

論文 / 著書情報
Article / Book Information

論題(和文)	話者認識における登録話者数の効果
Title(English)	
著者(和文)	古井 貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	昭和52年度電子通信学会情報部門全国大会, , , p. 228
Citation(English)	, , , p. 228
発行日 / Pub. date	1977,
URL	http://search.ieice.org/
権利情報 / Copyright	本著作物の著作権は電子情報通信学会に帰属します。 Copyright (c) 1977 Institute of Electronics, Information and Communication Engineers.

話者認識における登録話者数の効果

228

古井 貞昭 (日本電信電話公社 武蔵野電気通信研究所)

I まえがき 一般に話者認識の問題は、話者識別 (Identification) と話者照合 (Verification) にわけることができる。前者は、入力音声があらかじめ登録されている複数の話者のうちのいずれの話者によるものであるかを判定するものであり、後者は、入力音声がある登録話者のものであるか否かを判定するものである。いずれの場合も、登録話者の数が変わると認識率が変化する。この様子はすでに G.R.Doddington⁽¹⁾ のシミュレーション実験によって定量的に明らかにされているが、これは実際の音声データに基づいたものではなく、話者照合の判定の方法も、実際に広く用いられているものと異なっている。本報告は、筆者らが以前から検討している単語の統計的パラメータ⁽²⁾を用いた話者認識における、登録話者数と誤り率の関係を調べたものである。

II 実験方法

- ・ 音声サンプル: 4 単語 (/ko:geN/, /bakuoN/, /baNgo:/, /namae/)
- ・ 話者: 成人男性 116 名, 成人女性 40 名
- ・ 特徴パラメータ: 基本周波数および対数断面面積の統計量

各話者が各単語を 1 時期に繰返して 3 回発声したうちの 2 回のサンプルの特徴パラメータの平均値を標準パターンとし、残りの 1 サンプルを未知サンプルとした (3 通り)。標準パターンと未知サンプルの類似度には、特徴ベクトルの各成分が独立に、すべての話者に共通の分散を有する正規分布に従うものと仮定した重みつき距離を用いた。

認識の判定には 1 単語を用い、話者識別の場合は未知サンプルから最も近い標準パターンの話者によるものと判定し、話者照合の場合は標準パターンからの距離に閾値を設けて、判定を行った。この閾値は棄却誤り率と受入れ誤り率が等しくなる値に事後的に設定した。

III 実験結果 特徴ベクトルの次元数 (n) を 13 から 43 まで変えたときの、登録話者数 (N) と誤り率 (e) の関係を Fig. 1 に示す。男性についての結果を (a) に、女性についての結果を (b) に示した。いずれも、すべての話者から各条件の数の登録話者の組を選ぶ種々の組合せと、4 種類の単語に関する結果を平均したものである。これらの結果から、男性、女性に共通して、話者識別においては、ほぼ $N \geq 5$ の範囲で e は $\sqrt{N}$ に比例して増加すること、話者照合においては、ほぼ $N < 5$ の範囲では N の増加とともに e が増加するが、 $N \geq 5$ の範囲では N が増加しても e はほとんど変わらず、むしろやや低下することがわかる。また、女性の誤り率は、男性のその約 2 倍になっている。

どの n の条件においても、 N が小さい範囲では話者識別の誤り率よりも話者照合の誤り率の方がやや大きく、 N が大きい範囲ではこの傾向が逆転するために、この 2 種の誤り率はある N の値 (N_0 とする) で交差する。 n を変えたときの N_0 の値の変化を調べたところ、Fig. 2 のような結果が得られた。男性と女性で類似した結果を示し、 N_0 の値は n の増加とともに単調に大きくなる。これは、話者照合よりも話者識別の場合の方が n が増加することによる e の低下が著しいためである。

IV むすび 話者識別と話者照合における登録話者数と認識誤り率の関係を調べ、両者の比較を行った。なお認識の判定に 4 単語を組合せて用いる場合についても、同様の結果が得られている。

[謝辞] 日産自動車株式会社研究部長をはじめ研究室の諸氏、並びに実習生・後援者 (阪大) に感謝する。

[文献] (1) G.R.Doddington: Technical Report, RADC-TR-UI-963700-F, 1974 (2) 古井・板倉: 通研集報, 24, 5, p. 959, 1975

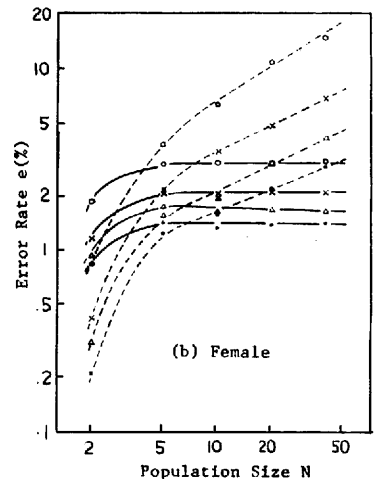
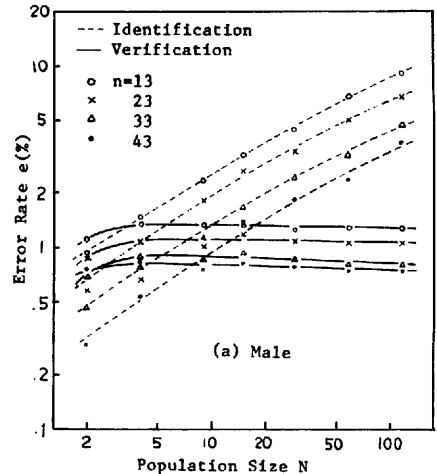


Fig. 1. Error Rate for Speaker Identification and Verification as a Function of Population Size.

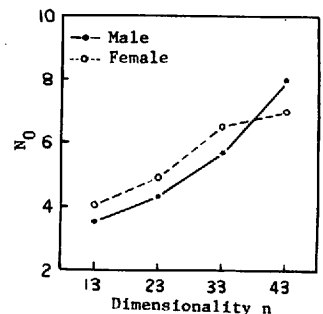


Fig. 2. Relation between the Feature Vector Dimensionality and N_0 .