

非言語音の音韻知覚における音韻記憶表象の役割

高山 智行

(2003年9月30日受理)

Contribution of phonemic representation in long-term memory to
phonemic perception of nonspeech sounds

Tomoyuki Takayama

To confirm the hypothesis that a phonemic perception is a process in which perceptual representation of speech signals is matched to phonemic representation in long-term memory, performances in identification learning of sinewave analogs of speech sounds were compared under three different instructions. Six sinewave analog stimuli were constructed, using three time-varying sinusoids, to imitate formant structures of vowel sounds /a/, /i/, and /u/ in a female and a male voice. These are usually not perceived as involuntarily speech sounds. Participants were randomly assigned to three instructions groups: speech instructions group with matched labels, who were told to identify analogs using vowel labels matched to original speech, speech instructions group with mismatched labels, who were told to identify them using mismatched vowel labels, and nonspeech instructions group, who were told to identify them as inharmonic nonmusical cords using arbitrary labels, "A", "B", and "C". It was shown that performance was improved through repetition of identification in all instructions groups, but that, with regard to performance level, speech instructions group with matched labels was superior to other instructions groups, and speech instructions group with mismatched labels was the most inferior. These results replicated previous findings (Takayama, 2002a). In addition, it was found that voice of original vowels or tokens of analog vowels influenced identification both in speech instructions group with mismatched labels and nonspeech instructions group, but not in speech instructions group with matched labels. Therefore, it could be concluded that matching to phonemic representation in memory might occur at more abstract or categorical level of representation, not at the level where acoustic details of speech sounds were preserved.

Key words: sinewave analogs of vowel sounds, speech and nonspeech perception, phonemic representation

キーワード：正弦波アナログ母音，言語音知覚と非言語音知覚，音韻記憶表象

本論文は，課程博士候補論文を構成する論文の一部として，以下の審査委員により審査を受けた。

審査委員：宮谷真人（主任指導教官），利島 保，
山崎 晃，深田博己

言語音とは、言語内容を伝達するために主として人間の発話器官により発せられた音を言い、特に、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を音韻と言う。過去多くの研究が、言語音の知覚と非言語音の知覚、あるいは音韻知覚と言語音そのものやその構成要素の音の大きさ・高さといった聴覚的側面の知覚との間に見られる知覚様式の質的な差異を指摘している(e.g., Liberman, 1982; Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967; Liberman & Mattingly, 1985; Repp, 1982)。

例えば、言語音の知覚は、音響的变化の弁別が言語音の音韻的同一性に左右される(e.g., Liberman et al., 1967)という意味においても、また音韻的同一性が複数の音響的手がかりの相補的な関係で実現されているため、音響構成の異なる言語音間で知覚的等価性が成立する(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 と呼ばれる, Best, Morrongiello, & Robson, 1981; Fitch, Halwes, Erickson, & Liberman, 1980; Repp, 1982)という意味においても、カテゴリー的である。他方、非言語音の知覚においては、識別されるカテゴリーの数よりもはるかに多くの信号が弁別可能である(e.g., Liberman et al., 1967)。この他にも以下に示すように、言語音知覚の特異性や言語音と非言語音の知覚的差異を示すさまざまな実験的証拠が報告されている。すなわち、言語機能の左大脳半球優位性との関連性が仮定される両耳分離聴取下での音韻知覚の右耳有利性(right ear advantage, REA と略す, e.g., Kimura, 1961; Studdert-Kennedy & Shankweiler, 1970)、単一の音響パターンが音響手がかりとして言語音の知覚を生起させると同時に、それとは別の非言語音としても知覚され、言語音の知覚と非言語音の知覚が別々のモジュールの機能であることを示唆する二重知覚 duplex perception (e.g., Mann & Liberman, 1983; Rand, 1974)、言語音の音韻内容の知覚が発話者の口の動きに影響されるという言語音情報の視聴覚的統合(McGurk 効果として知られている, McGurk & MacDonald, 1976)、そして言語音のフォルマント構造をその中心周波数に対応する周波数変調音で模擬した正弦波アナログ音の知覚において、それを言語音として聞くか非言語音として聞くかに依存して生じる知覚様式の違い(Best et al., 1981; Best, Studdert-Kennedy, Manuel, & Rubin-Spitz, 1989; Johnson & Ralston, 1994)、などである。

言語音の音響的側面そのものの知覚は、他の非言語音の知覚と同様、基本的には心理物理学的法則に従う聴覚的現象であるので(e.g., Mattingly, Liberman, Syrdal, & Halwes, 1971)、言語音知覚あるいは音韻

知覚に特異的な現象は、それに特殊化したメカニズムの働きを反映すると考えられてきた。例えば、言語音知覚の運動理論(e.g., Liberman et al., 1967)あるいはその改訂版(e.g., Liberman, 1982; Liberman & Mattingly, 1985)においては、言語音知覚が言語音生成と共通基盤をもつ生物学的に特殊化した過程の働きによるとされ、知覚は調音(発声)事象を反映すると主張される。また、直接知覚理論(e.g., Fowler, 1986)は、音の知覚は本質的に音源の知覚であり、その意味で知覚対象は調音動作であって、この対象についての情報は音響信号とそれが提示される環境から直接抽出されるとする。他方、情報処理モデル(e.g., Sawusch, 1986; Studdert-Kennedy, 1976)の立場では、言語音の知覚過程が、音響信号に対する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からなり、言語的処理よりも前の段階、すなわち音響的特徴の分析や一般的な聴覚的処理は言語音処理と非言語音処理に共通する、とされる。

ここで注目すべきは、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自動的に導かれるというわけではなく、聴者の聴取方略や注意のあり方によって左右されることがあるという点である(e.g., Nusbaum & Schwab, 1986)。すなわち、言語音と非言語音との区別は、信号の音響特性あるいはその内的表象と、それをどのように処理するかに関わる聴者の知覚様式とにより決定される。

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を示す代表的な刺激が、先述した正弦波アナログ音である。これは、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的複合音であるが、教示等で聴者にその複合音を言語音と聞かせることにより、従来から言語音に特異的とされてきた知覚的効果を確認することができる。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)。加えて、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ている。

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をもたないで、言語音とは非常に異なって聞こえる。事実、Remez, Rubin, Pisoni, & Carrell (1981)による文を模擬した正弦波アナログ音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、文全体あるいは文中の多くの音節を正しく識別できた。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言語音的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、その後の研究において、アナログ音の音韻知覚をもたらす言語音関連特性が、主に周波数スペクトルの時間的変動であり、ある程度振幅次元の変動も寄与していることを示した (Remez & Rubin, 1990)。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。

しかし、アナログ音に内在する言語音関連特性がスペクトル次元や振幅次元での時間的変動であるにしても、注意あるいは構えを媒介としてそれらが言語音的に処理されるか、非言語音的に処理されるかが決まるのであれば、その音韻知覚過程が刺激規定的とは言えないことは明らかである。そこには、それら音響的音韻関連特性を長期記憶内に保持されている音韻表象に結びつける過程が想定されなければならないであろう。すなわち、聴者がアナログ音から音韻内容を知覚できるのは、それがもつ音響的な言語音関連特性あるいはその内的表象(知覚表象)と、長期記憶内の言語音表象との対応づけが可能であるためであると推測される。そのような対応づけができなければ、非言語音として知覚されるはずである。

高山(2002a)は、有声破裂子音-母音音節「バ、ダ、ガ」を模した正弦波アナログ音の識別学習過程を、言語音教示と非言語音教示のもとで比較した結果に基づいて、この音韻記憶表象との対応づけ過程の存在を主張している。言語音教示には2通りの条件が設けられ、別群の参加者が、アナログ音をもとの言語音に対応する音韻名を用いて識別するよう求められる条件(言語音教示一致群)と、アナログ音ともとの言語音の音韻名との対応を全く変えた状態で識別が求められる条件

(言語音教示不一致群)で識別学習を行った。他方、非言語音教示群に割り当てられた参加者は、歪んだ和音として予め実験者が用意したラベルを用いてアナログ音の識別を学習した。なお、いずれの聴取群も識別に対する正誤のフィードバックは与えられなかった。

実験の結果、言語音教示一致群は、アナログ音と音節名との連合学習の繰り返しにともない識別成績も上昇し、非言語音教示群もわずかながら学習効果を示した。他方、言語音教示不一致群の成績は偶然レベルで推移していた。

高山(2002a)はこれらの結果を、音韻知覚過程が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るパターン認識過程である、と仮定して解釈した。すなわち、言語音教示のもとでは聴者の注意はアナログ音の言語音関連特性に向けられることになり、そこで識別に適切な音韻ラベルが与えられていると、アナログ音の知覚表象は、識別すべき候補として長期記憶内で活性化されている言語音の記憶表象と照合可能となるであろう。結果として適切な音韻知覚が可能となって、識別学習が着実に進行したものと考えられる。しかし、不適切な音韻ラベルが与えられた場合、アナログ音の知覚表象と音韻ラベルにより活性化されている記憶表象との照合が混乱することになる。その結果、言語音教示不一致群では、アナログ音の識別が非常に困難あるいは不可能となって、学習系列を重ねても偶然レベルの識別成績しか得られなかったであろう。他方、非言語音教示のもとでは、照合されるべき内的表象は不定であり、学習が成立するためには与えられたラベルとアナログ音(の知覚表象)との新たな連合が形成される必要がある。何らフィードバックが与えられない状況ではこの連合の形成は困難であると予想されるので、識別学習は困難な課題となり、結果的に非言語音群の識別学習は、言語音教示一致群ほどには進まなかったと考えられる。

高山(2002a)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、ということ強く示唆している。本研究の目的は、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ことである。

音韻知覚において、高山(2002a)のアナログ音が模擬する破裂子音は、その手がかりとなる音響的特徴が後続の母音の影響を強く受けるという意味で文脈依存的であり、強く符号化された言語音であるとされる(Liberman et al., 1967)。破裂音とは対照的に、相対的に発話文脈によるスペクトル変動の少ない母音は、

符号化されていないか符号化の程度が弱いとされる。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きた。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、REAを生じやすいが、母音のように文脈変動の少ない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(e.g., Eimas, 1963; Liberman et al., 1967; Studdert-Kennedy & Shankweiler, 1970)。

高山(2002a)の実験でも、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のかもしれない。それ故、本研究では、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確認することを第1の目的とした。

また、高山(2002a)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と考えられる。本研究の第2の目的は、この点を確認することにあった。すなわち、高山(2002a)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、本研究では課題場面に主に2つの変更を加えた。高山(2002a)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。これは、一部、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に起因したのかもしれない。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。このことは、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せる。すなわち、フィードバックを与えることは、高山(2002a)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、両者は必ずしも相関しない。高山(2002a)が解釈したように、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。

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、高山(2002a)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。

方 法

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。これらの参加者のいずれにも実験実施上問題となるような聴力障害はなかったが、参加者のうち4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。

刺激 刺激音は、男声と女声の合成母音「ア、イ、ウ」(高山, 2002b)のフォルマント構造を、その第1フォルマント~第3フォルマントの中心周波数に対応する正弦波で模擬した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1に示す。

これらアナログ音は、パーソナルコンピュータと自作ソフトウェアを用いて11.025kHz, 16bitの精度で合成した。持続時間は200ms, 波形包絡線の立ち上がり減衰は20msとなるようにした。

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。用いたカテゴリー名は、各アナログ音のもととなった母音名(言語音教示一致群), もとの言語音と一致しないように対応を変えた母音名(言語音教示不一致群), あるいはA, B, Cという

表 1. 女声と男声の正弦波アナログ母音の構成

	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

任意のカテゴリ名（非言語音教示群）のいずれかであり、参加者はこれらの教示群にランダムに振り分けられた。

言語音教示一致群と言語音教示不一致群の聴者には、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、非言語音教示群の聴者には、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群の聴者に対して、学習すべき各カテゴリには2つの同名異音があることも併せて教示した。

被験者は、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(300-500msの範囲で100ms間隔)で刺激音が提示された。

ここでの被験者の課題は、刺激音を聞いて、その刺激音のカテゴリ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、カテゴリ名のキーへの割り当ては被験者毎にランダムであった。被験者の選択したカテゴリ名と選択反応時間、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、被験者の反応直後に、正誤のメッセージと、200ms以下(尚早反応)あるいは1000ms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1s間CRT上に提示された。これらのメッセージが消えた2.5s後に次の試行が始められた。試行系列の間には少なくとも10s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。

実験は基本的にはコンピュータにより制御され、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、予めプログラミング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。また、刺激音は、コンピュータ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提示された。

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1に、各群における各系列での平均識別率を図2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×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1参照)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 $F_{(2,186)} = 47.888, p < .001$), 多重比較(Ryan法)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、アナログ音をそのもと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た。

教示群×系列の交互作用(図2参照)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

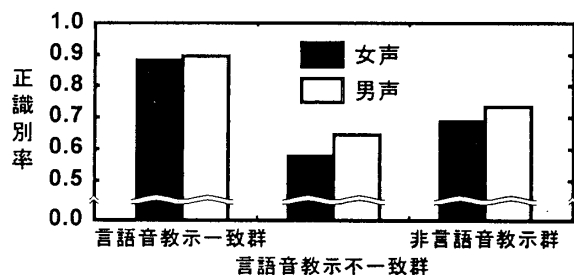


図 1. 各教示群における声質別の正識別率

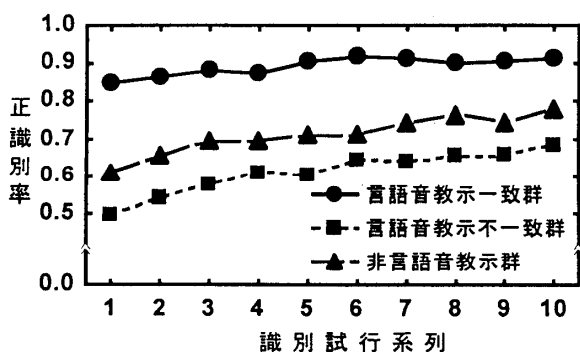


図2. 各教示群における識別学習にともなう正識別率の変化

$F_{(9, 837)} = 4.550, p < .001$; 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Scheffe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, 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であり($MSE=0.020, df=837, p < .05$), 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, さらに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。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反応時間について, 200ms以下のものを尚早反応, 1000ms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名の被験者の結果が以下の分析で用いられた。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3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 $F_{(9, 765)} = 7.53, p < .001$; 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3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あるいは第3系列以降の反応時間が大きく改善されていることを反映している。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

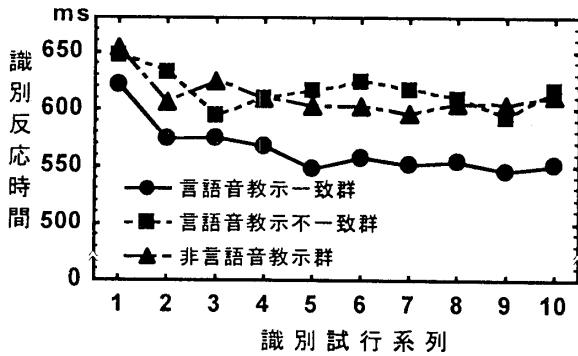


図3. 各教示群における識別学習にともなう反応時間の変化

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 ($F_{(16,680)} = 1.743, p < .05$) が有意となり、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 Scheffe 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。

考 察

正弦波アナログ音は、言語音のフォルマント構造をその中心周波数に対応する周波数変調音で模擬した非調波複合音である。通常、このような複合音からもとの言語音の音韻内容が自発的に知覚されることはないが、歪んだ言語音として聞くと教示や期待のもとでは、多くの場合、もとの言語音の音韻内容を知覚することができる (e.g., Remez et al., 1981)。このことは、音韻知覚が、信号の発生源や信号の音響特性に規定される生得的に固定された受容システムというより

も、刺激のスペクトル的・時間的構造とそれに対する処理との相互作用により規定される能動的な過程であることを示唆する。

高山(2002a)は、破裂子音音節に関する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。この結果から、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推測した。

本研究では、高山(2002a)において破裂子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、その結果を解釈するために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。その結果、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。これは、高山(2002a)の結果を再現するものである。

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。他方、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。

加えて、言語音教示不一致群と非言語音教示群では、もとの母音の声質が識別成績に影響したが、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生じると考えられる。

ところで、本研究の結果あるいは高山(2002a)の結果は、厳密に言えば、アナログ音に対する音韻知覚と言語音に対する音韻知覚との関係について直接言及できるものではない。このことは、アナログ音を用いて音韻知覚の研究を行っている多くの研究にも当てはまることである(Best et al., 1981, 1989; Grunke & Pisoni, 1982; Johnson & Ralston, 1994; Remez et al., 1981; Schwab, 1981; Tomiak et al., 1987)。Takayama(2000)はこの点に着目し、アナログ音に関する言語音教示と非言語音教示のもとで、アナログ音と言語音との間のプライミング効果を比較した。プライミングは、先行刺激がそれに関連のある後続刺激の認知を促進する現象であり、先行刺激が記憶表象を活性化することによると仮定されている。実験の結果、アナログ音が言語音として知覚されている状況で、アナログ音とそのもととなった言語音との間に相互プライミングが見いだされ、Takayama(2000)はそれを、アナログ音と言語音が共通の音韻記憶表象にアクセスすることを示すものと解釈した。

過去多くの研究が、言語音の知覚と非言語音の知覚の質的な差異あるいは言語音知覚の特異性について論じている(e.g., Liberman, 1982; Liberman et al., 1967; Mann & Liberman, 1983)。それらの多くは、それぞれの知覚的特質が聴者の知覚様式、すなわち言語音(音声)様式と非言語音(非音声)様式、を反映するものとし、中にはそれらがカプセル化された機能モジュールにより実現されていると論じるものもある(Liberman & Mattingly, 1985)。本研究並びに高山(2002a)の結果は、このような差異が音響信号と長期記憶内の音韻表象との照合可能性の有無によることを示唆している。他方、この音韻記憶表の性質については、運動理論が示唆するように、調音動作に関連するものであるのか、音響的な構造特性を維持したものであるかは明らかではない。この点は音韻知覚モデルの中核をなす問題であり、今後検討すべき課題の一つである。

【注】

- 1) 本研究の結果は、2003年度日本認知心理学会第1回大会において発表された。

【引用文献】

- Best, C. T., Morrioniello, 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 pattern. *Perception & Psychophysics*, 45, 237-250.
- Eimas, P. D. (1963). The relation between identification and discrimination along speech and nonspeech continua. *Language and Speech*, 6, 206-217.
- Fitch, H. L., Halwes, T., Erickson, D. N., & 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.
- Grunke, M. E., & Pisoni, D. B. (1982). Some experiments on perceptual learning of mirror-image acoustic patterns. *Perception & Psychophysics*, 31, 210-218.
- Johnson, K., & Ralston, J. V. (1994). Automaticity in speech perception: Some speech/nonspeech comparisons. *Phonetica*, 51, 195-209.
- Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Canadian Journal of Psychology*, 15, 166-171.
- Liberman, A. M. (1982). On finding that speech is special? *American Psychologist*, 37, 148-167.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431-461.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1-36.
- Mann, V. A., & Liberman, A. M. (1983). Some differences between phonetic and auditory modes of perception. *Cognition*, 14, 211-235.
- Mattingly, I. G., Liberman, A. M., Syrdal, A. K., & Halwes, T. (1971). Discrimination in speech and nonspeech modes. *Cognitive Psychology*, 2, 13-157.
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
- 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* (pp. 113-157). New York: Academic Press.
- Rand, T. C. (1974). Dichotic release from masking for speech. *Journal of the Acoustical Society of America*, 55, 678-680.
- Remez, R. E., & Rubin, P. E. (1990). On the perception of speech from time-varying attributes: Contributions of amplitude variation. *Perception & Psychophysics*, 48, 313-325.
- 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. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, 92, 81-110.
- 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* (pp.51-88). New York: Academic Press.
- Schwab, E. C. (1981). Auditory and phonetic processing for tone analog of speech. *Dissertation Abstracts International*, 42, 3853B.(UMI No. 8204110)
- Studdert-Kennedy, M. (1976). Speech Perception. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 243-293). New York: Academic Press.
- Studdert-Kennedy, M., & Shankweiler, D. (1970). Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, 48, 579-594.
- Takayama, T. (2000). Priming effects in speech and nonspeech mode of perception. Abstracts of the XXVII International Congress of Psychology, *International Journal of Psychology*, 35, 53.
- 高山智行 (2002a). 正弦波言語音アナログの識別と、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 (2002b). 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.
- 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.

(主任指導教官 宮谷真人)