

論文 / 著書情報
Article / Book Information

論題(和文)	音声の個人性パラメータの変動特性の解析と話者認識への応用
Title(English)	
著者(和文)	古井 貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1974年講演論文集, , , pp. 359-360
Citation(English)	, , , pp. 359-360
発行日 / Pub. date	1974, 6

3—2—3 音声の個人性パラメータの変動特性の解析と 話者認識への応用*

古井 貞熙 (日本電信電話公社 武蔵野電気通信研究所)

1. まえがき これまでに、音声の個人性パラメータの時期的変動の大きさを定量的に評価する尺度として時期的変化率を提案し、音声波から音源特性と声道特性を近似的に分離して、各パラメータについてその値を求めた結果、声道特性よりも音源特性の方が、時期的にゆっくり変化する傾向が強いことがわかった。⁽¹⁾ ここではその解析をさらに進めた結果について述べる。

2. 音源特性と声道特性の分離 ここでは、短文章の長時間平均スペクトルの概形は、放射特性を含む音源特性を近似的に表現しているものとする。10秒程度のサンプルを用いて、その長時間平均スペクトルの概形を、簡単のために1次系あるいは2次臨界動系(以後2次C.D.系と略記)で近似して求め、次に分析すべき音声波をこの特性の逆フィルタに通す。逆フィルタに通したのちの波形のPARCOR係数によって声道特性を抽出する。逆フィルタの伝達関数は、 $1 + \beta Z$ (1次系) および $1 + \alpha Z + \frac{\gamma}{4} Z^2$ (2次C.D.系) と書ける。

3. 分散分析

3.1. パラメータ別分散分析 単語の有声音区間における基本周波数(f_0)と、逆フィルタ後のPARCOR係数($K_1 \sim K_{10}$)を抽出し、それらの全区間に関する平均値を求める。男性9名が発声した2単語 / bango:/, / namae:/ のサンプルを用い、それらがほぼ10日間の4時期のものである場合と、ほぼ10か月間の4時期のものである場合について、逆フィルタの係数 β, γ を含む各パラメータ別に、話者要因に関する分散分析を行なった。危険率1%のときの有意水準で正規化した話者効果の大きさを図1に示す。2単語に関する結果を平均して示した。この結果から、音源スペ

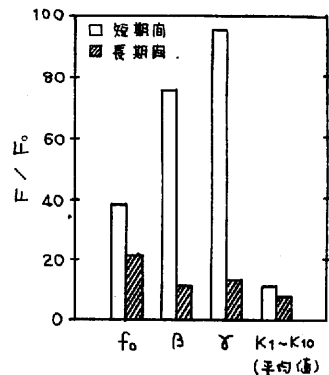


図1. 各パラメータの話者効果

クトルに関するパラメータ β, γ は、10日間程度の短期間では極めて大きい話者効果を有するが、10か月程度の長期間では、その効果は $1/8$ 程度に低下してしまうことがわかる。一方声道特性に関するPARCOR係数 $K_1 \sim K_{10}$ では、このような性質は少ない。

3.2. 多変量分散分析 音源特性を除去する逆フィルタに通した場合と通さない場合のPARCOR係数 $K_1 \sim K_{10}$ からなるベクトルについて、男性9名がほぼ3年間にわたって発声した単独5母音のサンプルを用いて、音韻要因、話者要因、および音韻—話者交互作用に関する多変量分散分析を行なった。危険率1%のときの有意水準で正規化した各効果の大きさを図2に示す。この結果から、話者効果の大きさは、音韻効果とともに、音源特性を除去してもほとんど変わらないことがわかる。

4. 認識実験 単語の有声音区間における基本周波数とPARCOR係数の統計量を持

* Analysis of long-term variation of parameters and its application to talker recognition. By Sadaaki FURUI (Musashino E.C.L., N.T.T.)

微パラメータとする認識方法⁽²⁾により、逆フィルタの効果を調べた。図3に示すように、共通の6時期の未知サンプル(A~F)を用いて、既知サンプルが10日間程度の短期間のものである場合(a)と、10か月程度の長期間のものである場合(b)の認識実験を行なった。(a)の場合は既知サンプルを固定して、未知サンプルまでの時間を3か月から21か月まで変え、(b)の場合は、未知サンプルまでの時間を3か月に固定して、既知サンプルを変えた。認識実験の結果、認識方法によらず、(a)で未知サンプルまでの時間が3か月から21か月まで増加しても、単調に認識誤りが増加する傾向はない

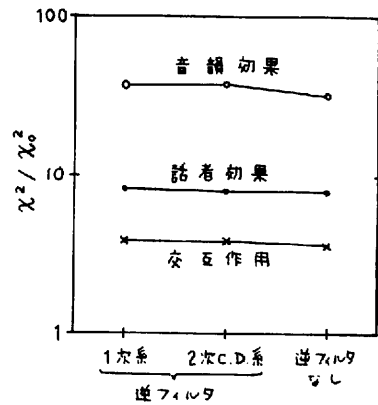


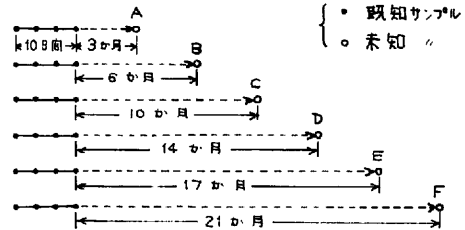
図2. PARCOR係数 $K_1 \sim K_{10}$ からなるベクトルの多変量分散分析結果

ことがわかったので、(a)(b)いずれの場合も、これら6時期の未知サンプルに関する認識結果を平均して表1に示した。この結果、既知サンプルが短期間のものである場合でも、音源特性を除去するとともに、2単語を組み合わせる方法を用いることにより、既知サンプルが長期間のものである場合と同程度の99%の認識率が得られることがわかった。

5. むすび 分散分析により、音源特性が有する個性は、短期間では大きいが長期間になるとかなり小さくなることを明らかにした。つぎに話者認識実験により、既知サンプルが短期間のものであっても、音源特性を除去することにより、既知サンプルが長期間のものである場合と同程度の認識率で、長期間たった後のサンプルを認識できることを示した。

〔謝辞〕日ごの御指導御鞭撻を頂く、野田基礎研究部長、斎藤才4研究室長、板倉調査員ならびに才4研究室の諸氏に感謝する。〔文献〕(1)古井、板倉：音学講義、2-3-1 (昭48-10) (2)古井、板倉：信学誌、56-A, 11, p.717

(a) 既知サンプルが短期間のものである場合



(b) 既知サンプルが長期間のものである場合

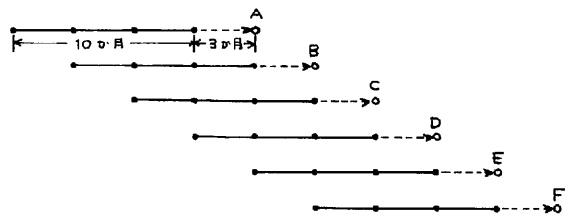


図3. 既知サンプルと未知サンプルの関係

表1. 認識実験の結果

既知サンプル	未知サンプルまでの時間	逆フィルタ	話者識別		話者照合(許容率8%)	
			1単語	2単語	1単語	2単語
短期間 (10日間)	3~21 か月	1次系	90.8%	99.4%	97.4%	99.0%
		2次C.D系	92.3	99.4	97.2	98.9
		なし	88.3	94.5	92.7	92.8
長期間 (10か月)	3か月	1次系	97.2	99.4	97.6	99.5
		2次C.D系	96.0	98.8	97.1	98.8
		なし	98.2	99.4	97.9	99.5

(昭48-11)