

アメリカ英語話者による日本語短母音・長母音・促音の学習 — 第2言語音声学習の基礎研究と教材開発に関わる諸問題 —

山田 恒 夫

第2言語音声の学習には、訓練刺激の特性、訓練方法の特性、学習者の特性等に関するさまざまな要因が作用する。そうした要因やその交互作用について、これまでも、いくつもの言語を対象に実験的研究がなされてきた (Strange & Jenkins, 1978; Flege, 1990; Strange, 1995)。しかしながら、これまでに十分検討されることのなかった視点・要因も少なくない。本研究では、はじめに、こうした問題点を概観する。次に、こうした論議をふまえわれわれが行った、音声知覚学習過程を記述するための実験パラダイムを報告する。日本語を第2言語として獲得する際の音声学習上の問題点については、これまでも数多くの報告があり、日本語に独特な拍と関係する音韻 (例えば、促音や撥音) が困難とされている。本研究では、「アメリカ英語話者による日本語短母音・長母音・促音の学習」過程を記述し、主に音声知覚学習に関する知見 (山田ほか、1993, 1994; Yamada et al., 1994, 1995) を検討する。最後に、こうした基礎的知見を日本語音声学習用教材にどう反映させていけばよいか、その応用に関わる問題点を、日本語教材のマルチメディア化、ネットワーク化への最近の動向と関連させながら考察する。

1. 背 景

日本語話者にとっての英語 /r-1/, アメリカ英語話者にとっての日本語拗音、促音のように、学習者の母語によって特徴的に困難な音声の学習がある。しかしながら、バイリンガルの

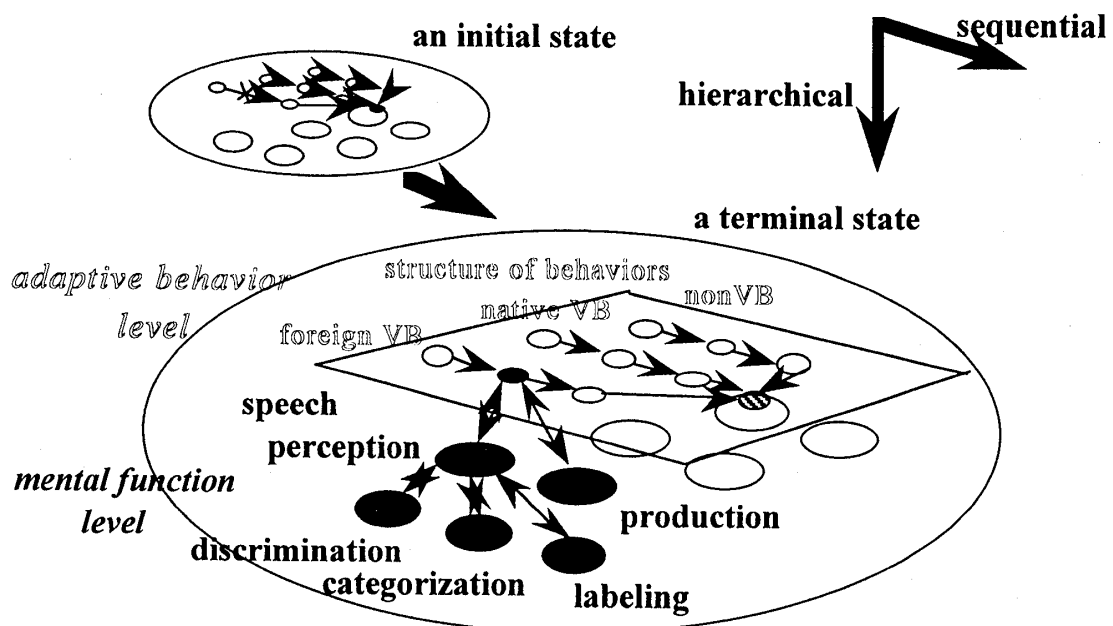


Figure 1 Hierarchical and Sequential Structures in speech perception learning (Yamada & Yamada, 1996)

音声言語能力が例示しているように、こうした音声知覚・成功特性の相違は、生得的に決定されているものではなく、個体発生の過程、すなわち、発達の過程で形成されたものである。したがって、音声知覚・生成のメカニズムを明らかにする上で、その獲得維持過程を記述することは不可欠な方法であるといえる。一般に、音声知覚や生成の研究において、それが個体発生の過程で獲得されたものであり、動的な学習過程としての側面を有するという視点も必要である（音声研究における学習パラダイム、Yamada & Yamada, 1996）。

第2言語学習者の音声知覚学習に関する実験的研究において見られた問題点を整理すると以下の3点にまとめられる。

- (1) 第2言語学習者の音声知覚学習に関しては、これまでもいくつかの言語を用いて実験的な研究がなされてきた。しかし、その目的は、どのような訓練方法が最適かという効率性の観点からの、訓練方法の比較にとどまるものが多く、獲得過程を記述するという観点は希薄であった。
- (2) こうした研究においては最適な訓練刺激および訓練手続の組み合わせが検討されたが、その評価には問題点が少なくない。音声学習も、他の学習行動同様、階層的時系列的構造を有し、階層間・階層内でさまざまな相互作用が生じうる。音声知覚学習といっても、実際には、個々の状況に対する全体的適応行動の一部として位置づけられるべきものであるから、音声知覚学習の比較に際しては適応行動レベルの学習行動が統制されていることが前提条件となる（Figure 1）。
- (3) どのような学習も白紙状態（*tabula rasa*）からではなく所与の初期状態から開始されるものであるが、第2言語学習においては、初期状態の重要性が理論的に明白である。こうした初期状態は、音声知覚学習の場合、訓練対象およびその周辺刺激に対する学習開始以前の知覚特性として記述できる。第2言語学習時の音声知覚におよぼす第1言語（母語）の影響は本質的とされ、これまでいくつかの仮説が提出されてきた（“Assimilation Model”, Best & Strange, 1992; “Speech Learning Theory”, Flege, 1990）。くわえて、それまでの外国語学習経験等によって、初期状態としての知覚特性は個体間で大きく変化する（cf. Yamada & Tohkura, 1992）。第2言語学習研究において、初期状態の相違を考慮することなく学習効果を論ずることには問題がある。

2. アメリカ英語話者による日本語短母音・長母音・促音の学習

第2言語として日本語を学習する際習得困難な音韻に、日本語拍と関連する特殊音素がある（村木・中岡、1990；小川・林ほか、1982）。本研究では、アメリカ英語を母語とし日本語経験のないアメリカ人を対象に、日本語短母音・長母音・促音の知覚訓練を行った。最近、複数の実験的研究が、母語にない分節音韻カテゴリーであっても、適切な方法さえ用いれば、成人でも学習可能であること

Table 1. Examples of the minimal triplets contrasting short vowel, long vowel and /Q/ used as stimulus materials.

Response Category	Structure	Examples
short V	C1V1-C2V2	/kato/
long V	C1V1-V1-C2V2	/kaato/
Q	C1V1-Q-C2V2	/kaQto/

を示した (Jamieson & Morosan, 1986 ; Logan, Lively & Pisoni, 1991 ; Lively, et al., 1994)。日本語促音の知覚あるいは長母音、短母音の区別においては、拍感覚に関わる長時間およびピッチの情報が大きな役割を果たしている。本研究では、こうした超分節的 (suprasegmental) な現象に関しその学習効果を検討した。

本研究では、まず、第2言語学習における音声知覚獲得過程をいかに記述すればよいか、方法論的に検討した。音声知覚学習の達成度はさまざまな刺激次元における般化で比較できる。訓練場面における正答率の推移は多用される指標ではあるが、適応行動としての全体的学習行動の反応遂行 (performance) を反映しており、動機づけや学習行動に関するさまざまなパラメータによっても影響をうける。そこで、オペラント条件づけをはじめとする、いくつかの技法を使用し、こうした全体的学習行動の反応遂行が条件間個体間で変動しないよう制御する必要がある。達成度テストとしては、従来の「プリテスト-ポストテスト」デザイン (Strange & Dittmann, 1984) に加え、訓練前半終了時の中間テスト、さらに毎セッション後のミニテストを用意し、般化過程をより詳細に記述できるようにした。異なる訓練刺激セットを用意し、獲得過程に差がないか検討した。

方法

(被験者) 日本語学習経験をもたないアメリカ人15名。いずれもアメリカ英語を母語とするモノリンガルで、方言は一般米語 (General American)。訓練は米国サウスフロリダ大学で実施された。被験者には実験参加に対する謝礼と成績に応じた報酬が支払われた。

(音声刺激) 音声刺激は、/C1V1C2V2/-/C1V1V1C2V2/-/C1V1QC2V2/ というミニマルトリプレットをなす2拍あるいは3拍の日本語自然音声 (例えば、かと、かあと、かっと ; cf. Table1)。V1として日本語5母音、C2として/k, s, t, h (F), g, z, d, b, p, sh, ch, dz/の12子音を用いた (Table 2)。このとき、V1、C2が同じでも、C1、V2の異なるトリプレットを複数組用意した。話者は日本語を母語とする男性1名、女性3名で、女性2名は関西方言、のこり2名は関東方言である。いずれの話者も長期間海外に滞在した経験はない。音声の収録はATR人間情報通信研究所 (京都府相楽郡精華町) 内で実施された。収録にあたり、同一トリプレットに属する音節はできる限り同じピッチ、アクセントで発声することが求められた。なお、日本人5名により評価テストを行ったところ、その平均正答率は99.4%であった。実験変数として操作されたのはV1、C2であり、訓練用刺激セットとして、

Table 2. Combinations of V1 and C2. Combinations for the V1-variable training were enclosed with a square, and those for the C2-variable training were enclosed with ellipses. Combinations marked with "f" were used in the full-test, and those marked with "m" were used in the mini-test.

C2	V1				
	a	i	u	e	o
k	f,m	f,m	f,m	f,m	(f,m)
s	f	f	f	f	(f,m)
t	f	f	f	f	(f,m)
h(F)	f	f	f	f	(f,m)
g	f	f	f	f	(f,m)
z	f	f	f	f	(f,m)
d	f	f	f	f	(f,m)
b	f	f	f	f	(f,m)
p	f	f	f	f	f
ʃ	f	f	f	f	(f,m)
tʃ	f	f	f	f	f
dʒ	f	f	f	f	(f,m)

V 1 は / O / に固定するが C 2 は 10 種類 (/ k , s , t , h (F) , g , z , d , b , s h , d z /) 用いる C 2 変動 (/ C1oC2V2 /) セットと、C 2 は / k / に固定するが V 1 として 5 母音用いる V 1 変動 (/ C1V1kV2 /) セットが用意された。訓練刺激として女性 3 名の発声がいられ、また V 1 変動セットでは各 V 1 につき C 1、V 2 の異なる 2 トリプレットがいられたので、各訓練用刺激セットは 30 トリプレットからなる。テスト用刺激セットとして、V 1 として日本語 5 母音、C 2 として 12 子音、4 話者用いるフルテストセット (240 トリプレット) と、その一部から構成したミニテストセット (/ kV1C2o / の一部、女性 3 名の発声、計 42 トリプレット) が用意された。訓練用刺激セットとテスト用刺激セットでは、C 1 および V 2 の異なる V 1 - C 2 トリプレットがいられ、構成要素は重複しない。音声は一旦 D A T に記録されたのち 16 ビット、44.1 あるいは 22.05 k H z で量子化された。

(実験装置) 音声刺激はヘッドフォン (スタックス、S R - A) により両耳的に呈示した。実験制御および反応記録は、コンピュータ (NeXT Cube Turbo) により自動的に行った。

(実験手続) 実験条件は、C 2 変動 (/ C1 o C2V2 /) セットを用いて訓練を行う C 2 変動 (C2-variable) 訓練条件、V 1 変動 (/ C1V1kV2 /) セットを用いて訓練する V 1 変動 (V1-variable) 訓練条件、テストのみ実施し訓練は行わない統制 (control) 条件の 3 条件である (各群独立、Table 3)。各条件とも、第 1 日に課題馴化およびプリテストセッション、第 2 - 5 日に訓練セッション (前半)、第 6 日に中間テストセッション、第 7 - 10 日に訓練セッション (後半)、第 11 日にポストテストセッションを行う。プリテスト、中間テスト、ポストテストの各セッションは、フルテストセット 1 シリーズ (720 試行、約 70 分) からなる。また、訓練セッションでは、訓練 2 群では同定訓練 (270 強化、約 20 分) の後ミニテスト (126 試

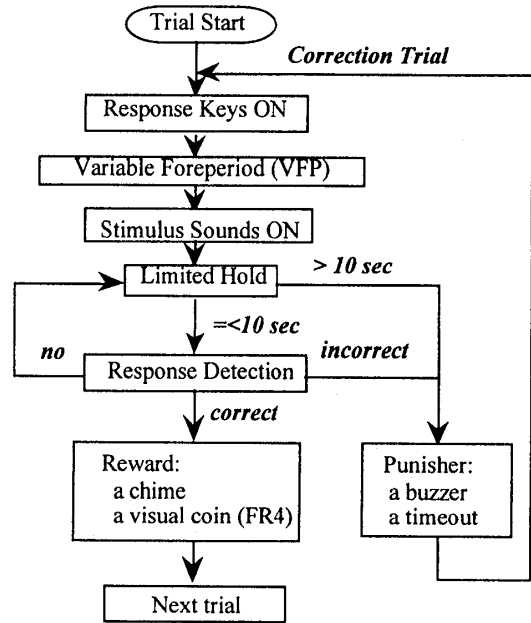


Figure 2. Flowchart of a training trial

Table 3 Schedule of training experiment

DAY		Conditions	
		C2 Variable V1 Variable	control
1	pretest	[1] habituation (10 trials) [2] full perception test (720 trials) [3] production test (30 trials)	
2	training 1	[1] identification training (270 trials) [2] mini perception test (126 trials)	[1] mini perception test (126 trials)
3	training 2		
4	training 3		
5	training 4		
6	interim test	[1] full perception test (720 trials) [2] production test (30 trials)	
7	training 5	[1] identification training (270 trials) [2] mini perception test (126 trials)	[1] mini perception test (126 trials)
8	training 6		
9	training 7		
10	training 8		
11	posttest	[1] full perception test (720 trials) [2] production test (30 trials)	

行、約10分)を行い、統制群ではミニテストのみ行う。同定訓練は、各訓練条件に対応した訓練刺激セット(C2変動セットかV1変動セット)3シリーズ、ミニテストはミニテストセット1シリーズからなる。訓練、テストとも、刺激の呈示系列はランダムに変更された。テスト課題はいずれも、刺激音声(繰返しなし)呈示後、3つの選択肢("short vowel", "long vowel", "short vowel + Q")の1つを選択する強制選択課題である。訓練もテストと同じ選択肢を用いた同定課題であるが、反応まで同一刺激が反復される点(刺激間間隔2秒、制限時間10秒)、正誤のフィードバックがある点(正答時;チャイムおよび現金に交換可能なトークン、誤答時;ブザーおよび矯正試行)が異なる。トークンはFR4スケジュールで呈示された。矯正試行として、正答が出現するまで同一の試行が反復された(Figure 2)。

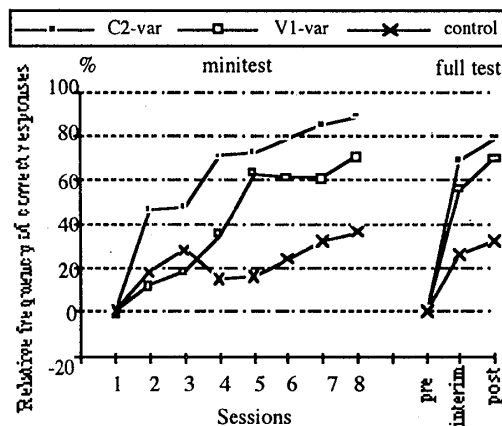


Figure 4. Generalization in minitests of each training session (left) and in two full tests, that is, interim test and posttest (right). Numbers on the ordinate show averaged relative frequency of correct responses. In each subject, the number of correct responses in the first session was the starting point (0%) for the rest of minitests and that in the pretest was the starting point for the interim test and posttest (Yamada, Yamada & Strange, 1994).

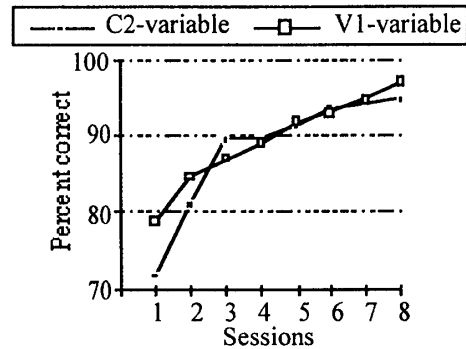


Figure 3. Learning curve of each training group in the identification training. Numbers of the ordinate show mean percentage of correct responses in each session (Yamada, Yamada & Strange, 1994).

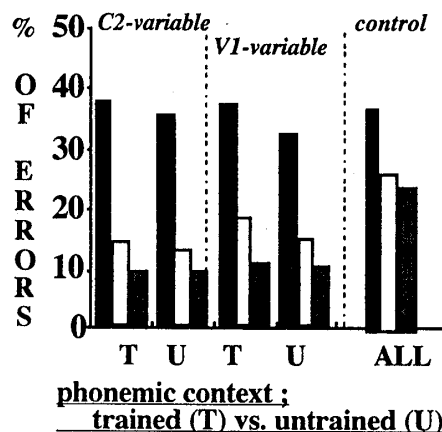
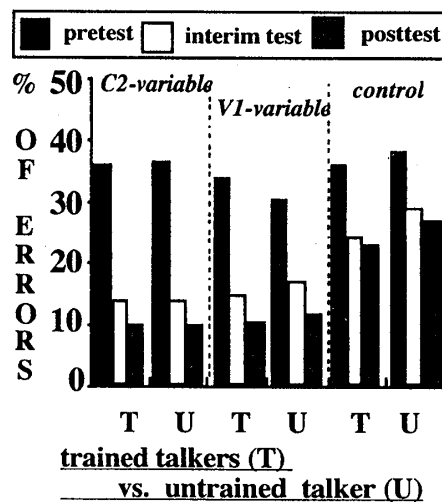


Figure 5 Generalization processes in interim test and posttest (Yamada, Yamada & Strange, 1995). The ordinate indicates mean percentage of errors in sub-categories, (1) talkers used in training vs. an unfamiliar talker (upper), and (2) trained phonemic contexts vs. untrained contexts (lower).

結果および考察

2. 1) 日本語経験のないアメリカ英語話者による日本語短母音・長母音・促音の知覚一ブリテスト

各音韻文脈別の平均正答率は60-70%で、チャンスレベル(33%)を上回った。これは長母音と短母音の区別が容易であったことによる。本実験では、訓練に先立ち、日本語特殊音韻に関する音声学的説明等を行わなかった。しかし、選択肢のラベルからの推測と2、3回の典型例の呈示で十分であったようである。C2が/s/, /s h/, /c h/, /k/の場合、誤答が多かった。以後、条件間で初期状態に偏りができないよう、被験者を各群に割り当てた。

2. 2) 同定訓練における学習曲線

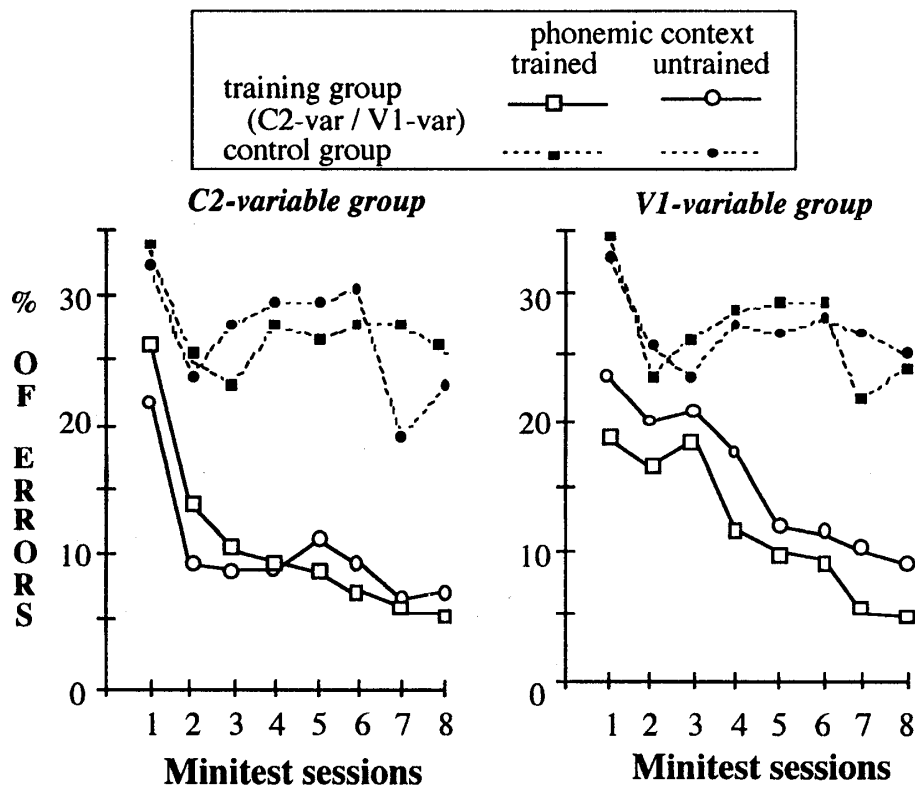


Figure 6. Generalization processes in the minitests. The ordinate indicates mean percentage of errors in each training condition, that is, (1) training group vs. control group and (2) trained phonemic contexts vs. untrained contexts. Constituents of each context condition in control group correspond to those in each training group.

両訓練群において、訓練の進行とともに、訓練場面における正答率の著しい向上がみられた。この弁別行動に見られる反応遂行の高さは、本実験で用いられた動機づけおよび訓練手続の有効性を示している。また、平均正答率では両訓練条件の学習曲線に差がなかった。全体的学習行動の反応遂行レベルの相違によって、音声知覚学習過程が異なる可能性がある。しかし、本訓練では音声知覚学習へのそうした影響が訓練条件間で大差なかったことを示唆するものといえる (Figure 3)。

2. 3) 達成度テストにおける般化

同定訓練の進行にともない、音韻文脈に関し新奇な刺激が呈示されるミニテストや、加えて新奇な話者の刺激が呈示される中間テスト、ポストテストにおいて、広汎な般化が視察された

(全体的傾向、Figure 4 ;フルテスト、Figure 5)。ほぼ完全な般化を示す被験者も複数あり、本課題は比較的容易に学習されたといえる。

2. 4) 訓練条件間の獲得過程の相違

各訓練セッションで行われたミニテストにより、C 2 変動訓練条件と V 1 変動訓練条件との間で、その獲得過程の異なることが明らかとなった。C 2 変動訓練条件では、最初の 3 セッションで急激に誤答数が減少するのに対し、V 1 変動条件では数セッションかけて徐々に減少した (Figure 6)。

促音は重ね子音 (geminata) といわれるが、音声分析によると、中間部が無音のもの (例、かっど) と雑音のもの (例、かっしゃ) がある。前者に属する /k, t, g, z, d, b, dz/ と後者に属する無声摩擦音、/s, h (F), sh/とで、その獲得過程を比較したところ、両訓練条件ともに、後者 (雑音) の獲得が悪く、その多くにおいて長母音と混同されることが明らかとなった。こうした C 2 の効果に比べ、V 1 の効果は明らかではなかった。

2. 5) 個人差

本実験は、アメリカ国内で日本語経験のない被験者を用いて実施したのにもかかわらず、その初期状態には大きな開きがあった。また、訓練群に、学習効果は見られたものの到達基準に到達しない被訓練者がいた一方、統制群にもほぼ完全な般化を示す被験者がいた。このような個人差が何に由来するのか、それぞれの状態に最適な訓練刺激構造とは何であったのかは今後検討していく必要がある。

本研究においては、アメリカ人日本語学習者における音声知覚獲得過程が、さまざまな般化テストを用いて記述された。後続子音に関し variability の大きい訓練刺激セットは特に訓練初期の学習を促進すること、無声摩擦音が後続子音である場合に学習が困難なことが明らかとなった。アメリカ英語話者に対する日本語短母音・長母音・促音の知覚訓練として、本研究で用いた方法は有効であり、音韻文脈・話者ともに広汎な般化効果を検証できた。同一訓練条件であっても、初期状態あるいは獲得過程の個体差は大きく、多様性に富んだ獲得過程を記述することが、その機制を明らかにし、また最適な訓練方法を開発する上で重要であると示唆された。

3. 教材開発に向けて

先の実験は、日本語音声教材制作に関し、いくつかの点において示唆的である。音声知覚学習において、学習者の初期状態や獲得過程には無視できない個体差があり、また速やかな般化のためには訓練用刺激音声の variation を大きくすることが有効であった。これを実現するには、学習者毎にきめこまやかな指導をおこなうことがのぞましく、一方で、より多くの話者の、さまざまな音韻文脈の音声が必要となるが、それには機械による個別学習が適している。実際、本研究ではコンピュータを利用した学習により短期間に広範囲の般化が観察された。本研究の刺激セットは最適化されているとはいえないが、こうした実験を繰り返し、訓練刺激セットの最適時系列構造が明らかになれば、その習得はさらに短期間で可能となるだろう。ただ、どのような内容でもコンピュータによる個別学習が適しているわけではない。反復を必要とする、

どちらかといえば低次の音声知覚や生成の学習は機械による代替に適した分野と考えられる。

本実験では、特別な訓練を施さなかった統制群においても、ある程度の成績向上が見られた。これは、単に外国語音声にさらされるだけでも、学習者に適度な動機づけがあれば、なんらかの学習が生じることを示唆している。しかし、その場合統制群各自によって学習されたものは、必ずしも日本語学習として適切なものとは限らない。最近エデュテイメントとよばれる新しいタイプの教材が出現しているが、本実験の結果を対応づけると、強力な動機づけ機能さえあれば教材内容が貧弱であっても、ある程度の学習効果も予想できるということになる。すでに単にマルチメディア教材を作ればよいという時期は過ぎ、その内容・効果から吟味すべき時期にさしかかっている。今後マルチメディア教材が学校や社会に普及するに際しては、その内容まで厳密に評価する必要があるだろう。

さまざまな教育分野において、教材の電子化、マルチメディア化、ネットワーク化が進展している。日本語教育も例外でなく、個別学習用、授業用、教師教育用等、さまざまな用途のメディア教材が開発提供されつつある。すでに海外では、授業資料やレポートの配布など授業支援のために、あるいはカリキュラムや教材所在情報など教師や学生向け情報提供のために、ネットワークが利用されている。そして、開発の主目標は、ネットワークを利用したマルチメディア教材に移りつつあるというのが実感である。今後、国内でも世界に発信できる優れた教材を開発するとともに、そのための教材・素材データベースセンター、著作権処理センターを整備することが望まれる。

補注

- 1) 本論文は、第17回東京音声言語研究会（1996年7月27日、国立国語研究所）で発表した論文に加筆したものである。
- 2) 本論文は山田玲子研究員（ATR人間情報通信研究所）およびWinifred Strange教授（サウスフロリダ大学）との共同研究をもとにしている。

謝辞

本研究の一部は、第1著者に対する、文部省科学研究費補助金（奨励研究A、一般研究C）および同在外研究員としての補助を受け実施した。実験遂行に際しては、Desiderio Saludes, Brett H. Fitzgerald 両氏の協力を得た。

参考文献

- Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20, 305–330.
- Flege, J. E. (1990). Perception and production: The relevance of phonetic input to L2 phonological learning. In C. Ferguson & R. Huebner (Eds.), *Crosscurrents in second language acquisition and linguistic theories*. Philadelphia: John Benjamins.
- Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English /ð / - /θ / contrast by francophones. *Perception & Psychophysics*,

40, 205–215.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/: III. Long-term retention of new phonetic contrasts. *The Journal of the Acoustical Society of America*, 96, 2076–2087.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, 89, 874–886.

村木正武・中岡典子 (1990). 撥音と促音－英語・中国語話者の発音－. 杉藤美代子 (編)、講座 日本語と日本語教育、第3巻 日本語の音声・音韻 (下). 明治書院. P p. 139–177.

小川芳男・林 大・他 (1982). 日本語教育事典、大修館書店.

Strange, W. (Ed. 1995). *Speech Perception and Linguistic Experience : Issues in Cross-language Research*. Timonium, Maryland : York Press.

Strange, W., & Jenkins, J. J. (1978). Role of linguistic experience in the perception of speech. In R. D. Walk & H. J. Pick, Jr. (Eds.), *Perception and Experience*. New York : Plenum. Pp. 125–169.

Strange, W., & Dittmann, S. (1984). Effects of discrimination training in the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics*, 36, 131–145.

Yamada, R. A., & Tohkura, Y. (1992). The effects of experimental variables on the perception of American English /r/ and /l/ by Japanese listeners. *Perception & Psychophysics*, 52, 376–392.

山田恒夫・山田玲子・Winifred Strange (1993). アメリカ人による日本語短母音・長母音・促音の知覚学習. 日本音響学会平成5年度秋季研究発表会講演論文集、395–396.

山田恒夫・Strange, Winifred・山田玲子 (1994). 第2言語学習者における音声知覚獲得過程の分析. 日本音響学会聴覚研究委員会資料 (H-94-61). P p. 1–8.

Yamada, T., Yamada, R. A., & Strange, W. (1994). Perceptual learning of Japanese mora syllables by native speakers of American English. *Proceedings of 1994 International Conference on Spoken Language Processing (Yokohama, Japan)*, 2007–2010.

山田恒夫・山田玲子・Winifred Strange (1995). 日本人英語学習者によるアメリカ英語母音・子音の知覚. 日本音響学会平成7年度春季研究発表会講演論文集、379–380.

Yamada, T., & Yamada, R. A. (1996). A review on experimental studies of second language speech perception training : Necessity of the learning paradigm in speech perception studies. *Bulletin of the Faculty of Human Sciences, Osaka University*, 22, 157–175.