

TR-A-0049

日本人とアメリカ人における  
英語 / r, l, w / 音知覚の手掛かり

山田 玲子 東倉 洋一  
Reiko Yamada Yoh'ichi Tohkura

1989. 4. 6

A T R 視聴覚機構研究所

© (株) A T R 視聴覚機構研究所

# 日本人とアメリカ人における 英語 /r, l, w/音知覚の手掛かり

山田玲子・東倉洋一

## はじめに

日本語には、英語(ここではアメリカ英語)にある /r/、/l/ という音素が存在しない。従って、日本人は英語を学習する際に、これらの音の聴き取り、調音を共に苦手としている。この日本人の英語 /r//l/ 音知覚は、言語環境が知覚に及ぼす影響や、聴く能力と調音能力の関係を調べる上でよい例題となるため、多くの研究報告がある。それらは、大別すると自然音声を刺激音としたものと、合成音を刺激音としたものに分けることができる。

自然音声をを用いた実験では、主として単語の中での位置、前後環境と聴き分けの関係を調べている。一方、合成音を用いる研究では、刺激音の音響的手掛かりを自由に設定することができるため、知覚様式に関する詳細な検討が可能である。そこで、合成音を用いた従来の特徴と問題点を次にまとめてみよう。

(1) [r]と[l]を区別する主な音響的特徴には、F2, F3周波数、遷移時間等が報告されている。しかし、どの特徴に着目するかは、研究者によって異なっている。英語を母国語とする被験者を用いた実験では、これらの音響的特徴と知覚との関係はかなり明らかにされている。しかし、日本人が /r/と/l/を聴き分ける際にどの手掛かりに着目するかという点に関する系統だった検討は不十分である。

(2) 従来の研究における知覚実験では、[r]と[l]を区別する音響的手掛かりのうちの一つまたは複数を変化させた[r-l]刺激連続体上の各刺激に対し、/r/か/l/かを強制判断させている。しかし、それらの刺激連続体上に/r/と/l/以外の音素は知覚されないのかという問題が生じる。

以上(1)(2)の問題点は、従来見落とされる傾向にあったが、異なる言語間での比較を行う場合には実験結果を左右する重要な要因である。

(3) 日本人の英語聴取能力は、その経験によって非常に個人差が大きいことは周知の事実である。従って、日本人の/r//l/音知覚の研究では被験者の選び方、特に被験者の英語聴取能力(hearing)には十分な注意を払う必要がある。この時、英米文科の学生か否かといった概念的な点に注目するのではなく、被験者の聴取能力を直接、客観的に測定することが重要である。

(4) 知覚実験の刺激として用いる合成音の品質は、明瞭性、自然性ともに高いものが望ましい。しかし、従来の研究において、とくに、F3周波数のみを変化させた場合にみられるように、刺激音の音響的特徴を単純化した例では、十分に高い品質の合成音を得ることは難しく、自然音声に対する知覚との不一致が考えられる。

以上の問題点に関して検討するため、十分に品質の高い合成音を作成し、これを使った知覚実験をアメリカ人と日本人を被験者として行った。その結果を以下の2部に分けて報告する。

#### 第1部：日本人の英語/r,l,w/音知覚における特徴

刺激音に与える手掛かり、強制選択課題における選択肢、被験者の分類方法等の差が結果に及ぼす影響について検討した。

本報告は、聴覚研究会資料 H-89-6, pp.1-10 (1989) に発表したものに基づく。

#### 第2部：日本人とアメリカ人の英語/r,l,w/音知覚における手掛かり

日本人とアメリカ人で、利用する手掛かりの差について検討した。

本報告は、信学技報 SP88-125, pp.25-32 (1989) に発表したものに基づく。

## 第 1 部

# 日本人の英語 /r, l, w/ 知覚における特徴

Perceptual characteristics of English syllable-initial /r,l,w/  
for Japanese listeners

内容概説： 合成音を用いて日本人におけるアメリカ英語/r//l/音知覚の性質について検討した結果、次のような点が明らかになった。(1)日本人は/r/と/l/を聴き分けるのに多くの手掛かりを必要とするため、手掛かりの与え方によって、反応パターンが大きく変化する。アメリカ人は、手掛かりの一部が欠如した条件でも主たる手掛かりによる知覚が明確で、/r/と/l/を聴き分けることができる。(2)日本人では[r-l]刺激連続体上に知覚される/w/の領域がアメリカ人より広く、/w/の知覚を考慮することにより、知覚に関する詳細な検討が可能となる。(3)日本人被験者において、英語自然音声の聴取テストの成績と合成音の知覚結果とは対応する。

Abstract: This study investigates the perceptual characteristics of American English /r,l/ for Japanese listeners using synthesized stimuli. Three major findings are obtained. (1) Japanese listeners identify the stimuli using a variety of acoustic cues, and their response patterns are strongly influenced by acoustic features of the stimuli. In contrast, American listeners can identify /r/ and /l/ as long as a primary cue remains, even under the condition where some of the acoustic cues are lacking. (2) As Japanese listeners tend to perceive more stimuli as /w/ than American listeners do, detailed analyses of the response patterns clarify the perception mode of the Japanese listeners by taking account of not only /r/ and /l/ perception but also /w/ perception. (3) A positive relationship between the identification of the natural /r,l,w/ spoken by native American and that of the synthesized /r,l,w/ is found for the Japanese listeners.

# 1. はじめに

日本人の英語(ここではアメリカ英語)/r/、/l/音知覚に関しては、合成音を用いた多くの知覚実験の報告がある。例えばF3開始周波数を変化させた[ra-la]刺激連続体を用い、アメリカ人はそれらを範疇的に知覚しているが日本人はそうではないことを示した実験<sup>(1,2)</sup>、日本人バイリンガルを用い、F3,F2開始周波数、F1遷移時間の全てを変化させた[rak-lak]刺激連続体による実験<sup>(3)</sup>、F3のみを変化させた[ra-la]連続体を用いた6カ国(日本を含む)の被験者に対する実験<sup>(4)</sup>等がある。これらの実験では、典型的な[r]に近い刺激から[l]に近い刺激までの音響的特徴を何段階かに等分し、それぞれの刺激を/r/か/l/か判断させている。これら[r]と[l]を区別する主な音響的特徴には、F2,F3周波数、遷移時間等が報告されている<sup>(5)</sup>。しかし、上記の実験においてはどの手掛かりを与えるかは、研究によって異なっており、手掛かりの種類の違いによる知覚の変化については検討されていない。山田(1988)<sup>(6)</sup>山田・東倉(1989)<sup>(7)</sup>は、手掛かりが異なる複数の刺激条件について子音開始定常部のホルマント周波数に関し、種々のF2-F3の組み合わせに対する反応パターンを調べた。この結果、アメリカ人は/r/と/l/を主としてF3開始周波数を用い、ホルマント遷移部に着目して知覚しているが、日本人は与えられた種々の手掛かり(F1遷移時間、F2、F3開始周波数等)を用いて聴き分けることを明らかにした。また、/r//l/と隣接した領域に/w/が知覚され、しかも日本人の方が/w/と知覚される領域が広いことがわかった<sup>(6,7,8)</sup>。そこで、本研究では[r-l]刺激連続体の知覚に関して、手掛かりの与え方の違いが結果にどのような影響を及ぼすか、また、この時/r//l/2肢強制選択課題でなく、/w/を加えた3肢強制選択課題にすることにより得られる結果について調べた。

また、従来の研究では日本人の被験者を分類せず結果はひとまとめに扱うか、被験者の分類を行う場合でも、大学で英語を専攻したかどうかということで分けることが多い。しかし、日本人の英語聴取能力は、その経験によって非常に個人差が大きい。従って、日本人の/r//l/音知覚の研究では被験者の選び方、特に被験者の英語聴取能力(hearing)には十分な注意を払う必要がある。そこで、英米文科の学生か否かといった概念的な点に注目するのではなく、被験者の聴取能力を直接、客観的に測定し、その結果と合成音を用いた結果の関係について検討した。

本論文では、以上の点に関して検討するために、F3開始周波数、F2開始周波数、F1開始周波数の3つの手掛かりに着目した[r-l]刺激連続体上に、十分に品質の高い合成音を作成し、これを使った知覚実験を行った。本論文の構成は以下の如くである。まず、実験Ⅰでは、手掛かりの欠如が日本人の知覚に及ぼす影響についてアメリカ人と比較しながら検討した。実験Ⅱでは、日本人においてより顕著に認められる[r-l]刺激連続体上での/w/の知覚に関し、同定実験における選択肢に/w/を加え、その影響について検討した。実験Ⅲでは、自然音声の英語聴取能力を直接測定し、この結果による被験者の分類を行い、/r//l/知覚と英語聴取能力の関係を明らかにした。

## 2. 実験Ⅰ

アメリカ人男性によって発声された/rait/, /lait/という二つの音声を分析した結果、[r]と[l]を区別する主な音響的特徴には F2、F3 開始周波数、F1の遷移時間があつた。そこで、これらの手掛かりの与え方により、アメリカ人と日本人で反応パターンにどのような影響があるか検討した。

### 2.1 刺激音

アメリカ人言語学者(男性)によって発声された単語、"right"、"light"を分析し、F1、F2、F3の時間パターンを求めた。ホルマントの抽出では、音声研究用ソフトウェア SPIRE<sup>(9)</sup>を用い、スペクトログラムと短時間スペクトル及びその包絡情報に基づいた視察による抽出を主体とした。また、これら抽出結果に基づいた合成と時間パターンの修正を繰り返し、語頭部の[r]と[l]に関し、明瞭性、自然性共に充分品質の高い合成音を与えるF1、F2、F3の時間パターンを決定した。Fig.1にこれらを模式的に示す。子音開始部からホルマント遷移終了までの子音区間は150msecとし、後続母音[ai](約200msec)は、なるべく忠実に元の音声のホルマントを模擬している。さらに語尾には元の音声の[t]の部分から切り取った音声波形を付加した。

Fig.1の[r][l]のホルマント周波数の時間変化パターンから、知覚実験に使用する[r-l]刺激連続体の合成パラメータを定めた。その際、以下の5つの条件につい

て、F2開始周波数、F3開始周波数、F1遷移時間の3つの手掛かりのうちのいくつかを[r]から[l]に連続的に変化させ、合成した。

(1)総変動条件(Condition with Full Acoustic events) : F2、F3開始周波数、F1遷移時間の3つの手掛かりをすべて[r]から[l]に19段階に変化させた。すなわち、F2開始周波数は920Hzから1280Hzまで、F3開始周波数は1200Hzから3000Hzまで、F1遷移時間は67msecから13msecまでをそれぞれ18等分して同時に変化させた。

(2)F1固定条件(Condition with Fixed F1) : F1遷移時間を[r]の音響的特徴である67msecに固定し、F2、F3開始周波数を[r]から[l]まで19段階に変化させた。

(3)ISS除去条件(Condition with Removed ISS) : (1)と同様F2、F3開始周波数、F1遷移時間をすべて19段階に変化させたが、それらすべてについて音源の立ち上がりを50msec遅らせることにより、子音開始定常部(Initial Steady State)を除去した。

(4)低呈示レベル条件(Condition with Low Presentation Level) : (1)と全く同じ刺激連続体(19種類)であるが、通常の呈示レベル(85dB SPL)より40dB低いレベル(45dB SPL)で呈示した。

(5)F2固定条件(Condition with Fixed F2) : F3開始周波数とF1遷移時間を同時に[r]から[l]に10段階に変化させた。すなわちF3は1200Hzから3000Hzまで、F1は67msecから13msecまで9等分して変化させた。その際、F2開始周波数は800、1000、1200Hzの3通りに固定した。ただし、(F2,F3)が(1200Hz,1200Hz)となる場合は除いたので、刺激の総数は29(10×3-1)となった。

ホルマントバンド幅は分析における実測値に基づいて定め、またF4およびF5周波数はすべて3400,3950Hzに固定された。これらの刺激はあらかじめアメリカ人1名と日本人7名に絶対判断させ、/r//l//w/の3つの音素が知覚され、その他の音は聴こえないことを確認した。

刺激音は、D.H.KlattによるCascade型のホルマント合成器を用い、F1~F5の5つのホルマントを使って合成した。合成音は、サンプリング周波数は20kHz、16ビットでD/Aされた。これらは、DAT(SONY DTC-1000ES)で録音、再生され、ヘッドフォン(STAX SR Λ Pro)を用いて両耳呈示された。呈示レベルはピーク値で約85dB SPL(低呈示レベル条件では45dB SPL)に固定され、これは被験者にとって聴き取りやすいレベルであった。実験は防音室内で行われた。



## 2.2 被験者

被験者は次の3つの群に分けた。

(1) A群：英語を母国語とする者で、アメリカで生まれ育った者、10名。

(2) J群：日本人で、海外に長期滞在経験の無い者、6名。

(3) JE群：日本人で、アメリカに長期滞在経験があり、日本、アメリカ以外の国には長期滞在経験がない者、2名。2名ともにアメリカ滞在経験は約2年であった。

## 2.3 手続き

(1)-(4)の4つの刺激条件において、条件ごとに同定テストとABX弁別テストを行った。(5)の条件では同定テストのみを行った。

同定テストでは、各条件は別々のセッションでテストした。(1)-(4)の条件場合、19種類の刺激を5回ずつランダムに並べたものの冒頭に5個のダミー刺激を付加し、1セッションとした。F2固定条件の場合も29種類の刺激音がランダムに並べられた。10試行を1ブロックとし、ITIは3秒、ブロック間間隔は6秒とした。各ブロックの冒頭には予告刺激として、ピープ音を呈示した。各試行毎に語頭の子音がなんという音に聞こえたか、/r//l//w/3つの選択肢から強制選択させ、選択肢を印刷した反应用紙に印を付けさせた。また、被験者には、各反応の出現頻度は一定であるとは限らず、しかもセッションにより変化する可能性もあるため、反応の偏りが起こりうることを教示した。

同定テストでは、A群の被験者では各刺激につき約5回、J群、JE群の被験者では約15回判断させた。

ABX弁別テストでは、19種類の刺激を4ステップで組合せた。組合せは60通り(1-5-1、1-5-5、5-1-1、…、19-15-15、19-15-19)となり、それらをランダムに並べ変えて、1セッション(60試行)とした。10試行を1ブロックとし、ISIは1sec、ITIは3sec、ブロック間間隔は6secとした。各ブロックの冒頭には予告刺激として、ピープ音を呈示した。判断は反应用紙に印をつけさせた。

4つの条件((1)-(4))のABX弁別テストは、A群の被験者では各条件2セッション、J群の被験者では各条件6セッション課した。

各テストにおいてJ群におけるセッション数を多くしたのは、反応が集中せず、個人差の多いJ群のデータを分析する際に、個人別にみても、ある程度安定したデータを得るためである。

## 2.4 結果

刺激総変動条件、F1固定条件、ISS除去条件、低呈示レベル条件における同定実験のA群、J群の結果をそれぞれ Fig.2に示した。

A群では典型的な[r][l]音の付近の刺激は確実にそれぞれ/r//l/と答え、/r/と/l/の間に/w/と答える領域が存在した。この/w/という反応は被験者の約半数ではほとんど出現せず、残りの半数の被験者で出現した。/w/という反応をした被験者では、/r/、/w/、/l/が明確に分かれ、範疇的に知覚していた。

/w/音のカテゴリーの有無の影響等によって、この刺激連続体上でのカテゴリーの境界は、A群のなかでも個人差があった。そのことを反映して、ABX弁別テストの結果をすべての被験者に関して平均した場合、ピークは[r]と[l]の間あたりの刺激を中心に広い範囲で広がった。また、A群の同定テスト結果は条件の違いによる反応の変化は少なく、低呈示レベル条件において、/w/が消失し、むしろ/l/と反応するようになった程度で、典型的な/r/近傍、/l/近傍ではどの条件においても確実に/r//l/と答えた。弁別テスト結果にも大きな変化はなかった。

J群では条件によって反応パターンに大きな変化があった。同定テストでは、総変動条件、F1固定条件、ISS除去条件では、/r/と/l/がある程度分化していたが、/w/の反応に差があった。総変動条件では、全体的に/w/の反応があり、/r/と/l/の間あたりで小さなピークがあったのに対し、F1固定条件では/w/の反応が多くなり、/r/と/l/の間に明確な/w/の範疇が出現した。ISS除去条件でも、総変動条件に比べ、/w/がやや多くなった。低呈示レベル条件では、大きな影響を受け、典型的な[r][l]近傍の刺激ですら判断が困難になった。弁別テストの結果は総変動条件、F1固定条件、ISS除去条件でアメリカ人の結果よりはひくいピークがみられ、条件によってその性質は異なった。特にISS除去条件では顕著な2つのピークが出現し、これらは/r//w//l/という3つの範疇の境界に対応する。しかし、低呈示レベル条件では、同定結果を反映して、弁別結果もフラットな形状となりチャンスレベル(50%)であった。

F2固定条件の結果をFig.3に示した。A群ではF2周波数が高くなるにつれて、/w/と知覚される領域が小さくなる程度で、F3周波数が低い刺激音は/r/、F3が高い刺激音は/l/という反応パターンには変化がなかった。これに対し、J群ではF2周波数が反応に大きな影響を与えた。すなわち、F2周波数が800Hzの場合、すべての刺激に対し、/r/の反応が多く、A群が/l/と判断する刺激についても、/r/の反応が優勢だった。また、F2が高くなるにつれて、/l/の反応が増加した。

J E群の結果のうち、5-6才の時にアメリカに滞在した被験者JE1では、条件(1

)-(4)においてアメリカ人の結果と類似していた。F2固定条件では判断は明確であったが、F2の値の影響を受け、F2が高くなるにつれ、/l/と判断する領域が広がった。もう一方のJ E群の被験者JE2は24-25才時にアメリカに滞在したが、実験結果は日本人の平均的な性質と類似していた。

## 2.5 考察

母国語にない音素にの知覚においては、Nativeの被験者とは、別の手掛かりに着目する等全く異なったストラテジーで判断が行われる可能性がある。日本人の/r//l/知覚においても、アメリカ人による知覚とは異なり、刺激音の特性におけるわずかな変化が反応パターンに大きな影響を与えることが明らかになった。すなわち、与えられる手掛かりによって結果が変わることが示唆された。また、低呈示レベル条件で日本人の反応が大きく影響を受けることから、今まであまり注意が払われていなかった刺激呈示レベル、S/N比、刺激音の品質などが実験結果を左右することが示された。したがって、日本人の/r//l/知覚の検討にあたってはこれらのことに十分留意して、検討をすすめなければならない。J E群の結果からは幼少時の言語経験の影響が示唆されるが、この点については今後の研究課題である。

# 3. 実験Ⅱ

[r]-[l]刺激連続体上に/r/,/l/以外の音素が聴こえる可能性がある。この問題は、異なる言語間での比較をする場合には特に注意を払う必要がある。予備実験から、/r/と/l/の間に/w/が知覚されることがわかったので、選択肢に/w/を付加することの影響について、実験を行った。

## 3.1 刺激音

実験Ⅰと同様の合成音を用いた。ただし、3つの手掛かりを同時に[r]から[l]に変化させた総変動条件のみを用いた。つまり、Fig.1の模式図に従って、F2開始周波数は920Hzから1280Hzまで、F3開始周波数は1200Hzから3000Hzまで、F1遷移時間は67msecから13msecまでを18等分して変化させた。刺激数は19個となり、[r]側

から順に、1、2、…、19という刺激番号で示すこととする。

### 3.2 被験者 日本人6名。

### 3.3 手続き

同定テストとABX弁別テストを行った。

同定テストでは、19種類の刺激を5回ずつランダムに並べたものの冒頭に5個のダミー刺激を付加し、1セッションとした。 $/r//l/$ の2選択肢から選択する条件(RL-Condition)と、 $/r//l//w/$ 3つの選択肢から選択する条件(RLW-Condition)の2つについて行った。実験方法は実験Iと全く同じであった。

各被験者に2つの条件の同定テストを主としてそれぞれ3セッション課した。各条件は別々の日に実験を行った。

ABX弁別テストは、実験Iと全く同じ方法で行った。各被験者につき、主として6セッション課した。

### 3.4 結果

6名の被験者の平均および6名のうち特徴がはっきりしていた3名(S1,S2,S3)のデータをFig.4に示した。平均値でみた場合、選択肢を $/r/$ と $/l/$ の2つにした条件では反応は $/r/$ から $/l/$ に徐々に変化している。選択肢が $/r//l//w/$ の3つの条件で判断させた場合、 $/r/$ と $/l/$ の間に $/w/$ が出現することがわかる。この $/w/$ の現れ方は、個人差が大きく、例えば被験者S1は、同定実験において $/r//w//l/$ の順に反応が分化しており、弁別テストにおいてもそれぞれの境界においてピークが出現した。つまり、 $/r//w//l/$ という範疇が存在した。S2はS1と範疇の順序が異なり、 $/w//r//l/$ という順で範疇が存在した。また、S3は $/l/$ という範疇ははっきりしているが、 $/r/$ と $/w/$ の区別がつかず、弁別実験のピークも1つで $/l/$ とそれ以外の音の境界に出現している。以上の3名の結果は $/r//l/$ の2選択肢による実験では、S1、S2の結果は、反応が $/r/$ から $/l/$ に緩やかに変化する連続的知覚のパターンと類似し、S3の結果は $/r/$ と $/l/$ の範疇的知覚のパターンと類似しており、誤った結果を導く可能性がある。

### 3.5 考察

$[r]-[l]$ 刺激連続体上には $/r//l/$ 以外に $/w/$ という音素が知覚される。 $/r//l/$ の2選択肢による実験において同定実験の結果だけでは $/r/$ から $/l/$ への変化がし、

連続的に知覚しているように見える場合でも、/w/を選択肢に加えることによって、被験者の個人差は存在するものの、/r//w//l/のうちのいくつかの範疇を持つ場合があることが示された。例えば被験者S1の場合は /r//w//l/、S3の場合は/l/の範疇がはっきりしている。また、実験Iの総変動条件は本実験の RLW-Condition に相当するため、そのデータから次のことが明らかになった。例えばJE群のある被験者では、/r/と/l/という2つの明確な範疇のみを持っていた。また、A群の被験者でも、/r/と/l/の2つの範疇を持つものと/r//w//l/3つの範疇を持つものがあった。これらの結果から、同定実験における選択肢の重要性が明かになり、/w/を選択肢に入れることにより、今後、知覚の手掛かり、発達、生成との関連等の興味深いトピックスについて、信頼性の高いデータを提供し、詳細かつ発展性のある検討が可能となる。また、/w/音が知覚される可能性がある場合に、/r//l/のうちから強制選択させることは、被験者に無理な課題を強いることになり、好ましい実験方法とはいえない。

さらに、S1、S2、S3における反応パターンの差に関しては、以下のように考察できる。Fig.5は、子音開始定常部のホルマント開始周波数F2、F3の組み合わせを表すF2-F3平面上に、知覚実験から得られた各被験者の/r//l//w/の範疇を模式的に示したものである<sup>(6,7)</sup>。また、このF2-F3平面上の直線が本実験で用いた刺激連続体を表す。本実験の刺激連続体は、S1では/r//w//l/の順、S2では/w//r//l/の順に各範疇を横切っている。また、S3では/w/と/r/の境界領域と/l/の範疇を横切っている。S1とS2における/r/の反応と/w/の逆転は、/r/と/w/の範疇の逆転ではなく各被験者の/r/と/w/の境界の差、つまり用いる手掛かりの差に起因することがわかる。また、S3では/r/と/w/が区別できないのではなく、本実験の刺激連続体が、S3の知覚における/r/と/w/の境界領域と一致した結果である。このように、各被験者の知覚に利用する手掛かりに依存したF2-F3平面上での反応パターンから本実験の結果を説明することができる。

これらのことから、日本人の/r,l,w/音の知覚様式に関しては/w/音を選択肢に入れ、さらに刺激音の特性を考慮した細かい検討が必要であることが示された。この結果は、本実験における刺激音に限った結果であるので、パラメータ、文脈等が変化した場合のすべてに必ずあてはまるとはかぎらない。しかし著者らによるF3開始周波数のみを変化させた[ra]-[la]刺激連続体を用いた予備実験からも同様の結果が得られた。また、VCV音声(aCa,iCi,uCu)を用いた聴取実験<sup>(10)</sup>においても、/w/と知覚される領域が、/r/または/l/と知覚される領域に隣接している。さらに実際の/r//l//w/音の分析結果<sup>(11)</sup>からもF2/F1-F3/F1平面上で/w/が/r//l

/に隣接している。これらを勘案すれば、本実験以外の音素環境においてもここで示した/w/の重要性が再認識されることが予測される。また、アメリカ人に比較して日本人において、/w/の範囲が大きいという結果は日本語に/wai/という音声が存在することとの関連を示唆する。しかし、本研究では/w/の問題は本研究では/r//l/知覚に付随して生じたものであり、この問題を積極的に論ずるには、刺激音の手掛かりをさらに考慮した実験、別の後続母音を用いた実験等による検討が必要であろう。

## 4. 実験Ⅲ

外国語の知覚では、被験者の経験と知覚の関係を考えることが必要であり、日本人被験者をどのように分類するかは実験結果を左右する重要な問題である。日本人の英語知覚は個人差が非常に大きいので、平均してしまうと、特徴的な結果が見えなくなる可能性があるため、何等かの有効な手段で被験者の分類をしなければならない。その際の客観的な指標として自然音声(英語)の聴取テストの成績をとりあげ、その妥当性について検討した。

### 4.1 刺激音

(1)自然音声：語頭の子音が/r//l//w/のみで区別される英単語のミニマル・ペア16対(各対間の後続母音は異なる)をアメリカ人男女各1名ずつに発声させた。これらを20kHz、16ビットでA/Dし、刺激音とした。

(2)実験Ⅱと同様の合成音。

### 4.2 被験者

A群 10名(実験Ⅰと同じ)。

J群 36名、JE群 3名。

### 4.3 手続き

すべての被験者に自然音声の同定テスト、合成音の同定テスト及びABX弁別テストを課した。

自然音声の同定テストでは、96種類の刺激音を各1回ずつランダムイズし、1セッションとした。各刺激について単語の先頭の音が/r//l//w/のうちのいずれに聞こえたか強制選択させた。

合成音の同定テストの手続きは、実験Ⅰ,Ⅱと同様で、選択肢は/r//l//w/の3つであった。ABX弁別テストも実験Ⅱと同様であった。ただし合成音のテストはA群の被験者の場合、実験Ⅰ総変動条件と同一テストとなるので、実験Ⅰのデータを用いた。

自然音声、合成音同定テスト、ABX弁別テストはそれぞれA群では1、1、2セッション、J群とJE群では2、3、6セッションが課された。

#### 4.4 結果

自然音声の同定テストの結果は、正当率がA群では99.8% J群では65.4% JE群では83.5%であった。また、合成音のテストにおけるJ群の結果の平均をFig.5に示した。J群における同定テストの結果はA群(Fig.2 Full)に比べ、反応が確実でなく、[r]音から[l]音に向かって、徐々に/r//w//l/と変化した。また、弁別実験の結果もA群に比べて、顕著なピークは認められなかった。

J群の自然音声の成績は50%から90%の間に分布していたので、J群を自然音声で50-60%、60-70%、70-80%、80-90%の4つのグループに分け、合成音声同定テストとABX弁別テストの結果を整理し、Fig.7に示した。各グループの人数と自然音声の成績の平均値は以下のとおりであった。

- (1)J-50 (50%以上、60%未満) 14名 54.9%
- (2)J-60 (60%以上、70%未満) 10名 64.6%
- (3)J-70 (70%以上、80%未満) 7名 73.1%
- (4)J-80 (80%以上、90%未満) 5名 84.1%

同定テストの結果は、J-50でもっとも反応の分化が乏しく、自然音声の成績が上昇するにつれて分化した。J-70とJ-80では、かなり分化しており、/r//l//w/の範疇が比較的はつきりしていることを示唆した。また、/w/の反応がJ-80で顕著になった。弁別テストの結果もJ-50では殆どチャンスレベル(50%)であった。しかし、自然音声の成績が高いグループでは弁別テストの結果にピークが認められた。J-70とJ-80の差は、同定テスト結果では、刺激番号5から13の領域で/w/の反応がJ-80の方が高いという程度である。しかし、弁別結果ではピークの大きさに差が認められ、J-80ではJ-70に比べて、より範疇的な知覚を示す。つまり、自然音声の聴き分け能力が高いグループほど、合成音/r//l//w/の同定テストにおける反応も

よく分化し、より範疇的に知覚するようになる。これらの結果から、本実験における合成音を用いた実験の結果は、自然音声の成績を反映していることが明らかになった。

JE群は被験者が3名と少数であったため、そのなかでのグループ分けはできなかったが、自然音声の成績(それぞれ、97%、81%、70%)は合成音の結果に反映されていた。

被験者の中に、英文科の大学生が7名いたので、それらの平均を求めたが、特に反応が分化していることもなく、弁別成績もチャンスレベルであった。この7名の自然音声による同定テスト成績は平均 61.0% であり、合成音のテストの結果は、自然音声の結果を反映していた。

#### 4.5 考察

合成音の知覚を自然音声の聴取能力と関連づけて分析することにより、日本人の /r//l//w/ に関する知覚様式の特徴がはっきりした。本実験で用いた自然音声は語頭の /r//l//w/ に関するペアなので、総合的な聴取能力を判定するには不十分かもしれない。今後自然音声の聴取能力の測定方法、自然音声の混同率の傾向等に関しても併せて検討する必要がある。

## 5. まとめ

合成音を用いて日本人における /r//l/ 音知覚の特徴について、総合的に検討を行った結果、次の様な点が明らかになった。

(1) アメリカ人では、手掛かりの一部が欠如した条件でも、反応には大きな変化はなく、典型的な [r][l] それぞれの近傍の音は確実に /r//l/ と反応する。しかし、日本人では与える手掛かりの違いによって、反応パターンが変化した。これは、日本人が F1 遷移時間、F2 開始周波数、F3 開始周波数、子音開始定常部など種々の手掛かりを用いて聴き分ける<sup>(6,7)</sup> ことに起因すると考えられる。これらの結果は言い替えれば、[r]-[l] 刺激連続体を作成する場合に用いる手掛かりによって、アメリカ人に近い結果が得られる場合もあれば、全く逆の結果を得る可能性もあるということになる。従って日本人を用いて実験を行う場合にはどのような手掛か



りを、どのような条件(呈示レベル、刺激音の品質、S/N比等)で与えたかということが重要な問題となる。

(2)日本人ではアメリカ人より/w/が知覚されることが多い。このため/w/を選択肢に入れた知覚実験によって、より明確な結果を得ることができる場合がある。言い換えれば、選択肢を/r//l/に限った実験から、日本人が/r/と/l/の範疇的知覚ができないことを説明することは難しい。/w/を考慮することにより、より詳細な検討が可能となった。また、アメリカ人においても/w/の影響が無視できないことが示唆された。刺激音の手掛かりや、文脈によっても/w/の知覚が異なると思われる。今後この/w/の問題はさらに検討していかねばならないと思われる。

(3)自然音声の聴取テストの成績と合成音の結果とはよく対応することが明らかになった。日本人を被験者として合成音を用いた実験をする場合には、併せて自然音声の聴取テストを行うことにより被験者の聴取能力を測定し、これによってグループ分けを行う必要がある。

以上の3点はいずれも実験結果に非常に大きな影響を及ぼすため、日本人における英語/r//l/音知覚について議論する場合に、考慮しなければならない必要不可欠な項目である。

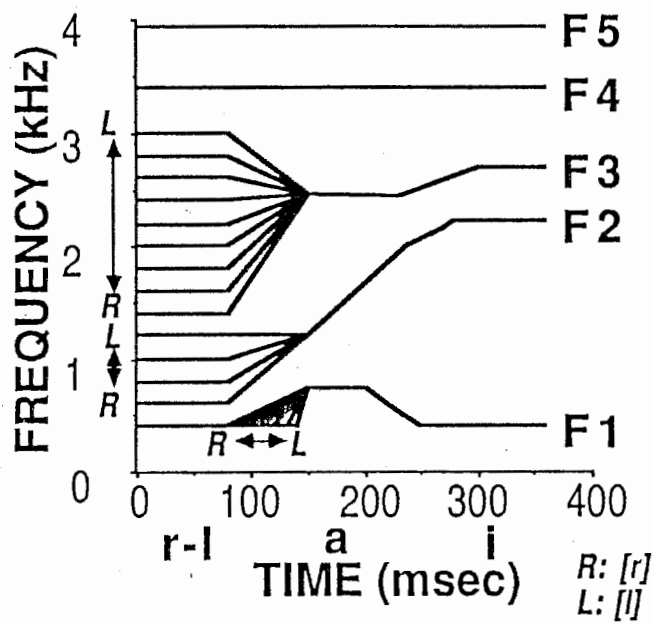


Fig.1 Schematic representations of spectrum for the part of the stimuli, which are extracted from /rai/-/lai/ in the /rait/-/lait/ series.

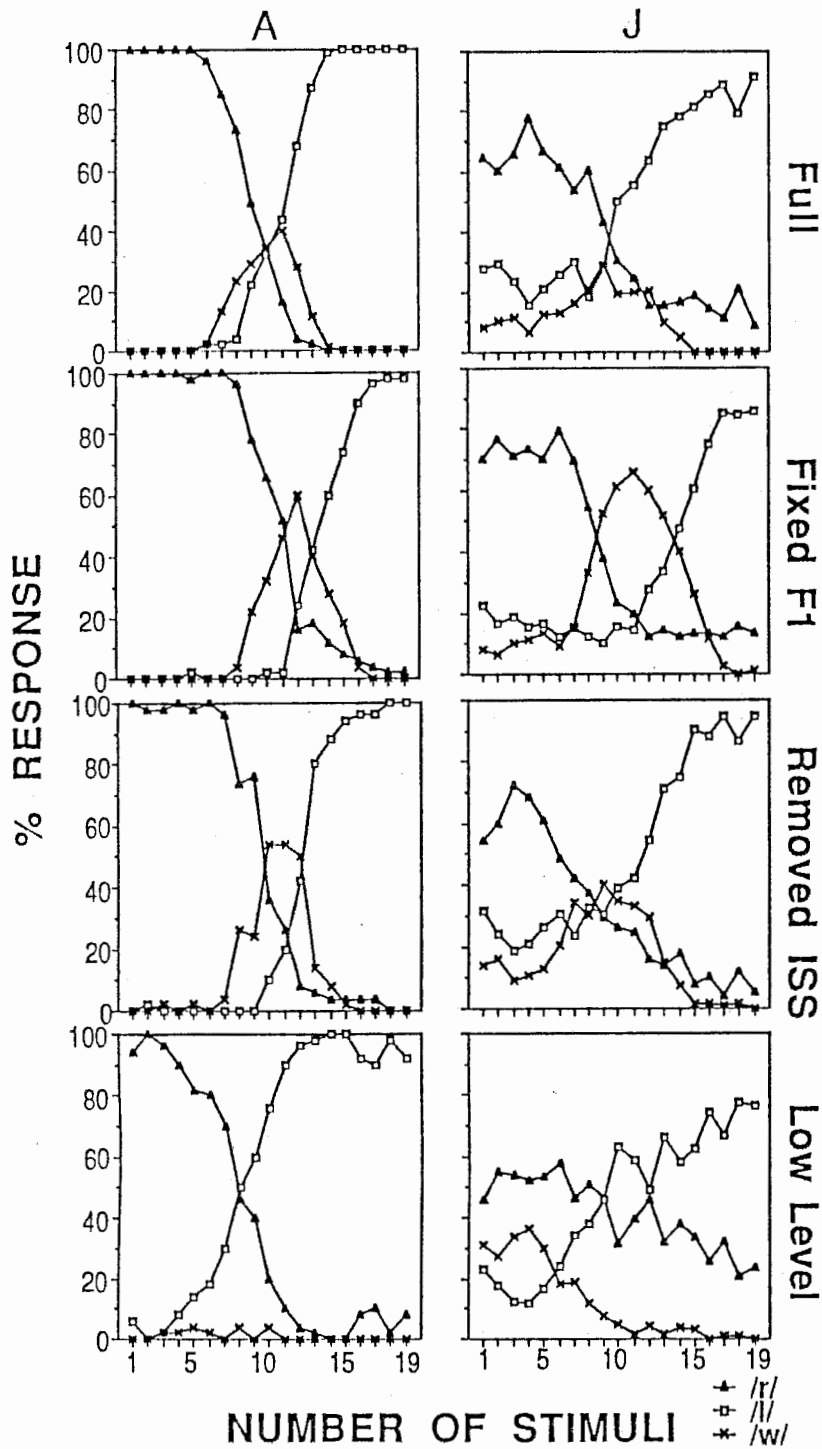


Fig.2 Response patterns of Native Americans (A) and Native Japanese (J) under following four conditions; condition with full acoustic events, fixed F1, removed ISS(Initial Steady State), and low presentation level. The abscissa represents the number of stimuli, and the ordinate represents the response rate.

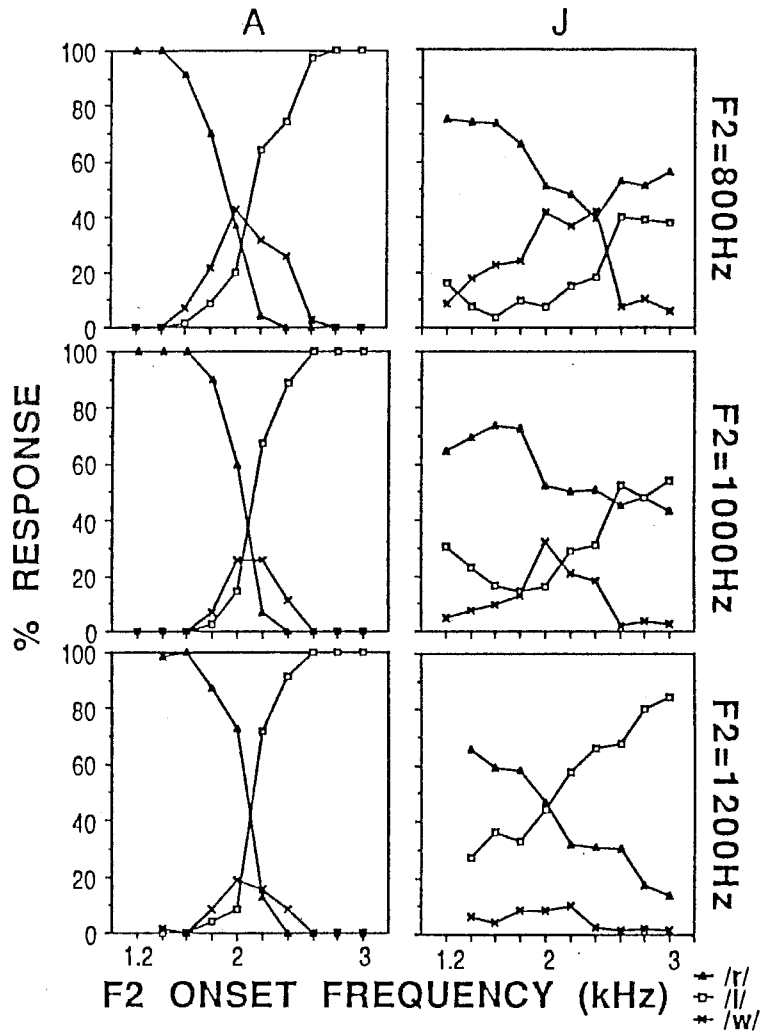


Fig.3 Response patterns of Native Americans (A) and Native Japanese (J). The value of F2 frequency was fixed to 800Hz, 1000Hz, or 1200Hz.

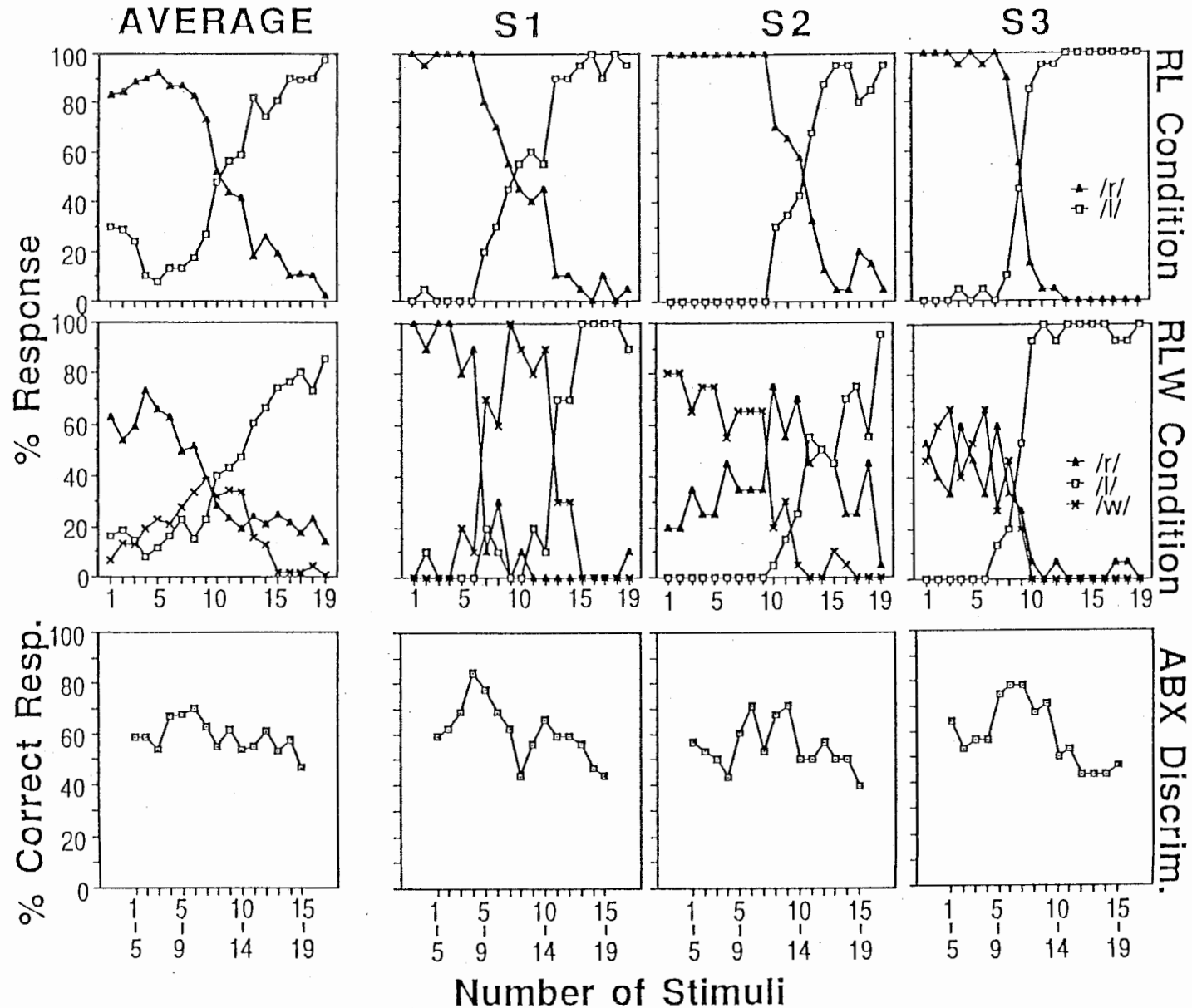


Fig.4 Pooled results of six Japanese subjects, and individual data obtained for three of the subjects; S1, S2, and S3. The results from the identification tests in the RL condition ( i.e., a forced choice from two alternatives; /r/, and /l/), and those in the RLW condition(i.e., a forced choice among three choices; /r/,/l/, and /w/) are shown in the upper and middle lows. The lower graphs represent the results from 4-step ABX discrimination tests.

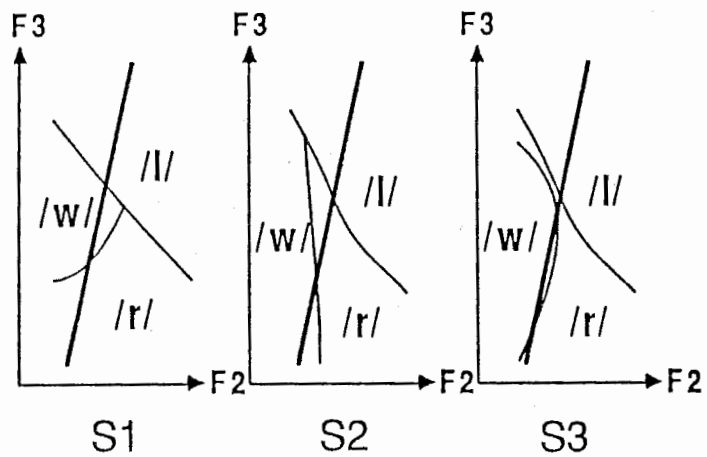


Fig.5 Schematic patterns representing three categories, /r/, //, and /w/ for each of three subjects; S1, S2, and S3. The stimulus continua in this experiment are described by the thick lines.

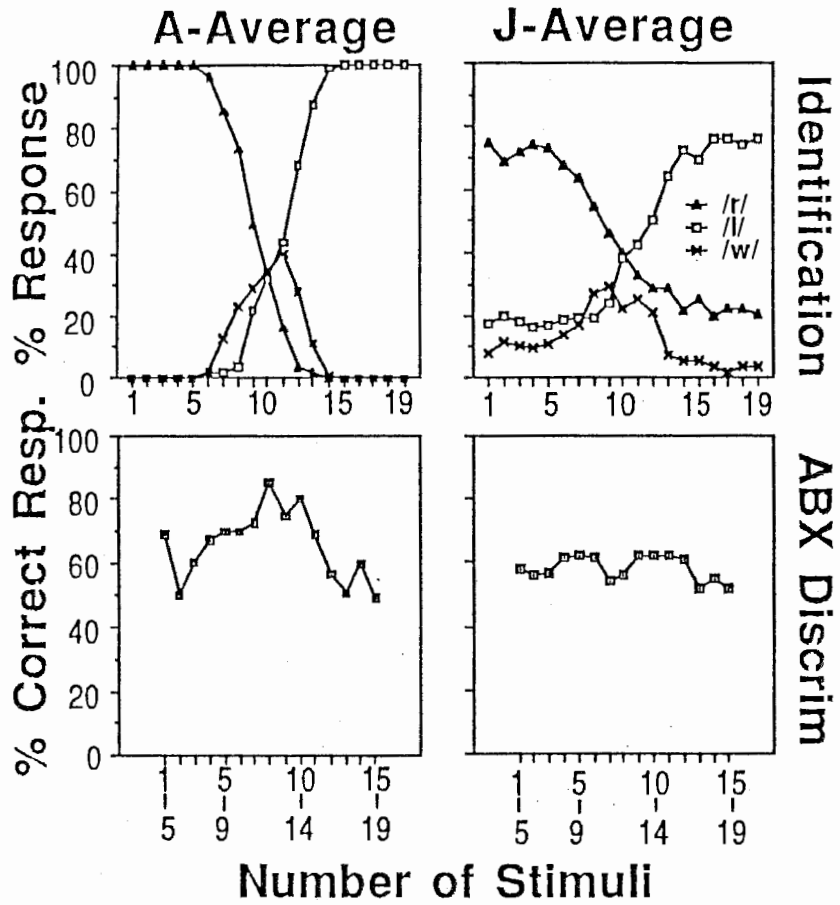


Fig.6 Results from identification tests(upper), and ABX discrimination tests(lower). Pooled data for two groups; Native Americans (A), and Native Japanese (J), are shown.

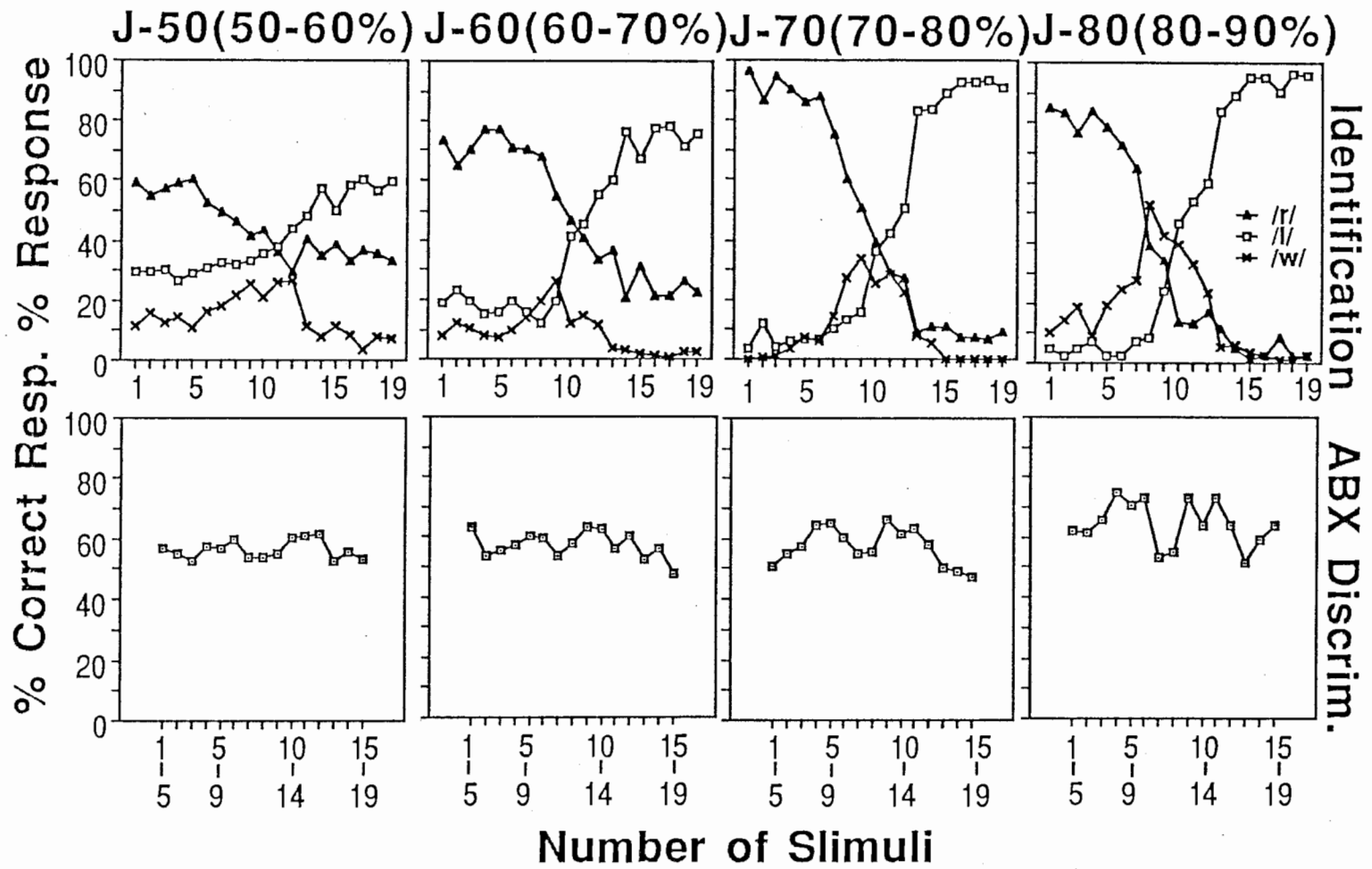


Fig.7 Data for four Japanese groups are classified according to the listening ability (i.e., identification rates) of natural English words.



## 文献

- (1) Liberman, A.M., Miyawaki, K., Jenkins, J.J., and Fujimura, O. (1973) Cross-Language Study of the Perception of the F3 Cue for [r] vs. [l] in speech- and nonspeech-like patterns. *J. Acoust. Soc. Jpn.*, 29, 315.
- (2) Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A.M., Fujimura, O., and Jenkins, J. (1975) An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Percept. Psychophys.*, 18, 331-340.
- (3) MacKain, K.S., Best, C.T., and Strange, W. (1981) Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Appl. Psycholinguist.*, 2, 369-390.
- (4) Shimizu, K., and Dantsuji, M. (1987) A cross-language study on the perception of synthetic speech sounds of [r-l]. *Proceedings of the Eleventh ICPHS.* 6, 71-74.
- (5) Dalstone, R.M. (1975) Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. *J. Acoust. Soc. Am.*, 57, 462-469.
- (6) 山田玲子 (1988) 日本人における米語 /r, l, w/ 音知覚様式とその手掛かり -F1の遷移時間、およびF2, F3の開始周波数-. 関西心理学会第100回大会論文集, pp.13.
- (7) 山田玲子、東倉洋一 (1989) 日本人の英語 /r/, /l/, /w/ 音知覚に関する再検討. 信学技報88, 25-32.

- (8) Mochizuki, M. (1981) The identification of /r/ and /l/ in natural and synthesized speech. J. Phonet., 9, 283-303.
- (9) Cyphers, D.S. (1985) Spire: a speech research tool. MS Degree Thesis, MIT.
- (10) Lisker, L. (1957) Minimal cues for separation /w, r, l, y/ in intervocalic position. Word 13, 256-267.
- (11) O'Connor, J.D., Gerstman, L.J., and Liberman, A.M. (1957) Acoustic cues for the perception of initial /w, j, r, l/ in English. Word, 13, 25-43.

## 第 2 部

# 日本人およびアメリカ人の英語 /r, l, w/ 音 知覚における手掛かり

Acoustic cues for the perception of English syllable-initial /r, l, w/  
by Japanese and American listeners.

内容概説： 合成音を用いて日本人における米語/r,l,w/音知覚の手掛かりに関して検討した。F1遷移時間、F2およびF3開始周波数の3つの音響的特徴をパラメータとした[r-l]刺激連続体を作成し、知覚実験を行った。アメリカ人では、知覚様式が明確で、主としてF3周波数を用いて/r/と/l/を聴き分けている。日本人ではF2、F3周波数、F1遷移時間を主体とする多くの手掛かりを用いてようやく典型的な/r/と/l/の聴き分けが可能となる。また、[r-l]刺激連続体上に/w/の知覚が出現し、その領域は日本人の方がアメリカ人より広がった。日本人では、子音開始定常部の除去や刺激呈示レベルの低下等、知覚の手掛かりが総体的に不明確になる条件において、大きな影響を受けることが示された。

Abstract: This study reports on the perception of American English /r,l,w/ by Japanese listeners using synthesized stimuli. Synthetic [r-l] continua have been generated using three parameters; duration of the first formant(F1) transition, onset frequency of the second formant(F2), and those of the third formant(F3). The American listeners perceive these stimuli barely identify typical /r/ and /l/ using various acoustic cues provided to them. Some stimuli on the [r-l] continua are perceived as /w/, which are more frequent in Japanese listeners than in American listeners. When some acoustic cues are absent, the identification scores in Japanese listeners are deteriorated.

# 1. はじめに

日本人は英語/r//l/音の聴き取り、調音を共に苦手としている。この日本人の英語/r//l/音知覚は、言語環境が知覚に及ぼす影響や、聴く能力と調音能力の関係を調べる上でよい例題となるため、多くの研究報告がある。自然音声を用いた実験では、主として単語の中での位置、前後環境と聴き分けの関係を調べている<sup>(1)(2)</sup>。合成音を用いる研究では、刺激音の音響的手掛かりを自由に設定することができるため、知覚様式に関する詳細な検討が可能である。例えば、Liberman, et al.<sup>(3)</sup>、Miyawaki, et al.<sup>(4)</sup>は F3開始周波数を変化させた [ra-la] 刺激連続体を用いた実験から、アメリカ人はそれらを範疇的に知覚しているが日本人はそうではないことを示した。また、日本人バイリンガルを用い、F3, F2 開始周波数、F1遷移時間の全てを変化させた [rak-lak] 刺激連続体による実験<sup>(5)</sup>、F3のみを変化させた [ra-la] 連続体を用いた6カ国（日本を含む）の被験者に対する実験<sup>(6)</sup>等がある。これら合成音を用いた従来の特徴と問題点を次にまとめてみよう。

(1) [r]と[l]を区別する主な音響的特徴には、F2, F3周波数、遷移時間等が報告されている<sup>(7)</sup>。しかし、どの特徴に着目するかは、研究者によって異なっている。英語を母国語とする被験者を用いた実験では、これらの音響的特徴と知覚との関係はかなり明らかにされている<sup>(8)(9)</sup>。しかし、日本人が/r/と/l/を聴き分ける際にどの手掛かりに着目するかという点に関する系統だった検討は不十分である。

(2) 従来の研究における知覚実験では、[r]と[l]を区別する音響的手掛かりのうちの一つまたは複数を変化させた [r-l] 刺激連続体上の各刺激に対し、/r/か/l/かを強制判断させている。しかし、それらの刺激連続体上に/r/と/l/以外の音素は知覚されないのかという問題が生じる。実際、Mochizuki<sup>(1)</sup>や、山田<sup>(10)</sup>の実験においては/r/と/l/の間に/w/が知覚され、しかも日本人の方がアメリカ人より/w/を知覚する領域が広いことが示された。

(3) 日本人の英語聴取能力は、その経験によって非常に個人差が大きいことは周知の事実である。従って、日本人の/r//l/音知覚の研究では被験者の選び方、特に被験者の英語聴取能力(hearing)には十分な注意を払う必要がある。この時、英米文科の学生か否かといった概念的な点に注目するのではなく、被験者の聴取能力を直接、客観的に測定することが重要である。

(4) 知覚実験の刺激として用いる合成音の品質は、明瞭性、自然性ともに高い

ものが望ましい。しかし、従来の研究において、とくに、F3周波数のみを変化させた場合にみられるように、刺激音の音響的特徴を単純化した例では、十分に高い品質の合成音を得ることは難しく、自然音声に対する知覚との不一致が考えられる。

本論文では、以上の点を考慮した上で、F3開始周波数、F2開始周波数、F1開始周波数の3つの手掛かりを変化させた[r-l]刺激連続体上に、十分に品質の高い合成音を作成し、これを使った知覚実験を行った。この時、/r/、/l/の2者強制選択課題ではなく、/w/という選択肢を加えた3者強制選択課題とし、知覚の手掛かりに関する詳細な検討を行った。また、アメリカ人と日本人の平均的な知覚様式の相違に着目するだけでなく、自然音声を使った実験による日本人被験者の英語聴取能力の測定を併せて行い、個人差の多い日本人の知覚を被験者の英語聴取能力に関連づけて検討した。

本論文の構成は以下の如くである。実験Ⅰでは、知覚の手掛かりについて系統的に検討するために、F2及び、F3開始周波数、F1遷移時間の各特徴を変化させた刺激連続体を用いて実験を行った。実験Ⅱでは、[r-l]の子音開始定常部除去、知覚実験における刺激音呈示レベル低下といった手掛かりの一部欠如が日本人の知覚にどのような影響を及ぼすかアメリカ人の場合と比較して検討した。

## 2. 実験Ⅰ

[r]と[l]を区別する特徴はF2、F3開始周波数、F1遷移時間等である。本実験では、これらの特徴が、日本人とアメリカ人による/r//l/知覚にどのような影響を及ぼすかを系統的に調べた。また、日本人被験者の海外長期滞在経験と知覚との関係についても併せて検討した。

### 2.1 刺激音

アメリカ人言語学者（男性）によって発声された単語、"right"、"light"を分析し、F1、F2、F3の時間パターンを求めた。ホルマントの抽出では、音声研究用ソフトウェア SPIRE<sup>(11)</sup> を用い、スペクトログラムと短時間スペクトル及びその包絡情報に基づいた視察による抽出を主体とした。また、これら抽出結果に基づ

いた合成と時間パターンの修正を繰り返し、語頭部の[r]と[l]に関し、明瞭性、自然性共に充分品質の高い合成音を与えるF1、F2、F3の時間パターンを決定した。Fig.1にこれらを模式的に示す。子音開始部からホルマント遷移終了までの子音区間は150msecとし、後続母音[ai](約200msec)は、なるべく忠実に元の音声のホルマントを模擬している。さらに語尾には元の音声の[t]の部分から切り取った音声波形を付加した。

Fig.1の[r]、[l]のホルマント周波数の時間変化パターンから、知覚実験に使用する[r-l]刺激連続体を以下のように作成する。連続的に変化させたパラメータはF2開始周波数、F3開始周波数、F1遷移時間の3つである。F2は、600Hz[r]から1400Hz[l]まで200Hzおきに変化(5通り)させ、F3は、それとは独立に、1200Hzから3200Hzまで200Hzおきに変化(11通り)させたが、3200Hzでは音の自然性が失われたので、3100Hzとした。また、F2の値がF3の値を超える組合せは排除した。従って、この実験で用いるF2、F3のすべての組合せはFig.1のタイプAの範囲で示され、全部で52条件(5x11-3)となる。F1遷移時間については次の2条件を用いた。

(1) Full Condition : F3の値が[r]から[l]に動くのと連動させてF1の遷移時間を70msec([r])から13msec([l])まで変化させた条件。

(2) -F1 Condition : F2、F3の値に関わらず、F1遷移時間を70msecに固定した条件。

したがって、F1、F2、F3の各条件を組み合わせた刺激音は全部で104種類(52条件x2条件)となる。ホルマントバンド幅は分析における実測値に基づいて定め、また、F4およびF5周波数はすべて3400,3950Hzに固定された。

また、(F2、F3)がそれぞれ(800Hz, 2600Hz)(1200Hz,1800Hz)という2つの組合せを選び、F1遷移時間を70msecから10msecまで10msecステップで変化させた14種類の刺激音を付加した。したがって、この実験で用いる刺激音は、総数で118種類(104+14)となる。

これらの刺激はあらかじめアメリカ人1名と日本人7名に絶対判断させ、/r//l//w/の3つの音素が知覚され、その他の音は聴こえないことを確認した。

刺激音は、D.H.KlattによるCascade型のホルマント合成器を用い、F1~F5の5つのホルマントを使って合成した。合成音はサンプリング周波数は20kHz、16ビットでD/Aされた。これらは、DAT(SONY DTC-1000ES)で録音、再生され、ヘッドフォン(STAX SR A Pro)を用いて両耳呈示された。呈示レベルは被験者にとって聴き取りやすいレベルに調整、固定された。このとき、刺激音のピーク値は約85dB SPLであった。実験は防音室内で行われた。

## 2.2 被験者

被験者は次の3つの群に分けた。

- (1) A群 : 英語を母国語とする者で、アメリカで生まれ育った者、8名。
- (2) J群 : 日本人で、海外に長期滞在経験の無い者、7名。
- (3) JE群 : 日本人で、アメリカに長期滞在経験(1-6年)があり、日本、アメリカ以外の国には長期滞在経験がない者、7名。

## 2.3 手続き

118種類の刺激をランダムに並べ、冒頭に12個のダミー刺激を付加して、1セッションとした。10試行で1ブロックとし、ITIは3秒、ブロック間間隔は6秒とした。各ブロックの冒頭には予告刺激として、ビーブ音を呈示した。各試行ごとに、語頭の子音が何という音に聴こえたか同定させ、判断は/r//l//w/の3者強制選択課題とした。反応は選択肢を印刷した反应用紙の選択肢に印をつけさせた。また、各反応の出現頻度は一定であるとは限らないので、反応の偏りは気にせず答えるように教示した。各被験者には主として15セッション課した。

## 2.4 結果

実験の結果を、A群、J群、JE群別に整理する。F2とF3を独立に変化させた場合の/r,l,w/知覚をF1の2つの条件(Full Condition と -F1 Condition)について別々にFig.2に表した。反応の集中度が濃淡で示され、各刺激に対してもっとも頻度の高かった刺激を示すように/r/ /l//w/の境界線が引かれた。この結果より、以下のことが示された。

(1) A群ではF3が2000-2200Hzあたりに/r/と/l/の境界がある。/w/の知覚はFull Condition においては、F2が低く、F3が2000Hz付近の場合に極在しているが、-F1 Condition では/w/の範囲がF3が高い方向に広がった。

(2) J群では Full Condition の場合、A群に比べて反応の集中度が低く、/r/と/l/の境界はF2、F3の両軸に依存する。また、F2が低い領域でA群と同様/w/が出現し、その範囲はA群より広い。さらに -F1 Condition においても、A群に比べて/w/の領域が広がり大きい。

(3) JE群では、反応の集中度、/r/と/l/の境界の傾き、/w/の出現する領域、ともにA群とJ群の中間の結果を示した。

F1遷移時間のみを70msecから10msecにと変化した場合の結果をFig.3に示す。



(F2, F3)が (1200Hz, 1800Hz)において、F1遷移時間が減少するとともに各群とも /r/から /l/に変化し、 /w/の知覚はほとんどみられなかった。特にA群、J E群では変化が鋭く、反応が分化していたが、J群では反応の分化が乏しかった。(800 Hz, 2600Hz)においては、A群とJ E群では、F1遷移時間の減少に伴って、反応が /w/から /l/に変化し、特にA群では /r/の知覚はほとんどみられなかった。しかし、J群では /w/から /r/と /l/の混在領域に変化した。また、 /w/という判断はA群、J E群、J群の順に多くなった。

## 2.5 考察

前節の実験結果をまとめ、考察を加える。

(1) A群すなわちアメリカ人では主としてF3周波数に着目して /r/と /l/を聴き分けている。とくに、F1の遷移時間をF3とともに変化させる Full Condition では、F1の遷移時間を70msecに固定した場合 (-F1 Condition)に比べ、F2周波数が低い領域に表れた /w/が減少し、 /l/の反応となる。つまり、F1遷移時間が /l/と /w/を区別する手掛かりとして使われている。このことは、F2, F3周波数を固定し、F1遷移時間だけを変化させた結果(Fig.3)からも明かである。すなわちF2が低い場合(Fig.3の下段の例)、F1の遷移の性質は /l//w/知覚の重要な手掛かりとなる。

(2) J群では /r/と /l/の区別に F3、F2の両方の手掛かりを用いている。F2開始周波数に対する依存度は個人差が大きく、F3よりもむしろ F2を手掛かりとして用い /r/と /l/を区別している被験者もあった。また、 /w/と知覚される領域がA群より広い。このことは、F1遷移時間を変化させた場合の知覚特性(Fig.3 下段)において一層顕著に認められる。

(3) J E群では、A群とJ群との中間的な結果を示した。しかし、これは平均的な結果であり、被験者の個人差は非常に大きい。すなわち、J E群の被験者は、2つの群に分かれ、おおむねA群あるいはJ群のどちらかと類似した結果を示す者に分かれる。A群と類似し、判断のカテゴリーがはっきりしていた被験者は、3名であった。その3名はすべて幼少時(2-3、4-5、6-8才)にアメリカに滞在した被験者であった。一方、これら3名以外の被験者のアメリカ滞在経験は13才以上の年齢であった。この結果から、幼少時における言語習得の経験が知覚に重大な影響を及ぼすことが示唆される。

## 3. 実験 II

実験 I の結果から F2、F3 開始周波数や F1 遷移時間の /r//l/ 知覚に果たす役割が明らかになった。すなわち、日本人が /r//l/ 知覚において用いている手掛かりはアメリカ人のものと異なり、それぞれの被験者によって多種多様な音響的手掛かりを利用してようやく /r//l/ を聴き分けることが可能となる。従って何等かの総体的な手掛かりの欠如が起これば、これが知覚に及ぼす影響はアメリカ人の場合に比較して非常に大きいことが予想される。本実験では、手掛かりの総体的な欠如の例として次の 2 つの条件を取り上げる。

(1) [r-l] の開始定常部を取り除く (実験 II A)。 (2) 刺激音の呈示レベルを低下させる (実験 II B)。

また、J 群の被験者数を増し、実験 I の各条件を加えた実験を行うことにより A 群と J 群が知覚に用いる手掛かりに関するさらに詳細な検討を行うことを目的とした。

### 実験 II A

本実験では、[r-l] 刺激連続体の各刺激の子音開始定常部を除去することによる知覚の変化を調べる。

#### 3.1.1 刺激音

F2 と F3 の開始周波数に関し、以下のような組合せを作成する。F2 は 800Hz、1000Hz、1200Hz の 3 種類とし、F3 は F2 とは独立に 1200Hz から 3000Hz まで 200Hz ステップで変化させ、10 種類とした。その際、F2 の値が F3 の値を超える組合せは排除した。従って F2 と F3 のすべての組合せは、Fig. 1 のタイプ A' の範囲で示される 29 種類 ( $3 \times 10 - 1$ ) となる。

これらとは独立に次の 3 つの条件を設けた。すなわち、

- (1) Full Condition : 実験 I と同様に、F3 の値が [r] から [l] に動くのと連動させて F1 の遷移時間を [r] (70msec) から [l] (16msec) まで変化させた条件。
- (2) -F1 Condition : F1 遷移時間を 70msec に固定させた条件。
- (3) -ISS Condition : 刺激音の子音開始定常部 (ISS: Initial Steady State)

を取り除いた条件。すなわち、Fig.1において、ホルマント遷移開始前の語頭を除く。したがって本実験で使用する刺激音の総数は87種類(29x3)となる。実験Iと比較してわかるように上記(1),(2)の条件は実験Iと全く同様である。

### 3.1.2 被験者

A群 10名、J群 36名。

### 3.1.3 手続き

87種類の刺激がランダムイズされ、1セッションを構成する。各被験者には主として18セッションが課された。反応は、/r//l//w/の3つの選択肢からの強制選択とした。

### 3.1.4 結果

実験結果をFig.4にまとめた。図の表示形式はFig.2と同じである。

(1) A群では実験Iの結果と同様、反応が集中しており、Full Condition と -F1 Conditionの結果は実験Iと同様である。すなわち、実験Iと本実験における被験者の差異はほとんど認められない。-ISS Condition では/w/の領域がF2の高い方向へわずかに広がっている。

(2) 日本人の/r//l/知覚が英語聴取能力に依存していることを考慮し、アメリカ人発話による自然音声を使った/r,l,w/聴取テストによって測定された被験者の英語聴取能力に基づいて、J群を聴取能力の高い J-Skilled群と聴取能力の低い J-Unskilled群に分類した<sup>(12)</sup>。Full Conditionと-F1 Conditionに関しては、実験Iと同様の結果であるが、J-Skilled 群は J-Unskilled群に比べ、A群の反応に近い結果を示している。しかし、実験Iと同様、J群全体について反応の集中度が低かった。また、-ISS Condition では/l/の領域が大きく広がっており、この傾向は特にJ-Unskilled群において著しい。

### 3.1.5 考察

実験II Aの結果は以下のように考察される。

(1) Full Condition および、-F1 Conditionの結果は実験Iの結果を再現した。実験Iの結果は、本実験のJ-Skilled,J-Unskilled両群の平均的なものになっている。

(2) 日本人の知覚に関しては、J-Unskilled群で反応の集中度が非常に低く、/

r,l,w/を縛然としか聴き分けられないことが明らかになった。

(3) -ISS Condition の結果は特徴的であった。子音の開始定常部を除去することにより、A群では/w/の領域がわずかに広がる程度で顕著な影響は被らない。すなわち、A群では子音開始定常部の存在とは無関係に後続母音へのホルマント遷移部に着目した知覚が行われていることを示す。それに比べてJ群では/l/の領域が著しく拡大し、Full Conditionや、-F1 Condition における結果と大きく異なった結果を示した。この傾向は J-Unskilled群で一層顕著である。つまりJ群ではホルマント遷移だけでは、知覚の手掛かりとして不十分であり、子音開始定常部を含めた多くの手掛かりに基づいた知覚を行っていることを示唆している。

## 実験 II B

本実験では刺激呈示レベルの/r//l/知覚への影響を調べた。

### 3.2.1 刺激音

F2、F3開始周波数、F1遷移時間の手掛かりを同時に[r]から[l]に変化させ、合成音刺激を作成した。すなわち、F2開始周波数は 920Hz から1280Hzまで、F3開始周波数は1200Hzから3000Hzまで、F1遷移時間は 67msec から 13msec までをそれぞれ18等分して同時に変化させた。これは、F2、F3に関していえば、Fig.1のタイプBの直線上の点に相当する。

本実験では刺激の呈示レベルに関し、次の2条件を設定する。実験IおよびII Aと同じ呈示レベル(85dB SPL Condition)、およびこれらの呈示レベルより約 40dB 低い呈示レベル(45dB SPL Condition)である。

### 3.2.2 被験者

A群 10名、J群 36名(実験II Aと同じ)。

### 3.2.3 手続き

1セッションでは19種類の刺激を各5回ずつ呈示することとし、ランダム化した。刺激呈示レベル条件の変更はセッション毎に行い、それぞれ主として3-4セッションを被験者に課した。反応は/r//l//w/3選択肢の強制選択とした。

### 3.2.4 結果

J群は実験ⅡAと同様、J-Unskilled群とJ-Skilled群のふたつのグループに分けた。各群の結果をFig.5に示した。Fig.5の横軸は[r]から[l]へと変化する刺激番号を表す。

(1) A群では、呈示レベルを低下させ45dB SPLにすることによって/w/の反応が消滅し、/r/と判断されるようになった。

(2) それに比べて、J-Unskilled、J-Skilled群では/l/の判断にはあまり変化がなかったが、/w/と/r/の判断が不明確となった。

### 3.2.5 考察

結果から以下のことが考察される。

(1) 通常の呈示レベルである85dB SPL Conditionで反応のしっかりしているA群では呈示レベル低下の影響が少ない。すなわち、狭い領域に限定されていた/w/の知覚が消失するだけで、/r//l/知覚の傾向に基本的な変化は認められない。

(2) 通常の呈示レベルで反応が不安定なJ群では、呈示レベル低下による影響は大きく、特にJ-Unskilled群では、刺激番号1~5、すなわち、F2、F3開始周波数共に低い部分における/r//w/の区別は全く行われていない。

## 4. まとめ

日本人における/r//l//w/音知覚の手掛かりを合成音を使った実験を通して総合的に検討した。本研究で得られた結果を以下にまとめる。

(1) 日本人とアメリカ人において/r,l,w/知覚の際に利用する手掛かりは異なる。すなわち、日本人ではF1遷移時間、F2、F3開始周波数など種々の特徴と手掛かりを用いてようやく典型的な/r/と/l/を聴き分けている。

(2) 従って、日本人の知覚には個人差が大きく、本研究で用いられたように、合成音を用いた実験とは別の方法で測定された英語聴取能力によって被験者を分類し、グループ別に知覚の性質を調べる必要がある。

(3) また、日本人は、アメリカ人のような明確な知覚様式を持たないため、各

刺激に対する反応が集中せず、子音の開始定常部や、刺激呈示レベル低下等、手掛かりの一部が失われた場合に、大きな影響を受けることが明らかになった。

(4) (1)-(3)は、日本人の知覚においては、刺激連続体の作成に用いる音響的特徴、被験者の選択、実験方法などによって全く異なった実験結果を得る可能性を示している。

(5) さらに、本研究において、知覚実験は/r//l/の2者強制選択課題とせず、/r//l//w/の3者強制選択課題とした。日本人ではアメリカ人より/w/が知覚されることが多く、/w/を加えることにより、知覚の手掛かりに関するより詳細な検討が可能となる。また、アメリカ人においても/w/の影響が無視できないことが示された。本研究における/w/の知覚は/r//l/知覚に付随して生じたものであり、/w/の知覚をより積極的に取り扱うには、今後さらに検討が必要である。

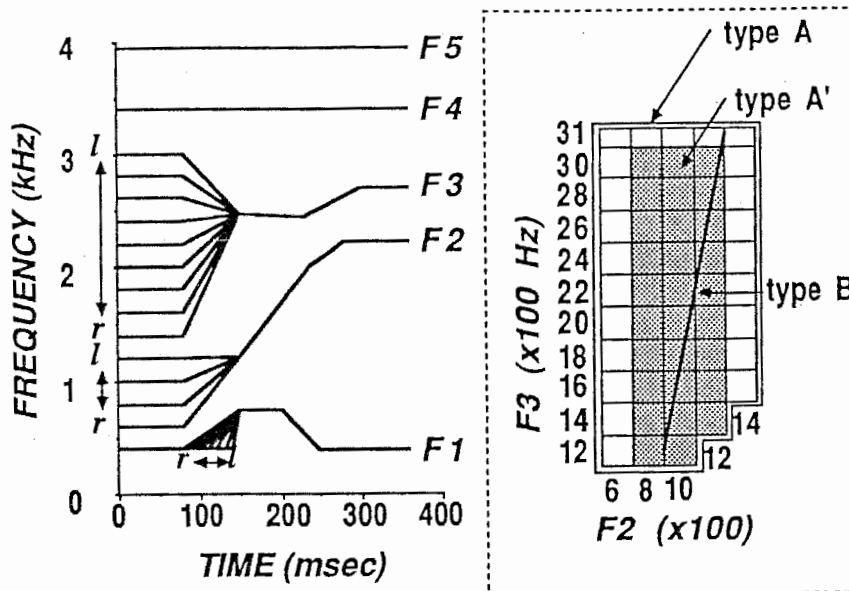


Fig.1 Schematic representations of spectrum for the part of the stimuli, which are extracted from /rai/-/lai/ in the /rait/-/lait/ series. Variations of F2/F3 frequency and F1 transition duration are shown.

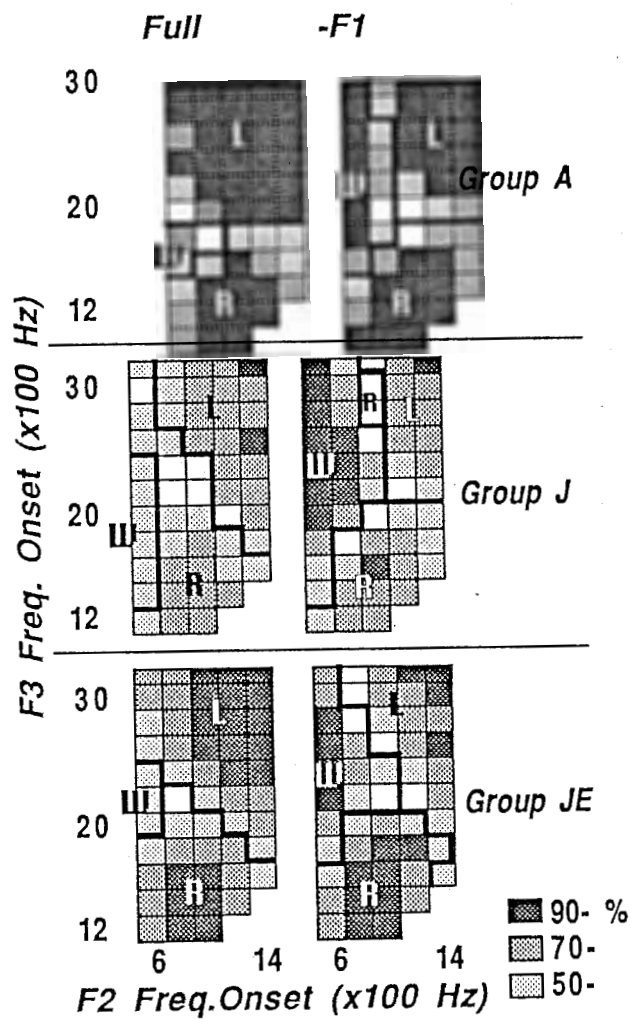


Fig.2 Response patterns of three listener groups; Native Americans (A), Native Japanese (J), and Japanese with living experience in the U.S.A (JE). F2 and F3 onset frequencies are represented by the abscissa and ordinate, respectively. Duration of F1 transition is varied from 70 to 13msec as F3 onset frequency is varied from 1200 to 3100Hz in the Full Condition (left), and duration of F1 transition is fixed to 70msec regardless of F3 frequencies in the -F1 Condition (right). The grey scale in the patterns indicates the concentration of responses. Bold lines representing the category boundaries are obtained based upon the most frequent responses to each category.



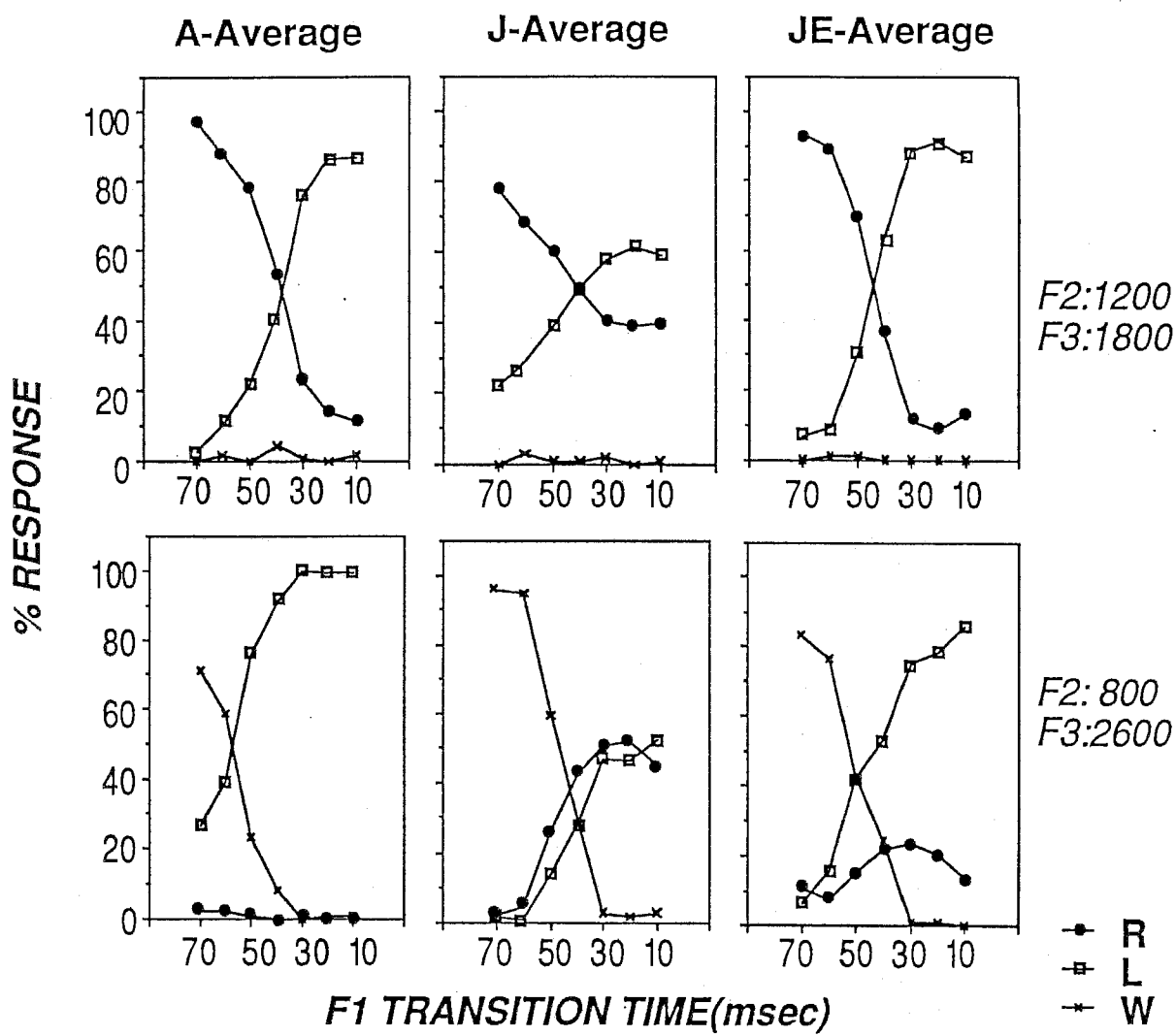


Fig.3 Response patterns in duration variations of F1 transition. F2 and F3 onset frequencies are selected for two conditions; (F2,F3) =(1200Hz,1800Hz) and (800Hz,2600Hz).

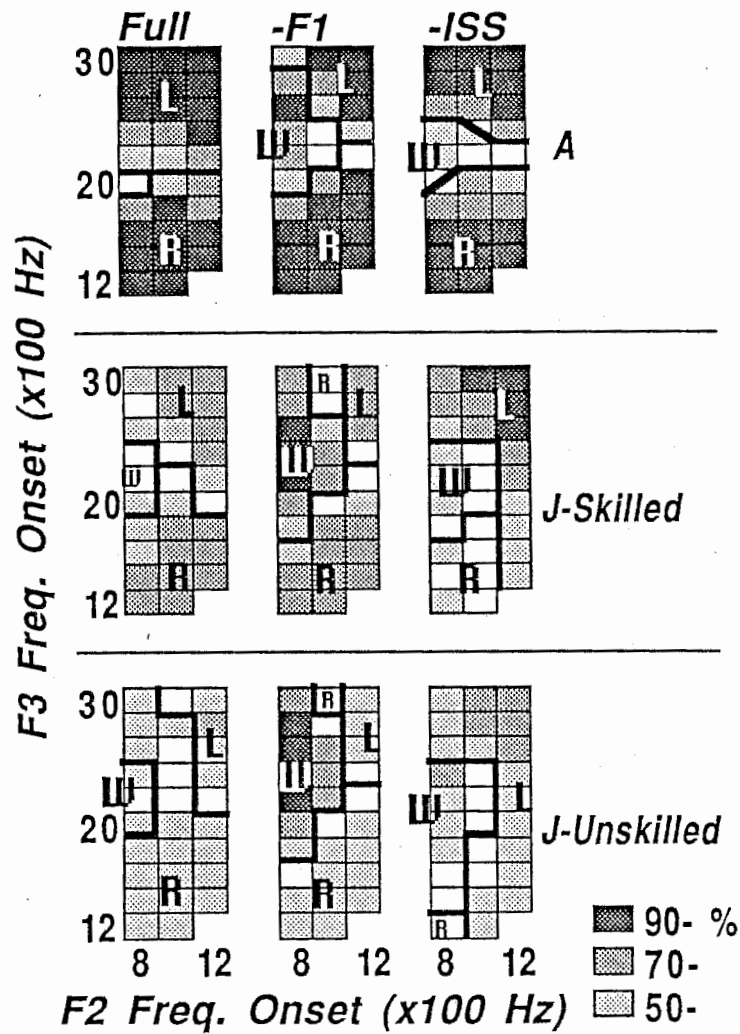
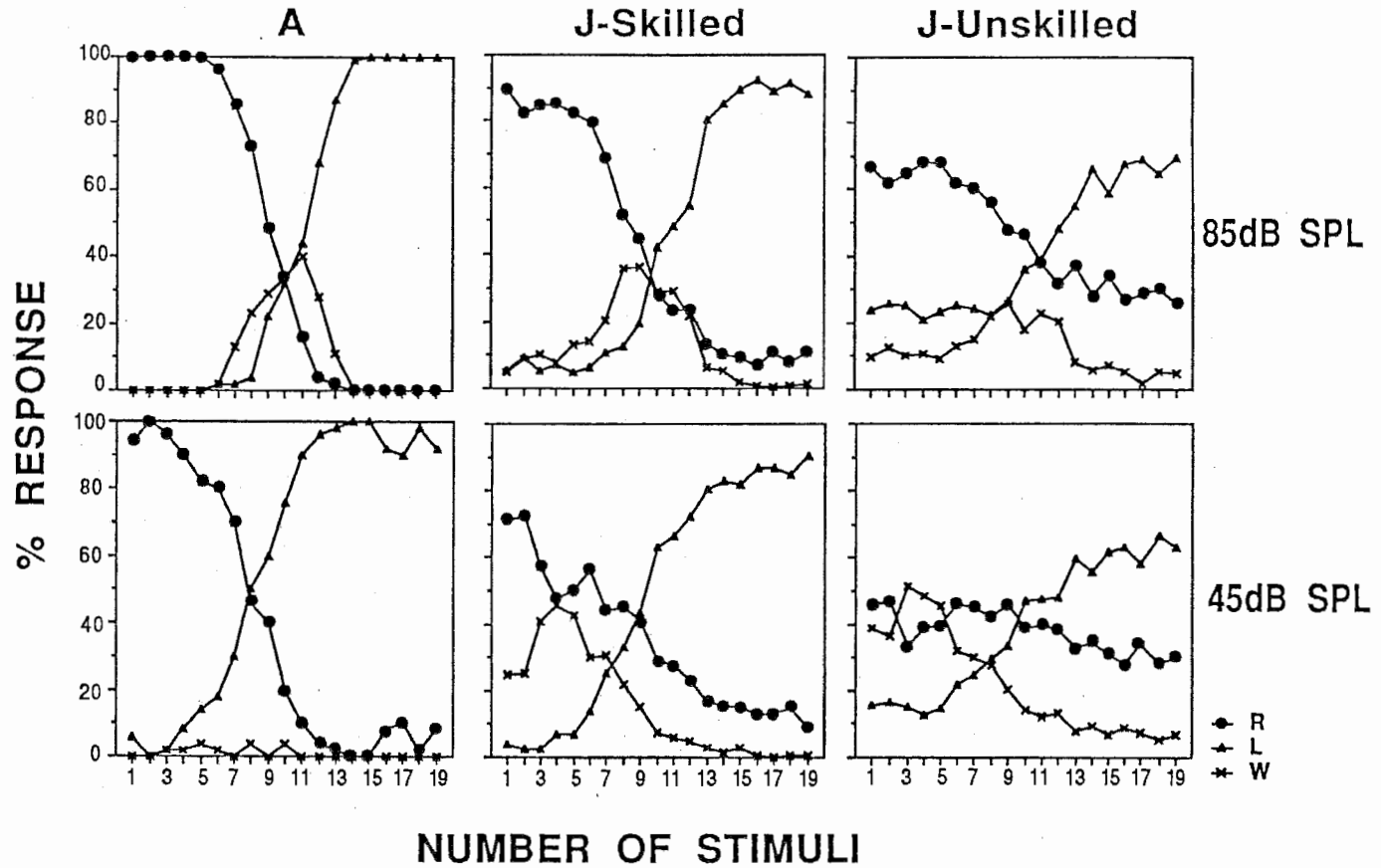


Fig.4 Response patterns of three listener groups; Native Americans(A), Japanese-Skilled(J-Skilled) with good perception in English word tests, and Japanese-Unskilled(J-Unskilled) with poor perception. The Full Condition is in the left, and -F1 Condition is in the middle. In the right is -ISS Condition, where initial steady state of the consonants is eliminated.



**Fig.5** Identification performance degradation with decrease of stimuli presentation level in each group. The results obtained with a decreased presentation level (45dB SPL) are compared with those obtained with the normal presentation level(85dB SPL).

## 文献

- (1) Mochizuki, M. (1981) The identification of /r/ and /l/ in natural and synthesized speech. *J. Phonet.*, 9, 283-303.
- (2) Sheldon, A., and Strange, W. (1982) The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Appl. Psycholinguist.*, 3, 243-261.
- (3) Liberman, A. M., Miyawaki, K., Jenkins, J. J., and Fujimura, O. (1973) Cross-Language Study of the Perception of the F3 Cue for [r] vs. [l] in speech- and nonspeech-like patterns. *J. Acoust. Soc. Jpn.*, 29, 315.
- (4) Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Fujimura, O., and Jenkins, J. (1975) An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Percept. Psychophys.*, 18, 331-340.
- (5) MacKain, K. S., Best, C. T., and Strange, W. (1981) Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Appl. Psycholinguist.*, 2, 369-390.
- (6) Shimizu, K., and Dantsuji, M. (1987) A cross-language study on the perception of synthetic speech sounds of [r-l]. *Proceedings of the Eleventh ICPhS.* 6, 71-74.
- (7) Dalstone, R. M. (1975) Acoustic characteristics of English /w, r, l/ spoken correctly by young children and adults. *J. Acoust. Soc. Am.*, 57, 462-469.

- (8) Lisker, L. (1957) Minimal cues for separation /w, r, l, y/ in inter-vocalic position. *Word* 13, 256-267.
- (9) O'Connor, J.D., Gerstman, L.J., and Liberman, A.M. (1957) Acoustic cues for the perception of initial /w, j, r, l/ in English. *Word*, 13, 25-43.
- (10) 山田玲子 (1988) 日本人における米語 /r, l, w/音知覚様式とその手掛かり -F1の遷移時間、およびF2, F3の開始周波数-. 関西心理学会第100回大会論文集、pp.13.
- (11) Cyphers, D.S. (1985) *Spire: a speech research tool*. MS Degree Thesis, MIT.
- (12) 山田玲子、東倉洋一 (1989) 日本人の英語 /r, l, w/知覚における特徴。日本音響学会聴覚研究会資料、H-89-6, 1-10.

# あとがき

日本人における /r//l/ 音知覚の特徴について、以下の2部に分けて総合的に検討を行った。

第1部：日本人の英語 /r, l, w/ 音知覚における特徴

第2部：日本人とアメリカ人の英語 /r, l, w/ 音知覚における手掛かり

その結果、次のようなことが明らかになった。

## (1) 反応の安定性：

アメリカ人は反応が安定しており、個人差が少ないのに対して、日本人は反応が集中せず、個人差も大きい。

また、アメリカ人は知覚が明確で、手掛かりの一部が欠如した条件でも、反応には大きな変化はなく、典型的な [r][l] それぞれの近傍の音は確実に /r//l/ と反応する。しかし、日本人では与える手掛かりの違いによって、反応パターンが変化する。つまり、[r]-[l] 刺激連続体を作成する場合に用いる手掛かりによって、アメリカ人に近い結果が得られる場合もあれば、全く逆の結果を得る可能性もあるということになる。日本人を用いて実験を行う場合にはどのような手掛かりを、どのような条件(呈示レベル、刺激音の品質、S/N比等)で与えたかということが重要な要因となる。

## (2) /w/ の知覚：

日本人ではアメリカ人より /w/ が知覚されることが多い。このため /w/ を選択肢に入れた知覚実験によって、より明確な結果を得ることができる場合がある。言い換えれば、選択肢を /r//l/ に限った実験から、日本人が /r/ と /l/ の範疇的知覚ができないことを説明することは難しい。/w/ を考慮することにより、より詳細な検討が可能となった。また、アメリカ人においても /w/ の影響が無視できないことが示唆された。刺激音の手掛かりや、文脈によっても /w/ の知覚が異なると思われる

る。今後この/w/の問題はさらに検討していかなければならないと思われる。

(3)日本人被験者の分類方法：

自然音声の聴取テストの成績と合成音の結果とはよく対応することが明らかになった。日本人を被験者として合成音を用いた実験をする場合には、併せて自然音声の聴取テストを行うことにより被験者の聴取能力を測定し、これによってグループ分けを行う必要がある。

(4)手掛かり：

日本人とアメリカ人において/r,l/知覚の際に利用する手掛かりは異なる。すなわち、アメリカ人では、主としてF3開始周波数、しかもそのホルマント遷移部を手掛かりとして/r/と/l/を聴き分けるのに対して、日本人ではF1遷移時間、F2、F3開始周波数など種々の特徴と手掛かりを用い、典型的な/r/と/l/を聴き分けている。しかも子音定常部に着目している。

従って、日本人の知覚は、子音定常部除去、刺激呈示レベル低下等、手掛かりの一部が失われた場合に、大きな影響を受ける。

また、これらの傾向は自然音声聴取能力の低い群程、顕著であった。

これらの結果に基づき、今後次のような課題に関して検討を加える予定である。

(1)音韻環境の考慮：

本実験では、刺激音が /ra<sup>ʌ</sup>t, la<sup>ʌ</sup>t, wa<sup>ʌ</sup>t/ と、語頭の子音で後続母音が /a<sup>ʌ</sup>/ に限られていた。今後、他の音韻環境（音節中の位置、後続母音）においても、上の様なことが成立するかどうか検討する。

(2)刺激セットの影響：

/r/と/l/をアメリカ人は絶対判断するのに対し、日本人では相対判断しているらしい。刺激セットによってどのような影響を受けるか、日本人において、/w/の知覚は/r//l/とは異なるのか等に関して検討する。

(4)知覚と調音の関係：

本報告の被験者により生成された英語 /r//l/音を分析することにより、知覚と調音の関係について考察する。

更に、学習実験等を通して、知覚様式の変化、調音の変化、およびそれらの関係について検討する。

(5)発達の側面：

帰国子女等、日本人で一時アメリカに滞在していた経験を持つものの知覚を調べ、滞在していた年齢との相関をみることにより、言語環境の影響を受ける年齢に関して調べる。

それと併せて、種々の年齢における知覚データを集積することにより、/r//l/音知覚の発達に関して検討する。

これらの結果から、アメリカ人と日本人の知覚様式の差、知覚と生成の関係、音声知覚の発達、環境が知覚に及ぼす影響等の問題点について明らかにしていきたい。