

論文 / 著書情報
Article / Book Information

論題(和文)	音声認識における個人差の正規化に関する検討
Title(English)	
著者(和文)	古井貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1975年講演論文集, , , pp. 137-138
Citation(English)	, , , pp. 137-138
発行日 / Pub. date	1975, 10

1—4—19 音声認識における個人差の正規化に関する検討*

古井 貞照 (日本電信電話公社 武蔵野電気通信研究所)

1. まえがき 音声スペクトルの個人差を正規化することは、音声認識における極めて重要な問題である。一般に音声スペクトルに含まれる個人差情報は、音韻の種類に無関係なもの(A)と、音韻の種類に依存したものに分けられ、後者はさらにすべての音韻に共通した構造を有するもの(B)と、音韻毎に異なる構造を有するもの(C)に分けられる。(A)に含まれる情報には、声帯波スペクトルの傾斜や声道長の個人差に起因すると考えられるものがあり、本報告ではこれらの正規化に関する検討結果について述べる。このような試みは、これまでにあいまいなされているが、個人差の正規化の効果について、定量的に検討されたものはほとんどない。(B)に含まれる情報を抽出し正規化する方法としては、前回に報告した声質空間の設定あるいは声質情報のクラスタ化による方法が有効である。⁽¹⁾

2. 方法 (1) スペクトルの傾斜の除去—スペクトル包絡を1次系または2次の臨界制動系(2次CD系)の特性で近似し、この逆フィルタに通すことにより、スペクトルの概形を平坦にする。スペクトル包絡の近似は、分析するフレーム毎に行なう場合([適応的])と、話者について固定する場合([固定])が考えられる。後者の場合については、ここでは全音韻の平均特性を用いる。(2) 声道長の正規化—スペクトルの周波数軸方向の伸縮を正規化するため、相関係数パターンと、基準パターンとの平均二乗誤差が最小になるように時間軸方向に伸縮する。

3. 分析

3.1. 実験に用いたサンプル 男性30名、女性30名、子供8名が発声した10数字音声を8kHzで標本化したもの、/a/, /i/, /u/, /e/, /o/, /n/, /N/, /g/, /r/の9音韻の区間を抽出し、 $\tau=1\sim10$ の相関係数の各区間の平均値を抽出する。上の112の処理はいずれも相関係数の領域で行うことができる。このことは、最尤法にもとづく認識法⁽²⁾を適用する際に有用である。

3.2. 多変量分散分析 音韻要因に関する多変量分散分析により、上の方法の個人差を正規化する効果を評価した。

2種類の正規化を行なう前の相関係数に含まれる音韻効果の大きさ(危険率1%のときの有意水準値の約10倍)を1として、種々の正規化を行なったときの音韻効果の大きさを求めた。図1に、男性・女性・

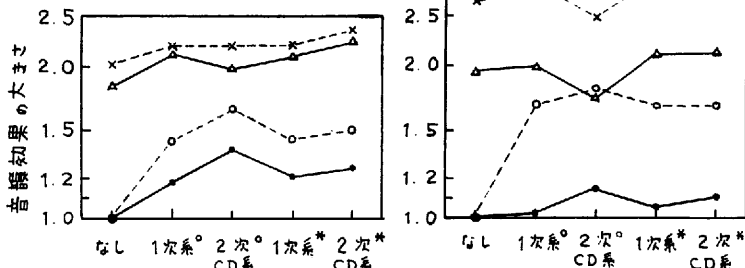


図1. 多変量分散分析の結果
(a) 男性・女性・子供の分析結果の平均 (b) 男性・女性・子供をひとまとめでした分析結果
○ 5母音 { 声道長正 } × 5母音 { 声道長正 }
△ 9音韻 { 正規化なし } ▲ 9音韻 { 正規化あり }

子供のサンプルを別々に分析した結果の平均値(a)と、男性・女性・子供をひとまとめでした

* Effects of the Interspeaker Normalization on Speech Recognition.

By Sadaoki FURUI (Musashino ECI, NTT)