知覚と非言語音知覚の差異をとらえるための有効な刺激材料として用いられてきた。例えば、アナログ音を音節として知覚する聴者はそれをカテゴリー的に知覚し (Best et al., 1989; Johnson & Ralston, 1994), また、音韻知覚の音響手がかり間の音声トレーディング関係を利用することができる (Best et al., 1981) ことが示されている。加えて、狭帯域破裂音ノイズの中心周波数を手がかりとする破裂子音の調音場所の知覚が後続母音によって左右されるという文脈効果が、日本語母音を模擬したアナログ音が破裂音ノイズに後続する場合にも認められることも示されている (高山, 1999)。さらに、非言語音と

しての知覚ではマスキングされているアナログ音の構成要素が、言語音としての知覚では音韻知覚の手がかりとして利用可能となり (Grunke & Pisoni, 1982; Schwab, 1981), 同様に, 非言語音として知覚されるときには相互独立的に処理される複数の構成要素が, 言語音として知覚されるときにはアナログ音節全体にわたって統合的に処理される (Tomiak, Mullennix, & Sawusch, 1987), ということも見いだされた。

何故正弦波アナログが言語音として知覚できるのかということについては, Remez et al. (1981) は, 正弦波アナログが提供するフォルマント中心周波数の情報は, 聴者の言語音知覚過程を自動的に始動させるには十分とはいえないが, 言語音教示により, 聴者の注意をアナログ音に含まれる音韻関連特性に向けさせることで, 言語音知覚を引き起こすことができる, としている。そして, その後の研究で (Remez & Rubin, 1990), このアナログ音に含まれる言語音的特性として, 主に周波数次元での, そしてある程度振幅次元での時間的変動が重要であることを示した。正弦波アナログ音のスペクトル構造が言語音に近似することが, アナログ音に関する音韻様のカテゴリー分類に重要であることは, Grunke & Pisoni (1982) や Schwab (1981) も指摘している。

正弦波アナログ音に対するこのような知覚的特徴は, 二重知覚の場合と同様に, 言語音知覚と非言語音知覚に関して異なる知覚モジュールあるいは知覚様式が存在する証拠と考えられている (Liberman, 1982; Liberman & Mattingly, 1985)。しかし, スペクトル構造の近似は正弦波アナログ音が言語音として知覚されるための必要条件であるとしても, 何故それが可能かということに関しては, 後述の運動理論や直接知覚理論の立場から, 言語音と同じ音韻関連特性を有しているから, と述べられるだけで, 必ずしも明らかにはなっていない。言語音知覚様式のもとでは,

正弦波アナログ音は言語音として知覚される、と論じても、言語音知覚様式そのものは現象の記述概念でしかなく、それを知覚メカニズムとして論じれば、循環論は免れない。

また、アナログ音を言語音として知覚できる程度にも音韻タイプによる違いが見いだされており、母音や半母音のように過渡的特性（例えば、フォルマント遷移）がないか、あるいは緩やかな音韻のアナログ音に対して、訓練をしなくても、もとの音韻が知覚されやすい、ということが示されている（高山，2002b）。符号化の程度の強い、あるいは過渡的特性の強い刺激において、カテゴリー的知覚や右耳有利性など言語音知覚に特異的な現象が確認されているが（Godfrey, 1974; Liberman et al., 1967; Pisoni, 1971, 1973; Shankweiler & Studdert-Kennedy, 1967; Studdert-Kennedy & Shankweiler, 1970; Weiss & House, 1973），アナログ音に対する音韻知覚の成立のし易さの違いはそれらとは反対方向に現れている。

あるいは、アナログ音の音韻様の知覚的効果自体、アナログ音が音韻的に知覚されていることを保証するものではないことを示す研究もある。Sawusch & Gagnon (1995) は、Best et al. (1989) が言語音教示によりカテゴリー的知覚が得られたのと同じ刺激連続体を用いて、十分な非言語音識別訓練を行うことで、非言語音として聞かれたアナログ音に関してもカテゴリー的知覚が得られ、さらに訓練の影響が母音文脈の異なる場合にまで般化される、ということを見いだした。彼らは、これらの結果から、言語音のカテゴリー化と非言語音のカテゴリー化との両者を支える抽象的な聴覚的表象の存在を主張している。また、文脈効果の項で述べたように、Pison et al. (1983) は、Miller & Liberman (1979) と同様にフォルマント遷移の持続時間と刺激全体の持続時間を操作した言語音の識別と正弦波アナログ音の非言語音としての識別の両者において類似の

効果を見だし、発話速度の標準化効果と解釈されていた結果が、すべての音響信号のカテゴリー化と弁別に影響する一般的な心理物理的法則を反映している可能性を示唆している。

1.7 視聴覚統合

日常の会話は対面で行われることが多いが、発話者の顔と唇の動きが、音韻知覚に強い影響を及ぼすことが知られている。この効果は、研究者の名前に因んで McGurk 効果と呼ばれている。McGurk & MacDonald (1976) は、人が /baba/ や /papa/ など 2 音節語を発話しているビデオ映像を用いて、一部の映像が、そこでは実際には発話されていない音声と同期するように編集し直し、参加者に発音されている音を報告するよう求めた。その結果、例えば、/gaga/ と発音している映像に /baba/ の音を同期させたところ、/dada/ と聞こえたという報告が多く、しかもそのほとんどの場合、映像と音声の矛盾にも気づかれていない、ということが見いだされた。図 1.14 には、/ga/ を発音する映像を見ながら、/ba/ の音を聞くと、/da/ の知覚が生じるという McGurk 効果の例を示す。

McGurk 効果は、言語音聴取と発話とに共通する段階での相互作用を示すものとして、Liberman et al. (1967) の言う言語音知覚に特殊化した過程の存在を支持する証拠と主張されることが多い (Liberman, 1982; Liberman & Mattingly, 1985; Repp, 1982)。しかし、ムーア (1994) は、「音響情報と視覚情報の組み合わせ方は複雑であり、その説明は必ずしも容易であるとは限らない (解説としては Summerfield, 1987 を参照)。... しかし、聴覚と視覚の統合は非音声音についても起こるということを心に留めておくべきである。」(p. 298) として、「聴覚と視覚の統合を根拠として、調音情報を利用するモード、音声に特殊化したモードが知覚に

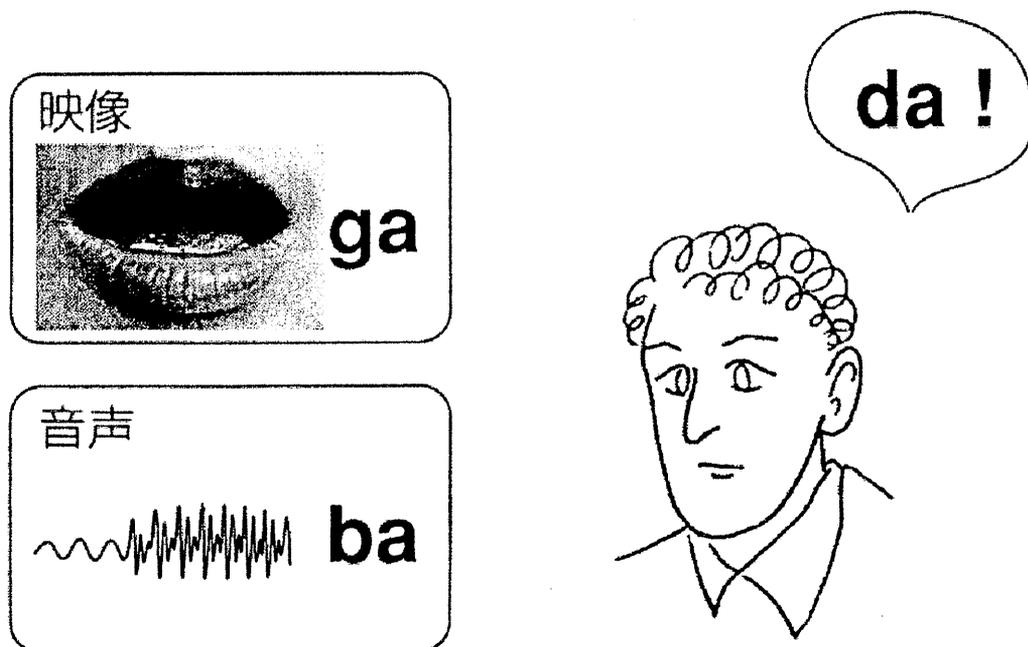


図 1.14 McGurk効果の例 (積山, 1997 より)

存在すると主張する」(ムーア, 1994, p. 298) ことには慎重であるよう注意を促している。また, 積山 (1997) も, 「... マクガーク効果を説明するには, 記憶内の音素の表象が, 単に音響的な特徴だけでなく, どうやって発音するかという調音上の特徴までも保持していると考えなければならぬ」(p.123-124) としながらも, 「ただし, マクガーク効果そのものを説明するには, 音素の表象にその調音的特徴が『筋肉運動的に』保持されていると考える必要は必ずしもなく, 調音的特徴は『視覚的に』保持されていると考えてよい」(積山, 1997, p. 124) と述べている。

2. 言語音知覚の理論

これまでに挙げた現象は, それらが見いだされた当初, 言語音知覚に

特異的なものと考えられ、言語音知覚に特殊化したメカニズムの働きを示す有力な証拠であると主張されていた。しかし、それぞれの項で述べたように、現在では、非言語音知覚においても類似の現象が見いだされたり、一般的な心理物理学的法則の枠組み内で説明が試みられているものもある。ここでは、これらの現象を説明しようとする3つの理論的立場を概観する。

言語音知覚メカニズムの特異性を主張し、その理論的立場を支持するものとして、精力的に言語音知覚に特異的な現象の発見に努めたのは、**言語音知覚の運動理論**の提唱者たちであった (Lieberman, 1982; Liberman et al., 1967; Liberman & Mattingly, 1985; Liberman & Whalen, 2000; Repp, 1982, 1984)。この理論では、言語音の知覚が、その発声（発声の意図から実際の発声までを含む）と共通基盤をもつ生物学的に特殊化した過程（モジュール）の働きによって、言語音自体ではなく、その発声の意図レベルでの調音動作、あるいは調音運動神経指令を直接対象として成立する、と主張されている。すなわち、聴者は、自身の言語音生成における調音結合の知識を用いて、「特殊化した言語音モジュール」(Lieberman & Mattingly, 1985) にカプセル化された内的な調音動作パターンに基づく抽象的な音素的表象として、言語音の音韻内容を解釈する。

この理論によれば、音韻知覚における音響信号の不変性の欠如と音韻知覚内容の恒常性は、それを発する調音動作に不変性があることによって説明される。同様に、カテゴリー的知覚も目的とする音韻を生じるための調音動作が離散的であることによって説明される。また、音声トレーディングや文脈効果、そして視聴覚統合についても、すべて調音事象という共通の次元でとらえれば、当然の現象であるとする。さらに、右耳有利性、二重知覚、正弦波アナログ音の知覚は、言語音知

覚に特殊化したモジュールあるいは知覚様式の存在を支持するものと解釈されている。

運動理論の延長として展開されてきた**直接知覚理論** (e.g., Fowler, 1986, 1996) も、知覚の対象はその言語音を生じさせた調音動作であるとする。ただし、運動理論では特殊化したメカニズムに重点が置かれていたが、直接知覚理論では、その伝達媒体である言語音信号の中に調音動作を特定するために十分な情報 (不変的特性) が含まれている、と主張する。すなわち、言語音信号そのものの特殊性が強調されている。したがって、先に挙げた言語音知覚に特異的な現象はすべて、調音動作の結果として、その起源を特定するために十分な情報を含む刺激により引き起こされていると考える。

それに対して、**情報处理的アプローチ** (e.g., Sawusch, 1986; Studdert-Kennedy, 1976) では、言語音知覚が、音響信号の聴覚的表象を言語的表象へと対応づけるための一連の処理段階を経て成立し、言語音処理と非言語音処理に共通する音響的特徴の分析や一般的な聴覚的処理の結果に対する言語的段階での統合的処理や評価に基づくものである、としている。このアプローチの中には、言語音知覚を、他の対象知覚と共通する一般的知覚過程として扱おうとするものもある (Massaro & Friedman, 1990)。

そもそも、前述したように、言語音知覚と非言語音知覚の違いを示すとされる現象のそれぞれについて、課題要請や注意・構えの影響を指摘する研究、あるいは一般的な心理物理学的法則の枠組み内で説明が可能であるとする研究など、言語音と非言語音に対して異なる知覚様式やメカニズムの存在を前提としない反論も提出されており、言語音知覚の特異性そのものについても、自明のことであるとは言えないように思える。

このように、言語音知覚が特殊化したメカニズムによって生じるのか、あるいは一般的な知覚的・認知的処理の枠組みの中でとらえることができるのかということは、現在でも未解決の問題である。

第3節 本研究の目的

言語音知覚が特殊化したメカニズムによって生じるのか、あるいは一般的な知覚的・認知的処理によって生じるのか、という問題を解決するための手がかりとなるのは、音韻知覚が、言語音信号の音響特性と生得的な受容メカニズムから受動的に生起するわけではなく、聴者の聴取方略や注意のあり方によって左右されるという事実である (Gordon, Eberhardt, & Rueckl, 1993; Francis, Baldwin, & Nusbaum, 2000; Nusbaum & Schwab, 1986)。Gordon et al. (1993) は、音韻識別課題遂行時に、同時に別の作業を課すと、識別に強く影響している音響手がかりの影響は弱められるが、弱い影響しか与えていない手がかりにはほとんど作用しない、ということを見いだした。また、Francis et al. (2000) は、語頭の破裂子音の調音場所に関する音響手がかりである破裂音ノイズとフォルマント遷移について、異なる調音場所を示す手がかり同士の組合せも含むあらゆる可能な組合せで合成した破裂子音-母音音節を用いて、たとえ手がかり同士が矛盾していても、注意が向けられる一方の手がかりのみに基づいて、破裂音（の調音場所）を識別するよう訓練することができる、ということを示した。これらの結果は、音韻知覚の過程が、処理の固定化された自動的な過程ではなく、聴者の聴取方略や注意のあり方によって左右される能動的な活動であることを示すものである。このことは、さらに、言

語音と非言語音との区別自体が、信号の発生源や音響特性によって一義的に規定されるのではなく、聴者がそれをどのように知覚し、どのように処理しようとするのか、という知覚的体制化や知覚様式に左右される内的過程の産物である、ということを示唆するように思える。

1. 正弦波アナログ音の知覚

言語音知覚と非言語音知覚の違いが必ずしも刺激規定的ではないことは、先述の正弦波アナログ音の知覚で例証される。この音は、言語音のフォルマント構造を2～4の周波数変調音で模擬した非言語音で、言語音のような調波構造も、フォルマントを中心とした広い周波数帯域へのエネルギー分布ももたないので、言語音とは非常に異なって聞こえる。そのため、通常、自動的に言語音様に知覚されることはない (Remez et al., 1981)。しかし、教示等で聴者にそれを言語音として聞くよう求めると、例えば、カテゴリー的知覚 (Best et al., 1989; cf. Sawusch & Gagnon, 1995)、音声トレーディング (Best et al., 1981)、文脈効果 (高山, 1999) などの言語音様の知覚が再現される。

非言語音である正弦波アナログ音に対して音韻知覚が可能となるのは、そのスペクトル的時間的パターンが音韻関連特性として機能しているからに他ならない。運動理論と直接知覚理論は、この音韻関連特性が自然言語音に内在する音響的音韻関連特性と等価であることを前提として、言語音教示等で聴者がこの特性に注意を向けることができれば、自動的、自律的にあるいは直接的に音韻知覚が成立する、と主張している (e.g., Best et al., 1989; Remez et al., 1981)。他方、情報処理アプローチの立場では、その音韻関連特性への注意はもちろんであるが、さらに音韻知

覚の成立には、それに対する聴覚的表象と長期記憶内の音韻表象との対応づけが必要であると考えられている (Sawusch & Gagnon, 1995; Tomiak et al., 1987)。音韻知覚における記憶要因については、カテゴリー的知覚のモデル化において取り入れられており (Fujisaki & Kawashima, 1969, 1970; Macmillan et al., 1988; Pisoni, 1971, 1973), 最近では音韻カテゴリー内におけるメンバー間の序列構造への関心から、音韻カテゴリーの内的標準あるいはプロトタイプへの参照過程の存在 (Kuhl, 1992; Miller, 1994) も指摘されている。

2. 本研究の目的

本研究では、音韻知覚に関する情報处理的アプローチの立場に立ち、「音韻知覚は、音響信号に関する知覚表象と長期記憶に保持された音韻表象との照合により成立するものであり、言語音知覚と非言語音知覚の差異は音韻記憶表象を参照するか否かによる」という仮説を、正弦波アナログ音を用いて検討することを目的とした。

この仮説を検証するために、3つの実験を行った。まず実験 I と実験 II では、Remez et al. (1981) が、正弦波アナログ音の音韻知覚は、教示等により聴者の注意をアナログ音の音韻関連特性に向けさせることによって可能となる、と述べていることに着目した。運動理論や直接知覚理論によれば、音韻関連特性に注意が向けられれば、音韻知覚が自動的、直接的に成立することになる。しかし、上記仮説が正しければ、音韻知覚が成立するためには、アナログ音の音韻関連特性に注意が向けられるだけでなく、その知覚表象が適切な音韻記憶表象と照合されなければな

らない。この点を検討するために、実験Ⅰでは破裂子音－母音音節アナログを用いて、実験Ⅱでは母音アナログを用いて、アナログ音に与えられるラベルを操作する識別学習実験を行った。アナログ音を歪んだ言語音として聞くよう求める言語音教示のもとで適切な音韻ラベルを用いた識別学習と、不適切な音韻ラベルを用いた識別学習、並びにアナログ音を歪んだ和音とする非言語音教示のもとでの任意のラベルを用いた識別学習という3つの識別学習の成績を比較するとき、音韻関連特性と音韻記憶表象との既存の照合関係を利用できない非言語音教示のもとでの識別学習が相対的に劣ること、そして、不適切な音韻ラベルを用いた識別学習では、既存の照合関係からの妨害的な影響を受け、学習が大きく損なわれることが予想される。

実験Ⅰと実験Ⅱは、正弦波アナログ音に対する音韻知覚が、その音韻関連特性の知覚表象と長期記憶における音韻表象との照合により成立することを検証する目的で実施されたが、それが検証されても、その結果は、アナログ音と言語音の音韻知覚が同じメカニズムによって成立することを直接支持するものではない。例えば、アナログ音の音韻知覚は音韻記憶表象との照合を媒介とするが、言語音の音韻知覚は本来直接的であるということも考えられる。そこで、実験Ⅲでは、正弦波アナログ音と言語音の音韻知覚が同じメカニズムによって成立するのか否かを明らかにするために、アナログ音と言語音の両者をターゲット刺激とした直接プライミング実験を行い、両者が共通の音韻記憶表象にアクセスする（音韻記憶表象を活性化させる）か否かを検討した。すなわち、アナログ音についての言語音教示条件と非言語音教示条件でプライミングの有無を調べることによって、言語音教示条件においてのみ言語音との相互プライミングが認められるという結果が得られるならば、それは、言語

音とアナログ音が，音韻知覚において共通の音韻記憶表象にアクセスしている（音韻記憶表象を活性化させている）こと，したがって両者の音韻知覚が同一のメカニズムによって成立していることを支持するものと考えられる。

第2章 言語音と非言語音の知覚的差異をもたらす内的過程に関する実験的検討

第1節 破裂音－母音アナログ音の識別学習（実験Ⅰ）

1. 目的

実験Ⅰでは、まず、正弦波アナログ音に対して音韻知覚が生じる過程において、実際に、その音韻関連特性の知覚表象と、長期記憶に保持された音韻表象との照合が行われているのか否かを検討した。

Remez et al. (1981) によれば、正弦波アナログ音が提供するフォルマント中心周波数の情報は、聴者の言語音知覚過程を自動的に始動させるには十分とはいえないが、言語音教示により、聴者の注意をアナログ音に含まれる言語音的特性に向けさせることで、音韻知覚を引き起こすことができる。すなわち、非言語音である正弦波アナログ音から音韻内容が知覚されるためには、アナログ音に元の言語音のスペクトル変動の時間構造が維持されていることが必要であるとともに、聴者がそのようなスペクトル的・時間的特性を音韻関連情報として適切に処理することが必要である。先の仮説に基づけば、後者の点は、アナログ音の音韻知覚にとって、音韻関連特性と音韻記憶表象との照合が不可欠であることを意味するものである。

また、言語音教示と非言語音教示のもとで正弦波アナログ音の識別成績を比較した研究では、アナログ音のスペクトル的・時間的構造が言語音のそれに近似するほど、言語音教示のもとでの識別成績が非言語音教示のもとでの識別成績よりも高くなる、という結果が得られている

(Grunke & Pisoni, 1982; Schwab, 1981)。このことは、正弦波アナログ音の音韻知覚において、スペクトル的・時間的構造やその知覚表象が音韻記憶表象に容易に対応づけられるほど、識別が容易になることを示唆する。

さらに、Remez et al. (1981) は、いったん言語音として知覚されると、非言語音として知覚するのは困難になることも指摘している。すなわち、アナログ音の音韻関連特性と音韻記憶表象との対応づけが学習される（あるいは見いだされる）と、その学習はほとんど不可逆的となる。このことは、アナログ音に対して音韻知覚が可能になる過程においては、アナログ音がもつ音韻関連特性と長期記憶内の音韻表象との間にある何らかの既存の関係が利用されるようになることを示唆する。

以上のことから、実験 I では、先の仮説を検証するために、次の予測を確かめることを目的とした。すなわち、言語音教示のもとで音韻ラベルを用いてアナログ音の識別を学習させることは、Remez et al. (1981) が述べるように、必然的に聴者の注意をアナログ音の（もとの言語音から引き継いだスペクトル的・時間的構造のような）音韻関連特性に向けさせ、それを音韻知覚の手がかりとして利用できるようにすると考えられる。アナログ音に与えられた音韻ラベルが、その音韻関連特性と音韻記憶表象との既存の関係に矛盾しない限り、識別学習は容易に成立するであろう。しかし、音韻ラベルがそのような関係に矛盾する場合は、音韻関連特性と音韻記憶表象との対応関係に混乱が生じ、学習が妨げられるであろう。他方、非言語音教示のもとでの音韻以外のラベルを用いた識別学習は、そのような既存の関係は利用できないので、通常の対連合学習として成立するはずである。したがって、その際の学習の容易さや速さは、言語音教示のもとで適切な音韻ラベルを用いて学習する場合と不適切な音韻ラベルを用いて学習する場合の中間に位置すると予想され

る。

2. 方法

2.1 刺激

刺激音は、合成言語音「バ、ダ、ガ」のフォルマント構造を、その第1フォルマント (F1) ~ 第3フォルマント (F3) の中心周波数に対応する周波数変調音で模した3周波成分の正弦波アナログ音であった。各アナログ音の成分構成を図2.1に示す。

これらアナログ音は、F1の起点周波数が250 Hz、定常部の周波数についてF1が750 Hz、F2が1250 Hz、F3が2500 Hzで共通しており、F2とF3の起点周波数の組合せで異なっていた。「バ」アナログ音のF2起点周波数は750 Hz、F3の起点周波数は1750 Hzであり、「ダ」ではそれぞれ1750

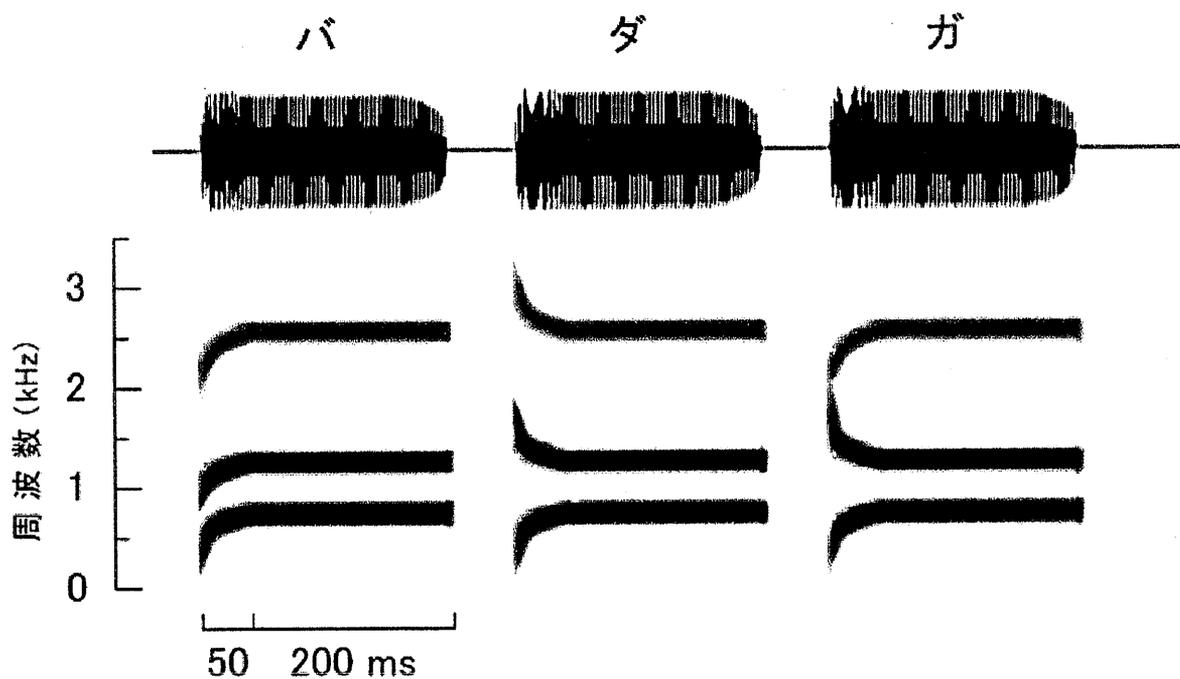


図 2.1 正弦波アナログ音の構成

Hz, 3250 Hz, 「ガ」ではそれぞれ 1750 Hz, 1750 Hzであった。3つの周波数成分は, その起点周波数から定常部の周波数に向けて 50 ms の間に対数関数的に変化した。また, 3つの周波数成分の相対振幅は F1 0 dB, F2 -6 dB, F3 -24 dB であった。

これらアナログ音は, パーソナルコンピュータ (IBM ThinkPad 235) を用いて, 自作ソフトウェアにより 16ビット, 10kHzの精度で合成された。

2.2 手続き

実験参加者は, 4～8名のグループで, 予め与えられたカテゴリー一名に基づいて前記3種のアナログ音を識別(分類)する実験を行った。与えられたカテゴリー一名に応じて参加者は3つの教示群に分けられた。言語音教示一致群では各アナログ音のもととなった言語音の音韻名が, 言語音教示不一致群ではもとの言語音との対応を完全に変えた音韻名が, 非言語音教示群では, 「A」, 「B」, 「C」という音韻とは無関連なカテゴリー一名が与えられた。

識別作業は2つのブロックで行われた。被験者は, 各ブロックにおいて, 1系列30試行, 6系列のアナログ音を聴取した。各ブロック第1系列では, 予め与えられたリストと照合しながら, 各10回ずつ現れるアナログ音の聞こえとカテゴリー名の対応を確認した。第2系列以後では, 初めに3つのアナログ音が一定の順序で2回繰り返され, 被験者はその聞こえとカテゴリー名の対応を再度確認した後, 残り24音の識別作業を行った。識別の正誤についてはいかなるフィードバックも与えられなかった。

アナログ音は, パーソナルコンピュータ (IBM ThinkPad 235) のオーディオ出力をアダプターを用いて各被験者のヘッドフォンに分岐させ, わ

ずらわしさを感ぜない程度の強さで提示された。各音の提示間隔は 2.5 s, 15 音毎に 10 s の無音区間を挿入し, 各 15 音の最初には, 開始の合図として 100 ms の 1 kHz 純音を提示した。ブロック間には 1 ~ 3 min 程度の休憩を設けた。

2.3 参加者

日本語を母国語とし, 健常聴力を有する近畿大学電子情報工学科 3 年生 63 名 (男 60 名, 女 3 名) が実験に参加した。いずれの参加者も, 以前, 言語音知覚に関する実験に参加したり, 正弦波アナログ音を聴取した経験はなかった。また, これらの参加者のうち 3 名の男子学生は, 主に教示を理解していなかったことにより, ほとんどすべての試行で, アナログ音を教示に従って正しく識別できなかつたので, 分析からは除外した。結果として, 60 名の参加者を分析の対象とした。

なお, 実験後のインタビューにおいて, 言語音教示不一致群の参加者の中で, 正弦波アナログ音に与えられた音韻ラベルがその音韻関連特性と矛盾することに気づいたと報告した者はおらず, 非言語音教示群の参加者の中で, アナログ音を言語音として聞いて識別しようとしたと報告した者もいなかった。

3. 結果

各群の試行系列毎の平均正識別率を図 2.2 に示す。全系列を込みにした平均識別率は, 言語音教示一致群で 58%, 言語音教示不一致群で 49%, 非言語音教示群で 39% と全般的に低く, また最も識別成績の良かった

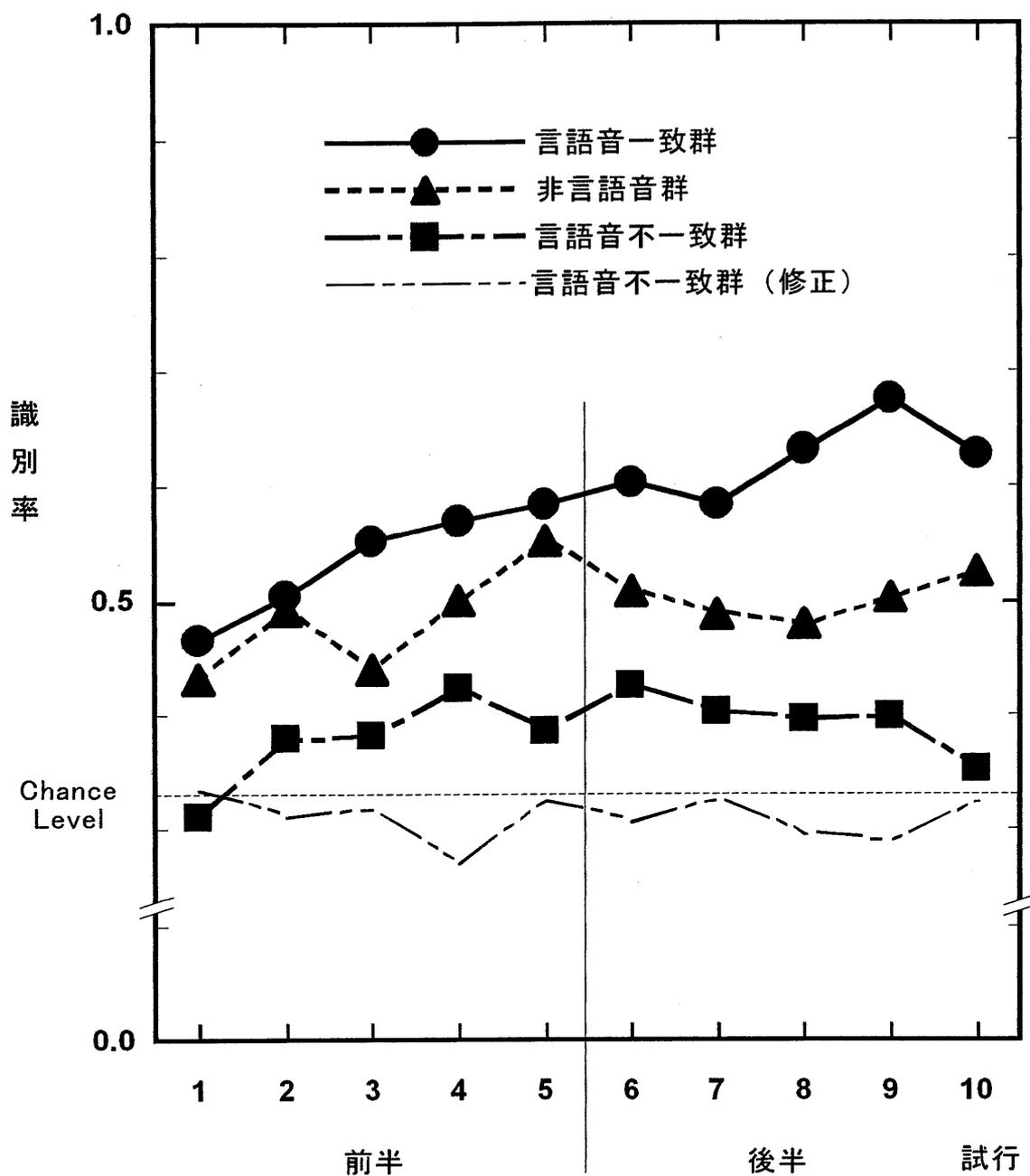


図 2.2 各試行系列における正識別率

言語音教示一致群においても系列間で最大 68% にしか達しておらず、実験 I で用いた識別課題は困難な課題であったと言える。特に言語音教示不一致群の成績は偶然レベルよりわずかに高い程度でしかなく、この点について、不一致群の分類を、アナログ音のもとの言語音の音韻名で識

別したものを正答として集計し直しても（図 2.2 の 2 点鎖線），識別成績が改善されるということにはなかった。

正識別率の角変換値を用いて教示群×ブロック×試行系列の 3 要因分散分析を行った。その結果，教示群 ($F(2,57) = 10.074, p < .001$)，ブロック ($F(1,57) = 9.685, p < .01$)，試行系列 ($F(4,228) = 4.041, p < .01$) それぞれの主効果，並びに教示群×ブロック ($F(2,57) = 3.307, p < .05$)，ブロック×試行系列 ($F(4,228) = 3.942, p < .01$) の交互作用が有意であった。また，教示群×試行系列の交互作用も有意水準には達しなかったが，傾向が認められた ($F(8,228) = 1.917, p < .10$)。

まず，教示群の主効果について，事前対比により教示群間で識別学習の全体的な成績を比較すると，言語音教示一致群が最も高く（言語音教示一致群 vs. 非言語音教示群， $F(1,57) = 4.144, p < .05$ ），言語音教示不一致群の学習成績が最も劣っていた（非言語音教示群 vs. 言語音教示不一致群， $F(1,57) = 5.985, p < .05$ ）。この効果は，教示群×ブロックあるいは教示群×試行系列に関わる交互作用においても維持されており，各ブロックにおける教示群の単純主効果（第 1 ブロック： $F(2,114) = 6.009, p < .01$ ，第 2 ブロック： $F(2,114) = 12.432, p < .001$ ），並びに各試行系列における教示群の単純主効果（第 1 試行： $F(2,285) = 5.529, p < .01$ ，第 2 試行： $F(2,285) = 4.745, p < .01$ ，第 3 試行： $F(2,285) = 8.708, p < .001$ ，第 4 試行： $F(2,285) = 9.318, p < .001$ ，第 5 試行： $F(2,285) = 12.370, p < .001$ ）のいずれも有意であった。各ブロックあるいは各試行における教示群間の識別学習成績について，Tukey の HSD 検定により比較したところ，すべての比較において言語音教示不一致群の成績は，言語音教示群と非言語音教示群の成績よりも有意に劣っており，また言語音教示一致群と非言語音教示群に関する 7 つの比較のうち，3 つで言語音教示一致群の成績は非言語音教示群の成績

よりも有意に高かった（各ブロックにおける教示群間の比較： $MSE=0.031$, $df=57$, 5%水準, 各試行系列における教示群間の比較： $MSE=0.016$, $df=228$, 5%水準）。

これらの結果は、正弦波アナログ音の識別において、音韻ラベルを用いることは相対的に識別を促進したが、それはアナログ音の時間的スペクトル構造と相関的な言語音の音韻ラベルが用いられるときに限られており、アナログ音のスペクトル時間的構造と矛盾する言語音の音韻ラベルを用いることは識別学習を妨げることを示すものであった。

ブロックあるいは試行系列の主効果について、このこと自体は識別学習の効果を支持するものであり、第1ブロックより第2ブロックで、また試行系列が進むにつれて、識別成績が上昇することを意味する。しかし、教示群×ブロックあるいはブロック×試行系列の交互作用、並びに教示群×試行系列の交互作用の傾向は、この学習効果が単調に進行するものではないことを示唆している。教示群×ブロックの交互作用に関して、ブロックの単純主効果から、前半のブロックよりも後半のブロックで識別成績が上昇したことが示されたのは言語音教示一致群だけであった（ $F(1,57)=15.185$, $p<.001$, 非言語音教示群, 言語音教示不一致群ともに $F<1$ ）。また、教示群×試行系列の交互作用傾向に関して、試行系列の単純主効果から、試行系列の進行にともなう識別成績の上昇を示したのは言語音教示一致群と非言語音教示群であった（言語音教示群： $F(4,228)=4.383$, $p<.01$, 言語音教示不一致群： $F<1$, 非言語音教示群： $F(4,228)=4.383$, $p<.05$ ）。

以上の結果は、正弦波アナログ音の識別学習において、学習結果に対するフィードバックが与えられない状況では、そのもとになった言語音の音韻ラベルを用いて学習する場合には一貫した学習効果が認められる

が、任意のラベルを用いたり、言語音ラベルであっても、もともとの対応とは異なるラベルで学習する場合には、十分な学習効果は得られない、ということを示す。特に、もとの言語音と矛盾する音韻ラベルを用いる場合には、全く学習効果が認められなかった。

なお、ブロック×試行系列に関して、試行系列の単純主効果は第1ブロックにおいてのみ有意であった（第1ブロック： $F(4,228)=6.766, p<.001$ ，第2ブロック： $F<1$ ）。図 2.2 から見て取れるように、これは主に言語音教示不一致群と非言語音教示群の成績を反映しているようである。すなわち、言語音教示群では、系列を繰り返すことで前半後半に限らず、ほぼ直線的に識別成績は向上しているが、言語音教示不一致群と非言語音教示群では、前半は識別成績が向上するものの、後半は成績の改善は認められなかった。

ところで、通常、正弦波アナログ音系列を用いたカテゴリー的知覚実験などで得られる識別成績は、実験 I で得られた言語音教示一致群や非言語音教示群の識別成績よりも高い値が得られている。実験終了後の両群の被験者からの報告では「ダ」と「ガ」あるいは「B」と「C」との区別が非常に難しかったとのことであつたので、刺激音別に識別学習の経過を比較した。図 2.3 は、用いられたラベル別に図 2.2 の結果を集計し直したものである。図中で同種の線が同じ刺激音を表しており、各線傍に示されたパラメーターは識別に用いられたラベルである。言語音教示不一致群については、識別の音韻ラベルとともに刺激音に正しく対応する音韻名を括弧内に記した。

図 2.3 から明らかなように、言語音教示一致群と非言語音教示群においては、アナログ音「ガ」（あるいは「C」）の識別は、他の 2 音に比較して困難であり、特に非言語音教示群では偶然レベルの識別率でしか

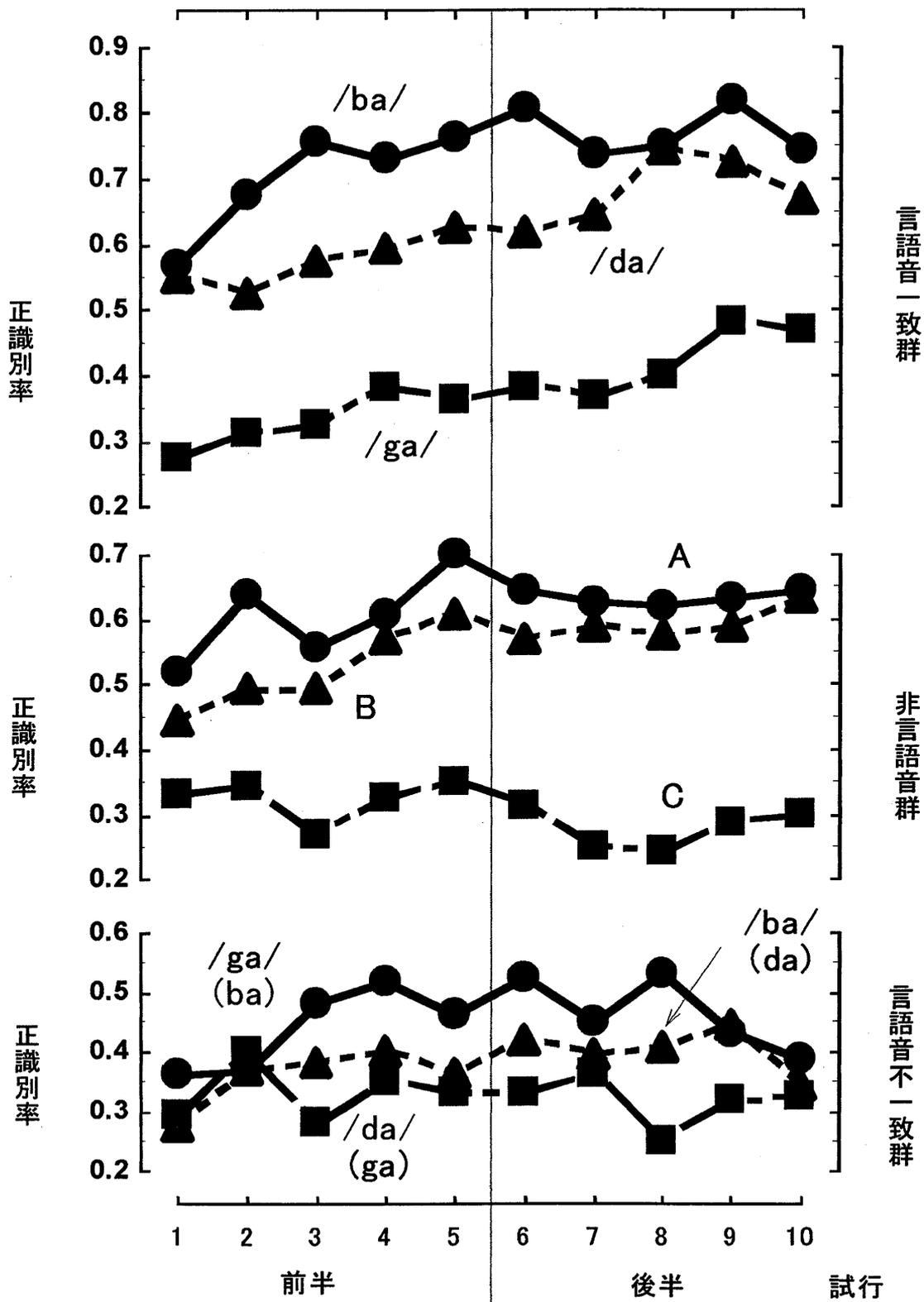


図 2.3 刺激音毎の識別学習経過

なかった。しかし、言語音教示一致群では、識別率は低いながら、ブロック、試行系列にともない成績は改善する傾向が見られた。他方、言語音教示不一致群では、学習系列を繰り返しても3つの刺激音とも偶然レベル近くの識別成績でしかなかった。これらは、アナログ音の識別のしやすさは、全体的な識別成績に影響を及ぼすものの、教示群毎に設定されたアナログ音に対するラベルの効果には影響を及ぼさないことを示唆するものである。

この点を確認するために、改めて、正弦波アナログ音毎に、教示群×ブロック×試行系列の3要因分散分析を行った。その結果、「バ」アナログにおいては、教示群 ($F(2, 57)=7.883, p<.001$), ブロック ($F(1, 57)=4.764, p<.05$) の主効果並びにブロック×試行系列の交互作用 ($F(4, 228)=6.325, p<.001$) が有意であり、試行系列 ($F(4, 228)=2.302, p<.10$) の主効果の傾向も認められた。また、「ダ」アナログにおいては、教示群 ($F(2, 57)=12.786, p<.001$), ブロック ($F(1, 57)=22.103, p<.001$), 試行系列 ($F(4, 228)=3.089, p<.05$) の主効果が有意であった。したがって、「バ」、「ダ」の両アナログ音に関しては、アナログ音に与えられるラベルの効果（すなわち教示群の効果）は維持されていたと言える。

他方、「ガ」アナログにおいては、試行系列の主効果 ($F(4, 228)=2.437, p<.05$) が有意であり、教示群×ブロックの交互作用の傾向 ($F(2, 57)=2.997, p<.10$) が認められた。交互作用についての下位分析の結果、言語音教示一致群においてのみ、ブロックの単純主効果が有意であった ($F(1, 114)=4.840, p<.05$)。(おそらく床効果 floor effect により) 全体的な識別成績では教示群の効果は検出できなかったが、識別学習の経過に教示群の違いが維持されていたと言うことができる。すなわち、「ガ」アナログにおいても教示群の効果は維持されていた。

4. 考察

正弦波アナログ音は、言語音のような調波構造も、フォルマントを中心とした広い周波数帯域へのエネルギー分布ももたないので、元の言語音とは非常に異なって聞こえる。実験 I では、この本来非言語音である正弦波アナログ音から言語音が知覚されるのは、元の言語音から引き継いだスペクトルの時間的特性を音韻関連特性として、その知覚表象と、長期記憶に保持された音韻表象との照合によるという仮説を、破裂音一母音音節のアナログ音を用いて検討した。

言語音知覚の運動理論や直接知覚理論の立場から行われてきた研究では、正弦波アナログ音を歪んだ言語音として聞かせることにより、アナログ音のもつ音韻関連特性に注意を向けさせることができ、その結果、音韻知覚が自動的、あるいは直接的に成立すると考えられている (e.g., Remez et al., 1981)。しかし、単音節言語音のアナログ音を用いた研究の多くは、安定した識別成績を得るために十分な識別訓練を行っていた (e.g., Best et al., 1989; Sawusch & Gagnon, 1995)。すなわち、アナログ音から音韻内容を知覚するためには、その音韻関連特性と音韻記憶表象との対応づけを学習しなければならないのかもしれない。もし、アナログ音の音韻知覚が、その音響特性と音韻ラベルとの単なる対連合にすぎないのであれば、アナログ音の音韻知覚はそれらの任意の組み合わせにより成立するはずである。他方、仮説で述べたように、元の言語音から引き継いだ音韻関連特性と音韻記憶表象との既存の照合関係に基づいて音韻知覚が成立するのであれば、音響特性と音韻ラベルの組み合わせはそのような関係を維持したものでなければならない。

実験では、正弦波アナログ音の音響特性とそれに与えるラベルの関係

により 2 つの言語音教示群と非言語音教示群を設けた。アナログ音を歪んだ言語音と教示する際、言語音教示一致群では、元の言語音の音韻名を用いてアナログ音を識別するよう求め、言語音教示不一致群では、元の言語音の音韻名との対応を全く変えた音韻ラベルを用いて識別するよう求めた。他方、非言語音教示群では、アナログ音を歪んだ非音楽的和音として教示し、任意のラベルを用いて識別学習を行わせた。

これら 3 教示群の遂行に関して、次のように予想された。アナログ音の音韻関連特性と音韻記憶表象との間に何らかの照合関係が存在するならば、言語音教示一致群における識別学習では、与えられた音韻ラベルはそのような関係と一致し、相対的に促進的に機能すると考えられる。しかし、言語音教示不一致群における識別学習では、音韻関連特性と音韻記憶表象との照合関係に矛盾するラベルが与えられるので、照合過程は混乱し、学習は相対的に損なわれるであろう。また、非言語音教示群は、そもそも音韻関連特性と音韻記憶表象との照合は行われず、識別学習は与えられたラベルによる対連合学習とみなされる。その意味では、識別学習に関して中立的な位置づけであり、識別成績についても、先の 2 つの教示群の中間の成績になると予想される。実験の結果は、これらの予想を支持するものであった。

さらに、この結果は、個々のアナログ音の識別のし易さには影響されなかったと考えられる。アナログ音毎に識別学習過程を検討した結果は、「バ」アナログと「ダ」アナログに関しては、教示群間の学習成績の関係は維持されていた。また、「ガ」アナログも、識別が非常に困難であったことで、床効果により教示群間の成績差は見られなかったが、言語音教示一致群においてのみ、前半ブロックよりも後半ブロックでの成績が高くなるという学習効果が認められ、その意味では教示群間の違いが

維持されていたと言える。したがって、個々のアナログ音に関する結果からも、正弦波アナログ音の識別学習において、もとの言語音に対応した音韻ラベルを用いることが、他の任意のラベルを用いるよりも有効であることが支持された。

第2節 母音アナログ音の識別学習（実験Ⅱ）

1. 目的

実験Ⅰの結果は、言語音と非言語音の知覚的差異が、言語音パターン認識過程において長期記憶内の音韻表象を参照可能か否か、あるいは入力音響信号に対する知覚表象と音韻記憶表象の照合が成立するか否かによる、ということを示唆していた。実験Ⅱは、これらの解釈の妥当性を別タイプのアナログ音を用いて確認する目的で行った。

実験Ⅰにおいて正弦波アナログ音によりそのフォルマント構造を模擬した破裂子音は、前後の音韻文脈や発話の状況などの影響で音響構造を大きく変えるため、符号化された音韻とされ、音韻知覚の特異性を例証する典型的な刺激として用いられてきた（Lieberman et al., 1967）。他方、文脈による音響的再構成が小さく、符号化に関して破裂子音の対極にあるとされるのが母音であり、その知覚は、非言語音の知覚と類似した心理物理的特性を示す（Eimas, 1963）。これまで言語音知覚の特異性を例証してきた研究の多くがこの破裂子音と母音とを対比的に用いている。例

えば、破裂音のように強く符号化された言語音はカテゴリー的に知覚され、右耳有利性（REA）を生じやすいが、母音のように文脈変動が少なく符号化の程度の弱い言語音はカテゴリー的には知覚されず、REA も生じない、ということが示されている（Eimas, 1963; Liberman et al., 1967; Studdert-Kennedy & Shankweiler, 1970）。実験 I においても、後続母音は一定に保たれており、そこで得られた結果は音韻知覚一般というよりも破裂子音に特異的なものであった可能性も考えられる。そこで、実験 I で示された結果が、破裂音や符号化された音韻の知覚に特異的なものであるのか、音韻知覚一般に普遍化できるものかを確認する必要がある。それ故、実験 II の第 1 の目的は、破裂子音－母音音節のアナログ音で得られた結果について、定常的な母音アナログ音を用いてその普遍性を確認することにあった。

また、実験 I の結果と解釈がアナログ音の音韻知覚一般に当てはまるのであれば、さらに次のことが予想される。すなわち、音韻知覚過程の最終出力は個々の言語音の音響的変動が捨象され抽象化された音韻内容であると考えられる。したがって、アナログ音の識別学習を促進するのは音韻ラベルとアナログ音に対する音韻記憶表象との対応であると仮定するならば、その対応は抽象化された表象レベルで生じているものと考えられる。実験 II の第 2 の目的は、この点を確認することにあった。すなわち、実験 I と同様の教示群の間で識別学習過程を比較する際、音響特性の異なる男声と女声の母音「ア、イ、ウ」を模した正弦波アナログ音を用い、音響特性の違いが識別学習へ及ぼす影響を調べた。

これに加えて、実験 II では課題場面に主に 2 つの変更を加えた。実験 I では、いずれの教示群においても識別成績の水準は低く、特に言語音教示不一致群で識別学習の繰り返しによる成績の改善がほとんど示され

ず、非言語音教示群においてもわずかに改善が示されたのみであった。このような低い識別成績の一因として、識別結果についてのフィードバックが与えられていなかったことがあったのかもしれない。本研究では、識別学習を促すために、識別結果について正誤のフィードバックを与えるようにした。

他方、正誤のフィードバックを与えることで、識別数を指標とした学習効果について教示群間の違いが縮小されることも考えられる。例えば、言語音教示不一致群と非言語音教示群における識別学習課題が、アナログ音の音韻関連特性と音韻記憶表象との既存の関係を利用できないという意味では同等であり、アナログ音と任意のラベルとの新たな対連合学習と見なすこともできることから、両群の結果に違いが認められなくなることも予想される。すなわち、フィードバックを与えることは、実験 I の解釈の妥当性を検討する上で重要な変更となる可能性もある。そのため、学習効果のもう一つの指標として、識別に要する反応時間を加えた。識別の正確さと速さとは、課題の難易度によってはトレードオフを生じることもあり、両者の関係は必ずしも正の相関ばかりではない。したがって、実験 I で解釈したように、言語音教示不一致群においてアナログ音の知覚表象と音韻記憶表象との照合の混乱が識別学習の成立を妨げているのであれば、正誤のフィードバックが識別数の増加に寄与するにしても、識別反応時間では非言語音教示群との違いが維持されるかもしれない。あるいは、識別学習過程について、正識別数とは質的に異なる側面をとらえることができるかもしれない。

いずれにしても、フィードバックを与えることで識別学習を促すという実験操作にもかかわらず、母音アナログの学習過程に関して教示群間で識別成績や反応時間に違いが認めらるのであれば、実験 I の解釈とそ

れに基づく本研究での仮説が強く支持されるものと考えられる。

2. 方法

2.1 刺激

刺激音は、男声と女声の合成母音「ア、イ、ウ」（高山，2002a）のフォルマント構造を，その第1フォルマント～第3フォルマントに対応する周波数変調音で模擬した3周波成分複合音（正弦波アナログ音）であった。各フォルマントに対応する成分の周波数と相対振幅を，第1成分について T1，A1，第2成分について T2，A2，第3成分について T3，A3として，それぞれの値を表2.1に示す。図2.4はそれらのスペクトログラムである。

これらアナログ音は，パーソナルコンピュータ（EPSON Direct Pro-900）と自作ソフトウェアを用い

	女 声			男 声		
	ア	イ	ウ	ア	イ	ウ
T1 (Hz)	950	500	500	790	200	250
A1 (dB)		0			0	
T2 (Hz)	1450	2930	1450	1180	2750	1180
A2 (dB)	-2	-2	-10	-2	-2	-10
T3 (Hz)		3300			3000	
A3 (dB)	-6	-6	-15	-6	-6	-15

て 11.025 kHz，16 bit の精度で合成した。そして，波形編集ソフトウェア（Syntrillium Software, Cool Edit 2000）を用いて，持続時間は 200 ms，波形包絡線の立ち上がりと減衰は 20 ms となるように編集した。

2.2 手続き

実験参加者は，予め決められたカテゴリー名に基づいて3種のアナロ

グ母音を分類する識別学習実験を個別に行った。実験Iと同様，与えられたカテゴリー名に応じて参加者は3つの教示群に分けられた。

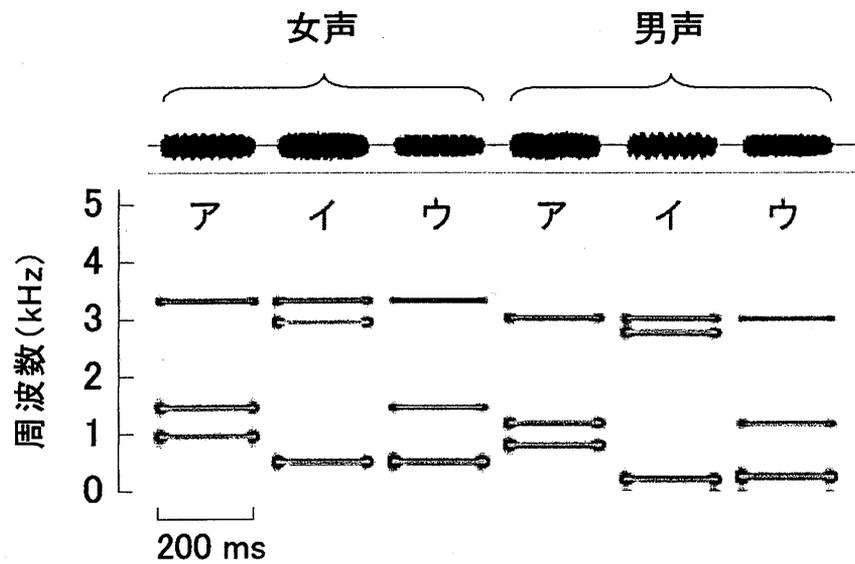


図 2.4 正弦波アナログ母音の構成

言語音教示一致群では，各アナログ音のもととなった母音名が，言語音教示不一致群では，もとの言語音と一致しないように対応を変えた母音名が，非言語音教示群では，「A」，「B」，「C」という母音とは無関連なカテゴリー名が与えられた。言語音教示不一致群においては，アナログ音に与えられる母音名は，もとの言語音の母音名に一致しないという制約のもとで，参加者間でランダムにされ，非言語音教示群においてもアナログ音へのカテゴリー名の割り当ては参加者間でランダムにされた。参加者はこれら3つの教示群にランダムに振り分けられた。

実験に先立って，言語音教示一致群と言語音教示不一致群の参加者には，3種の歪んだ人工母音を識別できるようになる過程について調べることが実験の目的であると教示し，非言語音教示群の参加者には，3種の非音楽的な和音を識別できるようになる過程について調べることが実験の目的であると教示した。さらに，3群の参加者に対して，学習すべき各カテゴリーには2つの同名異音があることも併せて教示した。

参加者は，上記の教示が与えられた後，まず学習試行として，6つの

刺激音を CRT 上に提示されるそのカテゴリー名と対応づけながら順々に 5 回繰り返して聞いた。続いて、10 系列の識別試行を行い、各系列では 6 音をカテゴリー名と対応づけながら順々に 2 回繰り返して聞いた後に、ランダムに計 5 回ずつ提示される 6 音を識別した。各試行では、CRT 上に刺激音がこれから提示される旨のメッセージが提示された後、ランダムな ISI (300 ~ 500 ms の範囲で 100 ms 間隔) で刺激音が提示された。

ここでの参加者の課題は、刺激音を聞いて、その刺激音のカテゴリー名に予め割り当てられたキーを、それぞれのキーに割り当てられた右手指でできるだけ速く正確に押すことであり、カテゴリー名のキーへの割り当ては参加者毎にランダムであった。参加者の選択したカテゴリー名と選択反応時間、並びに選択の正誤がコンピュータのディスクに記録された。なお、参加者の反応直後に、正誤のメッセージと、200 ms 以下 (尚早反応) あるいは 1000 ms 以上 (遅延反応) の反応時間を示した試行についてはそれに対する注意メッセージ (「不適切な反応時間」) が 1 s 間 CRT 上に提示された。これらのメッセージが消えた 2.5 s 後に次の試行が始められた。また、試行系列の間には少なくとも 10 s 以上の休憩時間が設けられ、被験者のペースで次の系列に進んだ。

実験は基本的にはパーソナルコンピュータ (FUJITSU FMV-6200 D7) により制御され、必要な教示の提示、刺激の提示とそのタイミング、反応の入力に対する反応時間の計測と反応の正誤のフィードバック等、予めプログラムされた手順で行われた。実験者は、実験開始に当たっての被験者プロフィールの入力、教示の補足、実験終了後の被験者の内省の記録を行った。

刺激音は、コンピュータの音源ボード (SoundBlaster 16) の出力をオーディオアンプ (Victor AX-S700)、ヘッドフォン (Panasonic RP-HT242) 経由

で被験者に煩わしさを感じさせない程度の強さで両耳に提示された。

2.3 参加者

近畿大学工学部在籍の1～4年生100名（男90名，女10名，平均年齢19.1歳）が実験に参加し，3つの教示群にランダムに割り当てられた。これらの参加者はいずれも日本語を母国語とし，実験実施上問題となるような聴力障害はなかった。参加者のうち4名（言語音教示一致群1名，非言語音教示群3名）は教示を十分に理解しておらず，各識別系列においてほとんど正しく識別できていなかったため分析から除外した。その結果，各教示群の参加者は32名となった。

なお，実験後のインタビューにおいて，言語音教示不一致群の参加者の中で，正弦波アナログ音に与えられた音韻ラベルがその音韻関連特性と矛盾することに気づいたと報告した者はおらず，非言語音教示群の参加者の中で，アナログ音を言語音として聞いて識別しようとしたと報告した者もいなかった。

3. 結果

まず，母音アナログの識別の正確さについて分析した。各教示群における母音の声質別の平均識別率を図2.5に，各群における各系列での平均識別率を図2.6に示す。識別率に関して，角変換値を用いて教示群×母音の声質×系列の3要因分散分析を行った。その結果，教示群 ($F(2,93) = 62.857, p < .001$)，声質 ($F(1,93) = 26.765, p < .001$)，系列 ($F(9,837) = 28.870, p < .001$) の主効果，並びに教示群×声質 ($F(2,93) = 3.108, p < .05$)，教示群

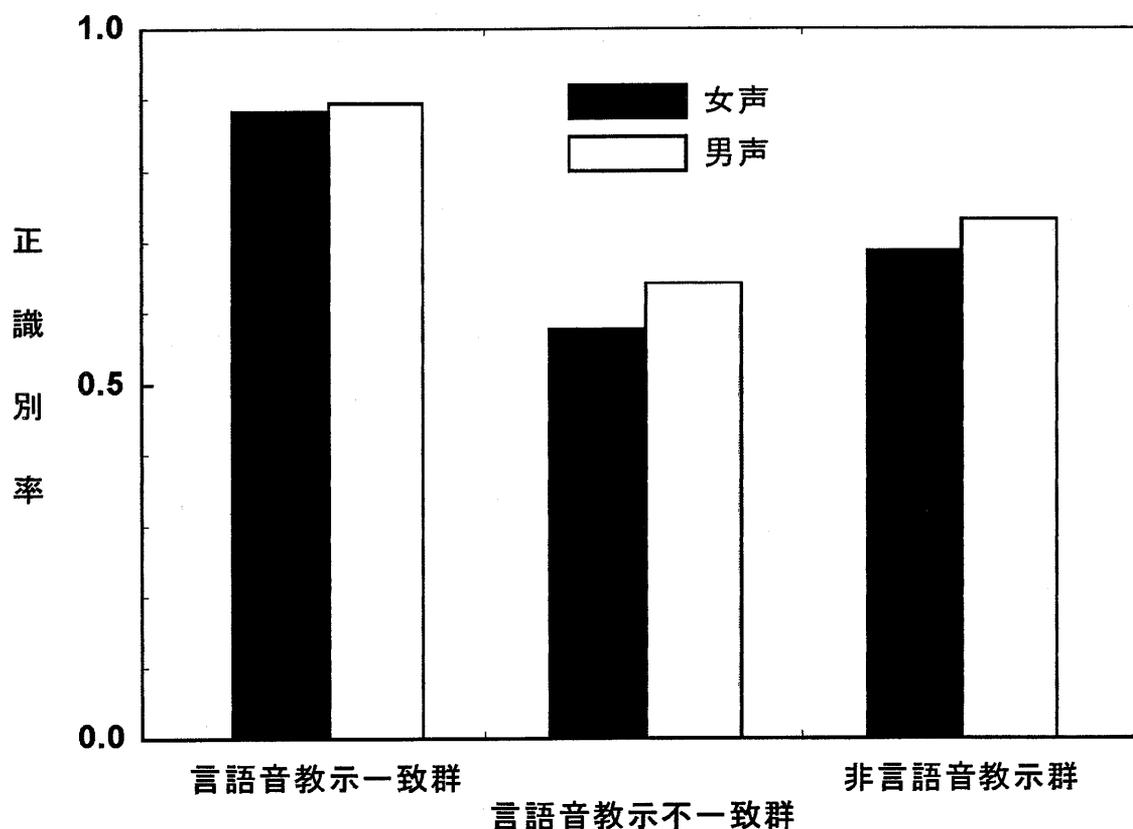


図 2.5 各教示群における声質別の正識別率

×系列 ($F(18,837) = 1.626, p < .05$) の交互作用が有意であった。

教示群 × 声質の交互作用 (図 2.5 参照) について下位分析を行ったところ、声質にかかわらず教示群の単純主効果は有意であり (女声: $F(2,186) = 65.121, p < .001$, 男声: $F(2,186) = 47.888, p < .001$), Ryan 法による教示群間での多重比較の結果は、いずれの声質においても言語音教示一致群, 非言語音教示群, 言語音教示不一致群の順で正識別率が有意に減少したことを示した ($MSE = 0.192, df = 186, p < .05$)。他方、声質の単純主効果は、言語音教示不一致群 ($F(1,93) = 20.513, p < .001$), 非言語音教示群 ($F(1,93) = 11.335, p < .01$) においてのみ有意であり、アナログ音をその元の母音名以外で識別することを学習する場合に、声質の違いが学習成績に影響していることが示された。

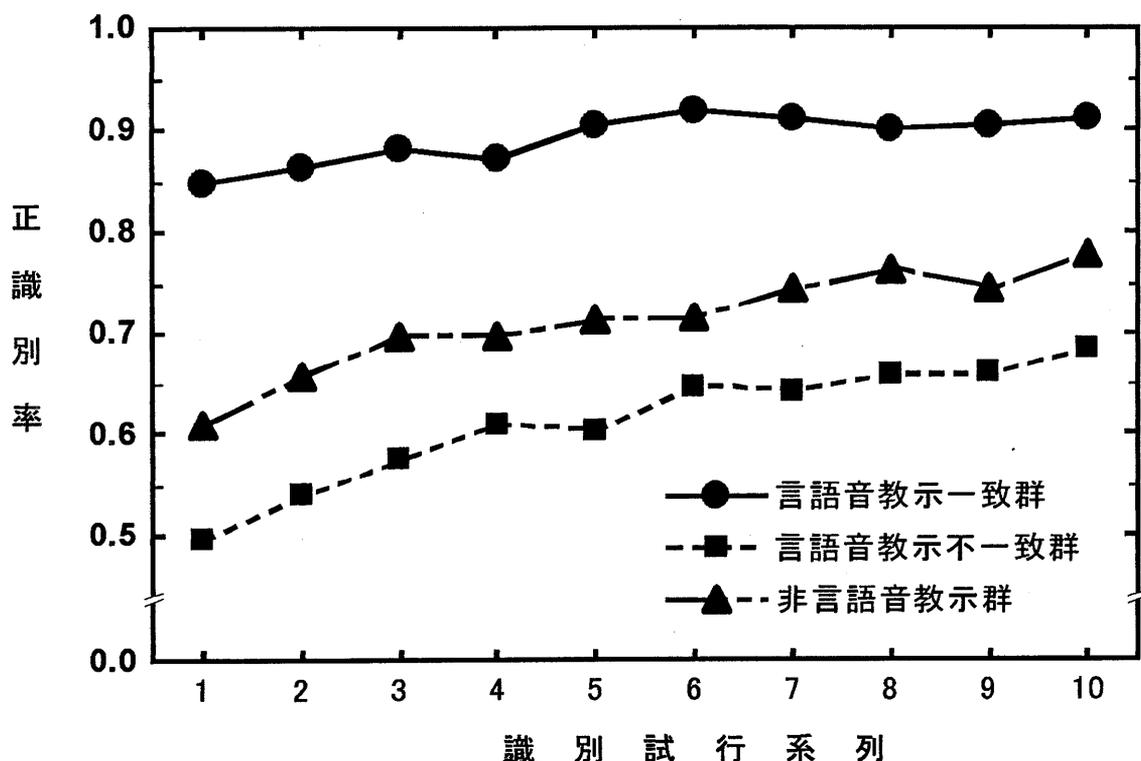


図 2.6 各教示群における識別学習にともなう正識別率の変化

教示群×系列の交互作用（図 2.6 参照）について下位分析を行ったところ、いずれの教示群においても系列の単純主効果が有意であった（言語音教示一致群： $F(9,837) = 4.550, p < .001$ ，言語音教示不一致群： $F(9,837) = 15.732, p < .001$ ，非言語音教示群： $F(9,837) = 11.851, p < .001$ ）。傾向分析の結果、各群とも 1 次の傾向が有意であり（言語音教示一致群： $F(1,837) = 29.205, p < .001$ ，言語音教示不一致群： $F(1,837) = 128.790, p < .001$ ，非言語音教示群： $F(1,837) = 97.283, p < .001$ ），さらに、言語音教示一致群と言語音教示不一致群においては 2 次の傾向も有意水準に達していた（言語音教示一致群： $F(1,837) = 6.849, p < .01$ ，言語音教示不一致群： $F(1,837) = 7.912, p < .01$ ）。

この両群について直線と 2 次曲線の当てはまりの良さを決定係数により比較したところ、言語音教示一致群では直線に対して 0.713，2 次曲

線に対して 0.880, 言語音教示不一致群で 0.910, 0.966 といずれも 2 次曲線への当てはまりが良かった。そこで, 各教示群における系列を, 1～3 系列, 4～7 系列, 8～10 系列の 3 系列群に分け, それぞれの平均正識別率の角変換値を Scheffé 法により比較したところ, 言語音教示一致群と言語音教示不一致群ではともに 1～3 系列と 4～7 系列の間と 1～3 系列と 8～10 系列の間で有意となり, また非言語音教示群では 3 つの系列群間のすべての組み合わせで有意となって ($MSE=0.020$, $df=837$, 5%水準), 傾向分析の結果を裏づけた。したがって, 両言語音教示群の識別成績は, その水準に関わらず上限に達しつつあったが, 非言語音教示群はさらに試行系列を繰り返すことで, 識別成績の改善が期待できるものと考えられる。

これらの結果にアナログ音の音響的差異が影響する可能性を検討するために, 系列を込みにして, 各教示群における各アナログ音 (女声と男声の「ア」, 「イ」, 「ウ」6 音) の識別率を比較した。図 2.7 に, 各教示群における各アナログ音の平均識別率を示した。

各被験者の各ターゲット音に対する識別率の角変換値を用いて, 教示群×アナログ音の分散分析を行った結果, 教示群 ($F(2,93)=62.845$, $p<.001$) とアナログ音 ($F(5,465)=62.904$, $p<.001$) の主効果, 教示群×アナログ音 ($F(10,465)=10.956$, $p<.001$) の交互作用いずれもが有意であった。教示群×アナログ音の交互作用について下位分析を行ったところ, いずれの教示群においてもアナログ音の単純主効果は有意であった (言語音教示一致群: $F(5,465)=3.540$, $p<.01$, 言語音教示不一致群: $F(5,465)=49.022$, $p<.001$, 非言語音教示群: $F(5,465)=32.255$, $p<.001$)。教示群毎にアナログ音の識別率を Ryan 法による多重比較 ($MSE=0.029$, $df=465$, $p<.05$) で比較したところ, 言語音教示一致群では女声「ア」と女声「ウ」の組合せでの

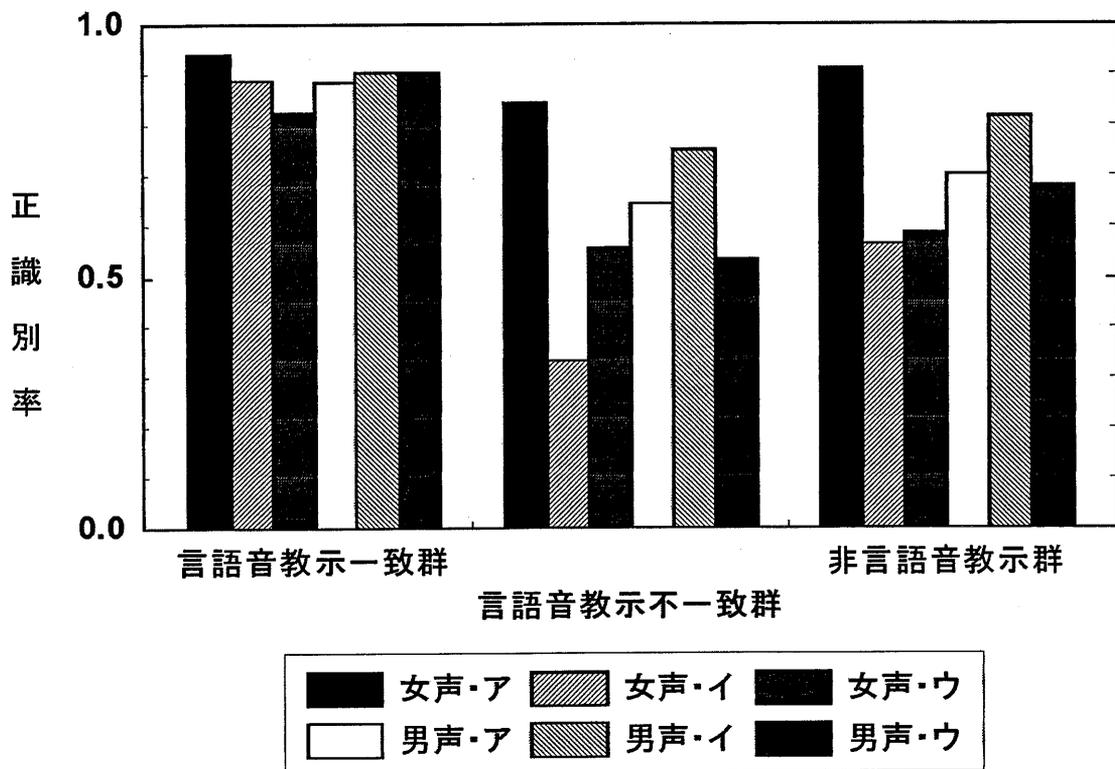


図2.7 各教示群における各アナログ音の正識別率

みアナログ音間に有意差が認められた。

他方、言語音教示不一致群では女声「ウ」と男声「ア」あるいは女声「ウ」と男声「ウ」の組合せを除くすべての組合せで、そして非言語音教示群でも男声「ア」と男声「ウ」あるいは女声「イ」と女声「ウ」の組合せを除くすべての組合せでその差が有意水準に達していた。すなわち、言語音教示不一致群や非言語音教示群では、おそらく音響的要因により、アナログ音間で識別率は大きく変動したが、言語音教示一致群ではそのような変動はほとんど生じなかった。

次に、母音アナログの識別に要した反応時間について分析を行った。まず、各被験者の各試行における識別反応時間について、200 ms以下のものを尚早反応、1000 ms以上のものを遅延反応として除いた。そして、分散分析を行う上で欠損値や外れ値が生じないように、1つ以上の系列に

おいて 2 種の声質それぞれに対して尚早反応と遅延反応を除く識別反応時間が 2 試行以上得られなかった被験者の結果を分析の対象から除外した。その結果、言語音教示一致群では 32 名、言語音教示不一致群では 24 名、非言語音教示群では 32 名の被験者の結果を以下の分析で用いた。これらの被験者の各系列における声質別の識別反応時間の幾何平均の対数変換値を用いて、教示群×母音の声質×系列の分散分析を行った結果、教示群 ($F(2,85) = 7.030, p < .01$)、系列 ($F(9,765) = 11.175, p < .001$) の各主効果、並びに教示群×系列の交互作用 ($F(18,765) = 1.689, p < .06$) が有意であった。

図 2.8 に各教示群の各系列における識別反応時間の幾何平均値を示した。教示群×系列の交互作用について、いずれの教示群においても系列の単純主効果は有意であった（言語音教示一致群： $F(9,765) = 7.53, p < .001$ 、

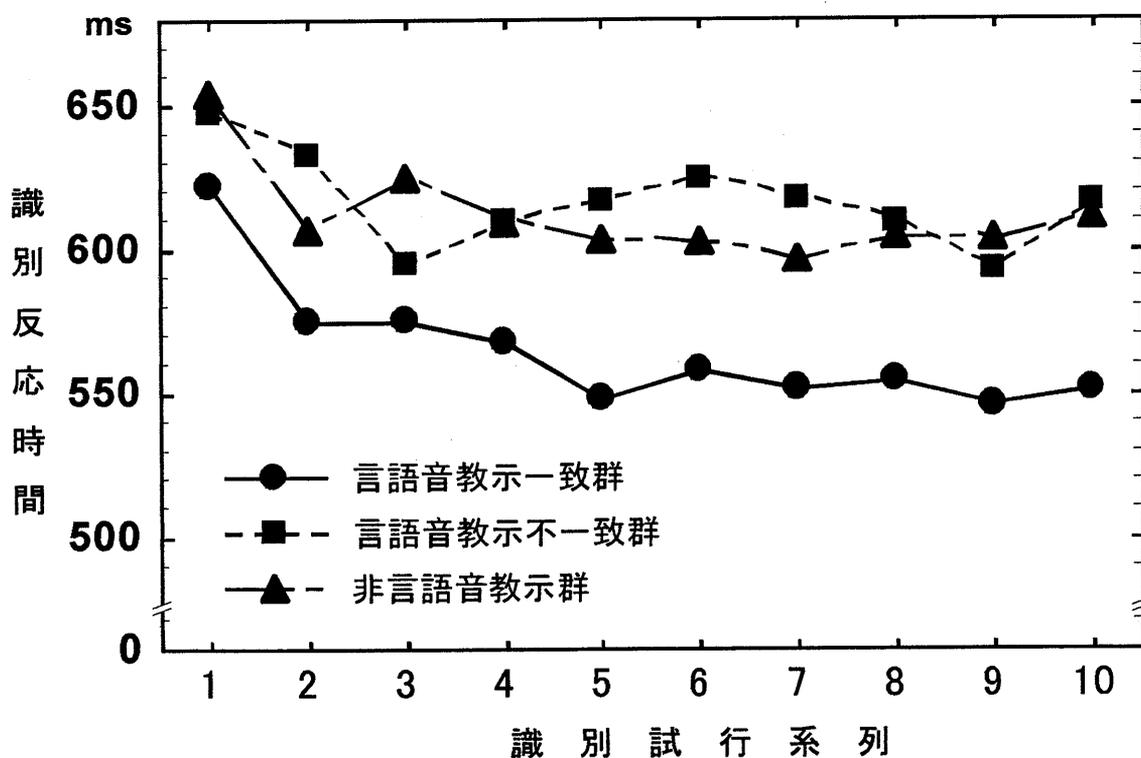


図 2.8 各教示群における識別学習にともなう反応時間の変化

言語音教示一致群： $F(9,765)=3.468, p<.001$ ，非言語音教示群 $F(9,765)=3.555, p<.001$ ）。

しかし，図 2.8 から明らかなように，この単純主効果は，第 1 系列での反応時間と比較して第 2 ないし第 3 系列以降の反応時間が大きく減少していることを反映している可能性がある。この点を確認するために，第 1 系列での反応時間に対する第 2 系列以降の反応時間の相対値について，その対数変換値を用いて改めて教示群×母音の声質×系列の分散分析を行ったところ，系列の主効果 ($F(8,680)=2.291, p<.05$)，並びに教示群×系列の交互作用 ($F(16,680)=1.743, p<.05$) が有意となった。交互作用の下位分析の結果は，言語音教示一致群 ($F(8,680)=2.289, p<.05$) と言語音教示不一致群 ($F(8,680)=2.526, p<.05$) で系列の単純主効果が有意であった。これら 2 群の系列の単純主効果について傾向分析を行ったところ，言語音教示一致群では 1 次の傾向 ($F(1,680)=14.111, p<.001$) が，言語音教示不一致群では 4 次の傾向 ($F(1,680)=14.050, p<.001$) が有意であった。そこで，両教示群における系列を，2～3 系列，4～7 系列，8～10 系列の 3 系列群に分け，それぞれの幾何平均反応時間の対数変換値を Scheffé 法により比較したところ，いずれの教示群においても有意水準に達した組合せはなく，唯一，言語音教示一致群の 2～3 系列と 8～10 系列の間で反応時間が減少する傾向が認められたのみであった ($MSE=0.002, df=680, p<.10$)。したがって，言語音教示一致群では第 2 系列以降も試行を繰り返すことでわずかながら識別に要する時間が短縮する傾向が認められたが，言語音教示不一致群では試行を繰り返しても識別反応時間は不安定であり，非言語音教示群では第 2 系列以降識別反応時間は全く短縮しなかった。すなわち，正弦波アナログ音の識別学習過程において，いずれの教示群においても，識別系列を繰り返すことで正

識別数は徐々に増加していたが、識別反応時間からその過程を検討するとき、順調に学習が進行していたといえるのは言語音教示一致群のみであった。

4. 考察

実験Ⅰの結果は、破裂子音－母音音節の正弦波アナログ音を用いて、アナログ音の音韻関連特性が適切な音韻カテゴリーと対応づけられるとき、その識別学習が達成されるが、任意のカテゴリーと対応づけられても学習はあまり進まず、誤った音韻カテゴリーと対応づけられると学習が妨げられることを示した。この結果は、音韻知覚が信号の知覚表象と既存の音韻記憶表象との照合にかかわるパターン認識過程であるとの仮説を支持するものであった。

実験Ⅱでは、実験Ⅰにおいて破裂子音－母音音節アナログに関して得られた結果を一般化するとともに、アナログ音の音韻知覚に関して提案された仮説の妥当性を検討するために、2つの声質を模擬した母音アナログを用いて、識別学習実験を行った。その際、課題場面に2つの変更を加えた。まず第1に、各識別試行において、識別結果に対するフィードバックを与えた。実験Ⅰでは、小集団で実施したこともあり、識別結果に対してフィードバックは与えられていなかったが、それが全体的な識別成績が低かったことの一因となったとも考えられる。第2の変更は、識別学習の指標として、新たに識別反応時間を追加したことであった。識別結果にフィードバックを与えることで、識別学習は促進されるものの、結果的に正識別数に関して教示条件間の差が縮小されることも予想

される。識別反応時間は、正識別数とは必ずしも正の相関を示さないの
で、正識別数ではとらえられない識別学習過程の側面をとらえることが
できるであろうと期待された。

実験の結果、予想されたように、識別結果のフィードバックにより、
いずれの教示群においても全体的な識別成績は向上していた。しかし、
教示群間の識別成績の相対的關係は、実験 I の結果を再現するものであ
った。すなわち、アナログ音が模擬する母音の名前を用いて識別を学習
する言語音教示一致群では学習の初期から識別成績が比較的高く、試行
を繰り返すことでさらに改善されたが、アナログ音ともとの母音名との
対応を変えて学習する言語音教示不一致群の成績は繰り返しとともに改
善されるものの、言語音教示一致群の識別水準には到達せず、また歪ん
だ和音として任意のラベルを用いて学習した非言語音教示群は両者の中
間の成績であった。これらの結果は、識別結果へのフィードバックの有
無は、教示条件間の遂行の差異に本質的に影響しなかったことを示すも
のである。

併せて記録された識別反応時間の結果は、3 群とも 2 ないし 3 系列の
繰り返しで反応時間は大きく短縮されるが、言語音教示一致群ではそれ
以降もわずかながら減少する傾向があることを示した。しかし、言語音
教示不一致群では、試行を繰り返すことで反応時間は大きく動揺してお
り、非言語音教示群では 2-3 系列以降は反応時間は改善しなかった。こ
れらの結果は、識別反応時間が、正識別数とは質的に異なる側面からア
ナログ音に対する識別学習過程をとらえる有効な指標であったこと示す
ものである。

加えて、もとの母音の声質に関して、言語音教示不一致群と非言語音
教示群では、声質の違いが識別成績に影響していたが、言語音教示一致

群ではそのような影響は認められなかった。さらに、アナログ音間で識別成績を比較した結果も、言語音教示不一致群と非言語音教示群ではそれぞれ 15 の組合せのうち 13 の組合せで識別成績の差が認められたが、言語音教示一致群では 1 つの組合せでしかそのような差異は認められなかった。これらのことは、言語音教示一致群における識別学習が個別のアナログ音そのものの学習というよりも、より抽象的な音韻レベルでの学習であったことを示唆している。

以上の結果から、破裂子音音節アナログで見いだされたもとの音韻名との対応による識別学習の進行と、不適切な対応による学習の妨害は、母音アナログに関しても確認され、それがアナログ音での音韻知覚一般に認められる現象であると結論できる。また、音韻知覚過程においてそのような学習と学習の妨害が生じる段階は、個別の信号に対応する知覚表象のレベルというよりも、それらを抽象化した音韻表象のレベルであると考えられる。

第 3 節 正弦波アナログ音に対するプライミング効果（実験Ⅲ）

1. 目的

実験Ⅰと実験Ⅱにおける正弦波アナログ音の識別学習の結果は、アナログ音の音韻知覚がその音韻関連特性（あるいはその知覚表象）と音韻記憶表象との照合によることを強く示唆する。しかし、この結果は、ア

ナログ音と言語音の音韻知覚が同じメカニズムによって成立することを直接支持するものではない。例えば、アナログ音の音韻知覚は音韻記憶表象との照合を媒介とするが、言語音の音韻知覚は本来直接的であるということもあり得ないことではない。

他方、これまで、正弦波アナログ音あるいはその変形を用いて言語音知覚（そして非言語音知覚）過程を検討した研究の多くは、言語音とのスペクトル的時間的構造の類似性故に、アナログ音を言語音の粗悪な代用刺激と見なし、言語音として知覚される限り、言語音と機能的に等価であり、同じように処理されると考えていた（Best et al., 1981, 1989; Johnson & Ralston, 1994; Remez et al., 1981; Tomiak et al., 1987）。実際、言語音として聞かれた正弦波アナログ音により言語音特異的とされる知覚的效果や現象が再現できることも示されており（Best et al., 1981, 1989; Johnson & Ralston, 1994; 高山, 1999; Tomiak et al., 1987）、言語音とその正弦波アナログ音の音韻知覚が同じ処理を受けて成立することが示唆される。「音韻知覚がその音韻関連特性（あるいはその知覚表象）と音韻記憶表象との照合によって成立する」という当初の仮説に基づけば、音韻知覚における両者の処理は、それぞれの信号の音韻関連特性が共通の音韻記憶表象にアクセスする（あるいはそれを活性化させる）ことにより達成されると推測される。

実験Ⅲの目的は、直接プライミング手続きを用いて、この推測の妥当性について検討することであった。プライミングでは、何らかの点でプライムする刺激が、プライムされる刺激に似ている限りにおいて、記憶内の表象を活性化させると仮定されている（Decoene, 1997; 太田, 1991）。それ故、プライムする刺激とプライムされる刺激の物理的同一性にかかわらず、言語音とそのアナログ音に共通する音韻関連特性が、音韻的

同一性に基づいてプライミング効果を及ぼすことが見いだされるならば、先の推測は支持されるものと考えられる。

したがって、実験Ⅲでは、次のような予測のもとに、アナログ音を歪んだ言語音として聞く教示のもとでのプライミング効果と、非言語音として聞く教示のもとでのプライミング効果を比較することによって、先の推測の妥当性を検討した。すなわち、音韻ラベルを用いてアナログ音を識別するよう教示された聴者は、言語音様式でアナログ音を知覚すると考えられる。その場合、言語音であろうとアナログ音であろうと、プライム刺激は音韻表象を活性化し、そして活性化された表象はターゲットの識別を促進するであろう。結果として、プライム刺激とターゲット刺激とが、物理的に同一であるか否かにかかわらず、音韻的に同一であるときプライミング効果が生じるであろう。他方、アナログ音を非言語音として識別するよう教示された聴者は、アナログ音を非言語音様式で知覚すると考えられる。その場合、言語音と非言語音アナログは相互に記憶表象を活性化させることはできないであろうから、プライミング効果が生じうるのはプライム刺激とターゲット刺激とが音響的に同一であるときに限られるであろう。

加えて、プライミング効果の時間経過についても検討した。アナログ音の音韻知覚は注意を要する過程である (Remez et al., 1981) ので、アナログ音の知覚とそのもととなる言語音の知覚とでは異なる過程を経て、音韻表象が活性化されるということはある程度である。連合プライミング実験において、本来的にターゲットと連合をもつプライム刺激は、短い提示間隔でターゲットの語彙判断を促進したが、獲得された連合によるプライミング効果はより長い間隔でのみ生じていた (Neely, 1977)。それ故、正弦波アナログによるプライミング効果が音韻表象との新たに

獲得された連合に基づくものであれば、アナログ音プライミング刺激は言語音プライミング刺激よりも時間的に遅れてターゲットの音韻識別に影響を及ぼすと期待されるであろう。他方、正弦波アナログによる音韻表象の活性化が言語音と共有されるスペクトル的時間的特性に起因するのであれば、プライミング効果の時間経過が言語音プライム刺激とアナログ音プライム刺激とで異なるとは期待できない。それ故、プライム刺激とターゲット刺激とのさまざまな時間間隔でのプライミング効果も併せて検討した。

2. 方法

2.1 刺激

ターゲット刺激として、第1セッションでは、2種の言語音をもとにした正弦波アナログ音が、第2セッションでは、2種の合成言語音「バ」/ba/と「ガ」/ga/ が用いられた。また、これら4つの音と1000 Hz正弦波純音が各セッションでのプライム刺激として用いられた。

これらの刺激音に関して、まず最初に、フォルマント音声合成ソフトウェア (Klatt, 1980; 高山, 1986) の直列合成部を用いて、言語音が合成された。合成に当たって、各音節の基本周波数 (F0)、並びに定常部の第1 (F1)、第2 (F2)、第3 (F3)、第4 (F4) および第5 (F5) フォルマントの中心周波数は、それぞれ 120 Hz, 790 Hz, 1180 Hz, 2750 Hz, 3300 Hz, 3750 Hz に固定され、また、F1 と F3 の定常部へ向かう遷移起点周波数も、それぞれ 200 Hz と 2000 Hz に固定された。2つの音節で異なるのは、F2 起点周波数と、F1, F2, F3 の起点周波数から定常周波数への遷移時間で

あり，音節「バ」では F2 起点周波数は 900 Hz，遷移時間は 30 ms，「ガ」では F2 起点周波数 2000 Hz，遷移時間 70 ms であった。

次に，言語音に対する 2 つの正弦波アナログ音が著者による音合成プログラムを用いてこれら合成言語音のフォルマント構造を模擬して構成された。それらは言語音の下から 3 つのフォルマント，すなわち F1, F2, F3, の中心周波数をたどる 3 つの周波数変調成分からなっていたが，識別を容易にするため遷移時間をわずかに長くし，「バ」で 40 ms，「ガ」で 90 ms であった。成分間の相対振幅は，言語音のフォルマントに対応させた。

これら 4 つの刺激と中立プライム刺激は，持続時間 250 ms で，波形編集ソフトウェア（Syntrillium Software, Cool Edit 2000）を用いて，その振幅包絡が最初の 20 ms で次第に立ち上がり，最後の 50 ms で徐々に減衰するよう波形編集されるとともに，すべての刺激の RMS 振幅が等しくなるよう振幅調整された。図 2.9 はこれら 2 つの合成言語音と 2 つの正弦波アナログ音のスペクトログラムを示す。

各セッションでの 5 つのプライム刺激は，これらの刺激音の音響的構造の類似性に基づいて，5 つのタイプに再分類された。同一プライム刺激はターゲット刺激と物理的に同じものであり，同種異音プライム刺激はターゲットと同じタイプではあるが異なる音，すなわち，第 1 セッションではもう一方のアナログ音であり，第 2 セッションではもう一方の言語音であった。異種同音プライム刺激はターゲット刺激に類似した周波数遷移成分をもつ音，すなわち，第 1 セッションではアナログ音ターゲット刺激の元の言語音であり，第 2 セッションでは言語音ターゲット刺激に対応する正弦波アナログ音であった。異種異音プライム刺激は刺激のタイプとフォルマント構造の両方でターゲット刺激と異なってい

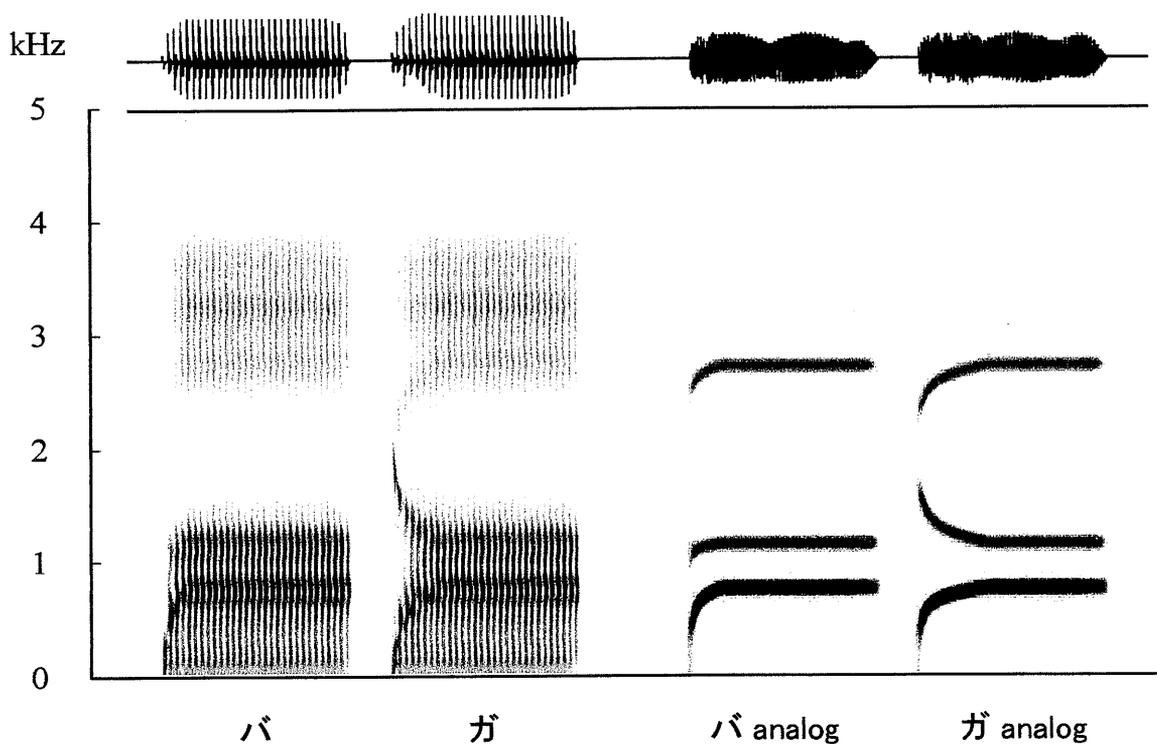


図 2.9 言語音刺激とアナログ音刺激のスペクトログラム

た。また、中立プライム刺激は 1000 Hz の純音であった。

すべての刺激はデジタル形式でコンピュータディスクに保存された。実験中、刺激はパーソナルコンピュータ（NEC PC-9801 VM2）に接続された D/A コンバーター（CANOPUS DAC-98）により、16 ビット、10 kHz で出力され、約 4.5 kHz の低域通過フィルター（エヌエフ回路設計ブロック抵抗同調フィルタ RT-8FLA2 -135 dB/oct.を用いて自作）、オーディオアンプ（Victor AX-S700）を經由し、ステレオヘッドフォン（Panasonic RP-HT242）を通して、煩わしくない程度の強さで再生された。

2.2 手続き

参加者は、ターゲットとして異なるタイプの刺激を用いた 2 つのセッションからなる識別プライミング実験に個別に参加した。ターゲットは、

第1セッションでは2つの正弦波アナログ音，第2セッションでは2つの合成言語音であった。これらのセッションは，各参加者の希望に応じて，同じ日か別の日に行われた。

各セッションの初めに，参加者は実験中に聞く音の性質について告げられた。第1セッションにおいて，参加者は言語音教示条件か非言語音教示条件のいずれかにランダムに割り当てられた。言語音教示条件の聴者は，2つの歪んだ合成言語音「バ」と「ガ」のどちらかを聞くと告げられ，他方，非言語音教示条件の聴者には，AとBと名づけられた2つのコンピュータ合成音のどちらかを聞くと告げられた。各群の聴者は，割り当てられた名前を用いてターゲットを識別するよう教示された。第2セッションにおいては，すべての参加者が合成言語音を「バ」あるいは「ガ」と識別するよう求められた。両セッションのこの段階では，2つのターゲット音が順々に5回ずつ1.5 s間隔で提示され，併せて聴者の前方に置かれたCRTモニター上にその名前が提示された。

続いて，識別練習が行われた。2つのターゲット音は，1ブロック10試行の中でランダムに5回ずつ提示された。各試行においては，まず100 msの1000 Hz純音が提示され，その500 ms後にいずれかのターゲット音が提示された。聴者は，ターゲットそれぞれに予め割り当てられたコンピュータのテンキーパッド上のキーをできるだけ速く正確に押すことによって，ターゲットを識別した。聴者が反応した後，その反応と正誤のフィードバックがCRT上に1000 ms提示された。反応キーへのターゲットの割り当ては試行の間中CRT上に示されていた。練習試行は，1ブロック内で100%の正識別に到達するまで，20ブロック以内で繰り返された。20ブロック内に100%の基準に到達しなかった聴者は，次の試行には進むことはできなかった。

識別プライミング段階では、各ターゲット音は、50, 100, 200, 400, 800, 1600ms のいずれかの試行間間隔 (ISI) で 5 つのプライム刺激の 1 つに続いて提示された。参加者は、ターゲット音がさまざまなタイミングで 5 つの合図音の 1 つの後に続いて提示され、ターゲットに対応する反応キーをできるだけ速く正確に押すことによって、その音を識別しなければならない、と教示された。識別時間がターゲット刺激の開始時から測定された。

2 つのターゲット、5 つのプライム刺激、6 つの ISI を組み合わせた 60 試行が 5 ブロック繰り返された。ブロック内の試行はランダム順であり、第 1 ブロックにおいてのみ、反応後、聴者にフィードバックが与えられた。各ブロックにおいては、ターゲットを正しく識別できなかった試行、あるいは刺激開始後 1500 ms 以内に反応できなかった試行がエラーおよび遅延反応として記録され、ブロックの終わりに繰り返された。ブロックの間とブロックの間では短い休憩時間を設けた。

各セッションにおいて、すべての試行後、被験者は音の聞こえや課題の困難さについての印象を尋ねられた。

2.3 参加者

近畿大学工学部の学生 15 名 (男 11 名, 女 4 名) が実験に参加した。すべての参加者は日本語を母国語とし、言語障害や聴覚障害はないと自己申告していた。また、いずれの参加者も、この実験以前に、合成言語音や正弦波アナログ音を用いた実験に参加したことはなかった。

彼らは第 1 セッションにおいて言語音教示群と非言語音教示群にランダムに割り当てられた。しかし、言語音群で 2 名、非言語音群で 1 名の計 3 名の参加者は、第 1 セッションでの練習試行において基準に到達で

きなかったので、それ以降の試行には進めなかった。したがって、各群6名（男4名，女2名）の参加者が残った。また，実験後のインタビューで，非言語音教示群の参加者の中で，正弦波アナログ音を言語音として聞いて識別しようとしたと報告した者はいなかった。

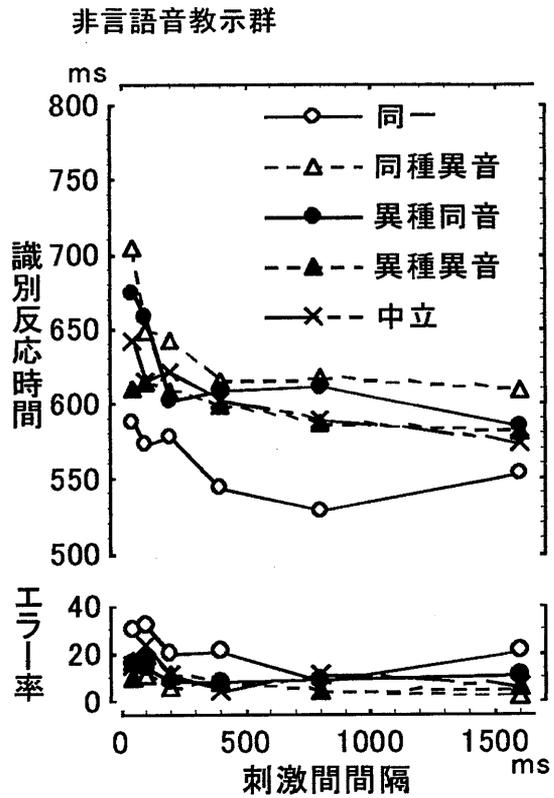
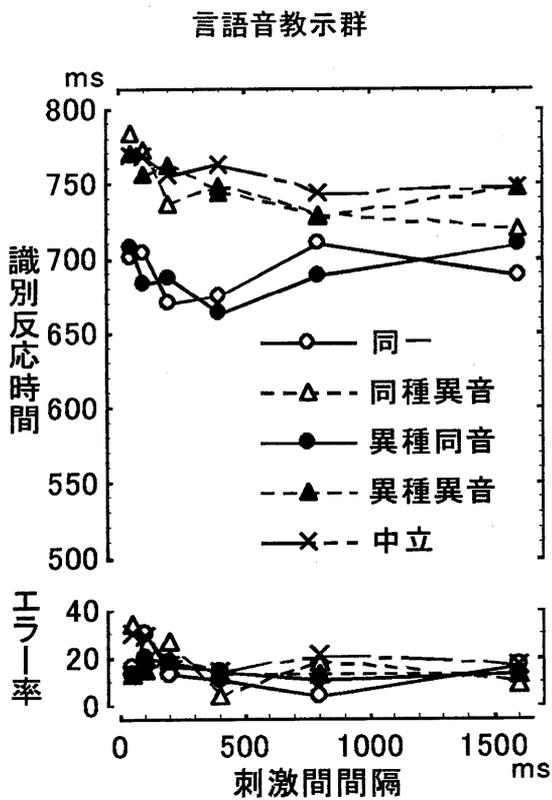
3. 結果

各聴者に関して，第1ブロックを除くすべての試行から，5つのタイプのプライミング刺激と6つのISIに関する識別反応時間の中央値とエラー率が計算された。図2.10は，各ターゲットセッションにおいて，参加者を込みにした各タイプのプライミング刺激に対する識別反応時間の幾何平均とエラー率を示す。第1セッションでのアナログターゲットが第2セッションでの言語音ターゲットより識別が困難であったことを除き，エラー率に関して統計的に一貫した傾向は認められなかった。それ故，以下の分析では，識別反応時間データのみを用いた。

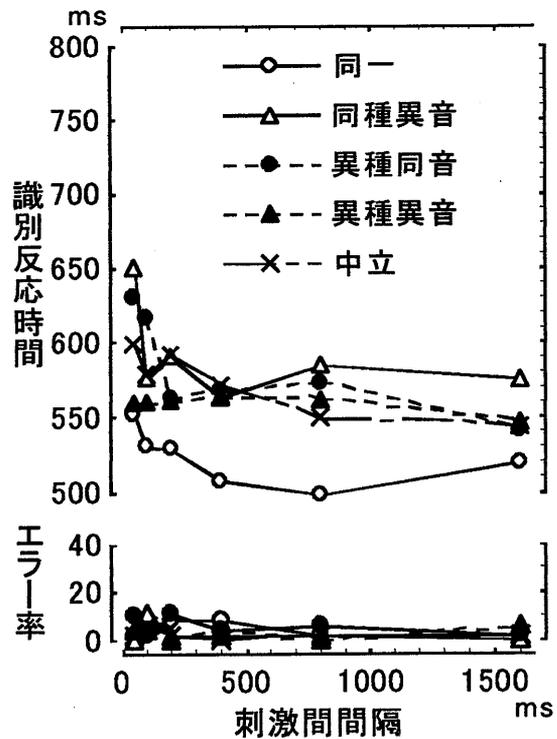
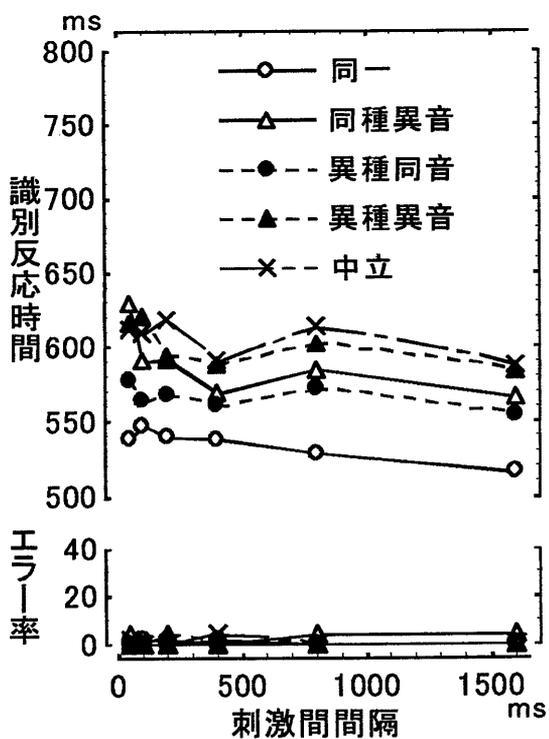
3.1 第1セッションにおけるアナログ音ターゲットに対する識別反応時間

言語音教示，非言語音教示におけるアナログ音ターゲットの識別反応時間にプライム刺激のタイプが及ぼす効果を検討するために，識別反応時間の中央値の対数変換値に関して，教示群×プライム刺激×ISIの3要因分散分析を行った。

教示群 ($F(1, 10) = 5.585, p < .05$) とプライム刺激 ($F(4, 40) = 15.015, p < .001$) の主効果が有意であり，またそれらの交互作用の傾向が認められた ($F(4, 40) = 2.262, p < .10$)。教示群の単純主効果はすべてのプライム刺激に



(a) アナログ音ターゲット



(b) 言語音ターゲット

図 2.10 各ターゲットセッション、各プライム刺激条件における識別反応時間とエラー率

関して有意かその傾向があり（同一プライム： $F(1, 50) = 6.682, p < .05$ ，同種異音プライム： $F(1, 50) = 5.850, p < .05$ ，異種同音プライム： $F(1, 50) = 3.100, p < .10$ ，異種異音： $F(1, 50) = 5.856, p < .05$ ，中立プライム： $F(1, 50) = 5.843, p < .05$ ），このことは，プライム刺激にかかわらず，非言語音群よりも言語音群で識別反応時間が長かったことを示していた。

両教示群においてプライム刺激の単純主効果が有意であった（言語音教示群： $F(4, 40) = 9.154, p < .001$ ，非言語音群： $F(4, 40) = 8.123, p < .001$ ）。どのプライム刺激に対してプライミング効果が生じたのかを調べるために，各教示群における中立プライムと他のプライム刺激との間で多重 t テストを用いた事前比較を行った。言語音教示群では，中立プライムと比較して，同一プライムと異種同音プライムがアナログ音ターゲットの識別時間を短縮した（同一プライム： $t(40) = 4.158, p < .001$ ，異種同音プライム： $t(40) = 4.263, p < .001$ ）。他方，非言語音群では，同一プライム刺激での識別時間のみが中立プライム刺激での識別時間よりも短かい，ということが示された（ $t(40) = 4.954, p < .001$ ）。

加えて，ISIの主効果が有意であった。TukeyのHSDテストによる多重比較の結果（5%水準）は，ISI 50 msの条件では200 msや400 ms条件においてより，ターゲットの識別に時間がかかった，ということを示している。しかし，ISIと他の要因，すなわち教示群とプライム刺激，との交互作用は有意水準には達しなかった。したがって，この効果は短い刺激間隔における一般的な反応の抑制によるように思える。すなわち，プライム刺激とターゲット刺激との時間間隔は，プライミング効果に影響しなかったと言える。

したがって，アナログ音ターゲットに関しては，言語音教示のもとでは，アナログ音であろうと言語音であろうと，音韻的に同じプライム刺

激に先行されるターゲットに対してプライミング効果が生じたが、非言語音教示のもとでは、物理的に同一のプライム刺激によって先行されるターゲットに対してのみプライミング効果が得られた。また、これらの効果は、プライム刺激とターゲット刺激との時間間隔には左右されなかった。

3.2 第2セッションにおける言語音ターゲットに対する識別反応時間

言語音ターゲットの識別反応時間にプライム刺激のタイプが及ぼす効果を検討するために、識別反応時間の中央値の対数変換値に関して、第1セッションでのアナログ音に関する教示群×プライム刺激×ISIの3要因分散分析を行った。

プライム刺激の主効果 ($F(4, 40) = 14.861, p < .001$) と、教示群とプライム刺激との交互作用 ($F(4, 40) = 3.622, p < .05$) が有意であった。プライム刺激の単純主効果は両教示群において有意であった(言語音教示群: $F(4, 40) = 9.546, p < .001$, 非言語音教示群: $F(4, 40) = 8.937, p < .001$)。中立プライム条件と他のプライム条件との事前比較は、両教示群にとも、物理的に同じプライム刺激に先行されるターゲットは、中立プライムに先行されるターゲットよりも速く識別された(言語音教示群: $t(40) = 5.316, p < .001$, 非言語音教示群: $t(40) = 5.613, p < .001$)。さらに言語音教示群では、異種同音プライムにより先行されるターゲットも中立プライムより速く識別される ($t(40) = 2.862, p < .01$) ということを示され、アナログ音プライムは、それらが言語音様に知覚される限り、言語音ターゲットの識別をプライムできたと言える。

また、ISIの主効果が有意であった ($F(5, 50) = 7.385, p < .001$)。TukeyのHSDテストによる多重比較の結果(5%水準)、相対的に短いISIにおいては

長い ISI においてよりもターゲットを識別するのに要する時間が長いことが示された (50 ms vs. 400 ms, 50 ms vs. 800 ms, 50 ms vs. 1600 ms, 100 ms vs. 1600 ms のそれぞれの比較において $p < .05$)。しかし、第 1セッション同様、この要因は他のいずれの要因とも交互作用がなく、プライミング効果の大きさは ISI には影響されなかった。

以上の結果から、正弦波アナログ音とその元の言語音とは、聴者がアナログ音を音韻的に知覚する限り、互いにプライミング効果を及ぼし合う、と結論づけられる。他方、アナログ音が音韻的に知覚されなければ、ターゲットとプライムが物理的に同一であるときのみプライミング効果が生じる。そして、これらの効果はプライムとターゲットとの時間間隔とは独立であった。

3.3 プライミング効果の大きさ

先の分析から、アナログ音であれ言語音であれ、音韻的に同じである限り、相互にプライミング効果を及ぼすことが示された。そこで、さらに、両者のプライミング効果の等価性を確認するために、聴者毎に、プライミング効果の指標を算出し、中立プライム刺激を除く 4 つのプライム刺激に対するプライミング効果の大きさを、アナログ音に対する教示群間並びにターゲット刺激間で比較した。下記にここで用いたプライミング効果の指標の算出式を示す。

$$\text{プライミング効果} = -\log \frac{\text{当該プライム刺激先行時の RT}}{\text{中立プライム刺激先行時の RT}}$$

この式に示されるように、用いた指標は、正のプライミング効果が正の値となるように、各プライム刺激条件での識別反応時間と中立プライム条件での反応時間との比を対数変換することにより求められた。その

際、各プライム刺激条件での反応時間は、先の分析において効果の認められなかった ISI については幾何平均値を用いて込みにされた。これらの指標について、第 1 セッションでのアナログ音に関する教示群×ターゲット刺激タイプ×プライム刺激タイプの 3 要因分散分析を行った。

その結果、教示群 ($F(1,10)=8.980, p<.05$)、プライム刺激 ($F(3,30)=27.213, p<.001$)、および両者の交互作用 ($F(3,30)=4.261, p<.05$) が有意であった。すなわち、プライミング効果の大きさは、アナログ音の性質に関して与えられた教示とプライム刺激とにより左右されたが、ターゲット刺激の種別による影響は認められなかった。

図 2.11 は、ターゲット刺激を込みにして、プライム刺激条件毎に教示群間でプライミング効果の大きさを比較したものである。図中の縦の区間表示は、95% 信頼区間を表している。各プライム刺激条件において、教示群の単純主効果が有意水準に達したのは、同種異音プライム ($F(1,10)=4.600, p<.05$) と異種同音プライム ($F(1,10)=18.445, p<.001$) であった。異種同音プライムに関しては、先の分析において、ターゲット刺激にかかわらず、言語音教示群で有意なプライミング効果が得られ、非言語音教示群ではプライミング効果は有意水準に達しなかったことから、当然予想された結果である。

他方、同種異音プライムにおける教示群間の違いは、プライミング効果が言語音教示群では相対的に促進的（正の方向）であったが、非言語音教示群では抑制的（負の方向）であったことによる。しかし、両者ともプライミング効果自体が有意水準には達していないので、特に意味のある結果ではないかもしれない。ちなみに、ターゲット音別にこの部分のプライミング効果を比較したところ、言語音ターゲットに関する結果がこの単純主効果に強く反映されており、特に第 1 セッションでアナロ

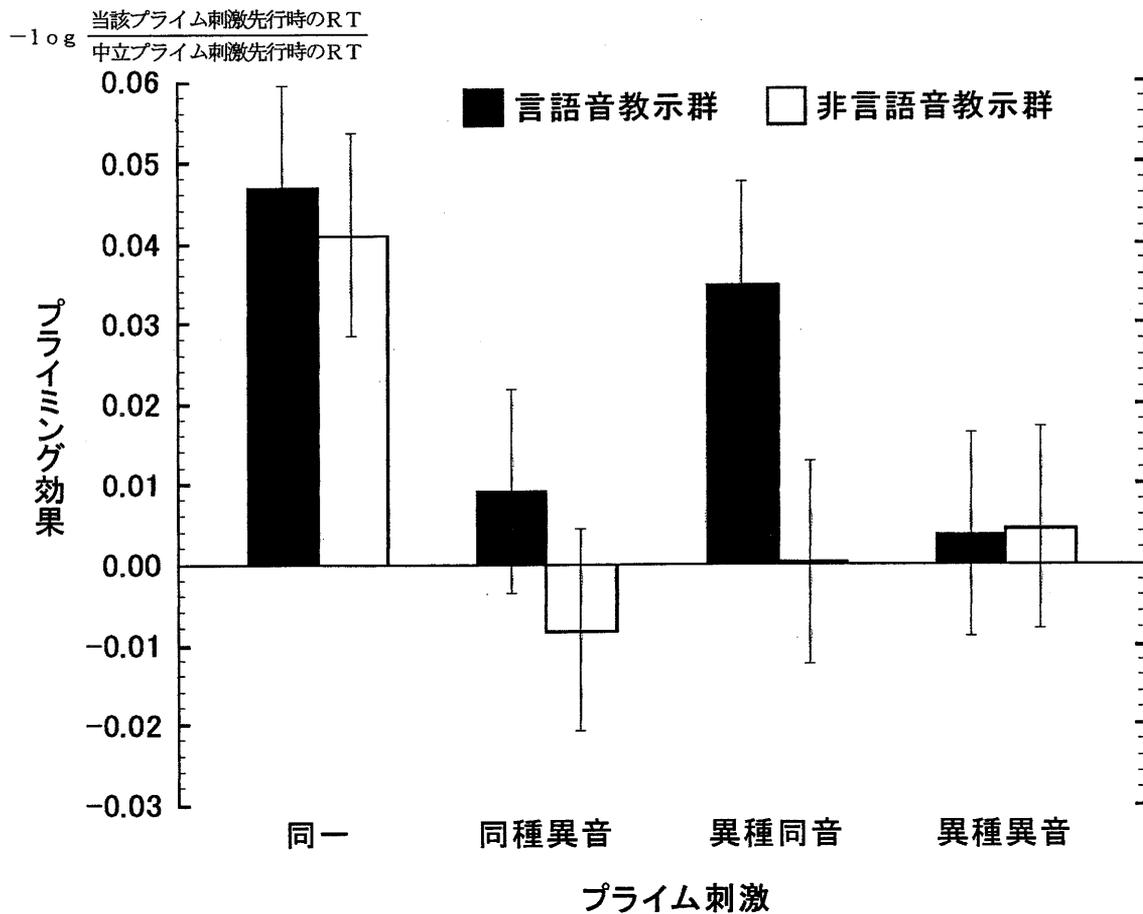


図2.11 言語音教示群と非言語音教示群におけるプライミング効果
(区間の表示は95%信頼区間を示す)

グ音を非言語音として教示された参加者において、言語音プライムに続く言語音ターゲットの識別で抑制傾向が認められた。

また、セッション別に行った先の分析からも明らかなように、各教示群におけるプライム刺激の単純主効果も有意であった（言語音教示群： $F(3,30) = 14.923, p < .001$ ，非言語音教示群： $F(3,30) = 16.551, p < .001$ ）。ここで特に関心があったのは、言語音教示群における同一プライム刺激と異種同音プライム刺激でプライミング効果の大きさに違いがあるか否か、という点であった。TukeyのHSDテストによる多重比較（ $MSE = 0.0003, df = 30, 5\%$ 水準）の結果は、言語音教示群において、同一プライムと異種

同音プライムでのプライミング効果が、他の2つのプライム刺激での効果より有意に大きいことを示した。しかし、同一プライムと異種同音プライムの違いは認められなかった。すなわち、物理的に同一であるか否かにかかわらず、プライム刺激とターゲット刺激が音韻的に同一であれば、同じ程度のプライミング効果を得ることができた。他方、セッション別の分析結果から明らかのように、非言語音教示群では同一プライムでのプライミング効果が他のプライム刺激での効果よりも有意に大きかった(5%水準)。

4. 考察

実験Ⅲは、音韻知覚が、言語音信号の知覚表象が長期記憶内の音韻表象と照合される過程であり、正弦波アナログ音の音韻知覚は、言語音と共通の音韻記憶表象にアクセスし(あるいはそれを活性化させ)、それと照合されることにより成立する、という推測の妥当性を検討するために行われた。実験では、正弦波アナログ音とその元の言語音の知覚における直接プライミング効果が、アナログ音を歪んだ言語音として聞くように求める言語音教示と、歪んだ和音として聞くように求める非言語音教示の間で比較された。

プライミングでは、プライムする刺激が、プライムされる刺激と何らかの点で似ている限りにおいて、記憶内の表象を活性化させる、と仮定されており、プライミング効果は表象レベルでの促進によるものとされる(Decoene, 1997)。それ故、言語音とその正弦波アナログ音との間に相互的にプライミング効果が認められるとすれば、両者で何らかの表象の

共有（あるいは、共通の表象へのアクセス）を意味することになる。また、当初の仮説が述べるように、言語音知覚と非言語音知覚との差異が、長期記憶内の音韻記憶表象を参照できるか否かによるとするならば、このような相互プライミングは言語音知覚様式のもとでのみ成立するであろう。

実験の結果は、参加者が正弦波アナログ音を音韻的に知覚するように求められるとき、ターゲットが正弦波アナログ音かそのもとの言語音かにかかわらず、またプライム刺激がターゲット刺激と物理的に同一か否かにかかわらず、ターゲットの識別は音韻的に同じプライム刺激により速められる、ということを示した。しかし、アナログ音が非言語音教示のもとで知覚される場合には、ターゲットのタイプにかかわらず、ターゲットとプライムとが物理的に同一であるときにのみ、プライミング効果が得られた。これらは、言語音教示によりもたらされた知覚様式と非言語音教示によりもたらされた知覚様式とでは、アナログ音を識別するメカニズムが異なっており、言語音知覚様式のもとでは、アナログ音と言語音とが共通の音韻記憶表象にアクセスする、という仮説を支持する。

加えて、ターゲット刺激が言語音であろうとアナログ音であろうと、プライムとターゲットとの時間間隔はプライミング効果に影響せず、また、その効果の大きさも、ターゲット刺激の種別には左右されなかったという結果は、アナログ音の音韻知覚が言語音と同じメカニズムにより成立していることを示唆するものである。

他方、同種異音プライムにおいて、プライミング効果の大きさ自体は有意水準には達しなかったものの、アナログ音に関する言語音教示群と非言語音教示群との間で効果の方向が逆になり、結果的に効果の相対値に有意差が認められた。この教示群間でのプライミング効果の方向の違

いは何によるものかを調べるために、アナログ音ターゲットと言語音ターゲットそれぞれについて、同種異音プライム条件におけるプライミング効果を調べたところ、言語音ターゲットの識別における両群の違いが反映されたものであった。しかし、このことが、ターゲット識別には直接関係しないものの、先行刺激として言語音と非言語音が提示される状況で、同種の言語音プライムが抑制的な影響を及ぼすことを示すものであるのか、音質にかかわらずいずれもが言語音として、音韻知覚にとっての準備的活性化をもたらす中で、音質の類似性による促進的影響を示すものであるのか、あるいはまた偶然の結果に過ぎないのか、ということについては、現段階では判断できない。

第3章 総合的考察と将来の課題

第1節 本研究の成果と意義

1. 実験結果の概要

本研究は、音韻知覚に関する情報处理的アプローチの立場に立ち、「音韻知覚は、音響信号に関する知覚表象と長期記憶に保持された音韻表象との照合により成立するものであり、言語音知覚と非言語音知覚の差異は音韻記憶表象を参照できるか否かによる」という仮説を、正弦波アナログ音を用いて検討することを目的とした。これらの仮説を検証するために、3つの実験を行った。

正弦波アナログ音に対して安定した音韻知覚を得るためには、識別を十分に訓練する必要があると知られている (e.g., Best et al., 1989; Sawusch & Gagnon, 1995)。実験Ⅰと実験Ⅱでは、この識別学習が、任意のラベルを用いて成立するような性質のものであるのか、音韻関連特性と音韻記憶表象との照合関係に基づいて成立しているのかを明らかにするために、アナログ音に与える識別ラベルを操作することによって、アナログ音の識別水準がどのような影響を受けるかについて、破裂音-母音音節のアナログ音 (実験Ⅰ) と単独母音のアナログ音 (実験Ⅱ) に関する識別学習実験において検討した。上記の仮説が正しいとするならば、アナログ音がもとの言語音の音韻ラベルを用いて識別できるようになるのは、音響信号とそれに与えられた識別ラベルとの単なる対連合学習によるのではない。そこには、もとの言語音の音韻関連特性とそれに対する長期記憶内の音韻表象との照合関係が引き継がれていると考えられる。

すなわち、アナログ音の音韻関連特性と長期記憶内の音韻表象との間に照合関係が成立している故に、アナログ音についても、その元の言語音と同様に、音韻知覚が可能になると考えられる。

したがって、アナログ音を言語音として教示し、学習すべき音韻ラベルがアナログ音のもととなった言語音の音韻名に一致するときには（言語音教示一致群）、アナログ音に内在する音韻関連特性と長期記憶内の音韻表象との間に既に成立している照合関係を利用できるので、識別学習は順調に進行すると予想される。しかし、与えられた音韻ラベルがもとの言語音と一致しない場合には（言語音教示不一致群）、学習すべき音韻ラベルが、音韻関連特性と音韻記憶表象との既存の照合関係と矛盾するために、識別学習は困難になるであろう。他方、アナログ音を非言語音として、その識別に任意のラベルが用いられるときには（非言語音教示群）、識別学習は信号とラベルとの対応づけに関する通常の対連合学習と考えることができ、識別水準は先の 2 つの間になると予想される。

実験の結果は、これらの予想を支持するものであった。実験Ⅰでは識別結果に対してフィードバックを行わず、実験Ⅱではフィードバックを行ったので、両実験の間で、全体的な識別成績の水準には違いがあったが、アナログ音に与えられるラベルに応じた相対的な識別成績には本質的な違いはなかった。すなわち、適切なラベルが与えられた場合、学習系列の初めから他のラベルが与えられた場合よりも高い識別成績を示し、系列を繰り返すことにより、さらに識別成績は改善された。しかし、不適切なラベルが与えられた場合には、ある程度の識別成績の改善は示されるものの、学習開始時から最も低い識別水準が維持されたままであった。また、無関連なラベルが与えられた場合には、適切なラベルが与

えられた場合ほどの識別成績には達しなかったが、識別系列の繰り返しにより識別成績は徐々に改善された。

実験Ⅱでは、識別学習の指標として、識別率に加えて、識別反応時間が測定された。学習系列の繰り返しにともなう識別反応時間の推移に関して、もとの言語音に対応した適切な音韻ラベルが与えられた場合には、識別反応時間は徐々に減少する傾向があった。しかし、もとの言語音と矛盾した不適切な音韻ラベルが与えられた場合には、識別系列を繰り返すと、識別反応時間は大きく動揺した。したがって、識別反応時間の結果も、与えられた音韻ラベルが音韻関連特性に対応した音韻記憶表象と一致する場合、識別反応の正確さとともに速さに関してもアナログ音の音韻識別学習は順調に進んだことを示す。他方、音韻ラベルが音韻記憶表象と矛盾する場合は、識別反応の正確さの側面では識別学習はある程度進むものの、速やかさの側面では学習効果は得られなかった。すなわち、フィードバックが与えられ、不適切なラベルでの識別が強いられる状況では、聴者がそのラベルによる識別反応を学習することは可能であるものの、おそらくアナログ音と本来的に対応した音韻記憶表象と強制された音韻ラベルとが競合するために、識別反応は注意を要する課題となり、正しい識別反応が出力されるタイミングが動揺したと考えられる。

実験Ⅰと実験Ⅱでの結果は、厳密な意味で、言語音とそのアナログ音の音韻知覚が同一のメカニズムにより成立する、ということをサポートするものではない。しかし、多くの研究において、言語音として聞かれた、あるいは、訓練の結果として聞くことが可能となったアナログ音に関して、カテゴリー的知覚 (Best et al., 1989; Johnson & Ralston, 1994) や音声トレーニング (Best et al., 1981) など言語音特異的とされている知覚的效果や現象が再現できることを示しており、両者が同じ音韻知覚メカニズム

により処理されることが示唆される。先の仮説に基づけば、言語音とアナログ音が共通の音韻関連特性を有しているが故に、共通の音韻記憶表象と照合されることになる。

そこで、実験Ⅲでは、その点を明らかにするために、音韻知覚において、言語音とそのアナログ音とが共通の音韻記憶表象にアクセスしているか否か（あるいは共通の音韻記憶表象を活性化させるか否か）を、言語音とアナログ音の間の相互プライミング実験によって検討した。その際、言語音とアナログ音とのプライミング効果の時間経過を比較するために、プライミング刺激からターゲット刺激までの時間間隔を6段階設定した。アナログ音の音韻知覚は、注意要請的な課題であるので、その意味で言語音の音韻知覚とは異なる過程を経て成立するかもしれない。その場合には、プライミング効果の時間経過が言語音とは異なると予想される。しかし、上記の仮説が正しければ、アナログ音が言語音として知覚されるとき（言語音教示群）、ターゲット刺激とプライム刺激として言語音とアナログ音がどのような組み合わせで用いられたとしても、プライム刺激がターゲット刺激と同じ音韻であれば、プライム刺激により活性化された音韻記憶表象がターゲット刺激の音韻知覚を促進し、正のプライミング効果が生じると予想される。他方、アナログ音が非言語音として知覚されるときには（非言語音教示群）、ターゲット刺激とプライム刺激が物理的に同一であるときにのみ、プライミング効果（反復効果）が生じる。さらに、これらのことは、プライム刺激とターゲット刺激の提示時間間隔には無関係に生じると予想される。

実験の結果は、これらの予想を支持するものであった。アナログ音が言語音として知覚されているとき、ターゲット刺激がアナログ音（第1セッション）であろうと言語音（第2セッション）であろうと、プライ

ム刺激がターゲット刺激と同じ音韻名であれば、プライム刺激の物理的同一性にかかわらず、また刺激提示間隔にかかわらず、プライミング効果が得られた。他方、アナログ音が非言語音として知覚されているときには、プライミング効果は、刺激提示間隔の影響を受けないものの、プライムとターゲットが同一刺激であるときにのみ生じていた。したがって、アナログ音が言語音として知覚されている限り、言語音とそのアナログ音の音韻知覚は、同一のメカニズムにより成立することが示唆された。それゆえ、音韻知覚は音韻関連特性の知覚表象と音韻記憶表象との照合によって成立し、言語音と非言語音との知覚的差異は、音響信号の知覚的処理において音韻記憶表象を参照するか(あるいは参照できるか)否かによる、という仮説は支持されたと言えよう。

2. 正弦波アナログ音を用いた音韻知覚研究における本研究の位置づけと成果

2.1 正弦波アナログ音を用いた音韻知覚研究と本研究の位置づけ

従来、正弦波アナログ音は、言語音と非言語音が異なる様式で知覚されることを例証する目的で用いられてきた。

Best et al. (1989) は、合成言語音 /ra/-/la/ 連続体を模擬した正弦波言語音アナログの連続体を作成し、一群の被験者には歪んだ言語音として、他群の被験者には歪んだ和音として識別するよう求めた結果、前者ではカテゴリー的知覚を示す結果を、後者では非言語音の連続的知覚を示す結果を得た。同様の結果は、半母音 /wa/-/ya/ あるいは /ay/-/aw/ を模した正弦波言語音アナログを用いた Johnson & Ralston (1994) においても得ら

れている（また，高山，1993 も参照）。

また，Best et al. (1981) は，摩擦音－母音音節 /sei/ と摩擦音－破裂音－母音音節 /stei/ の知覚的区別において破裂音の手がかりとなる，摩擦音部と有声音節部との間の無音の長さ，有声音節部の第 1 フォルマント (F1) 起点周波数との間の手がかりトレーディングが，言語音知覚に特異的なものか心理音響学的な効果であるのかを検討するために，ノイズと正弦波アナログ音からなる疑似音節を用いて識別並びに弁別実験を行った。その結果，この疑似音節を言語音として聞いた被験者は，無音区間と F1 起点周波数との間のトレーディングを示したが，非言語音と聞いた被験者は，無音区間の長さか F1 起点周波数のいずれかを手がかりとした識別と弁別を行っていた，ということが見いだされた。

これらの研究は，言語音知覚の運動理論や直接知覚理論の立場から行われた研究であり，言語音様式のもとでの特異的な知覚的効果を再現している。他方，情報处理的アプローチから行われた研究もある。

Tomiak et al. (1987) は，言語音知覚の統合的処理について検討するために，摩擦音－母音音節 /fæ/, /ʃæ/, /fu/, /ʃu/ を模擬したノイズ－正弦波アナログ音を用いて，選択的注意に関する Garner 課題^{注5)}を行った。これらの刺激には，自然言語音がもつ摩擦音から母音へのフォルマント

注5) Garner 課題とは，2つの刺激次元の組合せに関する3つの刺激セット（単一次元セット，次元直交セット，次元相関セット）それぞれにおける特定次元への選択的注意のもとでの二者択一分類反応時間を測定し，単一次元セットでの反応時間 (RT) を基準とした各刺激セットにおける RT の増減から，刺激次元間の依存関係を明らかにしようとする課題である (Garner, 1974)。単一次元セットでは，注意を向けられない次元の値は一定のままで，注意を向けられる次元の値が変化する。次元相関セットでは，注意を向けられる次元のそれぞれ値は，必ず注意を向けられない次元の特定の値と対にされる。次元直交セットでは，2つの刺激次元でそれぞれの値が独立に変えられる。次元相関セットでの RT の減少あるいは次元直交セットでの RT の増加は，両刺激次元が統合的であることを示し，前者で RT が減少しないか後者で RT が増加しなければ，刺激次元は独立である，と解釈される。

遷移は含まれておらず、ノイズ部と母音部それぞれの識別に関して、非言語音として教示された参加者は、両者を相互に独立して処理しているという結果が示された。他方、言語音として教示された参加者は、ノイズ部と母音部の両者の識別に関して相互依存的な処理を示した。これらの結果から、Tomiak et al. (1987) は、調音結合（とその音響的結果）についての知識が言語音知覚の基本的な要素であり、信号が調音結合に関する情報を含んでいない状況であっても、音韻的符号化においては調音結合の知識の利用が強えられる、と論じている。

また、Schwab (1981) は、言語音知覚における聴覚的過程と音声的過程との相互作用について検討するため、言語音（破裂子音－母音音節と母音－破裂子音音節）への近似の程度を操作した正弦波アナログ音を用い、5つの実験によって、音声的処理が自発的に生起するか否か、音声的処理は聴覚的分析の出力を利用できるのか否か、を検討した。その結果、聴者は意識的に正弦波アナログ音の音声的処理が可能であり、言語音への近似度が高まるほど、音声的ラベリング（言語音識別）の成績は向上することが示された。他方、言語音に近似するほど聴覚的ラベリング（非言語音識別）の成績が低下することが見いだされたが、それは、音声的処理が自動的に行われるようになったというよりも、周波数遷移の方向の異なる複数の成分が含まれるようになることで、信号内に競合する情報が増えることによるものであると解釈された。さらに、刺激の構成要素に対する感度について、処理様式により違いが認められ、聴覚的（非言語音）様式での知覚において、刺激開始時の破裂音の手がかりに対する逆行マスキングが生じていることが見いだされた。これらの結果から、Schwab (1981) は、聴覚的処理様式と音声的処理様式とが相互に排他的であると述べている。

言語音知覚に関して、知覚的体制化の特異性の側面からアプローチした研究もある。Remez と彼の共同研究者たちによる一連の研究 (Remez et al., 1981; Remez, Pardo, Piorkowski, & Rubin, 2001; Remez & Rubin, 1990; Remez, Rubin, Berns, Pardo, and Lang, 1994) はその代表である。彼らの主張は、基本的には、直接知覚理論の立場からのものであり、さまざまな要素から構成される言語音がまとまりをもって (coherent に) 知覚されるのは、一般的な聴知覚において述べられる Gestalt の法則や二次的な図式的過程 (Bregman, 1990) によるのではなく、音声的な知覚的体制化が、言語音の調音動作を反映した音響変動に対して特異的に鋭敏であることによって達成される (Remez et al., 1994), というものである。先に引用した Best et al. (1989) もこの流れにあると言って良いであろう。

しかし、このような体制化が、言語音知覚に特異的と言うよりも、一般的な聴覚的カテゴリー化の基礎にあることを示唆する研究もある。Grunke & Pisoni (1982) は、言語音知覚において、例えば語頭と語尾のように異なる位置に現れる音韻間の類似性の問題を取り上げ、子音-母音音節と母音-子音音節における破裂音のフォルマント遷移が時間的の互いの鏡映像になっていることに着目して、それらを模擬した正弦波アナログ音の知覚的類似性に関する研究を行った。知覚学習パラダイムを用いて、聴者は、周波数遷移の方向のようなスペクトル的時間的共通性に基づいて分類するよりも、鏡映像パターンに一貫して任意の反応カテゴリーを割り当てることを容易に学習することができる、ということが示された。したがって、異なる分類規則に応じてこれらのパターンを分類することを学習する際、学習者は音響パターンの個々の要素だけでなく、パターンの内的体制を利用できる、ということが示唆されている。

また、Sawush & Gagnon (1995) は、Best et al. (1989) が示した言語音教

示におけるカテゴリー的知覚と非言語音教示における連続的知覚という結果が、アナログ音の識別訓練によるアーティファクトであるとして、十分な非言語音識別訓練を行うことで、非言語音として聞かれたアナログ音に関してカテゴリー的知覚が得られ、訓練の影響が母音文脈の異なる子音にまで般化される、ということを示した。これらの結果から、彼らは、言語音のカテゴリー化と非言語音のカテゴリー化との両者を支える抽象的な聴覚的表象の存在を主張している。

以上、正弦波アナログ音やその変形を用いて行われてきた研究の代表的なものを列挙したが、その主要な関心は、知覚の言語音様式と非言語音様式それぞれにおける知覚内容からもたらされる効果や現象、あるいはそれらの特徴であって、そもそもどのようにしてそのような知覚内容がもたらされるのか、ということについてはほとんど関心が払われていなかったように思える。確かに、一部の研究者たちは、アナログ音の音韻知覚にとって、言語音にスペクトル的・時間的に近似している必要があることを指摘している（Grunke & Pisoni, 1982; Remez et al., 1981; Schwab, 1981）が、それはあくまでも実験条件の操作の上からの指摘でしかなかった。多くの研究に見られるように（Best et al., 1981, 1989; Johnson & Ralston, 1994; Remez et al., 2001）、本来、現象や効果の記述でしかない知覚の言語音様式と非言語音様式によって、言語音と非言語音との知覚的差異を説明しようとするれば、結局、循環論に陥ることになるであろう。むしろ、このような知覚様式の違いをもたらすものは何か、ということが明らかにされなければならない。本研究の関心はまさにその点にあった。

2.2 正弦波アナログ音を用いた音韻知覚研究に対する本研究の成果

本研究の成果としてまず第 1 にあげるべき点は、通常、正弦波アナロ

グ音の識別が自動的、あるいは直接的には成立せず、特に単音節アナログ音に関して安定した音韻知覚を得るためには十分に訓練する必要があるという事実 (e.g., Best et al., 1989; Sawusch & Gagnon, 1995) を重視し、そこから音韻知覚過程における処理をとらえたことにある。従来の研究では、アナログ音は言語音の粗悪な代用品であって、言語音として知覚されないのは、その粗悪さ故に、音韻知覚過程を自動的に始動させられないからであり、アナログ音の音韻関連特性に適切に注意を向けることができれば、音韻知覚を引き起こすことができるとされていた (Remez et al., 1981)。しかし、実験 I と実験 II の結果は、アナログ音を言語音として知覚するよう教示することで、その音韻関連特性に注意を向けさせたとしても、それだけで適切な音韻知覚 (あるいは識別) が成立するわけではなく、識別学習を通して徐々に可能になることを示している。しかも、実験 I の結果が示すように、適切なフィードバックが与えられなければ、カテゴリー的知覚などの音韻知覚研究で求められる識別水準には到達しない。アナログ音を用いたこれまでの音韻知覚研究においては、このことが自明のこととされ、そこに関わる内的過程が見落とされていたように思える。

本研究の成果の第 2 の点は、アナログ音に対して安定した音韻知覚が可能となるために十分な識別学習が必要であるとしても、それは単純な対連合学習によって成立するのではなく、元の言語音に対して成立していた (と考えられる) 音韻関連特性と音韻記憶表象との照合関係を正弦波アナログ音が受け継いでおり、それに基づいて音韻識別が成立しているということを明らかにしたことである。実験 I と実験 II の結果は、正弦波アナログ音の識別学習は、任意のラベルを用いても可能であるが、本来の音韻ラベルを用いて学習するよりは困難であり、元の言語音と矛

盾する音韻ラベルを用いるとさらに困難になることを示した。

このことに関連して、第3の成果は、正弦波アナログ音の音韻関連特性と音韻記憶表象との照合が、個々の信号の特徴を保持した段階ではなく、それらの変動に関して標準化された抽象的表象の段階で生じていることを示したことである。実験Ⅱの結果は、2種類の声質の母音に対して構成されたアナログ母音の識別学習において、適切な音韻ラベルを用いる限り、声質は識別成績にほとんど影響しないが、任意のラベルや不適切な音韻ラベルを用いる場合には、声質（したがって、おそらく音響的弁別性）が大きく影響することを示した。

本研究の第4の成果は、言語音とその正弦波アナログ音の音韻知覚において、両者が共通の音韻記憶表象にアクセスする（あるいは活性化させる）ことを実験的に示したことである。すなわち、実験Ⅲにおいて、直接プライミングを用いて、言語音とアナログ音との間での音韻的同一性に基づく相互プライミングを確認できた。このことは、言語音とその正弦波アナログ音の音韻知覚が同じメカニズムの働きにより成立することを示唆する。

そして、第5の成果が、アナログ音の音韻知覚が、何らかの認知的媒介を経て成立しているのではなく、音韻識別学習が成立すれば、言語音の知覚と等価な過程により達成されていることを示したことである。先にも述べたように、正弦波アナログ音は粗悪な言語音として扱われ、言語音とそのアナログ音の音韻知覚が同じメカニズムに基づいていることは自明のこととされていたように思える（Best et al., 1981, 1989; Johnson & Ralston, 1994; Remez et al., 1981）。しかし、それは、アナログ音が言語音のフォルマント構造を模擬していること、そして言語音様に知覚されるとき言語音特異的とされてきた知覚的効果や現象が得られることの2点に

依拠しているだけで、言語音と等価な処理を受けているか否かについて、直接検証されていたわけではない。実験Ⅲにおいて示されたように、言語音と正弦波アナログ音との間で相互プライミングが成立すること、そして正弦波アナログ音に対する音韻的プライミング効果が言語音に対する音韻的プライミング効果と等価であり、かつプライム刺激とターゲット刺激の刺激間隔に影響されないということは、言語音の音韻知覚過程とアナログ音の音韻知覚過程が等価であることを強く示唆している。

3. 運動理論および直接知覚理論による解釈と問題点

3.1 運動理論および直接知覚理論による本研究結果の解釈

前節で述べた本研究結果の解釈は情報処理的アプローチからものであり、アナログ音の音韻知覚を音韻関連知覚表象と音韻記憶表象との照合過程による結果として扱っていた。他方、本研究の結果を言語音知覚の運動理論 (Lieberman & Mattingly, 1985) や直接知覚理論 (Fowler, 1986, 1996) の立場からでも解釈することは可能であろう。

運動理論や直接知覚理論では、調音動作を反映するスペクトルの時間的特性が維持されていれば、そして聴者の注意をそのような音響特性に方向づけることができれば、自動的、直接的に（伝統的な音韻や調音的特徴に関係づけられる意図された調音動作を対象とした）音韻知覚が成立する、とされている (Best et al., 1989; Remez et al., 1981)。特に、運動理論の立場では、生物学的に特殊化したメカニズムの存在を仮定するので、教示による注意や期待の操作はそのメカニズム内で生じる過程には影響しないはずである。

したがって、言語音教示一致群と言語音教示不一致群との結果の違いをもたらした教示の効果は、音韻関連特性から自動的、直接的に音韻表象が得られた後に作用していると考えることができる。言語音教示不一致群において不適切な音韻ラベルを用いることで生じた識別の低下は、特殊化した過程での処理結果としての音韻出力と、アナログ音信号に与えられた音韻ラベルとの間での競合が、識別反応を出力する段階で生じることによりもたらされた、と解釈できるかもしれない。また、実験Ⅲで得られた言語音と、言語音として与えられたアナログ音との間で生じる相互プライミング効果も、両信号はともに言語音として処理されるので、結果的に同一音韻が繰り返されることにより生じる反復効果として解釈できるであろう。

3.2 運動理論および直接知覚理論の問題点

このような解釈に対して、いくつかの問題点も指摘できる。

まず第1に、実験Ⅰと実験Ⅱにおける言語音教示不一致群への参加者の中で、正弦波アナログ音に与えられた音韻ラベルが、アナログ音の音韻関連特性と矛盾することに気づいた者がいなかったことが指摘される。運動理論や直接理論の立場では、教示でアナログ音を言語音として知覚するよう求められれば、聴者の注意はその音韻関連特性に向けられ、結果的に、音韻知覚が自動的あるいは直接的に成立することになる。その際、知覚内容と異なる音韻ラベルで識別するよう求められれば、必然的に知覚と反応の矛盾が意識されたであろう。しかし、参加者の実験後のインタビューには、そのような報告はなかった。

たとえ意識されなかったとしても、反応競合を引き起こしていた可能性も考えられる。本研究での実験結果を運動理論や直接理論の立場から

解釈すれば、前述のように、識別反応出力段階での効果ということになる。この解釈が妥当であれば、例えば、実験Ⅰと実験Ⅱの参加者において、言語音教示不一致群の参加者自身が反応キーを押し間違えたと意識する割合が、他群の参加者よりも多かったと予想される。それに対する直接的な証拠を得る手続きを用意していなかったため、必ずしも断定はできないが、実験後の参加者へのインタビューからはいずれの教示群からもそのような報告は非常に少なく、群間でその割合に違いはないように思える。また、実験Ⅲの参加者のインタビューにおいても、言語音教示群の参加者は同一音韻が連続する場合の反応キーの押しやすさではなく、異なる音韻が連続する場合の聞きにくさのほうを主に報告していた。したがって、意識下での反応競合の可能性も考えにくい。

本研究では、音韻知覚に関する理論的対立を主要なテーマとして取り上げているわけではないので、ここにあげた以上には、実験結果から運動理論や直接知覚理論の主張の可否を論じることはできない。しかし、正弦波アナログ音に対する音韻知覚に関してもう1点問題点をあげるならば、そして何よりも問題となるのは、その理論的立場との矛盾であろう。運動理論にせよ直接知覚理論にせよ、音韻知覚を生じさせる音響信号あるいは処理メカニズムの生態学的、生物学的特殊性を主張する。そのような知覚システムにおいては、意識の介在なしに、入力信号に対して自律的、自動的に（何らかの特殊な）処理が遂行される。すなわち、Remez et al. (1981) が述べるように、「言語学的にプライムされた聴者は、ほとんどの場合、単に知覚の言語音様式で聞くよう教示するだけで、正弦波信号の音声的特性に注意を向けることができた」(p. 949) というのは、結果から推測される現象記述でしかない。本来、運動理論や直接知覚理論では、音韻知覚の成立に注意が介入する余地はない。あるいは、

たとえその余地があったとしても、本来の音韻知覚を変更させるとは考えられない。したがって、アナログ音信号がそのもととなる言語音の音韻関連特性を維持しているのであれば、言語音として聞くよう教示されようと、非言語音として聞くよう教示されようと、自律的、自動的に、あるいは直接的に音韻知覚が成立しなければならないはずである。アナログ音を用いて知覚の言語音様式と非言語音様式のもとで識別と弁別（すなわちカテゴリー的知覚）を比較した研究のいくつかは、前もって言語音とも非言語音とも教示せず実験を行っており、参加者が自発的にアナログ音を言語音様に聞くか非言語音として聞くかに応じて、教示を与えた場合と同様の結果を得ている（Best et al., 1981; Johnson & Ralston, 1994）。しかし、本来的な運動理論や直接知覚理論の立場からすれば、実験参加者の自発的な聴取方略は、自律的、自動的に、あるいは直接的な音韻知覚の成立を支持する証拠にはならない。

さらに、音韻知覚全般に関わる運動理論・直接知覚理論の問題点を指摘しておく。本来の運動理論や直接知覚理論では、たとえ音韻知覚の成立に注意が介入することを認めたとしても、その影響については説明できない点が多い。特に、音韻知覚に関する音響手がかりの寄与が、注意の操作で変動する事実を説明する正当な根拠はない。これらの理論では、音韻知覚に関する音響的手がかり（あるいは音韻知覚関連特性）は、調音動作の結果として言語音信号に組み込まれたものである。そのため、言語音信号あるいはそれを模した正弦波アナログ音信号は、音韻内容（あるいは伝統的な音韻や調音的特徴に関係づけられる意図された調音動作）を特定するために必要な情報をすべて含んでいる、あるいは、特殊化された過程の働きにより、それらの信号からそのような音韻内容（あるいは調音動作）を特定できる、とされる（Lieberman & Mattingly, 1985;

Fowler, 1986, 1996)。特定の音韻知覚に関する音響の手がかりは複数あり、それらの手がかりの間に知覚的な相補・相殺関係が成立することを示すのが音声トレーディングである (Best et al., 1981; Fitch et al., 1980) が、通常、それぞれの音響手がかりのセットは、協同して音韻内容 (あるいは調音動作) を特定するために用いられている。したがって、注意の状態によって音響の手がかりの処理が影響され、識別結果が変化することはない。

しかし、音韻知覚への注意の割り当てに応じて、音韻知覚に関する複数の音響の手がかり間の相対的寄与が変化する、ということが見いだされている。Gordon, Eberhardt, & Rueckl (1993) は、音声の基本周波数 (ピッチ) の起点周波数をパラメータとして声開始時間 (VOT) を段階的に変えた有声破裂音—無声破裂音系列、あるいは持続時間をパラメータとして F1 ~ F3 中心周波数を段階的に変えた母音 /i/—/I/ 系列を用い、同時作業の有無で注意の配分を操作した音韻識別実験を行った。その結果、同時作業を課さない場合、破裂音系列では VOT が、母音系列では F1 ~ F3 中心周波数が識別反応を強く支配したが、同時作業が課せられるとそれらの影響は弱まる、ということを見いだした。

さらに、音韻識別に関する訓練によって、注意を向けるべき音響手がかりを変えることも可能であることも見いだされている。Francis et al. (2000) は、有声破裂子音—母音音節の子音部の調音場所の手がかりであるフォルマント遷移と破裂音ノイズの組み合わせを操作し、異なる母音環境 (/a/と/i/) のもとで、フォルマント遷移手がかりに基づいて破裂音を識別するよう訓練する群と、破裂音ノイズに基づいて識別するよう訓練する群に関して、訓練の効果を比較した。先行研究 (Walley & Carrell, 1983) により、調音場所の決定に関して、フォルマント遷移と破裂音ノ

イズが競合するとき、フォルマント遷移手がかりが優勢である、ということが知られている。訓練の結果、単純な音韻カテゴリーのフィードバックにより、どちらの手がかりであっても、訓練した手がかりを用いて識別する傾向が強まる、ということが示された。

これらの議論は、音韻知覚に関して運動理論が主張する生物学的に特殊化したシステムの存在を必ずしも否定するものではない。しかし、言語音知覚の過程において、音響信号に対する聴覚的表象を言語音表象へ対応づける認知的過程が一定の役割を果たしていること、その過程において注意や方略など他の聴知覚過程と同じ要因が影響を及ぼしていることは紛れもない事実である。本研究は、改めてその点に着目し、実証したことに意義があると考えられる。

第2節 将来の課題

1. 音韻記憶表象の内容と形式

本研究は、アナログ音、言語音にかかわらず、それらの音韻知覚において、長期記憶内の音韻表象が参照される、ということを示唆する実験結果を得た。しかし、この音韻記憶表象そのものの性質については、本研究では特定できなかった。言語音知覚の主要な問題の一つは、聴者がどのようにして連続的な音響信号を離散的なカテゴリー的表象に対応づけるのか、を明らかにすることである (Miller & Eimas, 1995)。すなわち、

カテゴリー的な音韻記憶表象がどのような性質をもつものであるのか、あるいはそれがどのように形成されるのか、ということは、音韻知覚メカニズムの基本仮説に関わる問題である。

言語音のカテゴリー化に関する初期の研究の多くは、カテゴリー的知覚という現象に関心があった (e.g., Harnad, 1987; Liberman et al., 1967; Repp, 1984)。先にも述べたように、カテゴリー的知覚の本質は、音韻知覚の過程において、音韻カテゴリーの識別とは無関連な音響変動は捨象され、音韻的同一性に関する情報のみが保持される、ということである。

初期の研究では、カテゴリー的知覚の問題は、物理的差異が同一であっても、同じカテゴリーに分類される刺激同士の間での弁別は、不可能ではないにしても、かなりの程度困難であるのに対して、別のカテゴリーに分類される刺激の間での弁別は容易である、という刺激弁別における音韻的同一性の影響、という側面からとらえられていた。しかし、その後、記憶負荷等の課題状況を操作することによって、カテゴリー内弁別の精度を高めたり低めたりすることができるということが見いだされ、カテゴリー的知覚と連続的知覚の違いが、課題場面に関わる2つの記憶システムの相対的寄与の違いとして説明できるとの考えが主流を占めるようになった (Fujisaki & Kawashima, 1969, 1970; Pisoni, 1971, 1973; Macmillan et al., 1988)。すなわち、刺激の感覚的痕跡に基づき、短時間のうちに減衰するとともに、時間的に近接して生じる他の音の影響を受ける痕跡的あるいは聴覚的な記憶と、長期的な言語経験の中で安定的に形成され、カテゴリー化 (あるいはラベリング) の参照点を構成する音声的記憶 (Fujisaki & Kawashima, 1969, 1970; Pisoni, 1971, 1973) あるいは文脈的記憶 (Macmillan et al., 1988) が区別された。

カテゴリー化に関する長期記憶要因への関心から、認知心理学におい

ては、対象や事象をカテゴリー化する能力を説明するために仮定した、長期記憶内のある種の認知的参照点あるはプロトタイプという考え方が、言語音知覚研究に導入されている。そのような研究で見いだされたことは、同一カテゴリー内のメンバーが弁別可能であるだけでなく、カテゴリーのあるメンバーは他のメンバーよりも良い事例、あるいはよりプロトタイプの的であると知覚されるという事実である (Kuhl, 1992; Massaro, 1987; Miller & Volaitis, 1989; Samuel, 1982)。すなわち、音韻カテゴリーは序列的な構造を持っており (cf. Medin & Basalou, 1987; Rosch, 1978)、例えば、あるカテゴリー内のメンバーがランダムに提示されるとき、聴者はカテゴリーの良さ (category goodness) について非常に微細な判断を行うことができる (Kuhl, 1992; Miller & Volaitis, 1989)。また、刺激が判断されたカテゴリーの良さについて異なるとき、両耳分離聴 (Repp, 1977) や選択的順応 (Samuel, 1982) のような課題での知覚的有効さも同様に異なることが示されている。加えて、カテゴリーの中央にある刺激 (すなわち、カテゴリープロトタイプ) は、近隣の刺激を同化させ、プロトタイプのまわりの知覚的空間を縮めるとされている (知覚的マグネット効果, Kuhl, 1992; cf. Samuel, 1982)。

本研究で想定する音韻記憶表象も、プロトタイプの的な表象であった。実験 II では、声質の影響が大きいと考えられる母音を模擬した正弦波アナログ音を用いて、その識別学習に及ぼすラベリングの効果を検討した。女声と男声のフォルマント構造を模擬した 2 種類の音質の母音アナログの識別学習は、もとの母音名と一致するラベルが与えられたときには、アナログ音間でほぼ同程度の水準で学習が成立したが、もとの母音名との対応を変えたラベルによる学習は、非言語音として任意のラベルを用いた対連合学習と同様に、個々のアナログ音間の弁別性の影響を受けた。

すなわち、音韻記憶表象との照合が期待される識別学習においては、アナログ音の音響的弁別性の影響は受けないことが示された。これは、個々の信号の音響変動が捨象された段階で、入力信号の知覚表象と音韻記憶表象との照合が生じていることを示唆するものであり、カテゴリーの中心的位置を占めるプロトタイプの存在を支持するものと考えられる。

言語音知覚の運動理論も、必ずしもプロトタイプ的な表象の存在を否定するものではない。この理論においては、特殊化したモジュールが、言語音信号をその調音が意図されたレベルでの調音運動神経指令により解釈することによって、音韻知覚が成立する、とされている (Lieberman et al., 1967; Liberman & Mattingly, 1985)。すなわち、調音プロトタイプとも呼ばれるべき音韻記憶表象との照合を想定しているとも考えられる。しかし、Lieberman & Mattingly (1985) は、特殊化したモジュールの中でどのような処理が行われているのかについて、何ら特定していない。

他方、情報处理的アプローチでは、その理論構成において、プロトタイプあるいは鋳型的な記憶表象への照合という考え方が一般的であるように思える。例えば、Sawusch (1986) は、音韻知覚のコンピュータシミュレーションにおいて、音響信号から得られた聴覚的符号を音声カテゴリーに対応づける仕組みとして、言語音信号の鋳型的な音響構造 (Blumstein & Stevens, 1979; Searle et al., 1979; Kewley-Port, 1983) との照合を用いた。また、Oden & Massaro (1978) は、音韻知覚の過程を、言語音信号内の音響手がかりが独立に検出され、音韻特徴の存在の尤もらしさが評価された後に、特徴情報が統合され、それと長期記憶内のプロトタイプとの照合の程度が評価される、とモデル化している。

これらの研究は、いずれも単一の抽象的なプロトタイプを想定している。しかし、Miller & Eimas (1995) によると、「カテゴリーの中心はそれ

自体，単一の抽象的プロトタイプ（原型）として表象されているのではなく，むしろ言語の音響的－音声的形式についての微細な詳細を維持した構造として表象されている」（p. 471）。彼らは，音響信号と音韻カテゴリーとの対応づけの複雑さに寄与する主要な要因と考えられる文脈効果や，音韻的同一性を特定する音響的手がかりの多重性（すなわち，手がかりトレーディング）の効果が，従来，カテゴリーの成員性があいまいな周辺メンバーのみに影響することによるカテゴリー境界位置の移動としてとらえられていたが（Repp & Liberman, 1987），実際にはカテゴリーの中心にまでその影響が及んでおり，結果として，最も良いカテゴリー事例の位置が組織的に移動するということが示されている（Miller, 1994 参照），と述べている。そして，このような記憶表象の形式に関して，「ある種の文脈に依存した要約的記述に基づいているのか（Rosch, 1978），あるいは，おそらく頻度に関して重みづけられた，貯蔵されたカテゴリー事例によるのか（Nosofsky, 1988），ということは，まだ解決されていない（Li & Pastore, 1992; Medin & Barsalou, 1987）。」（Miller & Eimas, 1995, p. 472）としている。加えて，「詳細な表現が最もうまく記述されるレベル，例えば，特徴カテゴリーによるのか，セグメントカテゴリーによるのか，音節カテゴリーによるのか，ということは不明である。」（Miller & Eimas, 1995, p. 472）とも述べている。

いずれにしても，本研究で想定した音韻記憶表象がどのような内容を保持したものであるのか，そしてその内容がどのような形式で長期記憶に保持されているのかということについて，今後明らかにすべき重要課題である。

2. 音韻記憶表象の活性化

実験Ⅲにおいて、プライミング効果に関して言語音知覚と非言語音知覚の違いが明確に示されたのは、プライム刺激とターゲット刺激とが、音響的には異なるが、音韻的には同一である組合せにおいてであった。すなわち、プライミング効果が先行刺激による記憶表象の活性化に由来するものであるとすると (Decoene, 1997; 太田, 1991), 刺激の物理的同一性にかかわらず、音韻的同一性が維持されている限り、アナログ音と言語音は、共通の音韻記憶表象を活性化させる、ということが示された。

他方、プライミング効果の大きさ自体は有意水準に達しなかったものの、(アナログ音に関する) 言語音教示群と非言語音教示群の違いは、プライム刺激とターゲット刺激が同種の異なる音、すなわちともに言語音かともにアナログ音、であるときにも有意となった。この点について更に検討したところ (図 3.1 参照), アナログ音ターゲットに対しては両群のプライミング効果の大きさはほぼ等しく、言語音ターゲットに対して、言語音教示群で正のプライミング傾向、非言語音教示群で負のプライミング傾向が

得られ、前者では 6 名中 4 名の、後者では 6 名中 5 名の参加者においてそれぞれの方向へのプライミング効果が示されていた。すなわち、言語音

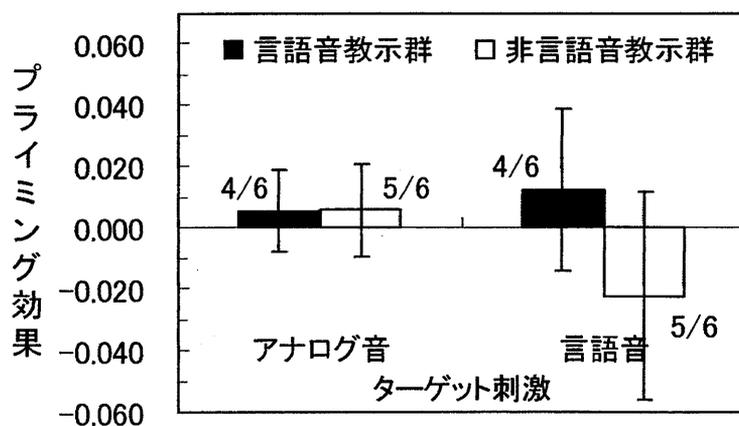


図 3.1 実験Ⅲの同種異音プライム条件における教示群とターゲット刺激に関するプライミング効果の比較

ターゲット刺激の識別に関して、言語音プライム刺激は、言語音教示群ではやや促進的な効果を、非言語音教示群では抑制効果をもたらしたと言えるかもしれない。

これらの効果が偶然の結果でないとするならば、音韻知覚における音韻記憶表象の活性化の役割を知る上で、重要な示唆が得られるかもしれない。例えば、実験Ⅲの課題場面を考えると、言語音教示群では、ターゲット刺激が言語音であれアナログ音であれ常に音韻識別課題を行うことになり、またその状況で提示される刺激は中立プライム刺激以外すべて音韻刺激である。したがって、課題場面と大部分の刺激によりもたらされる文脈によって、プライミング刺激が提示される以前に、音韻記憶表象は準備的活性状態にあるかもしれない。他方、非言語音教示群では、言語音識別課題であっても、プライミング刺激がもたらす文脈は言語音知覚と非言語音知覚の混在したものであろう。そのため、音韻記憶表象の準備的活性状態も弱いかもしれない。ここで、ターゲット刺激と類似した同種のプライム刺激によりもたらされる音韻記憶表象の活性化が、ターゲット刺激の識別に対して抑制的な効果を及ぼすと仮定すれば、言語音教示群ではこの抑制効果と準備的活性状態が相殺されることにより、プライミング効果は現れないであろうし、非言語音教示群では弱い準備的活性状態とプライム刺激による抑制的効果とが相殺されず、プライミング効果は抑制方向に現れることになるであろう。

音韻記憶表象の準備的活性化についてはあくまでも推測であり、直接それを支持する証拠はないようである。しかし、言語音文脈から切り離されて非言語的に知覚されるはずの音韻関連刺激（音響的手がかり）から、音韻識別やカテゴリー的知覚が可能であるという実験結果は、間接的に、このような音韻記憶表象の活性化の可能性を示唆するように思え

る。例えば、Nusbaum et al. (1983) は、破裂子音の調音場所の手がかりとなる第 2 フォルマント (F2) 遷移について、遷移部単独であっても、通常の言語音刺激に匹敵する程度に識別が可能であることを示した。Repp et al. (1983) は、破裂子音の調音場所の手がかりとなる第 3 フォルマント (F3) 遷移について、遷移部単独での識別を検討し、Nusbaum et al. (1983) の主張を否定しているが、このような結果の違いは、高山 (1989a) が破裂子音の構成要素に対する音韻識別の程度を比較した結果から示唆したように、音韻情報の伝達にとって F2 遷移は F3 遷移より重要であることを反映しているのかもしれない。

また、準備的活性化という考えにとって、大脳半球機能の非対称性に関して提唱された Kinsbourne (1973) の注意理論の主張が参考になるかもしれない。Kinsbourne の注意理論は、本来、認知的ラテラリティ現象を、「左右 (大脳) 半球の機能的非対称性は認めるものの」、「ある実験的文脈の中で主体が形成する刺激、課題、反応に対する期待や認知セット、あるいは課題に対して主体が選択する情報処理ストラテジーにより、それに応じた半球が刺激提示以前に選択的に賦活され、同時に賦活半球により逆側の半球活動が抑制されるという意味での半球間相互作用を仮定して」(利島他, 1980, p. 137) 説明しようとする理論である。この考えを借りれば、先の解釈の中で述べた、音韻識別の課題文脈における音韻記憶表象の準備的活性化状態と、言語音知覚と非言語音知覚の混在した中での音韻記憶表象の準備的活性化状態との違いをもたらすものは、言語音知覚システムと非言語音知覚システムとの相互作用ということになるのかもしれない。

いずれにしても、本研究においてプライミング効果の背後にあると想定した音韻記憶表象の活性化がどのようにしてもたらされるのか、すな

わち刺激の音響的特性（の知覚表象）と音韻記憶表象との照合が可能となるのはどのような条件においてであるのか，ということについては，今後明らかにすべき課題の一つである。

第3節 まとめ

本研究は，音韻知覚が，信号の音韻関連特性に関する知覚表象と，長期記憶に保持された音韻記憶表象との照合により成立し，この音韻記憶表象を参照するか否かが言語音と非言語音の知覚的差異をもたらす一因である，との仮説をたて，言語音のフォルマント構造を周波数変調信号で模擬した正弦波アナログ音を用いて検討した。破裂子音－母音音節（実験Ⅰ）あるいは単母音（実験Ⅱ）の正弦波アナログ音に関する識別実験の結果は，アナログ音が模擬する言語音の音韻関連特性（の知覚表象）と音韻記憶表象との既存の照合関係が識別学習の成績に影響を及ぼす，ということを示した。また，そのような音韻関連特性と音韻記憶表象との照合関係は，音響変動を捨象した抽象的レベルで成立していることも示唆された。さらに，音韻知覚における正弦波アナログ音と言語音との相互プライミング効果を示すことで，正弦波アナログ音の音韻知覚において支持された音韻関連特性と音韻記憶表象との照合関係が，言語音の音韻知覚においても成立していることが示唆された（実験Ⅲ）。

他方，本研究の仮説における中心的概念である「音韻記憶表象」がどのような内容のものなのか，どのような形式で長期記憶に保持されてい

るのか，そしてそのような「音韻記憶表象」がどのようにして活性化されるのか，ということについては推測の域を出ていない状況である。今後，この点についての実証的な研究を行う必要がある。

引用文献

- Bailey, P. J., & Summerfield, A. Q. 1980 Information in speech: Observations on the perception of [s]-stop clusters. *Journal of Experimental Psychology: Human Perception and Performance*, **6**, 536–563.
- Berlin, C. I., Lowe-Bell, S. S., Cullen, J. K., Jr., Thompson, C. L., & Loovis, C. F. 1973 Dichotic speech perception: An interpretation of right ear advantage and temporal offset effects. *Journal of the Acoustical Society of America*, **53**, 699–709.
- Best, C. T. 1995 A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience*. Timonium, MD: York Press. Pp.171–206.
- Best, C. T., Morrongiello, B., & Robson, R. 1981 Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics*, **29**, 191–211.
- Best, C. T., Studdert-Kennedy, M., Manuel, S., & Rubin-Spitz, J. 1989 Discovering phonetic coherence in acoustic patterns. *Perception & Psychophysics*, **45**, 237–250.
- Bever, T. G., & Chiarello, R. J. 1974 Cerebral dominance in musicians and nonmusicians. *Science*, **185**, 537–539.
- Blumstein, S. E. 1974 The use and theoretical implications of the dichotic technique for investigating distinctive features. *Brain and Language*, **1**, 337–350.
- Blumstein, S. E., Isaacs, E., & Mertus, J. 1982 The role of the gross spectral shape as a perceptual cue to place of articulation in initial stop consonants. *Journal of the Acoustical Society of America*, **72**, 43–50.
- Blumstein, S. E., & Stevens, K. N. 1979 Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants.

- Journal of the Acoustical Society of America*, **66**, 1001–1017.
- Blumstein, S. E., & Stevens, K. N. 1980 Perceptual invariance and onset spectra for stop consonants in different vowel environments. *Journal of the Acoustical Society of America*, **67**, 648–662.
- Blumstein, S. E., Tartter, V. C., Michel, D., Hirsch, B., & Leiter, E. 1977 The role of distinctive features in the perception of vowels. *Brain and Language*, **4**, 508–520.
- Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. 1999 Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception & Psychophysics*, **61**, 977–985.
- Bregman, A. S. 1990 *Auditory scene analysis*. Cambridge, MA: MIT Press.
- Bryden, M. P. 1982 *Laterality: Functional asymmetry in the intact brain*. New York: Academic Press.
- Bryden, M. P. 1988 An overview of the dichotic listening procedure and its relation to cerebral organization. In K. Hugdahl (Ed.), *Handbook of dichotic listening: Theory, methods and research*. Chichester: John Wiley & Son. Pp. 1–43.
- Cleary, M., & Pisoni, D. 2001 Speech perception and spoken word recognition: Research and theory. In E. B. Goldstein (Ed.), *Blackwell handbook of perception*. Malden: Blackwell. Pp. 499–534.
- Decoene, S. 1997 The representational basis of syllable categories. *Perception & Psychophysics*, **59**, 877–884.
- Eimas, P. D. 1963 The relation between identification and discrimination along speech and non-speech continua. *Language and Speech*, **6**, 206–217.
- Eimas, P. D., Cooper, W. E., & Corbit, J. D. 1973 Some properties of linguistic feature

- detectors. *Perception & Psychophysics*, **13**, 247–252.
- Eimas, P. D., & Corbit, J. D. 1973 Selective adaptation of linguistic feature detectors. *Cognitive Psychology*, **4**, 99–109.
- 江副裕二 2000 DAF 両耳シミュレーションを用いた課題使用頻度の差異に現れる注意配分の変化 平成 11 年度近畿大学工学部電子情報工学科卒業研究 (未公刊)
- Fairbanks, G. 1955 Selective vocal effects of delayed auditory feedback. *Journal of Speech and Hearing Disorder*, **20**, 333–346.
- Fant, G. 1960 *Acoustic theory of speech production*. The Hague: Mouton.
- Fitch, H. L., Halwes, T., Erickson, D. M., & Liberman, A. M. 1980 Perceptual equivalence of two acoustic cues for stop consonant manner. *Perception & Psychophysics*, **27**, 343–350.
- Fowler, C. A. 1986 An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, **14**, 3–28.
- Fowler, C. A. 1996 Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America*, **99**, 1730–1741.
- Francis, A. L., Baldwin, K., & Nusbaum, H. C. 2000 Effects of training on attention to acoustic cues. *Perception & Psychophysics*, **62**, 1668–1680.
- Fujisaki, H., & Kawashima, T. 1969 On the modes and mechanisms of speech perception. *Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokyo*, **28**, 67–73.
- Fujisaki, H., & Kawashima, T. 1970 Some experiments on speech perception and a model for the perceptual mechanism. *Annual Report of the Engineering Research Institute, Faculty of Engineering, University of Tokyo*, **29**, 207–214.
- 府川昭世 1980 朗読課題の熟知度と音声遅延フィードバック効果一言

- 語運動の外在・内在フィードバックモデルの観点からみた DAF 効果
(I) — 音声言語医学, 21, 103-108.
- 古井貞熙 1994 音声の基本的性質と音韻知覚 大山正・今井省吾・和
気典二(編) 新編 感覚・知覚心理学ハンドブック 誠信書房 Pp.
1118-1128.
- Garner, W. R. 1974 *The processing of information and structure*. Potomac, MD:
Erlbaum.
- Godfrey, J. J. 1974 Perceptual difficulty and the right ear advantage for vowels. *Brain
and Language*, 1, 323-335.
- Gordon, P. C., Eberhardt, J., & Rueckl, J. G. 1993 Attentional modulation of the
phonetic significance of acoustic cues. *Cognitive Psychology*, 25, 1-42.
- Grunke, M. E., & Pisoni, D. B. 1982 Some experiments on perceptual learning of
mirror-image acoustic patterns. *Perception & Psychophysics*, 31, 210-218.
- Halperin, Y., Nachson, I., & Carmon, A. 1973 Shift of ear superiority in dichotic
listening to temporally patterned nonverbal stimuli. *Journal of the Acoustical Society
of America*, 53, 46-50.
- Harnad, S. (Ed.) 1987 *Categorical Perception*. New York: Cambridge Univ. Press.
- Hayden, M. E., Kirstein, E., & Singh, S. 1979 Role of distinctive features in dichotic
perception of 21 English consonants. *Journal of the Acoustical Society of America*,
65, 1039-1046.
- House, A. S., Stevens, K. N., Sandel, T. T., & Arnold, J. B. 1962 On the learning of
speechlike vocabularies. *Journal of Verbal Learning and Verbal Behavior*, 1,
133-143.
- Jusczyk, P. 1986 Speech perception. K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.),
Handbook of perception and human performance Vol.2 *Cognitive processes and*

- performance*. New York: John Wiley and Sons. Pp. 27-1-27-57.
- Johnson, K., & Ralston, J. V. 1994 Automaticity in speech perception: some speech / nonspeech comparisons. *Phonetica*, **51**, 195-209.
- 河原英紀 2003 聴覚的フィードバックの発声への影響—ヒトは自分の話し声を聞いているのか?— 日本音響学会誌, **59**, 670-675.
- Kewley-Port, D. 1983 Time-varying features as correlates of place of articulation in stop consonants. *Journal of the Acoustical Society of America*, **73**, 322-335.
- Kewley-Port, D., & Luce, P. A. 1984 Time-varying features of initial stop consonants in auditory running spectra: A first report. *Perception & Psychophysics*, **35**, 353-360.
- Kewley-Port, D., Pisoni, D. B., & Studdert-Kennedy, M. 1983 Perception of static in dynamic acoustic cues to place of articulation in initial stop consonants. *Journal of the Acoustical Society of America*, **73**, 1779-1793.
- Kimura, D. 1961 Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, **15**, 166-171.
- Kimura, D. 1964 Left-right differences in the perception of melodies. *Quarterly Journal of Experimental Psychology*, **16**, 355-358.
- Kimura, D. 1967 Functional asymmetry of the brain in dichotic listening. *Cortex*, **3**, 163-178.
- Kinsbourne, M. 1973 The control of attention by interaction between the cerebral hemispheres. In S. Kornblum (Ed.), *Attention and performance IV*. New York: Academic Press. Pp.239-256.
- Klatt, D. H. 1980 Software for a cascade / parallel formant synthesizer. *Journal of the Acoustical Society of America*, **67**, 971-995.

- Kuhl, P. K. 1992 Psychoacoustics and speech perception: Internal standards, perceptual anchors, and prototypes. In L. A. Werner, & E. W. Rubel (Eds.), *Developmental Psychoacoustics*. Washington, D. C.: American Psychological Association. Pp. 293–332.
- Ladefoged, P., & Broadbent, D. E. 1957 Information conveyed by vowels. *Journal of the Acoustical Society of America*, **29**, 98–104.
- Lee, B. S. 1950 Effects of delayed speech feedback. *Journal of the Acoustical Society of America*, **22**, 824–826.
- Li, X., & Pastore, R. E. 1992 Evaluation of prototypes and exemplars in perceptual space for place contrast. In M. E. H. Schouten (Ed.), *The auditory processing of speech: From sounds to words*. Berlin: de Gruyter. Pp. 303–308.
- Lieberman, A. M. 1970 The grammars of speech and language. *Cognitive Psychology*, **1**, 301–323.
- Lieberman, A. M. 1982 On finding that speech is special. *American Psychologist*, **37**, 148–167.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. 1967 Perception of the speech code. *Psychological Review*, **74**, 431–461.
- Lieberman, A. M., Delattre, P. C., & Cooper, F. S. 1952 The role of selected stimulus variables in the perception of the unvoiced stop consonants. *American Journal of Psychology*, **65**, 497–516.
- Lieberman, A. M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. 1954 The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs*, **68**, 1–13.
- Lieberman, A. M., & Mattingly, I. G. 1985 The motor theory of speech perception revised. *Cognition*, **21**, 1–36.

- Liberman, A. M., & Whalen, D. H. 2000 On the relation of speech to language. *Trends in Cognitive Science*, **4**, 187–196.
- Macmillan, N. A., Goldberg, R. F., & Braida, L. D. 1988 Resolution for speech sounds: Basic sensitivity and context memory on vowel and consonant continua. *Journal of the Acoustical Society of America*, **84**, 1262–1280.
- Mann, V. A., & Liberman, A. M. 1983 Some differences between phonetic and auditory modes of perception. *Cognition*, **14**, 211–235.
- Massaro, D. W. 1987 Categorical partition: A fuzzy-logical model of categorization behavior. In S. Harnad (Ed.), *Categorical Perception*. New York: Cambridge Univ. Press. Pp. 254–283.
- Massaro, D. W., & Friedman, D. 1990 Models of integration given multiple sources of information. *Psychological Review*, **97**, 225–252.
- Mattingly, I. G., Liberman, A. M., Syrdal, A. K., & Halwes, T. 1971 Discrimination in speech and nonspeech modes. *Cognitive Psychology*, **2**, 131–157.
- McGurk, H., & MacDonald, J. 1976 Hearing lips and seeing voices. *Nature*, **264**, 746–748.
- Medin, D. L., & Barsalou, L. W. 1987 Categorization processes and categorical perception. In S. Harnad (Ed.), *Categorical Perception*. New York: Cambridge Univ. Press. Pp. 455–490.
- Miller, G. A., & Taylor, W. G. 1948 The perception of repeated bursts of noise. *Journal of the Acoustical Society of America*, **20**, 171–182.
- Miller, J. L. 1994 On the internal structure of phonetic categories: a progress report. *Cognition*, **50**, 271–285.
- Miller, J. L., & Eimas, P. D. 1995 Speech perception: from signal to word. *Annual*

- Reviews of Psychology*, **46**, 467–492.
- Miller, J. L., & Liberman, A. M. 1979 Some effects of later-occurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics*, **25**, 457–465.
- Miller, J. L., & Volaitis, L. E. 1989 Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics*, **46**, 505–512.
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., & Fujimura, O. 1975 An effect of linguistic experience: The discrimination of [r] and [l] by native speaker of Japanese and English. *Perception & Psychophysics*, **18**, 331–340.
- ムーア B. C. J. 大串健吾（監訳）1994 聴覚心理学概論 誠信書房
 (Moore, B. C. J. 1989 *An introduction to the psychology of hearing*. 3rd ed. London: Academic Press.)
- Natale, M. 1977 Perception of nonlinguistic auditory rhythm by the speech hemisphere. *Brain and Language*, **4**, 32–44.
- Neely, J. H. 1977 Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, **106**, 226–254.
- Nosofsky, R. M. 1988 Similarity, frequency and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **14**, 54–65.
- Nusbaum, H. C., & Schwab, E. C. 1986 The role of attention and active processing in speech perception. In E. C. Schwab, & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Vol. 1. Speech perception* New York: Academic Press. Pp. 113–157.
- Nusbaum, H. C., Schwab, E. C., & Sawusch, J. R. 1983 The role of "chirp" identification in duplex perception. *Perception & Psychophysics*, **33**, 323–332.

- Oden, G. C., & Massaro, D. W. 1978 Integration of featural information in speech perception. *Psychological Review*, **85**, 172–191.
- 太田信夫 1991 直接プライミング 心理学研究, **62**, 119–135.
- Papçum, G., Krashen, S., Terbeek, D., Remington, R., & Harshman, R. 1974 Is the left hemisphere specialized for speech, language and / or something else ? *Journal of the Acoustical Society of America*, **55**, 319–327.
- Parker, E. M., Diehl, R. L., & Kluender, K. R. 1986 Trading relations in speech and nonspeech. *Perception & Psychophysics*, **39**, 129–142.
- Pastore, R. E., Li, X. F., & Layer, J. K. 1990 Categorical perception of nonspeech chirps and bleats. *Perception & Psychophysics*, **48**, 151–156.
- Pisoni, D. B. 1971 On the nature of categorical perception of speech. *Supplement to status report on speech research* (SR-27). New haven, CT: Haskins Laboratories.
- Pisoni, D. B. 1973 Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, **13**, 253–260.
- Pisoni, D. B. 1976 Speech perception. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes*. Hillsdale, NJ: Erlbaum. Pp. 167–233.
- Pisoni, D. B., Carrell, T. D., & Gans, S. J. 1983 Perception of the duration of rapid spectrum changes in speech and nonspeech signals. *Perception & Psychophysics*, **34**, 314–322.
- Rand, T. C. 1974 Dichotic release from masking for speech. *Journal of the Acoustical Society of America*, **55**, 678–680.
- Remez, R. E., Pardo, J. S., Piorkowski, R. L., & Rubin, P. E. 2001 On the bistability of sine wave analogues of speech. *Psychological Science*, **12**, 24–29.
- Remez, R. E., & Rubin, P. E. 1990 On the perception of speech from time-varying

- acoustic information: Contributions of amplitude variation. *Perception & Psychophysics*, **48**, 313–325.
- Remez, R. E., Rubin, P. E., Berns, S. M., Pardo, J. S., & Lang, J. M. 1994 On the perceptual organization of speech. *Psychological Review*, **101**, 129–156.
- Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. 1981 Speech perception without traditional speech cues. *Science*, **212**, 947–950.
- Repp, B. H. 1977 Dichotic competition of speech sounds: The role of acoustic stimulus structure. *Journal of Experimental Psychology: Human Perception and Performance*, **3**, 37–50.
- Repp, B. H. 1982 Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, **92**, 81–110.
- Repp, B. H. 1984 Categorical perception: Issues, methods, findings. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice*. New York: Academic Press. Pp. 243–335.
- Repp, B. H., & Liberman, A. M. 1987 Phonetic category boundaries are flexible. In S. Harnad (Ed.), *Categorical Perception*. New York: Cambridge Univ. Press. Pp. 89–112.
- Repp, B. H., Milburn, C., & Ashkenas, J. 1983 Duplex perception: Confirmation of fusion. *Perception & Psychophysics*, **33**, 333–337.
- Rosch, E. 1978 Principles of categorization. In E. Rosch, & B. B. Lloyd (Eds.), *Cognition and Categorization*. Hillsdale, NJ: Erlbaum. Pp. 28–48.
- 佐久間章・高山智行 1986 言語音声の知覚における選択的順応について 言語科学（九州大学教養部言語研究会）， **21**, 57–74.

- Samuel, A. G. 1982 Phonetic prototypes. *Perception & Psychophysics*, **31**, 307–314.
- Sawusch, J. R. 1986 Auditory and phonetic coding of speech. In E. C. Schwab, & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Vol. 1. Speech perception*. New York: Academic Press. Pp. 51–88.
- Sawusch, J. R., & Gagnon, D. A. 1995 Auditory coding, cues, and coherence in phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 635–652.
- Sawusch, J. R., & Jusczyk, P. 1981 Adaptation and contrast in the perception of voicing. *Journal of Experimental Psychology: Human Perception and Performance*, **7**, 408–421.
- Schouten, M. E. H. 1987 Speech perception and the role of long-term memory. In M. E. H. Schouten (Ed.), *The psychophysics of speech perception*. Dordrecht: Nijhoff. Pp. 66–79.
- Schwab, E. C. 1981 Auditory and phonetic processing for tone analog of speech. *Dissertation Abstracts International*, **42**, 3853B. (UMI No. 8204110)
- Searle, C. L., Jacobson, J. Z., & Kimberley, B. P. 1980 Spectral as patterns in the 3-space of time and frequency. In R. A. Cole (Ed.), *Perception and production of fluent speech*. Hillsdale, NJ: Erlbaum. Pp. 73–102.
- Searle, C. L., Jacobson, J. Z., & Rayment, S. G. 1979 Stop consonant discrimination based upon human audition. *Journal of the Acoustical Society of America*, **65**, 799–809.
- 積山薫 1997 異種モダリティ情報の統合 長縄久生・椎名乾平・川崎
恵里子（編） 認知心理学の視点——理論と測定法—— ナカニシ
ヤ出版 Pp. 110–137.

- Shankweiler, D., & Studdert-Kennedy, M. 1967 Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology*, **19**, 59–63.
- Sparks, R., & Geschwind, N. 1968 Dichotic listening in man after section of neocortical commissures. *Cortex*, **4**, 3–16.
- Springer, S. P., & Deutsch, G. 1989 *Left brain, right brain*. 3rd ed. New York: W. H. Freeman and Company.
- Stevens, K. N. 1975 The potential role of property detectors in the perception of consonants. In G. Fant, & M. A. A. Tatham (Eds.), *Auditory analysis and perception of speech*. New York: Academic Press. Pp. 303–330.
- Stevens, K. N., & Blumstein, S. E. 1978 Invariant cues for place of articulation in stop consonants. *Journal of the Acoustical Society of America*, **64**, 1358–1368.
- Stevens, K. N., & House, A. S. 1972 Speech perception. In J. V. Tobias (Ed.), *Foundations of modern auditory theory*. Vol. 2. New York: Academic Press. Pp. 3–62.
- Studdert-Kennedy, M. 1976 Speech Perception. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics*. New York: Academic Press. Pp. 243–293.
- Studdert-Kennedy, M., & Shankweiler, D. 1970 Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, **48**, 579–594.
- Summerfield, Q. 1987 Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd, & R. Campbell (Eds.), *Hearing by eye: The psychology of lip reading*. Hillsdale, NJ: Erlbaum. Pp. 3–51.
- 高山智行 1986 16 ビットパーソナルコンピュータを用いた音声合成システム 言語科学 (九州大学教養部言語研究会), **21**, 33–55.
- 高山智行 1989a スペクトルの／時間的融合による言語音の知覚 九州心理学会第 50 回大会発表論文集, 17.

- 高山智行 1989b 言語音知覚における手掛り間の相互作用——摩擦音—
破裂音—母音音節における破裂音の知覚について—— 日本心理学
会第 53 回大会発表論文集, 693.
- 高山智行 1990 スペクトルの／時間的融合による摩擦音—破裂音—母
音音節の知覚 日本音響学会聴覚研究会資料, **H-90-61**, 1-8.
- 高山智行 1993 言語音構造を模した非言語音刺激の知覚と教示の効果
日本心理学会第 57 回大会発表論文集, 274.
- 高山智行 1996 両耳分離提示された言語音の知覚と示差的特徴の役割
近畿大学工学部研究報告, **30**, 155-168.
- 高山智行 1999 正弦波アナログによる音韻知覚——正弦波アナログの
音韻的識別, 後続アナログ母音による文脈効果—— 日本心理学
会第 63 回大会発表論文集, 356.
- 高山智行 2002a 音韻知覚とプライミング効果 (2) ——プライミング
効果に及ぼす声質の効果—— 中国四国心理学会論文集, **35**, 3.
- 高山智行 2002b 正弦波アナログ音の識別と, その識別学習に及ぼす知
覚様式の効果 近畿大学工学部研究報告, **36**, 51-60.
- 高山智行・佐久間章 1990 二分聴言語音刺激の識別におけるラテラリ
ティ効果 言語科学 (九州大学言語文化部言語研究会), **25**, 126-147.
- Tomiak, G. R., Mullennix, J. W., & Sawusch, J. R. 1987 Integral processing of
phonemes: Evidence for a phonetic mode of perception. *Journal of the Acoustical
Society of America*, **81**, 755-764.
- 利島保・富永大介・高山智行 1980 Ear Asymmetry と大脳半球における
機能的非対称性 広島大学教育学部紀要, 第 1 部, **29**, 133-143.
- 津崎実 1994 第 5 章 音声の知覚・認知 5.2 音韻の知覚とモデル ATR
国際電気通信基礎技術研究所 (編) 視聴覚情報科学—人間の認知の

本質にせまる — オーム社 Pp. 209–229.

Walley, A. C., & Carrell, T. D. 1983 Onset spectra and formant transitions in adult's and children's perception of place of articulation in stop consonants. *Journal of the Acoustical Society of America*, **73**, 1011–1022.

Warren, R. M., Obusek, C. J., Farmer, R. M., & Warren, R. P. 1969 Auditory sequence: Confusion of patterns other than speech or music. *Science*, **164**, 586–587.

Weiss, M. S., & House, A. S. 1973 Perception of dichotically presented vowels. *Journal of the Acoustical Society of America*, **53**, 51–58.

謝 辞

広島大学大学院教育学研究科博士課程後期に社会人入学して以来，本論文の作成にあたり，多くの方々に支えていただきました。

主任指導教官の宮谷真人教授をはじめ，利島保教授，山崎晃教授，深田博己教授には，論文作成でご指導いただきましたことに加えて，あらためて研究の厳しさを実感し，今後の研究への姿勢を見直す機会を与えていただきました。心より御礼申し上げます。また教育人間科学専攻の関係の先生方にも，研究に対して貴重なご意見を頂戴し，深く感謝いたしております。

学校法人近畿大学，近畿大学工学部におかれましては，社会人入学にあたり多大のご配慮をいただき，関係各位の皆様に厚く御礼申し上げます。所属する近畿大学工学部情報システム工学科 学科長 黒瀬能事教授をはじめ，学科の先生方にもあたたかく見守っていただき，心より感謝いたしております。

その他，広島大学大学院認知心理学研究室の学生の皆様，実験に協力してくれた近畿大学工学部の学生諸君にも感謝いたします。

これを新たなスタートとして，更なる目標を持ち，教育者として研究者として努力していく所存です。今後ともご指導のほどよろしく願いいたします。

平成16年1月

高山 智行