

論文 / 著書情報
Article / Book Information

論題(和文)	話者認識と音声認識におけるスペクトル変動の影響の比較
Title(English)	
著者(和文)	古井 貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1977年講演論文集, , , pp. 361-362
Citation(English)	, , , pp. 361-362
発行日 / Pub. date	1977, 4

3—2—7 話者認識と音声認識におけるスペクトル変動の影響の比較*

古井 貞熙 (日本電信電話公社 武蔵野電気通信研究所)

I まえがき 話者認識に関して最も大きな問題点は、その精度が特徴パラメータの時期的変動の影響を受けやすい点にあり、この点に関してはこれまでにいくつかの認識実験と定量的な分析を行って来た⁽¹⁾。一方音声認識においても、各話者の1時期の標準パターンを、時期がたった後にも用いるという場合が考えられ、このときの時期差と認識率との関係についても、時期差が3年までの場合についてはすでに報告した⁽²⁾。ここでは時期差を更に大きくした場合の認識実験と、スペクトルの変動量に因する分析を行い、話者認識と音声認識におけるスペクトルの時期的変動の影響の比較を行う。

II 認識実験 音声認識の対象は0~9の10数字で、実験には男性3名が約5年間にわたって種々の時期的間隔で40時期にわたり、各時期に100サンプルずつ繰返し発声した音声サンプルを用いた。標準パターンには、10数字を順に1回ずつ発声した学習サンプルから音韻単位に作成したものをを用いた。標準パターンと認識サンプルの時期差に因して、種々の組合せて認識を行った。認識サンプルののべ数は、約8万サンプルである。

時期差と認識誤り率との関係を図1に示す。単語の統計的パラメータ⁽³⁾による話者認識(話者照合)の結果とあわせて示した。両者の縦軸のスケールは5倍異なっている。この結果から、どちらの認識においても誤り率は時期差とともに増加し、時期差がある程度以上になるとほぼ飽和すること、時期差に対する誤り率の

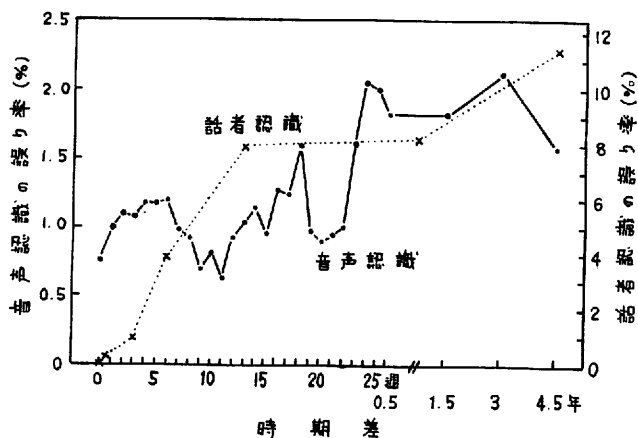


図1. 時期差と認識誤り率の関係

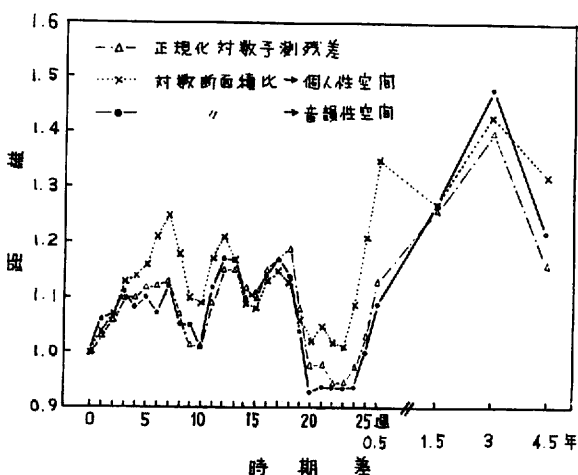


図2. 時期差とスペクトル変動量の関係

* Comparison between the Effects of the Temporal Variation of the Speech Spectrum on Speaker and Speech Recognition. By Sadaaki FURUI (Musashino ECL, NTT)

立上りは、音声認識よりも話者認識の場合の方が早いことなどがわかる。

Ⅲ スペクトル変動量との関連 認識実験に用いた各時期の各音韻の標準パターンを用いて、種々の時期差に対するスペクトルの変動量(距離)を求めた。距離尺度としては、音声認識の判定に用いた尺度に対応する正規化対数予測残差(NLPR)と、PARCOR係数を arctanh 変換した対数断面係比(LAR)をさらに判別分析におとづいて個人性空間および音韻性空間に写像した値のユークリッドノルムの二乗を用いた。時期差とこれらのスペクトル距離(時期差が3日以内のときの距離を1として正規化)の関係を図2に示す。音韻(5母音と/ɒ/, /ɒ/, /ɒ/)と3名の話者について平均した結果を示した。この結果から、LARの音韻性空間への写像とNLPRはよく似た結果を示し、LARの個人性空間への写像は、時期差による立上りが比較的早いことがわかる。比較のために男性30名の数字音声の中の5母音に関して、個人性空間における話者間平均距離と、音韻性空間における音韻間平均距離をもとめたところ、図2の縦軸の値で、前者は3.3、後者は10.2となり、約3倍の差があることがわかった。2種の認識における誤り率の差は、この差に起因していると考えられる。

つぎに、種々の時期差におけるNLPRと音声認識の誤り率との関係を調べたところ、図3のような結果が得られ、両者はよく対応していることがわかった。

Ⅳ 多数話者のスペクトル変動の分析 上記の検討に用いた3名の話者を含め、男性10名が約7年間にわたって発声した音声サンプルを用いて、時期差とスペクトル距離(NLPR)との関連を調べた。前節と異なり、10数字をランダムに発声した中から各数字音声を1つずつとり出し、各音韻のスペクトルを抽出した。話者と音韻に関して平均した結果を図4に示す。12か月までは時期差とともに距離が増加し、これ以上時期差が大い場合にはほぼ一定値をとることがわかる。この特性を指数関数とみなしてあてはめを行うと、時定数は約5か月となり、この値は個人性パラメータに関して得られている値⁽³⁾の約2.5倍になっている。

Ⅴ 標準パターン平均化の効果 音声認識において、異なる3日間の音声サンプルから平均的な標準パターンを作成して用いたところ、3年後の誤り率が1.4%となり、1時期のみのものを用いる場合の2/3に低下した。

Ⅵ むすび 話者認識と音声認識における標準パターンと認識サンプルの時期差の影響の比較を行い、どちらにも時期差とともに誤り率が増加して、ある一定値で飽和する性質があり、これはスペクトルの変動量と対応していること、音声認識の場合の方がクラス間距離が大いためにその影響が小さく、しかも用いる情報の変動の時定数が大きいことが明らかになった。

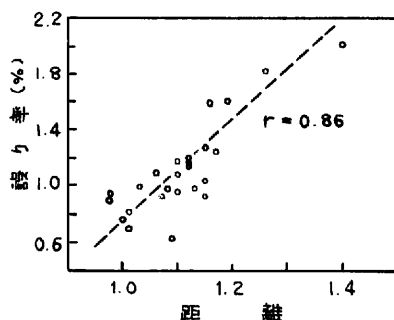


図3. スペクトル変動量と誤り率の関係

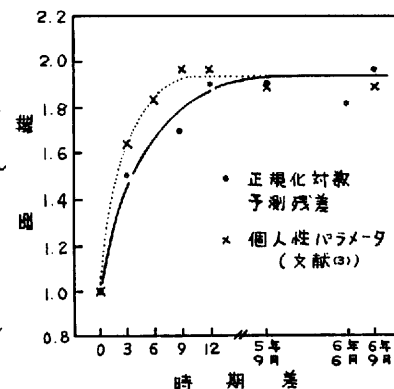


図4. 時期差とスペクトル変動量の関係 (多数話者)

[謝辞] 日頃御指導頂く斎藤研究室長、資料等を提供して頂いた好田調査員をはじめ、熱心に討論頂く研究室の諸氏に感謝を表す。

[文献] (1) 古井: 通研集報, 25, 2 p.425 (昭51) (2) 古井: 音学雑誌, 2-4-2 (昭51-10) (3) 古井: 被査: 通研集報, 24, 5, p.959 (昭50)