

音韻知覚と音韻記憶表象

高山智行

A phonemic perception as a matching to phonemic representation in long-term memory

Tomoyuki TAKAYAMA

Previous studies (Takayama, 2002, 2003) indicated that, in the identification learning of sinewave analogs of stop CVs or vowels, performance under speech instructions using phonemic labels matched to analogs was superior to performance under nonspeech instructions using arbitrary labels, which was better than performance under speech instructions using mismatched phonemic labels. Present study aimed to replicate these findings using analogs of other speech sounds, Japanese "わ" and "ら" in the first experiment and American English /la/ and /ra/ in the second experiment. It was predicted that the superiority of matched speech instructions over other instructions (and nonspeech instructions over mismatched speech instructions) should be found in learning on analogs of native syllables, but that, in learning on analogs of nonnative syllables, no difference should be among instructions groups. These predictions were mostly confirmed. The hypothesis that a phonemic perception is a process in which perceptual representation of signals is matched to phonemic representation in long-term memory was discussed.

Keywords: sinewave speech, identification learning, phonemic perception, phonemic representation

1. はじめに

これまで多くの研究が、長期記憶要因が言語音知覚に影響を及ぼすことを指摘している^{1),2),3)}。それらの中には、内的標準あるいはプロトタイプが、言語音信号と知覚された音韻カテゴリーとの間に介在していると主張するものもあり^{4),5),6)}、その主張は、さらに、音韻記憶表象を参照できるか否かが、言語音と非言語音の知覚を規定する要因であることを示唆する。

音韻知覚における音韻記憶表象の寄与を示唆する現象の一つは、正弦波アナログ音に対する言語音知覚の成立である。正弦波アナログ音は、言語音のフォルマント構造を周波数変調音で模した非言語複合音であり、聴者の構えに応じて言語音様にも非言語音様にも知覚される⁷⁾。

そのため、この種の音は言語音知覚メカニズムを探る有力な刺激音として用いられてきた^{8),9),10),11),12),13)}。しかし、何故、同じ非言語音刺激の知覚内容が構えによって変わるのか、ということについては十分に説明されているとは言えなかった。

高山^{14),15)}は、アナログ音の音韻知覚が、刺激音の音響的音韻関連特性あるいはその内的表象(知覚表象)と長期記憶内の音韻表象との照合により成立すると仮定して、識別ラベルを操作したアナログ音の識別学習実験を行った。実験では、破裂子音-母音音節(破裂音 CV音節と略す)¹⁴⁾あるいは母音¹⁵⁾を模したアナログ音を用い、アナログ音の識別ラベルとしてもとの音韻名を与える被験者群(言語音教示一致群)、アナログ音に本来の対応を変えた音韻ラベルを与える被験者群(言語音

教示不一致群), 非言語音として任意のラベルを与える被験者群(非言語音教示群)の3群の間でアナログ音の識別学習過程が比較された。その結果, 言語音教示一致群で最も高い学習成績が示され, ついで非言語音教示群での成績が言語音教示不一致群よりも高いことが示された。また, 声質の違いによる音響変動は, 適切な音韻ラベルが与えられた場合(言語音教示一致群)には識別成績にほとんど影響しなかった¹⁵⁾。これらの結果は, アナログ音の知覚表象と音韻記憶表象とが照合可能であるときに限り適切な音韻知覚が成立すること, そしてその照合は刺激の音響変動が捨象された段階で生じることを示すものと解釈された。

本研究の目的は, 新たな音韻セット(日本語音韻「わ」・「ら」: 実験Ⅰ)を用いてこれまでの結果を確認するとともに, 日本語にない音韻セット(米語音韻 /la/・/ra/: 実験Ⅱ)を用いて, 先の効果が聴者の言語経験に基づくことを示すことであった。

2. 実験Ⅰ 日本語音韻「わ」・「ら」を模した正弦波アナログ音の識別学習

2.1 目的

まず, 日本語音韻「わ」, 「ら」を模したアナログ音を合成し, これまでの研究^{14), 15)}同様, 聴取の構えと識別ラベルを操作した3つの教示群間でアナログ音の識別学習過程を比較した。その際, 高山¹⁵⁾で見いだされたように, 発話者の性別の違いにより生じた識別変動が, 言語音教示一致群で縮小もしくは消失するか否かも併せて検討した。

2.2 方法

2.2.1 刺激

日本人男女が発声する「わ」と「ら」を参考にして, その第1~第3フォルマントに近似する周波数変調音で模擬した4種の正弦波アナログ音を刺激として用いた。図1にその周波数構成を図式的に示す。女声「ら」を模したアナログ音の3つの構成音の起点周波数は, 下から450, 1550, 3250Hz, 女声「わ」の場合はそれぞれ450, 1150, 2850Hzで, 130msの間にそれぞれの定常部の870, 1400, 3100Hzまで直線的に変化した。また, 男声「ら」ではそれぞれ300, 1300, 2500Hz, 男声「わ」では300, 1000, 2200Hzで, 130msの間にそれぞれの定常部の740, 1200, 2400Hzまで直線的に変化した。これらのアナログ音は, 3つの構成音の相対振幅を下から0, -3, -6dBに設定し, 持続時間が300msとなるよう, 16bit, 11.025 kHzの精度でソフトウェアにより合成した。なお, 合成後, 4音の150~250ms区間の実効値振幅が等しくなるように, また波形の終末部の50msが0に収束するように, 波形編集を行った。

2.2.2 手続き

実験参加者は, 高山^{14), 15)}同様, 識別ラベルとしてアナログ音のもとなった言語音の音韻名が与えられる言語音教示一致群, 異なる言語音の音韻名が与えられる言語音教示不一致群, 非言語音としてAとBというラベルが与えられる非言語音教示群のいずれかにランダムに割り当てられた。

実験での課題は, あらかじめ与えられた2つのラベルによるアナログ音の分類を学習することであった。まず最初に, 参加者は, 3.7sの間隔で順々に呈示される4つのアナログ音を, そのラベル名を確認しながら5回繰り返し聞いた。アナログ音のラベル名は, アナログ音呈示直前から参加者正面に置かれたモニター画面上に呈示され, アナログ音呈示後1.4s後まで表示されていた。

続いて, 10系列の識別学習を行った。各系列では, 初めに4つのアナログ音がラベル名とともに2回ずつ3.7sの間隔で順々に呈示され, 参加者はアナログ音とラベル名との連合を確認した。それから, 4つのアナログ音がランダムに6回ずつ呈示された。各識別試行では, 刺激音がこれから提示される旨のメッセージがモニター上に提示された後, 300, 400, 500msのいずれかのISIでアナログ音が提示された。

参加者は, アナログ音を聞いて, その音のラベル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よう求められた。ラベル名のキーへの割り当ては参加者毎にランダムであった。参加者の選択したラベル名と選択反応時間, 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, 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間モニター上に提示された。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。

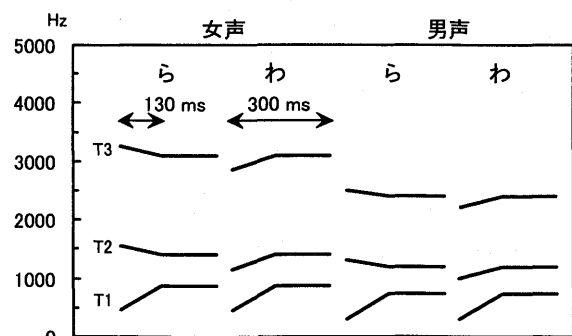


図1 日本語音韻「ら」・「わ」アナログのスペクトル構造

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、アナログ音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参加者プロフィールの入力、教示の補足、実験終了後の参加者の内省の記録を行った。また、アナログ音は、コンピュータの出力をヘッドフォンアンプ (Audio-technica AT-HA2)、ヘッドフォン (Audio-technica ATH-A900) 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。

2.2.3 参加者

近畿大学工学部在籍の1-4年生60名(男44名,女16名,年齢範囲18~22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も,正弦波アナログ音を聞いた経験はなく,また実験実施上問題となるような聴力障害はなかった。

2.3 結果

まず,アナログ音の識別の正確さについての分析を行った。教示群および声質ごとに学習経過を比較したものが図2である。正識別率の角変換値を用いて教示群×声質×系列の分散分析を行った結果,教示群 ($F(2,57) = 13.891, p < .001$) と系列 ($F(9,513) = 19.989, p < .001$) の主効果,教示群×系列 ($F(18,513) = 1.699, p < .05$) と声質×系列 ($F(9,513) = 2.780, p < .01$) の2つの交互作用が有意であった。

教示群の効果について,教示群×系列の交互作用における単純主効果を検討したところ,第2系列を除くすべての系列において有意であった(第1系列 $F(2,570) = 5.505, p < .01$; 第2系列 $F(2,570) = 1.086, ns$; 第3系列 $F(2,570) = 4.781, p < .01$; 第4系列 $F(2,570) = 9.891, p < .001$; 第5系列 $F(2,570) = 8.636, p < .001$; 第6系列 $F(2,570) = 8.639, p < .001$; 第7系列 $F(2,570) = 17.989, p < .001$; 第8系列 $F(2,570) = 7.026, p < .001$; 第9系列 $F(2,570) = 12.353, p < .001$; 第10系列 $F(2,570) = 7.651, p < .001$)。そして,教示群の単純主効果が有意となった各系列における多重比較 (Ryan 法) の結果は,すべての系列において言語音教示一致群の識別成績が他の2群よりも有意に高い,ということを示した ($MSE = .062, df = 570, p < .05$)。

次に,教示群×系列の交互作用について,各教示群における系列による変動の傾向分析を行った。その結果,いずれの群で直線傾向が有意であり(言語音教示一致群 $F(1,513) = 95.463, p < .001$; 言語音教示不一致群 $F(1,513) = 32.038, p < .001$; 非言語音教示群 $F(1,513) = 52.081, p < .001$), 言語音教示一致群では2次の傾向も有意であった ($F(1,513) = 8.925, p < .01$)。直線傾向について傾きの程度を比較するために,教示群ごとに系列間で識別率に関する多重比較 (Ryan 法) を行った。言語音教示一致群では,第4系列以降の成績が第1・2系列よりも有意に高く,第7系列以降は第3系列よりも有意に識別成績が上昇していた ($MSE = .031, df = 513, p < .05$)。他方,言語音教示不一致群では,第7系列以降でのみ第1系列での識別成績を有意に上まわったにすぎなかった ($MSE = .031, df = 513, p < .05$)。このことは,言語音教示不一致群は言語音教示一致群ほどには学習による識別の改善が進まないことを示している。

非言語音教示群での同様の分析では,識別成績は,第8系列以降で第1・2系列を,第9系列以降で第3系列を,第10系列で第4・5系列を有意に上まわっており ($MSE = .031, df = 513, p < .05$), 比較的順調に識別が改善されていることが示された。

声質の効果については,声質×系列の交互作用における単純主効果を検討したところ,一部の系列で有意となったが(第4系列 $F(1,570) = 7.765, p < .01$; 第7系列 $F(1,570) = 6.312, p < .05$; 第10系列 $F(1,570) = 5.079, p < .05$), 優位になる声質に一貫した傾向は認められなかった。

続いて識別反応時間の分析を行った。識別反応時間を分析するに当たって,まず,各参加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参加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17名,言語音教示不一致群では14名,非言語音教示群では15名の被験者の結果が以下の分析で用いられた。

これらの参加者の各系列における声

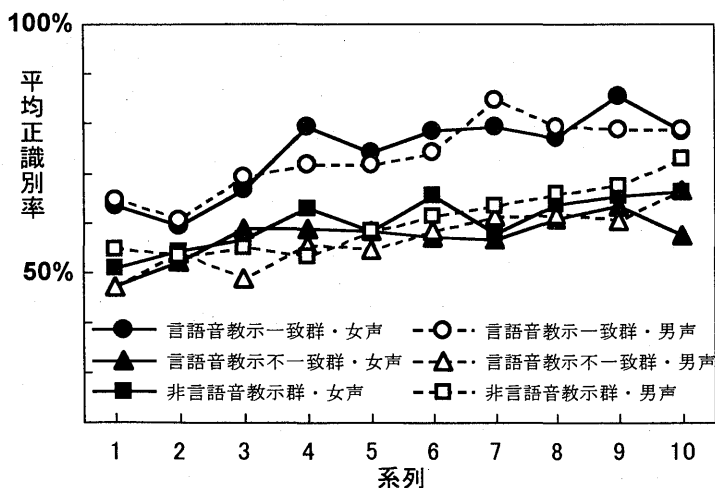


図2 実験Iの各教示群における声質毎の平均正識別率の推移

これらの参加者の各系列における声

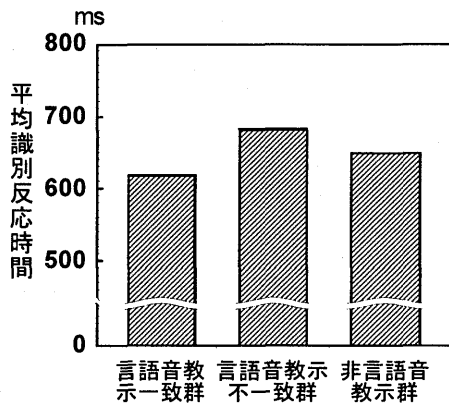


図3 実験Iの各教示群の平均識別反応時間

質別の識別反応時間の中央値の対数変換値を用いて、教示群×声質×系列の分散分析を行った結果、教示群の主効果のみ有意であった(図3, $F(2,43)=3.505, p<.05$)。教示群間の多重比較(Ryan法)を行ったところ、言語音教示一致群の識別反応時間が言語音教示不一致群に比較して有意に短かった($MSE=0.041, df=43, p<.05$)。

2.4 考察

日本語音韻「わ」、「ら」を模した正弦波アナログ音の識別学習において、これまでの結果^{14),15)}同様、正識別率、識別反応時間ともに言語音教示一致群における識別学習の優位性が認められた。他方、平均値としては非言語音教示群の識別率は言語音教示不一致群よりもわずかに高い値であったが、その差は有意水準には達しなかった。また、以前の研究¹⁵⁾で認められた教示群と声質との交互作用は、今回の結果では認められず、識別率について示された声質と系列との交互作用においても、声質に関する効果は明確なものではなかった。

以前の研究結果^{14),15)}で見いだされてきた言語音教示不一致群と非言語音教示群との差、あるいは言語音教示不一致群と非言語音群における声質による差が今回の実験で認められなかった理由の一つとして、これまでは3カテゴリーでの識別を求めているのに対して、今回はカテゴリーを2つに絞ったことがあげられるかもしれない。

3. 実験II 米語音韻 /la/・/ra/を模した正弦波アナログ音の識別学習

3.1 目的

これまでの研究では、用いた日本語音韻の範囲で、アナログ音の音韻知覚が、その知覚表象と音韻記憶表象の照合が可能であるとき成立する、という仮説を支持する結果が得られた。他方、この仮説のもとでは、日本語話

者にとって、日本語にない音韻を模したアナログ音の識別に関しては、異なる予測がなされる。すなわち、アナログ音が呈示される時、照合可能な音韻表象が長期記憶内に存在しないので、与えられる音韻ラベルがアナログ音のもとになった言語音と一致するしないにかかわらず、その識別学習は非言語音の識別学習と等価であり、3つの教示群間に識別学習成績の違いは生じないであろう。

実験IIでは、上記の予測を確かめるために、米語音韻 /la/・/ra/を模した正弦波アナログ音を用いて、これまでと同様に、聴取の構えと識別ラベルを操作した3つの教示群間でアナログ音の識別学習過程を比較した。

3.2 方法

3.2.1 刺激

米国人男女が発声する /la/・/ra/ の分析結果ならびに Sawusch & Gagnon¹⁶⁾ が用いた刺激を参考にして、/la/ と/ra/それぞれの第1～第3フォルマントに対応する周波数変調音で模擬した4種の正弦波アナログ音を識別学習のための刺激材料として用いた。図4にその周波数構成を図式的に示す。女声/la/を模したアナログ音の3つの構成音の起点周波数は、下から500, 1400, 3400Hzで、60ms間その値を保ち、次の40ms間で定常部の870, 1400, 3100Hzにそれぞれ直線的に変化した。女声/ra/の場合は、それぞれ500, 1200, 2100Hzで始まり、20ms間その値を保ち、次の80ms間で定常部の870, 1400, 3100Hzにそれぞれ直線的に変化した。また、男声/la/では、構成音はそれぞれ400, 1200, 2900Hzを起点周波数とし、女声/la/と同じ時間経過を辿って、定常部の740, 1200, 2400Hzに変化した。男声/ra/の場合は、400, 1000, 1400Hzで始まり、女声/ra/と同じ時間経過で、定常部の740, 1200, 2400Hzに変化した。これらのアナログ音は、3つの構成音の相対振幅を下から0, -3, -6dBに設定し、持続時間が300msとなるよう、16bit, 11.025kHzの精度でソフトウェアにより合成した。なお、合成後、4音の100～250ms区間の実効値振幅が等しくなるよう

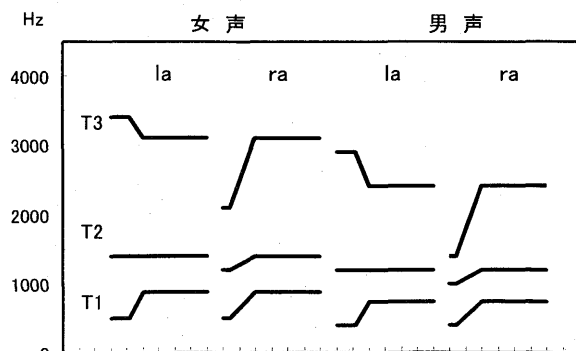


図4 英語音韻 /la/・/ra/のアナログのスペクトル構造

に、また波形の終末部が 50ms 間で 0 に収束するように、波形編集を行った。

3.2.2 手続き

用いたアナログ音と 2 つの言語音教示群に与える識別ラベルを /la/ と /ra/ としたことを以外、実験 I とまったく同様の手続きで実験を行った。

3.2.3 参加者

近畿大学工学部在籍の 1-4 年生 60 名 (男 57 名, 女 3 名, 年齢範囲 18 ~ 24 歳) が実験に参加し、実験 I と同様、言語音教示一致群、言語音教示不一致群、非言語音教示群の 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も、これまで正弦波アナログ音を聞いた経験はなく、また実験実施上問題となるような聴力障害はなかった。

3.3 結果

アナログ音に対する正識別率に関して、角変換を行った後、教示群×声質×系列の分散分析を行った。その結果、声質 ($F(1,57)=11.253, p<.01$) と系列 ($F(9,513)=44.473, p<.001$) の主効果、ならびに教示群×声質の交互作用 (図 5, $F(2,57)=5.355, p<.01$) が有意であった。

教示群×声質の交互作用について単純主効果を検討したところ、いずれの声質においても教示群の効果は認められなかったが (いずれも, $F<1$), 声質の効果は言語音教示一致群において有意であった ($F(1,57)=19.058, p<.001$)。これは、男声に対する識別が女声に対してより良かったためであり、この傾向は非言語音教示群においても認められた ($F(1,57)=2.848, p<.10$)。

また、系列の主効果に関して傾向分析を行ったところ、1 次 ($F(1,513)=376.790, p<.001$) と 2 次 ($F(1,513)=18.827, p<.01$) の傾向が有意であり、教示群にかかわらず正識別率は単調増加していることが示された。

識別反応時間については、実験 I と同様の処理をし、教示群×声質×系列の分散分析を行った。その結果、系列の主効果のみ有意であり ($F(9,423)=7.537, p<.001$)、傾

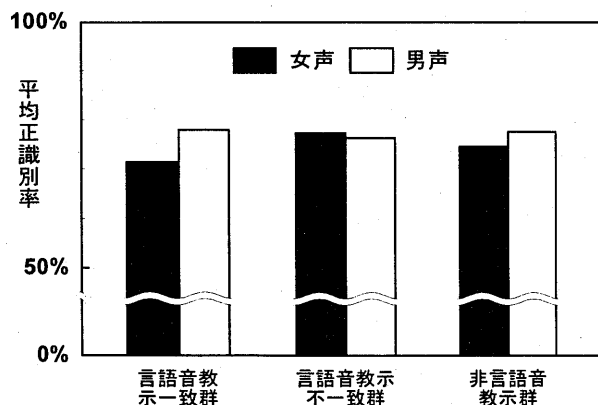


図 5 実験 II の各教示群における声質毎の平均識別率

向分析は 1 次 ($F(1,423)=45.00, p<.001$) と 2 次 ($F(1,423)=4.582, p<.05$) の成分が有意であることを示した。すなわち、いずれの教示群においても、系列を繰り返すとともに識別反応時間は単調に減少した。

3.4 考察

これまでの日本語音韻を用いた実験結果と異なり、日本語にない音韻を用いた場合には、教示群による識別学習の差異は認められなかった。また、声質の効果に関しては、全体として有意な効果が見いだされたが、教示群との交互作用の中で、言語音教示一致群と非言語音教示群においてそれぞれ有意な単純主効果ならびにその傾向が示された。

教示群に関する主効果や単純主効果が認められないということは、聴者の言語にない音韻については、照合されるべき音韻記憶表象自体が存在しないことによるものと解釈できるであろう。

他方、教示群×声質の交互作用における声質の単純主効果に関していえば、この解釈は高山¹⁵⁾の解釈と矛盾する。高山¹⁵⁾の研究では、母音アナログの識別学習において、言語音教示不一致群と非言語音教示群で声質の効果が示され、言語音教示一致群では声質の効果は見いだされなかった。それまでの結果で、識別学習に関する言語音教示一致群の優位性が確認され、アナログ音の音韻知覚がその知覚表象と音韻記憶表象の照合により成立するとの仮説は支持されたと言えるので、言語音教示一致群で声質差がなかったことは、音響変動が捨象され抽象化された段階でアナログ音の知覚表象と音韻記憶表象と照合が生じていることを示すものと解釈された。

この解釈は、声質による識別差が既存の音韻記憶表象との照合が成立しない条件で生じると予想するので、先の解釈が妥当であれば、今回の実験ではいずれの教示群においても声質による識別差が生じることが期待される。しかし、今回の結果では、言語音教示一致群において声質の効果が生じており、声質差が生じなかったのは言語音教示不一致群においてであった。高山¹⁵⁾が述べるように、声質による識別差の無いことが内的表象の抽象性の指標であるとすれば、当該音韻記憶表象とは異なる何らかの内的表象との照合が抽象的段階で生じていたのであろうか。あるいは、全くの偶然によるものであろうか。

ここで、実験手続きとアナログ音のスペクトルパターンから一つの可能性が示唆される。識別学習に関して与えられた音韻ラベルは、実験参加者前方のモニタ画面上に、実験が終了するまでほとんど常に表示されている。今回の実験では、“La”、“Ra”が用いられた。これらをローマ字として読めば、前者に対応する日本語音韻は存在しないが、後者は日本

語音韻の「ら」である。他方、図1と図4のアナログ音のスペクトルパターンを比較すると、日本語音韻「ら」のフォルマント遷移の方向は英語音韻/ra/よりもむしろ/la/と似ており、英語音韻の/ra/は日本語音韻「わ」に似ている。そうすると、言語音教示一致群というよりも言語音教示不一致群において、アナログ音の知覚表象と抽象化された音韻記憶表象の照合あるいは対応づけが生じていたと考えられないだろうか。すなわち、言語音教示不一致群において、アナログ音/la/の知覚表象と、それに与えられた音韻ラベル“Ra”から誘導された音韻記憶表象(ら)とが近似的に対応し、それを手がかりとして識別学習が進行したのかもしれない。両者の対応は識別学習を促進するほど十分ではないものの、対応づけは抽象化された表象を用いて行われたのであろう。この解釈の可能性を検討するために、追加実験を行った。

4. 追加実験 米語/la/・/ra/アナログの日本語音韻による識別学習

4.1 目的

実験Ⅱの言語音教示不一致群の識別学習過程において、意識的か無意識的かにかかわらず、アナログ音のスペクトル構造の時間的変動に類似したフォルマント構造をもつ日本語音韻が手がかりとして用いられていたとすれば、それらの手がかりを明示的に与えることによって学習過程を促進するとともに、声質の影響を抑えることができるかもしれない。追加実験では、米語音韻/la/・/ra/アナログに対して日本語音韻「ら」・「わ」を用いる識別学習の過程を検討した。

4.2 方法

4.2.1 刺激

実験Ⅱで用いた米語/la/、/ra/に関する4種のアナログ音を刺激として用いた。

4.2.2 手続き

4種のアナログ音を日本語「わ」、「ら」として学習させる以外は、実験Ⅱと同様であった。ただし、教示群については、非言語音教示群は実験Ⅱでのそれと等価であるので、刺激音とラベルの対応を入れ替える2群の言語音教示群のみを設け、フォルマント遷移の類似性から、米語/la/を日本語「ら」、米語/ra/を日本語「わ」と識別するよう教示する群を類似言語音教示群、逆の組合せで教示する群を非類似言語音教示群とした。

4.2.3 参加者

近畿大学工学部在籍の1-4年生40名(男33名、女7名、年齢範囲18~22歳)が実験に参加し、2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も、これまで正弦波アナログ音を聞いた経験はなく、また実験実施上問題となるような聴力障害はなかった。

4.3 結果と考察

正識別率について、教示群×声質×系列の分散分析を行った結果、声質($F(1,38)=5.903, p<.05$)と系列($F(9,342)=18.136, p<.001$)の主効果が有意であったが、教示群の効果は認められなかった。すなわち、日本語にない音韻をもとにしたアナログ音の識別学習に関して、識別ラベルとして与えられる日本語音韻のフォルマント遷移のパターンがアナログ音のスペクトル変化の時間的パターンに類似するか否かは、学習効率に影響しなかった。

他方、先の実験と同様の処理をした識別反応時間に関する分散分析の結果においては、教示群×系列の交互作用が有意であった($F(9,324)=2.920, p<.05$)。教示群ごとに系列の単純主効果を検討したところ、非類似言語音教示群で有意であり($F(9,324)=2.074, p<.05$)、傾向分析では1次の減少傾向のみ有意となった($F(1,324)=9.859, p<.01$)。すなわち、図6に示すように、類似言語音教示群においては、学習開始当初から識別反応時間はほぼ一定水準で推移したが、非類似言語音教示群においては、試行系列を重ねるにつれて識別反応時間が減少し、最終的に類似言語音教示群と同じ水準に達した。

この追加実験は、実験Ⅱにおける言語音教示不一致群で声質による識別差が見いだされなかったことに関してなされた推測を確かめるために実施された。正識別率に関する結果はその推測に反するものであったが、識別反応時間に関しては逆にその推測を支持するような結果であった。正識別率と識別反応時間のどちらかが内的な過程に鋭敏な測度であるかは判断が難しいが、実験Ⅰの考察で述べたように、二者択一的な反応カテゴリーを用いたことが正識別率にも現れるはずの差を縮小させたのかもしれない。反応カテゴリーを増やして同様の検討を繰り返す。

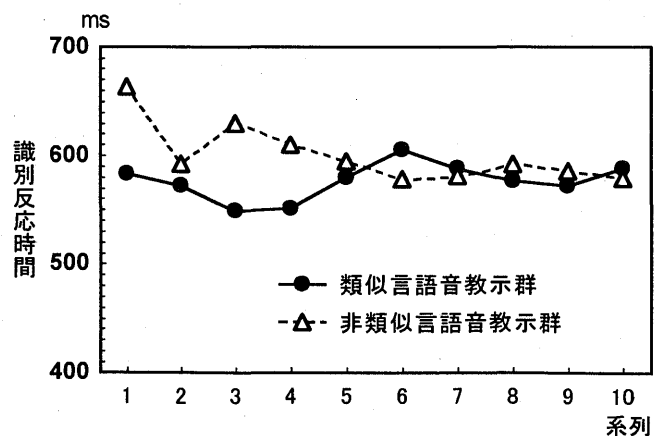


図6 追加実験の各教示群の識別反応時間の推移

返す必要があるであろう。

5. 全体的考察

5.1 正弦波アナログ音に対する音韻知覚

正弦波アナログ音は、その音響エネルギーの周波数-時間変動がもとの言語音と近似するものの、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、言語音とは非常に異なって聞こえる。Remez ら⁷⁾は、男性話者が発声した文のフォルマント構造を分析して、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周波数変調音で模擬した複合音を作成し、音の性質について何の教示も与えず、その音の印象を自由に報告するよう求めたところ、ほとんどの被験者が非言語音としての知覚内容を報告することを示している。他方、同じ実験において、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 ら⁷⁾はそれを、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たことによる、としている。すなわち、正弦波アナログ音から音韻内容を知覚するためには、アナログ音にもと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り、それとともに、聴者の注意がそれに向けられ、言語音情報として適切に処理されることが必要である。その意味で、アナログ音の音韻知覚過程は、決して Johnson & Ralston¹¹⁾が述べているような自動的な過程ではあり得ない。

本研究は、以前の研究^{14),15)}と同様、アナログ音の音韻知覚が、刺激の知覚表象と長期記憶内の音韻表象との照合により成立すると仮説し、アナログ音に与えられるラベルを操作した識別学習実験を通して、その仮説の検証を試みた。これらの実験の論理は次のとおりである¹⁷⁾。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 ら⁷⁾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係に矛盾しない限り(言語音教示一致群)、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係に矛盾する場合(言語音教示不一致群)は、音韻関連特性と音韻記憶表象との対応係に混乱が生じ、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(非言語音教示群)は、そのような既存の係は利用できないので、通常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(言語音教示一致群)と不適切な音韻ラベルを用いて学習する場合(言語音教示不一致群)の中間に位置すると予想される。

加えて、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棄され抽象化された音韻内容であると考えられる。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であろう。すなわち、与えられた音韻ラベルがアナログ音の音韻関連特性と音韻記憶表象との照合において矛盾しないものであれば、声質等の音響的差異はアナログ音の識別のし易さに影響しないであろう。しかし、矛盾した音韻ラベルであったり、非言語音として任意のラベルを用いる場合には、抽象化された表象レベルでの照合ができず、音響変動の影響を受けるかもしれない。

日本語音韻「ら」と「わ」のアナログ音を用いた実験Iの結果は、上記の予想を一部確認した。すなわち、これまでの結果同様、言語音教示一致群での識別学習が他の2群の識別学習よりも優位であった。他方、言語音教示不一致群と非言語音教示群との間には差は認められなかった。加えて、いずれの群においても声質による識別学習の難易度の差も認められなかった。

実験IIでは、さらに、日本語にない音韻である英語音韻/la/と/ra/のアナログ音を用いて、上記の予想を検討した。このようなアナログ音では音韻記憶表象との照合は期待されないので、どのようなラベルが与えられるかは識別学習の結果を左右しないと予想される。結果は予想を支持し、3つの教示群の識別学習の効率に関して違いは認められなかった。他方、音韻記憶表象との照合が期待されないとき、声質による音響的変動の効果が予想され、3つの教示群とも声質による識別学習への影響が期待されたが、言語音教示一致群と非言語音教示群で声質の効果が認められ、言語音教示不一致群では認められなかった。

後者の結果については、本来矛盾した音韻ラベルが逆に日本語音韻の記憶表象と対応づけられると仮定することで、解釈できるかもしれない。この点を確かめるために、英語音韻/la/、/ra/に比較的類似したフォルマントパターンをもつ日本語音韻「ら」、「わ」を音韻ラベルとして明示的に与える追加実験を行った。そこでは、アナログ音のスペクトル構造の時間的変動に類似したフォルマント構造をもつ日本語音韻をラベルとして与えることで、識別学習を促進し声質変動の効果を抑えることができるかもしれないと期待された。実験の結果、フォルマ

ント構造が類似した日本語音韻をラベルとして与えるか否かは、識別の正確さから見る識別学習過程には影響せず、声質変動の効果も抑えることはできなかった。他方、識別の速さから見ると、フォルマント構造の類似した音韻ラベルが与えられる場合には学習当初から比較的速やかに識別がなされるが、類似していない音韻ラベルが与えられる場合は識別が徐々に速やかになされるようになり、最終的に両者が同等の速さで識別されるようになるということが示された。これらの結果は、実験Ⅱの結果に対する後付的解釈を必ずしも支持するものではないが、新奇な音韻の学習における既存の音韻記憶表象の影響の仕方を示唆するものであるかもしれない。更なる検討が必要であろう。

いずれにしても、これらの結果は、アナログ音の音韻知覚が、刺激の知覚表象と長期記憶内の音韻表象との照合により成立するとの仮説を概ね支持するものと言えよう。今後は、この照合過程にかかわる表象の性質を明らかにすることが必要である。

5.2 音韻知覚における正弦波アナログ音と言語音との関係

ところで、本研究も高山^{14),15)}も、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^{9),11)}や音声トレーディング⁸⁾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。先に述べたように、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のであろう。

Takayama¹⁸⁾は、音韻知覚において、言語音とそのアナログ音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。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に何らかの点で似ているとき、記憶内の表象を活性化させると仮定されている^{19),20)}。したがって、プライムする側とされる側の物理的同一性にかかわらず、言語音とそのアナログ音との間に音韻的同一性に基いた相互プライミングが成立するならば、両者が共通の音韻記憶表象にアクセスしていると推測される。実験の結果、この推測は支持された。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であろうと言語音であろうと、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、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、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。すなわち、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と考えられる。

5.3 音韻知覚研究における正弦波アナログ音の意義

先述したように、正弦波アナログ音は、言語音のフォルマント構造を周波数変調音で模した非言語複合音であり、聴者の経験や構えに応じて言語音様にも非言語音様にも知覚される。これまでの多くの研究の主要な関心は、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の時間的に近似している必要があることを指摘している^{7),10),12)}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^{8),9),11),21)}、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。われわれのこれまでの一連の研究の関心はまさにその点にあった。

本研究ならびにこれまでの一連の研究^{14),15),18)}は、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するものであり、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、とする仮説を検討するために行われた。そのような研究目的にとって正弦波アナログ音を用いる利点は、正弦波アナログ音が、言語音をその主要なスペクトルの時間的特性に還元した非言語音であり、聴者の注意や期待・構えあるいは経験に応じて非言語音様にも言語音様にも知覚される、という点にある。

先に述べたように、これまで多くの研究では、アナログ音を、知覚の言語音様式と非言語音様式による知覚内容やその効果の違いを明らかにするための手段として用いてきた。その意味では、言語音知覚過程と非言語音知覚過程を分離する二重知覚²²⁾やラテラルリティ効果²³⁾と同じような位置づけでしかなかった。しかし、同一の刺激が聴者の知覚的処理のあり方に応じて異なる知覚内容あるいは効果をもたらすというアナログ音の特性は、知覚過程やそこでの処理様式を分離するだけでなく、その知覚的処理の内容にまで踏み込んでいけるという可能性を示唆する。すなわち、アナログ音は、通常の言語音では自動化され、隠蔽されているかのように思える言語

音処理過程を、注意により制御された過程²⁴⁾として分析・解明する手段を提供する。この刺激を用いた研究の今後の展開を期待したい。

引用文献

- 1) 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.
- 2) Pisoni, D. B. 1973 Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception & Psychophysics*, 13, 253-260.
- 3) 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.
- 4) 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.
- 5) 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.
- 6) Samuel, A. G. 1982 Phonetic prototypes. *Perception & Psychophysics*, 31, 307-314.
- 7) Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. 1981 Speech perception without traditional speech cues. *Science*, 212, 947-950.
- 8) Best, C. T., Morrongiello, B., & Robson, R. 1981 Perceptual equivalence of acoustic cues in speech and nonspeech perception. *Perception & Psychophysics*, 29, 191-211.
- 9) Best, C. T., Studdert-Kennedy, M., Manuel, S., & Rubin-Spitz, J. 1989 Discovering phonetic coherence in acoustic patterns. *Perception & Psychophysics*, 45, 237-250.
- 10) Grunke, M. E., & Pisoni, D. B. 1982 Some experiments on perceptual learning of mirror-image acoustic patterns. *Perception & Psychophysics*, 31, 210-218.
- 11) Johnson, K., & Ralston, J. V. 1994 Automaticity in speech perception: some speech / nonspeech comparisons. *Phonetica*, 51, 195-209.
- 12) Schwab, E. C. 1981 Auditory and phonetic processing for tone analog of speech. *Dissertation Abstracts International*, 42, 3853B. (UMI No. 8204110)
- 13) 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.
- 14) 高山智行 2002 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.
- 15) 高山智行 2003 非言語音の音韻知覚における音韻記憶表象の役割 広島大学大学院教育学研究科紀要第三部第 52 号, 273-281.
- 16) 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.
- 17) 高山智行 2004 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博士論文 (未公刊) .
- 18) Takayama, T. 2004 Priming effects in speech and nonspeech modes of perception. *Acoustic Science & Technology*, 25, 196-202.
- 19) Decoene, S. 1997 The representational basis of syllable categories. *Perception & Psychophysics*, 59, 877-884.
- 20) 太田信夫 1991 直接プライミング 心理学研究, 62, 119-135.
- 21) 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.
- 22) Mann, V. A., & Liberman, A. M. 1983 Some differences between phonetic and auditory modes of perception. *Cognition*, 14, 211-235.
- 23) Studdert-Kennedy, M., & Shankweiler, D. 1970 Hemispheric specialization for speech perception. *Journal of the Acoustical Society of America*, 48, 579-594.
- 24) 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.