

論文 / 著書情報
Article / Book Information

論題(和文)	単音節入力による大語い単語認識の高度化
Title(English)	
著者(和文)	古井貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会1981年秋季講演論文集, , , pp. 119-120
Citation(English)	, , , pp. 119-120
発行日 / Pub. date	1981, 10

3—1—2 単音節入力による大語い単語認識の高度化*

古井 貞照 (日本電信電話公社 武蔵野電気通信研究所)

I まえがき 音声スペクトル時系列の大
局的・局所的両特徴の組合せによる単音
節認識法と、この方法を用いて、音節毎
に区切って発声された大語いの単語音
声を認識する方式の有効性につ
いてはすでに報告した。⁽¹⁾⁻⁽³⁾ こと
では、この認識法をさらに改良し、
実際の音節区切りによる単語音声
に適用した結果について報告する。

II 改良点と認識実験

2.1 音節候補の予備選択 単音
節認識のブロック図は図1に示す
通りで、フーリエ展開係数による
距離にもとづき、上位の5位または
10位までの候補のみについて語頭部の(シフト)マッチングを行うように変更した。
有効性の評価には男性10名の68単音節
の音声を用い、50音順に5回発声したう
え、初めの3回で学習して残りの2回を認
識する実験と、後の3回で学習して初めの2
回を認識する実験を行った。その結果、
10位まで予備選択すれば正しい音節がそ
の中に入る割合は99.8%で、語頭部の
マッチング後の音節認識率は、オ1位から
オ5位の累積認識率まで、予備選択を
行わない場合と変わらず、5位まで予備選択しても、最終的なオ1位の認識率は変わらないこ
とがわかった。このことにより、全体としての計算量を大幅に低下させることがわかった。

2.2 拗音を含む認識 上記の話者のうちの5名について、拗音を含む全101音節の認識
実験を行った。実験方法は上述と同じで、実験の結果、拗音と拗音以外の音節との間の
誤認識は全くなく、5名の平均認識率は、68単音節とき92.6%、101単音節のとき
92.5%で、101音節に拡張しても認識率はほとんど低下しないことがわかった。拗音
節間の誤認識行列を表1に示す。hyu→kyu, sya→tya, ryo→gyo の誤りが比較的多い。

2.3 音節数の自動決定と促音の検出 単語音声区間全体の振幅パターンに対してフーリ
エコサイン/展開係数を計算する。図2に/koQkaʃikeN/ (国家試験)と発声したときの振幅パ

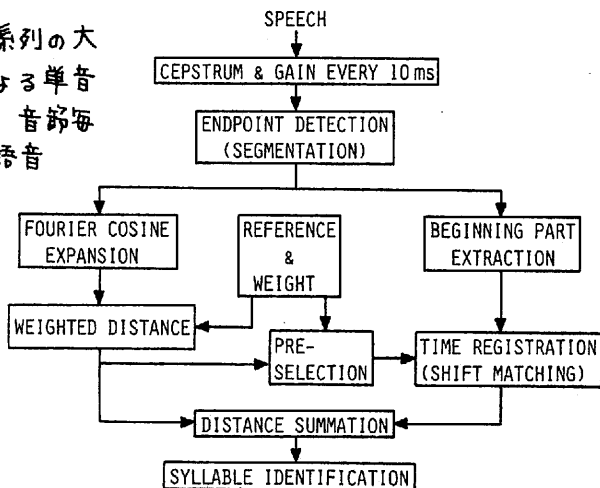


Fig.1-Block diagram of the improved syllable recognition method.

Table 1-Confusion matrix for the /CjV/ syllables.

Out In	y	ky	sy	ty	ny	hy	my	ry	gy	zy	by	py	RATE(%)
y	59						1						98.3
ky		56					4						93.3
sy			57	3									95.0
ty			3	57									95.0
ny					55		4		1				91.7
hy						9	50					1	83.3
my							1	59					98.3
ry									54	4		2	90.0
gy									4	53		1	88.3
zy									1	2	56		93.3
by									1	1	56	1	93.3
py												58	96.7
												AVERAGE	93.1

* Improvement in the Recognition of Spoken Japanese Words Uttered by Isolated Syllables.
By Sadaaki FURUI (Musashino Electrical Communication Laboratory, N.T.T.)

タンを、図3に /waraibanaʃi/ (笑い話) を含む2つの場合の展開係数の抽出例を示す。この値が急激に低下する位置から音節数(モーラ数、促音を含む)を決定し、振幅値とディップによって音節のセグメンテーションと促音の検出を行う。男性2名が発声した、促音を含む4-6音節(モーラ)の210単語音声を用いて実験を行ったところ、各音節数に対し促音の検出率はそれぞれ91%、92%、97.5%であった。

2.4. 標準パターンの適応化と単語音声認識

4文字(促音・拗音を含む)からなる3203単語を対象として、そこから無作為に抽出した321単語を男性1名が平均125文字/分の速度で発声した音声を用いて、認識実験を行った。各音節の標準パターンは、2.1で用いたものを初期値とし(未知音声との時期差は6か月)、未知音声を発声する前に、68単音節を各2回、学習用単語として定めた10単語を1回ずつ発声して、適応化(フーリエ係数は平均化、語頭時系列は置き換え)を行った。

実験の結果、音節数の決定誤りとセグメンテーションの誤りは全くなく、音節毎及び単語としての認識率は図4に示す通りであった。第1位の音節認識率は90.5%、単語認識率は99.1%で、調音結合のために、音節認識率は、同じ話者が68単音節を50音順にゆくり発声したときよりも低いが、単語認識率はシミュレーション実験⁽³⁾の結果とほとんど変わらず、理論値とのずれは小さい。誤認識単語の内訳は、/zaigaku/→/zaitaku/、

/seiran/→/heitan/、/onbou/→/ontou/ である。第10位までに正しい音節が入らなかったのは、1284音節中ba, ga(2回), ro, tuの5音節(0.4%)であった。発声者が変わった場合、発声の訓練効果等については、今後さらに検討を行う必要がある。

Ⅲ むすび 単音節認識法に関するいくつかの改良を行い、実際の音節区切りによる単語音声等を用いた実験により、この有効性を確認した。今後さらに、音節標準パターンの適応化方法、促音の検出方法、セグメンテーション法等の改良を進めて行きたい。

【謝辞】日頃御指導頂く当所 野柳基礎研究部長、橋本統祐氏、坂倉孝四郎研究員長、ならびに研究室の皆様へ感謝する。

【文献】(1) 古井：音学秋季講義 1-1-18 (昭55)、(2) 同：伝学総論 1351 (昭56)、(3) 同：音学春季講義 3-1-12 (昭56)

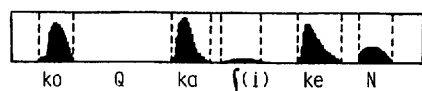


Fig.2-An example of the amplitude pattern of a spoken word.

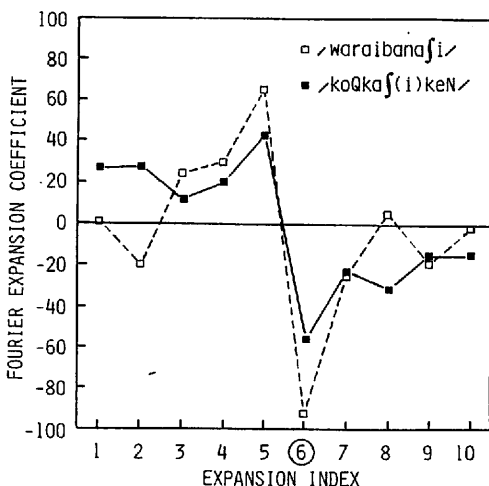


Fig.3-Examples of the Fourier cosine expansion coefficients.

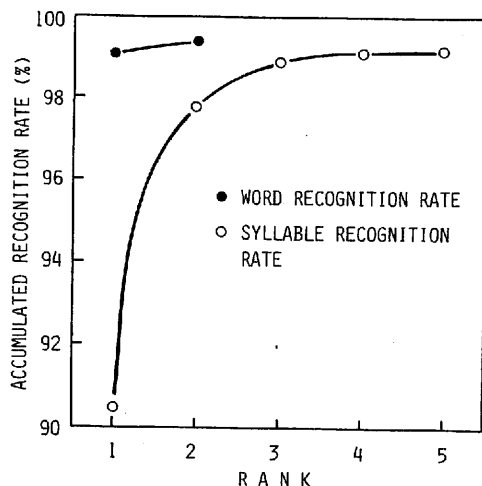


Fig.4-Accumulated syllable and word recognition rates.