

論文 / 著書情報
Article / Book Information

論題(和文)	音声スペクトルの動的特徴による話者認識の検討
Title(English)	
著者(和文)	古井 貞熙, 斉藤 収三
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1978年講演論文集, , , pp. 369-370
Citation(English)	, , , pp. 369-370
発行日 / Pub. date	1978, 10

2—2—9 音声スペクトルの動的特徴による話者認識の検討*

古井 貞照 斎藤 収三 (日本電信電話公社 武蔵野電気通信研究所)

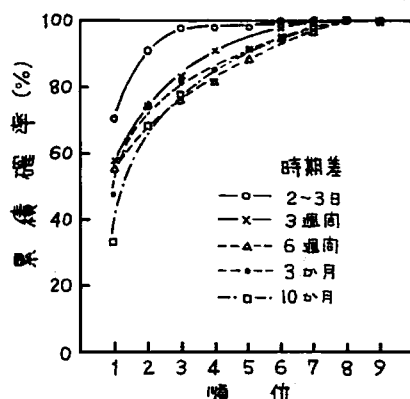
I まえがき 同じ話者が同じ単語を繰返し発声しても、そのスペクトルの時間構造は様々に変化するが、各話者に特有で比較的安定した部分と変動しやすい部分があると考えられる。筆者らはこれまでに、多数話者が発声した単語音声と対数断面係比(LAR)と基本周波数の時系列で表現し、同一話者内および異話者間の音声相互間の時間軸非線形伸縮マッチング(DPマッチング)を行ったときの対応づきの結果によって、この点を調べて来た。その結果、同一話者内においては、時間軸伸縮のおこりにくい音声区画が単語内に何か所もあり、この比較的安定した時間構造を有する区画(特定領域)は、概して音韻の過渡部に対応していることが明らかになった。⁽¹⁾⁽²⁾ そこで、入力音声と登録話者の標準パターンとのマッチングを行ったときに、標準パターンの特定領域全体のうちで、時間軸が伸縮せずに対応づけられたフレームの割合(安定度)を標準パターンとの類似度の一つとして用いて話者認識を行ったところ、この有効性が確かめられた。

ここでは、学習サンプルの収集期間が入力音声との時期差に比べて短い場合と長い場合について、DPマッチングを行ったときの時系列の類似度とこの安定度を組合せる効果を比較し、同時に、DPマッチングの類似度に単語音声区画の特徴パラメータの平均値を組合せる場合の結果との比較を行う。

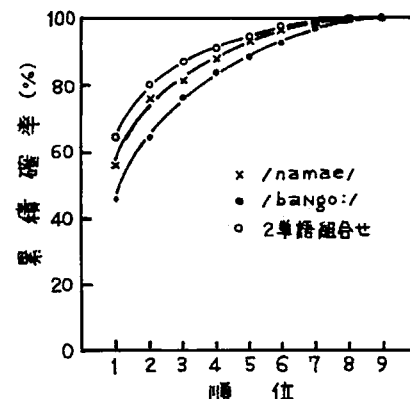
II 実験方法 男性9名が発声した単語音声 /namae/ と /bangō:/ を用いて話者識別実験を行う。特徴パラメータは1~12次のLARと基本周波数(フレーム周期: 10ms)である。学習サンプルと入力音声の时期的関係は図1に示す通りで、各話者は各時期に各単語を繰返し3回発声している。長期間学習の場合については、3か月おきの異なる8時期の入力音声について、同様の実験を行った。

標準パターンと入力音声の類似度

図1. 学習サンプルと入力音声の时期的関係 { 学習サンプル, 入力音声 }



(a) 短期間学習の場合 (2単語平均結果)



(b) 長期間学習の場合

図2. 安定度による識別で正しい話者の選ばれる割合

* Talker Recognition Using Dynamical Features of Speech Spectra.
By Sadaoki FURUI and Shuzo SAITO (Musashino ECL, NTT)

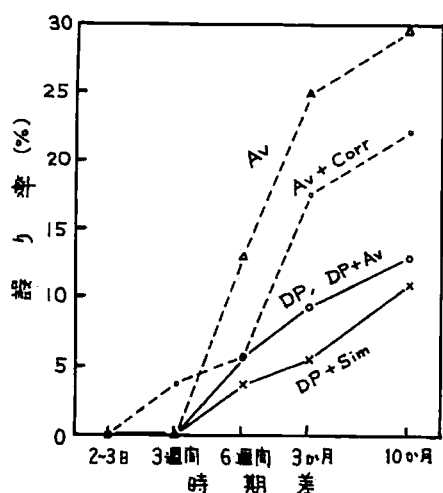


図3. 短期間学習の場合の時期差と認識誤り率の関係(2単語平均結果)

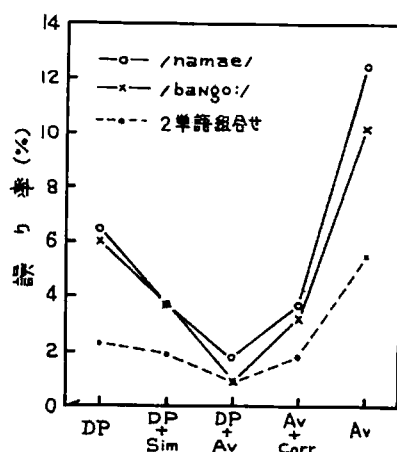


図4. 長期間学習の場合の種々の類似度による認識結果

としては、重みつきDPマッチング⁽¹⁾を行ったときの2つの時系列の類似度(DPと略記)を基本とし、これに特定領域における安定度(Sim)、あるいは単語全体の特徴パラメータの平均値の類似度(Av) — DPと同じ重みを使用 — を組合せる場合について検討した。比較のために、Avのみの場合、およびAvに特徴パラメータ相互間の相関行列(Corr)を組合せる場合(従来の「統計的パラメータ」⁽³⁾に相当)についても実験を行った。

III 実験結果 Simを単独に用いるときに正しい話者の選ばれる割合を図2に示す。短期間学習の場合、入力音声との時期差が2~3日のときと3週間以上との間で、結果に大きな差が見られるが、長期間学習の場合の結果と比較すると、時期差が3か月程度の場合には、学習期間の長さの影響はあまり受けにくいことがわかる。

種々の類似度を用いたときの認識結果を図3~5に示す。これらの結果から次のことがわかる。①学習期間の長さに拘らずSimを組合せることにより誤り率が低下し、時期差が3か月の場合には、共通してほぼ2/3に低下する。②DPやSimのように動的特徴を用いる場合の方が、学習期間や時期差の影響を受けにくい。長期間学習の場合にはAv+Corrの方がDPよりも誤り率が小さいが、短期間学習の場合には逆転する。③長期間学習の場合にはDPにAvを組合せる効果が大きく、Simの組合せ効果よりも大きい。短期間学習の場合には効果がない。④長期間学習の場合、いずれの類似度によっても、2単語を組合せると、誤り率は1単語のときのほぼ1/2に低下する。

IV むすび 今後さらに有効な動的特徴の抽出法の検討を進めて行く予定である。

[謝辞] 日頃熱心に討論して頂く研究室の諸氏に感謝す。[文献] (1) 古井・斎藤：音声学雑誌 4-1-25 (昭53-05) (2) 古井・斎藤：音声学雑誌 PRL78-30 (昭53-07) (3) 古井・斎藤：音声学雑誌 56-A, 11, p.717 (昭48-11)