

オハイオ州立大学言語学部音声学教室

研究開発部特別試験研究部門 内 田 照 久

はじめに

研究開発部の特別試験研究部門での研究テーマの一つに、外国語リスニング・テストをめぐる実施上の問題点の検討、及び問題内容や問題形式の在り方に関する基礎研究がある。

以前の大学入試フォーラムでの報告にもあるように、大学入試センター試験では、主に設備・技術上の問題から、少なくとも平成11年度までは外国語リスニング・テストは実施されない（大学入試センター、1997）。しかし、高校の英語科でのオーラル・コミュニケーション科目の導入もあり、リスニング・テスト実施上の問題点や問題内容に関連した基礎研究は現在も地道に続けられている。

今回、1997年3月から1998年1月まで、文部省の在外研究員として、アメリカのオハイオ州立大学言語学部の音声学教室に滞在した。そこで、日本語話者に特有な、言語音声に対する聴覚的知覚の特徴を研究する機会に恵まれた。ここに研究成果の一部を報告する。

1. 言語音声のための研究用話速変換プログラムPICOLA plus2の開発

言語音声の発話速度は、その発話内容の聴取理解度に対して大きな影響を与える。高齢者において聴力の低下は、最小可聴閾の上昇ばかりでなく、聴覚経路全般に渡る機能劣化が生じるため、語音識別速度が低下する。すなわち、早口で話された音声に対して、音自体は聞こえるものの、その内容が分からないといった現象が報告されている（中村・清山・池沢・都木・宮坂、1994）。

また、外国語教育における聞き取り学習の場面では、ゆっくりとした聞き取り易い発話速度の教材から、かなり早口な自然な会話速度までの、幅広い発話速度の教材が必要とされている。さらに今後は、学習者が自分の聴取能力の程度に応じた発話速度の教材を自由に選び、聴取訓練などに利用するといった自由度も求められる。

ここで、もし音声のピッチやスペクトル構造を変化させることなく、自然

な音声のまま、その発話速度のみを自由に变化させることができれば、上記の問題などへの対処も可能となろう。すなわち、高齢者用の放送補助機器や、話速変換型の補聴器、また単一の音声材料を自由な発話速度で利用できる新しい外国語教材用のツールも開発できる。

しかしそのためには、このような様々な用途や目的に適した発話速度の制御のために、その適切な制御範囲や、適否を吟味することが必要となる。そこで、発話速度に関する聴覚実験用の刺激作成にも利用できる、研究用の話速変換プログラムPICOLA plus2を開発した。

1-1 PICOLA plus2の概要

このPICOLA plus2は、パーソナル・コンピュータ上の音声ファイルに対し、その中の任意の複数区間に対し、個々に指定した割合で音声の時間軸上での伸縮を行うことができる。

また、摩擦音、破裂音など非周期的信号部分における変換時の音質劣化を低減するために、そこでの処理をスキップするための閾値を任意に設定するためのオプションも設けられた。

動作環境 本プログラムはMS-DOSアプリケーションであるので、一般のPC上で動作する。特別なハードウェアは特に必要としない。

対応フォーマット Windows用

Waveファイル、Complement Binary ファイル、Canopus 形式ファイル。いずれもデータ部分が16bit量子化のlinear PCMのモノラル・データであることが必要である。

なお、本プログラムは単独発声された言語音声のみをその処理対象としており、背景雑音を伴ったものや、重畳した会話音声、また音楽などには適さない。

1-2 言語音声の時間軸上での伸長圧縮技術の背景

本プログラムの方式は、Malar, Crochiere, & Cox (1981) の開発したTDHS (Time Domain Harmonic Scaling) の技術をその起源に持つ。それは音声の周波数軸上での情報圧縮を目的とし、それを時間軸上での操作によって実現したものである。この技術はその副産物として、音声のサンプリング周波数を一定にした場合には、音声の時間軸上での伸長圧縮を行うことができるという特徴を持つ。

森田・板倉 (1986) は、このTDHSの時間伸縮の用途に特化して改良を行い、より高音質なPICOLA (Pointer Interval Overlap and Add) のアルゴリズムを開発した。

本プログラムは、このPICOLAのアルゴリズムをエンジンとして、設計仕様を改良したものである。

近年では、河原・増田 (1996) らに

よるSTRAIGHTなどの優れた音声変換方式が開発されてきている。しかし、話速変換の操作のみに限定した場合には、計算機資源の面から鑑みた計算量の少なさと、その結果得られる音声の自然性とのバランスを考慮すると、PICOLA plus2ならではの利点もある。

1-3 聴覚実験への応用

本プログラムは話速操作の側面だけでなく、音声の特定区間の持続時間に対する制御にも応用できる。次に述べる外国人日本語学習者が不得手とする日本語特殊拍の聴覚的認知の特徴を調べるための実験においても、その音声刺激の作成に本プログラムが利用されている。

2. 日本語話者は無音、雑音、鼻子音、そして母音の中に何を聞いているか？

外国人日本語学習者にとっては、日本語特殊拍の長音 (long vowel) “一”, 促音 (geminate consonant) “っ”, 撥音 (moraic nasal) “ん”の聞き分けが非常に困難であることが知られている。特定の音韻の聴覚的な識別において、各国からの多くの外国人日本語学習者に支障が生じるということは、翻ってその現象の背後に、日本語話者に特異的な音声知覚上の聴覚的特徴があるものと想定される。

そこで、外国人学習者に特に識別困難が生ずる特殊拍の特徴を整理し、その音響的な特徴をまとめた。詳細は内田 (1998) に述べられているので、ここではその整理結果を述べる。

外国人日本語学習者において特殊拍の識別困難が生ずるのは、比較的、音響的に定常な音声区間の持続時間 (relatively steady-static speech sound duration) に対し、その音響的な手がかりが乏しい中で、心理的な情報処理過程によって区分化が必要とされる場面である。具体的には、促音を知覚するための閉鎖無音持続区間や摩擦雑音持続区間の区分化、撥音においては、撥音に同種の鼻子音が後続する条件下での鼻子音持続区間の区分化、さらに長音では、アクセント核が母音部に含まれない条件での母音持続区間の区分化である。このことは逆にとらえると、日本語話者は、無音や雑音の中に促音を聞き、鼻子音の一部に撥音を聞き、母音の継続中に長音を能動的に聞き取っていることを意味する。

これらの特殊拍に関し、この無音、雑音、鼻子音、母音を、音響的に定常的な性質を持つ音声区間として一元的に扱うこととした。そして、その持続時間の側面に焦点化することによって、異種の音韻にまたがって共通性の高い、日本語話者特有の聴覚的な知覚特徴の抽出を試みた。

ここでは、定常的音声区間の一次元的な操作が、それ単独で各音韻の知覚に十分な情報を有しているかを検討するため、カテゴリー的知覚の生起を指標として利用することとした。

このカテゴリー的知覚の生起を見出すには、各音韻境界付近での音声区間の持続時間に対する高い弁別力、及びそれ以外の領域での弁別力の低下をとらえる必要がある。そのため、まず識別課題実験 (discrimination task) が必要であり、さらにその音韻境界が音声刺激連続体上のどこにあるかを測定するための同定課題実験 (identification task) が必要となる。

2-1 実験音声刺激の作成

基準音声材料として無意味単語「ケソナト」/kesonato/が設定された。刺激作成のための原音声は、東京方言話者の成人男性が発音した「ケッソナト」/kessonato/が使用された。録音はオハイオ州立大学の言語学部の防音室で行われ、音声資料はDATに保存された後、あらためて量子化16bit、標本周波数22.050kHzでA/D変換された。

次に、その摩擦雑音区間/s/, 鼻子音区間/n/, 母音区間/a/, 及び破裂閉鎖無音区間/t/の持続時間がPICOLA plus2によって操作され、4種の刺激連続体が作成された。詳細はFigure 1.に示す。

2-2 被験者

実験には、24名の日本語母語話者が参加した。年齢が21歳から45歳までの聴力健常な成人で、男性14名、女性10名であった。

2-3 識別課題実験 (実験1)

実験刺激 各音声区間ごとに、連続体から3種の刺激対が、連続体上のshort, middle, 及びlongの範囲から選ばれた。各刺激対はターゲット区間の時間差が等しくなるように選定された。

手続き 実験はABX法による識別課題で、Xについての判断とその確信度の報告が求められ、防音室でのヘッドホン提示により個別に実施された。実験は、4種の音声区間ごと、連続体上の3領域の刺激対に関し、4回の反復測定を行う計48試行から構成された。

結果 識別課題の正答率について、被験者内2要因による音声区間要因 (4水準; /t/, /s/, /n/, /a/) × 領域要因 (3水準; short, middle, long) の分散分析を行った。

その結果、音声区間要因には有意な差は見出されず ($F(3, 69)=0.51, p>.05$), 音声区間の種類の違いによる、正答率への影響は見られなかった。

一方、領域要因に関しては有意な差が見出され ($F(2, 46)=74.59, p<.001$), さらに、音声区間要因と領域要

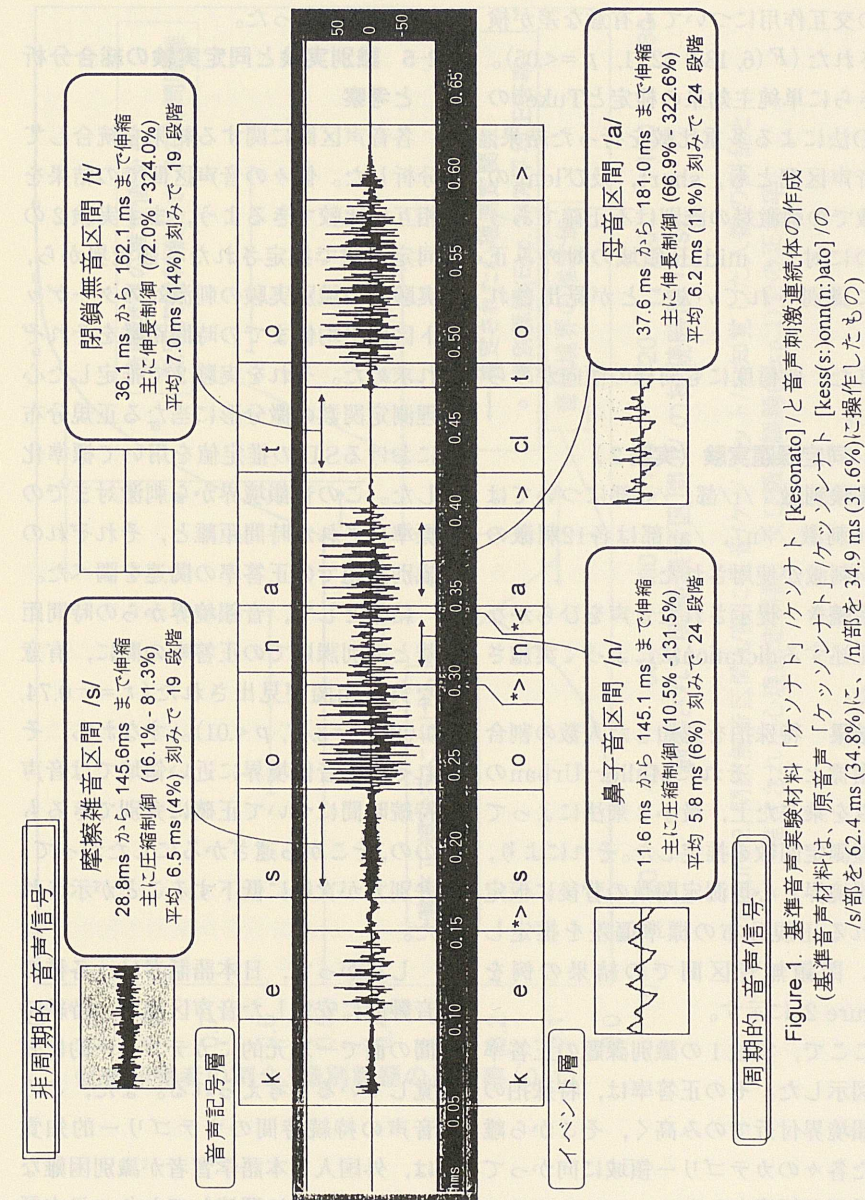


Figure 1. 基準音声実験材料「ケソナト」/kesonato/と音声刺激連続体の作成
(基準音声材料は、原音声「ケッソナト」/kessonato/の
/s/部を 62.4 ms (34.8%)に、/n/部を 34.9 ms (31.6%)に操作したもの)

因の交互作用についても有意な差が検出された ($F(6, 138)=2.71, p<.05$)。

さらに単純主効果の検定とTukeyのHSD法による多重比較を行った結果、各音声区間とも、short、及びlongの領域での刺激対の識別は不正確であったのに対し、middle領域の対のみ正確に識別されていたことが見出された。

また、確信度にも同様の傾向がみられた。

2-4 同定課題実験 (実験2)

実験刺激 /t/部、/s/部については各10刺激、/n/、/a/部は各12刺激の計44刺激が使用された。

手続き 提示された音声をひらがなで記述するdictation法によって実施された。

結果 特殊拍を認知した人数の割合を指標とし、それにMüller-Urbanの重みを乗じた上、最小2乗法によって心理測定関数を推定した。それにより、音韻境界と心理測定関数の背後に仮定される正規分布の標準偏差を推定した。閉鎖無音区間での結果の例をFigure 2. に示す。

ここで、実験1の識別課題の正答率も図示した。その正答率は、特殊拍の音韻境界付近でのみ高く、そこから離れた各々のカテゴリー領域に向かっては低下が観察された。

この傾向は、すべての音声区間にお

いて同様であった。

2-5 識別実験と同定実験の総合分析と考察

各音声区間に関する結果を統合して分析した。個々の音声区間での結果を相互に比較できるように、まず実験2の同定実験で推定された音韻境界から、実験1の識別実験の刺激対のターゲット区間平均値までの時間距離をそれぞれ求めた。それを実験2で推定した心理測定関数の微分形に当たる正規分布におけるS.D.の推定値を用いて標準化した。この音韻境界から刺激対までの標準化された時間距離と、それぞれの識別課題での正答率の関連を調べた。

結果として、音韻境界からの時間距離と識別課題での正答率の間に、有意な負の相関が見出された ($r=-0.74, t(10)=-3.47, p<.01$)。すなわち、それぞれの音韻境界に近い領域では音声持続時間について正確に弁別できるものの、そこから遠ざかるにしたがって、弁別力が次第に低下することが示された。

したがって、日本語話者は、各種の音響的に安定した音声区間を、持続時間の面で一次的にカテゴリー的に知覚していると考えられる。また、この音声の持続時間のカテゴリー的知覚は、外国人日本語学習者が識別困難な特殊拍の知覚に関連しており、日本語話者に特有であることが示唆されよう。

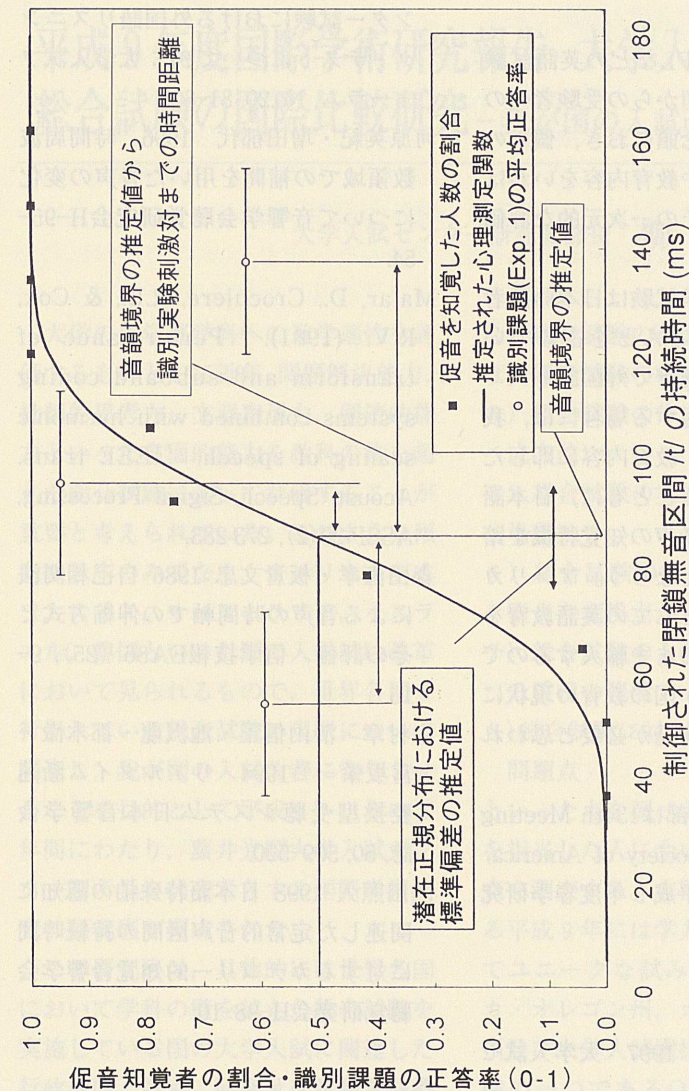


Figure 2. 閉鎖無音区間/t/部に促音/ヶ/を認知した人数の割合と推定された心理測定関数、及び識別課題(Exp.1)での正答率

3. まとめ

アメリカのTOEFLなどの英語試験の場合は、多くの国からの受験者への対処をはじめから念頭におき、個々の国々での教育課程や教育内容をいわば捨象することで、その一次元的な評価を可能にしている。

一方、大学の入学試験は日本語話者の受験生がそのほとんどを占めている。このような状況下で外国語リスニング・テストを検討する場合には、我が国での教育課程、教育内容に即した問題作成に留意するとともに、日本語話者に特有な言語音声の知覚特徴を踏まえることも重要となろう。アメリカなどでの第2言語としての英語教育を指向した試験をそのまま輸入するのではなく、今後は我が国の教育の現状に即した形のテスト開発が必要と思われる。

なお、本研究の一部は134th Meeting of the Acoustical Society of America, 及び日本音響学会平成9年度春季研究発表会で発表した。

引用文献

大学入試センター 1997 大学入試セ

ンター試験における外国語リスニング・テストについて'97: 大学入試フォーラム, No.20, 81-82.

河原英紀・増田郁代 1996 時間周波数領域での補間を用いた音声の変化について音響学会聴覚研究会H-95-54.

Malar, D., Crochiere, R.E., & Cox, R.V. (1981). "Performance of transform and subband coding systems combined with harmonic scaling of speech, "IEEE trans. Acoust. Speech, Signal Processing, ASSP-29(2), 273-283.

森田直孝・板倉文忠 1986 自己相関法による音声の時間軸での伸縮方式とその評価 信学技報EA86 (25), 9-16.

中村章・清山信正・池沢龍・都木徹・宮坂栄一 1994 リアルタイム話速変換型受聴システム日本音響学会誌, 50, 509-520.

内田照久 1998 日本語特殊拍の認知に関連した定常的音声区間の持続時間に対するカテゴリー的知覚音響学会聴覚研究会H-98-10.