

TR-I-0226

確率モデルを用いた言語識別の検討  
Language Identification Using Probabilistic Modeling

吉田 哲也 杉山 雅英

Tetsuya YOSHIDA Masahide SUGIYAMA

内容梗概

現在、音声認識や音声理解、さらに言語間の自動翻訳が活発に研究されている。しかしながら、発話された音声を用いてその属する言語を識別する研究は多くはない。既に筆者等は、ベクトル量子化の際の量子化歪みといった静的な特徴を用いて、バイリンガル話者によって発話された音声に対して日・英の言語識別を行ない、その有効性を報告した。本報告では同一の音声資料及びタスクに対し、それら静的な特徴に加え動的な特徴を利用することの効果を検討し、その基本的な性能の把握を行なう。ここで動的な特徴の表現には、N-gram や ergodic HMM を用いる。また、それらの各手法について雑音環境下での性能の検討を行なう。

© ATR Interpreting Telephony Research Labs.

© ATR 自動翻訳電話研究所

## 目次

1	はじめに	1
2	準備	1
2.1	bilingual 音声データベースについて	1
2.2	音響分析と距離尺度について	1
3	言語識別法について	1
3.1	VQ 符号帳を用いる方法	1
3.2	Universal 符号帳とその確率密度関数を用いる方法	2
3.3	Universal 符号帳と N-gram を用いる方法	2
3.3.1	0-gram ( unigram ) を用いる方法	2
3.3.2	1-gram ( bigram ) を用いる方法	2
3.4	ergodic HMM を用いる方法	3
4	言語識別実験	3
4.1	入力単語数との関係	3
4.2	符号帳の大きさとの関係	3
4.3	雑音環境下での識別	3
4.3.1	フレーム毎の付加雑音と識別率との関係	4
4.3.2	単語全体にわたる雑音との関係	4
5	むすび	5
A	日・英語の識別率	10
B	識別実験に用いたソフトウェア	27

## 表目次

1	bilingual 音声データベースについて	1
2	音声分析条件 ( VQ 符号帳 )	1
3	音声分析条件 ( Universal 符号帳, N-gram, ergodic HMM )	2
4	Universal VQ 符号帳 ( Euclid 距離 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ) )	11
5	Universal VQ 符号帳 ( K-L 情報量 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ) )	11
6	0-gram ( unigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ) )	11
7	1-gram ( bigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ) )	12
8	ergodic HMM ( 2 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ) )	12
9	ergodic HMM ( 3 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ) )	12
10	Universal VQ 符号帳 ( Euclid 距離 ) による日・英の識別率 ( 話者: 女性 1 名 ( FCF ) )	13
11	Universal VQ 符号帳 ( K-L 情報量 ) による日・英の識別率 ( 話者: 女性 1 名 ( FCF ) )	13
12	0-gram ( unigram ) による日・英の識別率 ( 話者: 女性 1 名 ( FCF ) )	13
13	1-gram ( bigram ) による日・英の識別率 ( 話者: 女性 1 名 ( FCF ) )	14
14	ergodic HMM ( 2 states ) による日・英の識別率 ( 話者: 女性 1 名 ( FCF ) )	14
15	ergodic HMM ( 3 states ) による日・英の識別率 ( 話者: 女性 1 名 ( FCF ) )	14
16	Universal VQ 符号帳 ( Euclid 距離 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 0dB )	15
17	Universal VQ 符号帳 ( Euclid 距離 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 5dB )	15
18	Universal VQ 符号帳 ( Euclid 距離 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 10dB )	15
19	Universal VQ 符号帳 ( K-L 情報量 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 0dB )	16
20	Universal VQ 符号帳 ( K-L 情報量 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 5dB )	16
21	Universal VQ 符号帳 ( K-L 情報量 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 10dB )	16
22	0-gram ( unigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 0dB )	17
23	0-gram ( unigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 5dB )	17
24	0-gram ( unigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 10dB )	17
25	1-gram ( bigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 0dB )	18
26	1-gram ( bigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 5dB )	18
27	1-gram ( bigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 10dB )	18
28	ergodic HMM ( 2 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 0dB )	19
29	ergodic HMM ( 2 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 5dB )	19
30	ergodic HMM ( 2 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 10dB )	19
31	ergodic HMM ( 3 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 0dB )	20
32	ergodic HMM ( 3 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 5dB )	20
33	ergodic HMM ( 3 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 10dB )	20
34	Universal VQ 符号帳 ( Euclid 距離 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 20dB )	21
35	Universal VQ 符号帳 ( Euclid 距離 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 30dB )	21
36	Universal VQ 符号帳 ( Euclid 距離 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), segmental SNR 40dB )	21
37	Universal VQ 符号帳 ( K-L 情報量 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 20dB )	22
38	Universal VQ 符号帳 ( K-L 情報量 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 30dB )	22
39	Universal VQ 符号帳 ( K-L 情報量 ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 40dB )	22
40	0-gram ( unigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 20dB )	23
41	0-gram ( unigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 30dB )	23
42	0-gram ( unigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 40dB )	23
43	1-gram ( bigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 20dB )	24
44	1-gram ( bigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 30dB )	24
45	1-gram ( bigram ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 40dB )	24
46	ergodic HMM ( 2 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 20dB )	25
47	ergodic HMM ( 2 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 30dB )	25
48	ergodic HMM ( 2 states ) による日・英の識別率 ( 話者: 男性 1 名 ( MAO ), SNR 40dB )	25

49	ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO),SNR 20dB) . . . . .	26
50	ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO),SNR 30dB) . . . . .	26
51	ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO),SNR 40dB) . . . . .	26

図目次

1	2 状態の ergodic HMM . . . . .	3
2	3 状態の ergodic HMM . . . . .	3
3	Universal 符号帳を用いた各種の言語識別法 . . . . .	4
4	VQ 符号帳による識別結果 . . . . .	6
5	Universal 符号帳とその確率密度による識別結果 (Euclid 距離) . . . . .	6
6	Universal 符号帳とその確率密度による識別結果 (K-L 情報量) . . . . .	6
7	0-gram (unigram) による識別結果 . . . . .	6
8	1-gram (bigram) による識別結果 . . . . .	7
9	ergodic HMM (2 states) による識別結果 . . . . .	7
10	ergodic HMM (3 states) による識別結果 . . . . .	7
11	フレーム毎に雑音が重畳した場合の識別結果 (ヒストグラム法 (K-L 情報量)) . . . . .	8
12	フレーム毎に雑音が重畳した場合の識別結果 (ergodic HMM (2 states)) . . . . .	8
13	フレーム毎に雑音が重畳した場合の識別結果 . . . . .	8
14	単語全体に雑音が重畳した場合の識別結果 (ヒストグラム法 (K-L 情報量)) . . . . .	9
15	単語全体に雑音が重畳した場合の識別結果 (ergodic HMM (2 states)) . . . . .	9
16	単語全体に雑音が重畳した場合の識別結果 . . . . .	9

## 1 はじめに

現在、音声認識や音声理解、さらに言語間の自動翻訳が盛んに研究されており、発話された音声に属する言語を識別することは重要である。しかし、これまでこのような言語識別に関する研究は多くない [1]-[6][8]。既に筆者等は、バイリンガル話者によって発話された音声を用いた言語識別の検討を報告した [9][10]。バイリンガル話者の音声を用いることにより、話者毎の音響的な特徴の違いを考慮せずに、言語の違いによる特徴の違いのみに注目して識別することができる。しかし、これまでの報告では、特徴量を量子化する際の量子化歪み、などの静的な特徴のみを利用するにとどまっている。

一方、話者認識の研究においても、静的な情報を用いた認識から動的な情報を用いた認識へと発展させ、最近、エルゴード HMM を用いた話者識別の検討 [11] が報告されている。

本報告では、バイリンガル話者によって発話された音声を対象に、これまでの静的な特徴の利用に加え、動的な特徴量を利用する言語識別の検討を行なう。そのために、時間情報を利用できる N-gram や ergodic HMM を言語識別に適用し、その基本的な性能について検討する。

さらに、これらの静的、動的な特徴を利用した各手法について、雑音が重畳した入力音声に対する識別能力について検討を行なう。

## 2 準備

### 2.1 bilingual 音声データベースについて

音声データベースは bilingual 話者の男女 1 名の発声した単語から成り立っている。表 1 に示すように、日本語 216 単語 (音素バランス 216 単語)、英語 500 単語である。識別実験においては先頭から 100 単語で学習を行ない、識別には 101 番から 200 番までの 100 単語を用いた。100 単語を 10 単語ずつ 10 のグループに分け、10 回の識別実験を行なってその平均で識別率とした。識別実験における入力単語の数が 1 の場合は 101, 111, 121, ..., 191 の 10 回の入力に対する識別率となる。従って、学習用の単語と識別用の単語の集合は独立である。

表 1: bilingual 音声データベースについて

	Japanese	English
FCF	216	500
MAO	216	500

### 2.2 音響分析と距離尺度について

VQ 符号帳を用いた方法の分析条件を表 2 に示す。算出特徴量は LPC 相関 (COR)、LPC ケブストラム係数 (CEP) である。また、評価したスペクトル距離尺度は CEP, WLR である。通常の対数スペクトル距離に対して、スペクトルのピーク重み付けに対応している。各尺度の定義式及びその計算式を式 (1), (2), (3), (4) に示す。

表 2: 音声分析条件 (VQ 符号帳)

自己相関分析	13 次
LPC 分析	10 次
標本化周波数	12 kHz
窓長	256 点 (21.3ms)
窓更新周期	学習時: 128 点 (10.7ms) 認識時: 64 点 (5.3ms)
高域強調	$(1 - 0.98z^{-1})$

$$d_{CEP}^2(f, g) = \int_{-\pi}^{\pi} (\log f(\lambda) - \log g(\lambda))^2 \frac{d\lambda}{2\pi} \quad (1)$$

$$d_{CEP}^2(f, g) = \sum_{i=1}^N (c_i^{(f)} - c_i^{(g)})^2 \quad (2)$$

ここで  $(c_1^{(f)}, \dots, c_N^{(f)})$  及び  $(c_1^{(g)}, \dots, c_N^{(g)})$  はケブストラム係数のベクトルである。打ち切り次数は  $N = 16$  とした。

$$d_{WLR}^2(f, g) = \int_{-\pi}^{\pi} (f(\lambda) - g(\lambda)) (\log f(\lambda) - \log g(\lambda)) \frac{d\lambda}{2\pi} \quad (3)$$

$$d_{WLR}^2(f, g) = \sum_{i=1}^N (r_i^{(f)} - r_i^{(g)}) (c_i^{(f)} - c_i^{(g)}) \quad (4)$$

次に、Universal 符号帳、N-gram、ergodic HMM を用いた方法の分析条件を表 3 に示す。算出特徴量は同様に LPC 相関係数と LPC ケブストラム係数である。また、Universal 符号帳は WLR 距離尺度により作成した。

## 3 言語識別法について

### 3.1 VQ 符号帳を用いる方法

各言語は VQ 符号帳  $V_j$  によって特徴付けられる [7]。入力音声は各  $V_j$  で量子化され、その累積量子化歪み  $d_j$  が算出され、 $d_j$  が最小である符号帳  $V_j$  を選択し、その添字  $j$  に対応する言語を識別結果とする。

表 3: 音声分析条件 (Universal 符号帳, N-gram, ergodic HMM)

自己相関分析	16 次
LPC 分析	14 次
標本化周波数	12 kHz
窓長	256 点 (21.3 ms)
窓更新周期	符号帳作成時: 144 点 (12 ms) 量子化時: 36 点 (3 ms)
高域強調	$(1 - 0.97z^{-1})$

### 3.2 Universal 符号帳とその確率密度関数を用いる方法

各言語は、すべての言語より作成された Universal 符号帳の符号の出現確率を表す密度関数により特徴付けられる。入力音声からも、Universal 符号帳の符号の出現確率を表す密度関数を算出する。この入力音声の密度関数に対し、各言語の密度関数との距離を算出し、最小の距離を与える密度関数の言語を識別結果とする。

確率密度関数の比較の距離尺度としては、Euclid 距離 (式 (5)) と、Kullback-Leibler の情報量 (式 (6)) を用いた。

$$d^2(f, g) = \sum_i (p_i^{(f)} - p_i^{(g)})^2 \quad (5)$$

$$d(f, g) = \sum_i p_i^{(f)} \log \frac{p_i^{(f)}}{p_i^{(g)}} \quad (6)$$

ここで  $p_i$  は、符号  $i$  の出現確率を表す。

### 3.3 Universal 符号帳と N-gram を用いる方法

#### 3.3.1 0-gram ( unigram ) を用いる方法

Universal 符号帳により量子化された特徴系列中の各符号の出現確率を各言語毎に学習単語より推定する。この推定された確率により、入力音声から観測された符号系列の出現確率を各言語毎に求め、最大の確率を与える言語を識別結果とする。

なお、推定された各符号の出現確率が極めて小さい場合には、その符号を含む符号系列の出現確率が極めて小さくなるのを避けるため、小さな値 ( $10^{-6}$ ) に置換した。

この 0-gram ( unigram ) は、以下のように定式化される。

- 学習

$W_1, W_2, \dots$  : 学習単語

$$P(x_i) = p_i = \frac{m_i}{M} \quad (7)$$

$P(x_i)$  : 符号  $x_i$  に対する出現確率 (=  $p_i$ )

$m_i$  : 符号  $x_i$  の出現頻度

$M$  : 符号の出現総数

- 認識

$W = w_1 w_2 \dots w_n$  : 入力符号系列

$$\log_2 P(W) = \sum_j \log_2 P(w_j) \quad (8)$$

$$= \sum_i n_i \log_2 p_i \quad (9)$$

$$= N \sum_i q_i \log_2 p_i \quad (10)$$

$q_i$  : 単語  $W$  における符号  $x_i$  の出現確率

$n_i$  : 符号  $x_i$  の出現数

$N$  : 単語  $W$  における符号の出現総数

#### 3.3.2 1-gram ( bigram ) を用いる方法

Universal 符号帳により量子化された特徴系列中の二つ組の符号の出現確率を各言語毎に学習単語より求める。この確率により、入力音声から観測された特徴系列の出現確率を各言語毎に求め、最大の確率を示した言語を識別結果とする。なお、0-gram ( unigram ) と同様に、推定された各二つ組の符号の出現確率が極めて小さい場合には、小さな値 ( $10^{-6}$ ) に置換した。この 1-gram ( bigram ) は以下のように定式化される。

- 学習

$W_1, W_2, \dots$  : 学習単語

$P(x_i, x_j)$  :  $x_i$  と  $x_j$  との出現確率

$P(x_i)$  :  $x_i$  の出現確率

- 認識

$W = w_1 w_2 \dots w_n$  : 入力符号系列

$$P(W) = P(w_1 | \text{start}) P(w_2 | w_1) \dots \dots P(w_n | w_{n-1}) P(\text{end} | w_{n-1}) \quad (11)$$

$$P(x_j | x_i) = \frac{P(x_i, x_j)}{P(x_i)} \quad (12)$$

$$\log_2 P(W) = \sum_i \log_2 P(w_i | w_{i-1}) \quad (13)$$

### 3.4 ergodic HMM を用いる方法

2 状態の ergodic HMM を図 1 に、3 状態の ergodic HMM を図 2 に示す。ergodic HMM は、全ての状態から、全ての状態へ遷移することを許す HMM である。HMM を特徴付けるパラメータは以下のようになる。

$$M = (A, B, \Pi) \quad (14)$$

ここで、 $A$  は状態間の遷移確率、 $B$  は状態における符号の出現確率、 $\Pi$  は初期状態の確率である。ergodic HMM を用いることにより、音素情報などのコンテキストに依存せずに動的な特徴を利用することができる。

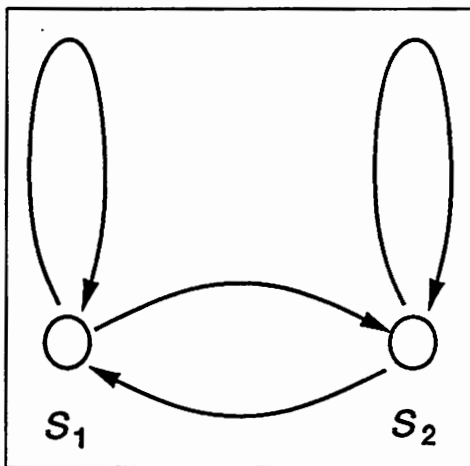


図 1: 2 状態の ergodic HMM

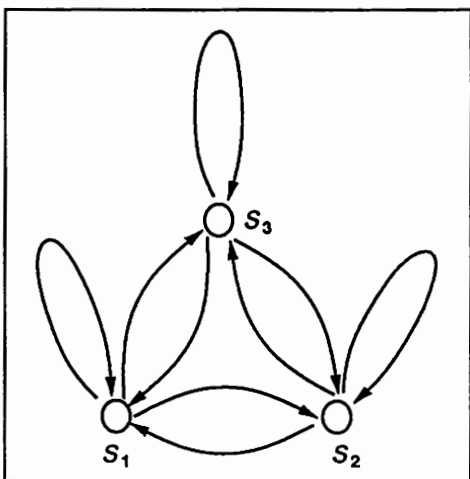


図 2: 3 状態の ergodic HMM

Universal 符号帳により各言語毎に離散 ergodic HMM を作成し、入力音声から観測された特徴符号系列の出力確率を各言語の ergodic HMM 毎に求め、最大の確率を与える ergodic HMM の言語を識別結果とする。

なお ergodic HMM の各状態の符号出力確率  $B$  の初期値として、全ての確率を等しくした場合には学習の効果が見られなかった。そこで  $B$  の初期値として、Universal 符号帳による確率密度関数を擾動させた関数を用いた。状態遷移確率  $A$  の初期値としては、学習の効果に与える影響は少なかったが、似た特徴が連続する間は自己遷移をするものと考え、学習単語の符号系列内において、同符号が連続する確率を自己遷移確率に、異符号が連続する確率を他の状態へ遷移する確率とした。また、全ての状態の初期状態確率  $\Pi$  は単純に等しくした。

## 4 言語識別実験

一連の各手法による識別方法を図 3 に示す。学習時では、学習単語を音声分析し、Universal 符号帳を作成する。その後、各言語毎に確率密度関数などの参照用のモデルを作成する。認識時には、入力音声进行分析後、あらかじめ作成した各々のモデルと参照し、より近いモデルを持つ言語を識別結果とする。

### 4.1 入力単語数との関係

入力単語数と識別率の関係を図 4-10 に示す。入力単語数は 1 ~ 10 まで変化させた。入力単語の数が増えるに従って、認識率は上昇する。VQ 符号帳を用いる場合、16 か 32 符号の符号帳を使用し、4,5 単語入力すればほとんど日英の識別ができることがわかる。また、他の Universal VQ 符号帳を基にする方法の場合、32 符号以上の符号帳を使用すれば、3 単語の入力でほとんど識別できることがわかる。ただし Euclid 距離尺度で確率密度関数を用いる方法の場合は、64 符号以上の符号帳を必要とする。また、ergodic HMM (3 states) を用い、512 符号の符号帳を用いれば 1 単語の入力で識別できる。以上の結果から静的な情報のみでなく動的な情報を用いる方が識別性能が高いことが分かった。

### 4.2 符号帳の大きさとの関係

符号帳の大きさと識別率の関係を図 4-10 に示す。Universal 符号帳は 2 から 512 まで変化させた。Universal 符号帳を用いる場合、8 符号の符号帳を使用すれば、ある程度入力単語が存在する場合、ほとんど識別できることがわかる。しかし、ergodic HMM (3 states) を用いる場合は、95% 以上識別するためには 16 符号の符号帳を必要とする。

いずれの方法の場合にも、符号帳が大きくなるほど識別率が良くなり、100 単語程度の学習で十分学習できている。しかし、bigram による識別は 512 符号の符号帳で 5% の識別率の低下を示している。

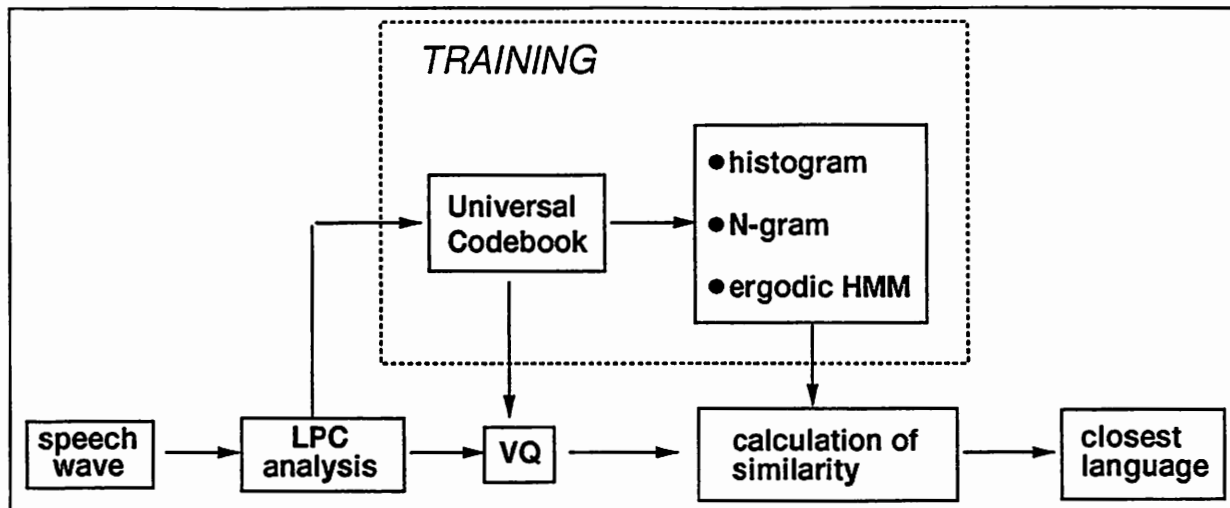


図 3: Universal 符号帳を用いた各種の言語識別法

### 4.3 雑音環境下での識別

Universal 符号帳を基にする各手法は、無雑音環境下においてはほとんどその性能に差異は見られなかった。しかし、実際には無雑音でシステムが動作されることは皆無であり、雑音が重畳した入力に対して識別率を劣化させないことは重要である。そこで、次に、フレーム毎に雑音が重畳した場合と、単語全体に雑音が重畳した場合のそれぞれについて、各手法の雑音に対する耐性を調べた。

#### 4.3.1 フレーム毎の付加雑音と識別率との関係

フレーム毎に雑音が重畳した場合について、静的な特徴を利用する方法の代表として Universal 符号帳とその確率密度関数を用いる方法 (K-L 情報量) による識別率と、動的な特徴を利用する方法の代表として ergodic HMM (2 states) による識別率を図 11,12 に示す。入力単語は 1 ~ 10 まで、符号帳の大きさは 2 ~ 64 まで変化させた。簡単化するために雑音は波形上で加えるのではなく相関係数上で加えることとした。ここで、segmental SNR ( $R$ ) は式 (15) のように定義され、雑音が重畳した音声の相関係数は式 (16) の様に算出される。

$$R = 10 \log_{10} \frac{u}{v} \quad (15)$$

$$\hat{r}_n = \frac{r_n}{1 + \frac{v}{u}} \quad (16)$$

ここで、 $u$  はそのフレームの信号のパワー、 $v$  はそのフレームの雑音のパワー、 $r_n$  は相関係数である。

どちらの方法による識別率も、8 符号、16 符号、といった中程度の大きさの符号帳を用いた場合に高い値を示している。両者の 8 符号と 16 符号の場合を比較すると、ergodic

HMM (2 states) の方が識別率がよく、動的な特徴を用いる方法の方が雑音に対する耐性が強いといえる。

また各手法において雑音が重畳した入力音声に対する識別率を図 13 に示す。8 符号の符号帳を用い、10 単語入力し、SN 比は 0dB ~ 10dB まで変化させた。SN 比が 10dB では、ほとんど雑音の影響は見られず、0dB では ergodic HMM (2 states) を除き、ほとんど識別できていない。SN 比が 5dB の場合において各手法を比較すると、ergodic HMM (3 states), 1-gram (bigram) による識別が高い値を示しており、動的な特徴の利用は静的な特徴の利用に比べ、雑音に対する耐性が強いことがわかる。

#### 4.3.2 単語全体にわたる雑音との関係

単語全体に雑音が重畳した場合について、静的な特徴を利用する方法の代表として Universal 符号帳による確率密度関数を用いる方法 (K-L 情報量) による識別率と、動的な特徴を利用する方法の代表として ergodic HMM (2 states) による識別率を図 14,15 に示す。入力単語数は 1 ~ 10 まで、符号帳の大きさは 2 ~ 64 まで変化させた。ここで、SNR ( $R$ ) は式 (17) の様に定義される。

$$R = 10 \log_{10} \frac{\hat{u}}{v} \quad (17)$$

$$v = \hat{u} 10^{-R/10} \quad (18)$$

$$\hat{r}_n = \frac{r_n}{1 + \frac{v}{u}} \quad (19)$$

雑音の重畳した音声の相関係数  $\hat{r}_n$  は、入力の平均パワー  $\hat{u}$  を求めた後に、SNR より雑音のパワー  $v$  を式 (18) の様に算出し、各フレームのパワー  $u$  に従い、式 (19) の様に算出される。



どちらの識別率も 32 符号の符号帳を用いた時に最も良い識別率を示しており、雑音環境下においては 32 程度の中程度の大きさの符号帳が高い識別率を与えることがわかる。

各方法について 重畳させる雑音の SN 比を 20dB ~ 40dB まで変化させた時の識別率を図 16 に示す。32 符号の符号帳を使用し、10 単語入力した。SN 比が 20dB の場合は全く識別できないが、SN 比が 40dB の場合には完全に識別できる。各手法の雑音に対する耐性を SN 比が 30dB の場合において比較すると、1-gram (bigram), ergodic HMM (2 states) が識別率を悪化させず強い耐性を示しており、動的な特徴を用いる方法は静的な特徴を用いる方法に比べて雑音に対する耐性が強いことがわかる。しかしながら、ergodic HMM (3 states) による識別率は他の方法に比べ低い。

## 5 むすび

これまでの静的な特徴を利用する方法、動的な特徴を利用する方法による識別実験を通じて、以下に挙げられる結果が得られた。無雑音環境下では、男女各一名について識別実験を行なったが、結果に顕著な差異は見られなかった。

- Universal 符号帳に基づく方法では 3 単語程度の入力で日・英語の識別が可能である。
- 無雑音環境下においては動的・静的特徴を用いる方法の識別性能の差はほとんど見られない。
- 雑音環境下においては動的な特徴を用いる方法 (bigram, ergodic HMM) は静的な特徴を用いる方法 (histogram 法) に比べて耐性が強い。
- 無雑音環境下においては、大きな符号帳 (256,512 符号) を用いる方がより高い識別率を与える。しかしながら、雑音環境下においては中程度の大きさ (8, 16, 32 符号) の符号帳が高い識別率を与える。

今回の識別実験ではフレーム周期については全く検討を行なわなかった。静的な特徴を利用する際には、大きな影響はないと考えられるが、動的な特徴を利用する際にはその影響は小さくないと考えられ、音声分析のフレーム周期について今後検討される必要があるであろう。今後の課題としては、より多数の言語間での識別や、異話者間での識別が挙げられる。

- フレーム更新周期の影響
- 多数の言語間での識別
- 異話者間での識別

## 謝辞

研究の機会を与えて頂いた樽松社長、貴重な御助言、御検討頂いた嵯峨山研究室長、ergodic HMM software を提供して頂くと同時に有益な助言をいただいた村上氏をはじめとする音声情報処理研究室の皆様へ感謝致します。

## 参考文献

- [1] A.E.House and E.P.Neuberg, Toward automatic identification of the language of an utterance, I. Preliminary methodological considerations, Journal of the Acoustical Society of America 62(3), pp.708-713 (1977).
- [2] R.A.Cole, J.W.T.Inoue, Y.K.Muthusamy and M.Gopalakrishnan, Language identification with neural networks: a feasibility study, IEEE Pacific Rim Conference on Commu., Comp. and Signal Processing (June 1989).
- [3] J.T.Foil, Language identification Using Noisy Speech, ICASSP86, 17.1 (1986-03).
- [4] F.Goodman, A.Martin, Improved Language Identification in Noisy Speech, ICASSP89, 35.S10b.4 (1989-05).
- [5] 杉山, 音響特徴量を用いた多言語音声識別の検討, 音響学会講演論文集, 3-3-6, pp.81-82 (1990-03).
- [6] M.Sugiyama, Automatic Language Recognition Using Acoustic Features, ICASSP91 (1991-05).
- [7] F.K.Soong, A.E.Rosenberg, On the Use of Instantaneous and Transitional spectral Information in Speaker Recognition, ICASSP86, 17.5 (1986-04).
- [8] 杉山, Automatic Language Recognition Using Acoustic Features, TR-I-0167.
- [9] Hiroya FUJIWARA, Masahide SUGIYAMA, パイリンガル音声を用いた日・英語の言語識別, TR-I-0212 (1991-04).
- [10] 杉山, Bilingual 音声を対象とした日英言語識別の検討, 音響学会講演論文集, 3-6-12 (1991-10).
- [11] 松井, 古井, エルゴード的 HMM による話者認識, 音響学会講演論文集, 3-6-14 (1991-10).

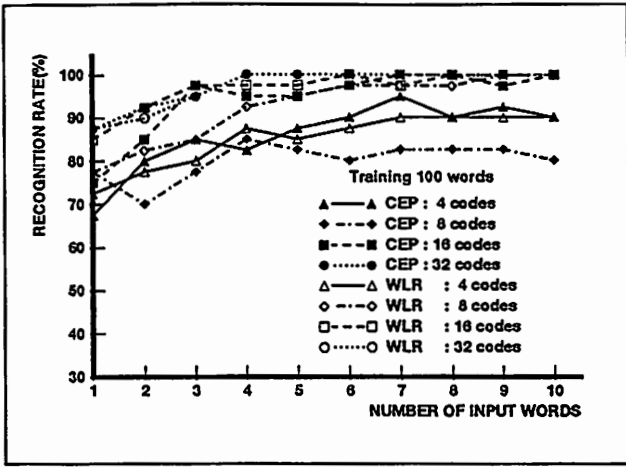


図 4: VQ 符号帳による識別結果

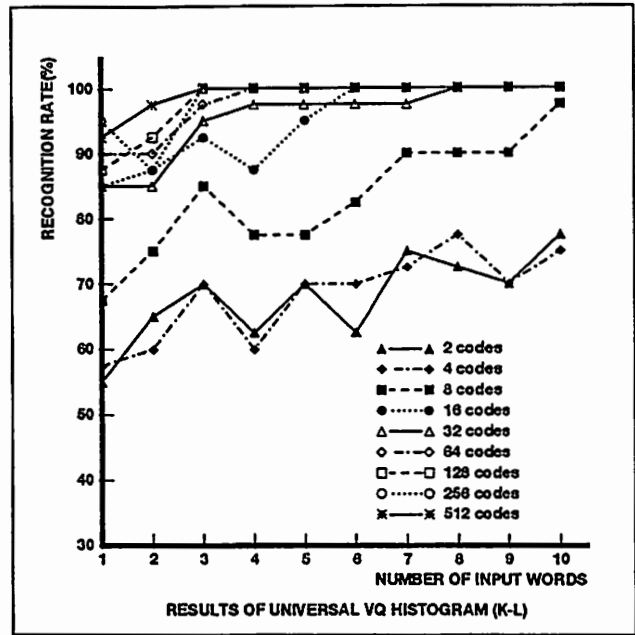


図 6: Universal 符号帳とその確率密度による識別結果 (K-L 情報量)

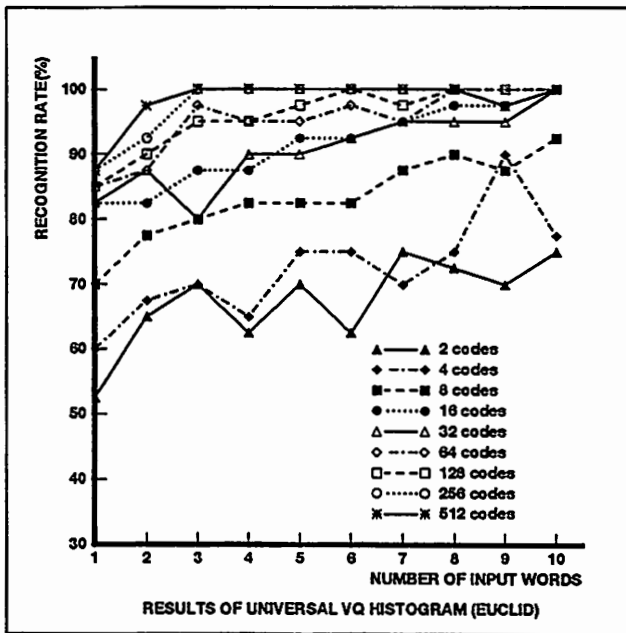


図 5: Universal 符号帳とその確率密度による識別結果 (Euclid 距離)

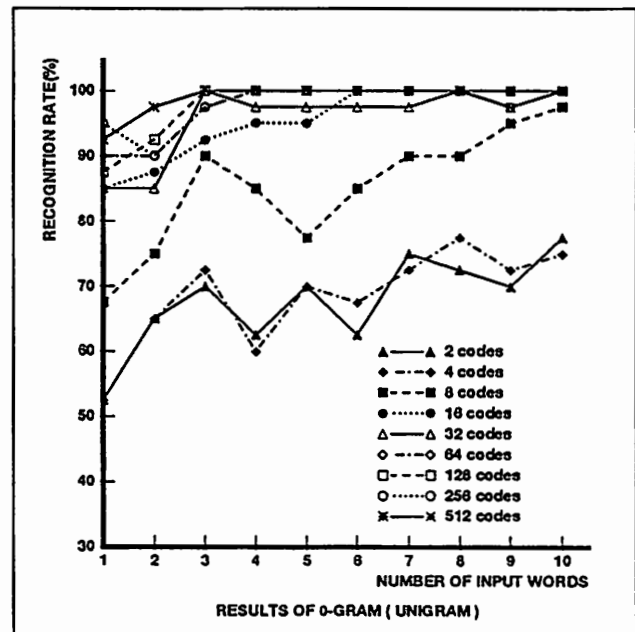


図 7: 0-gram (unigram) による識別結果

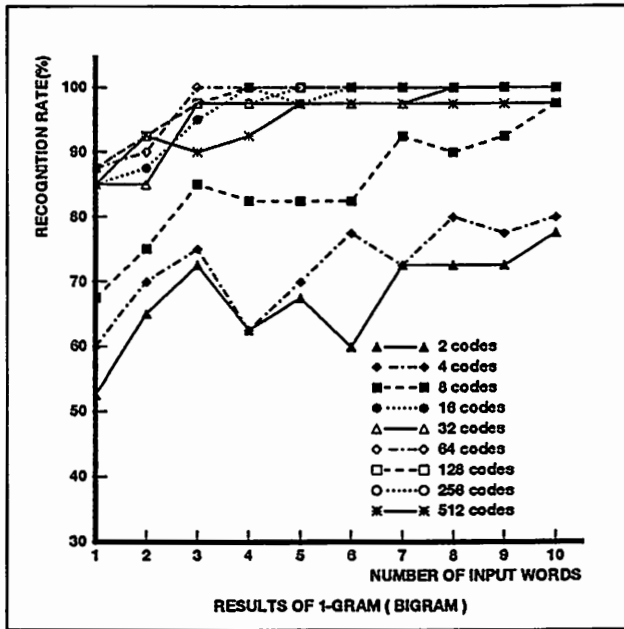


図 8: 1-gram (bigram) による識別結果

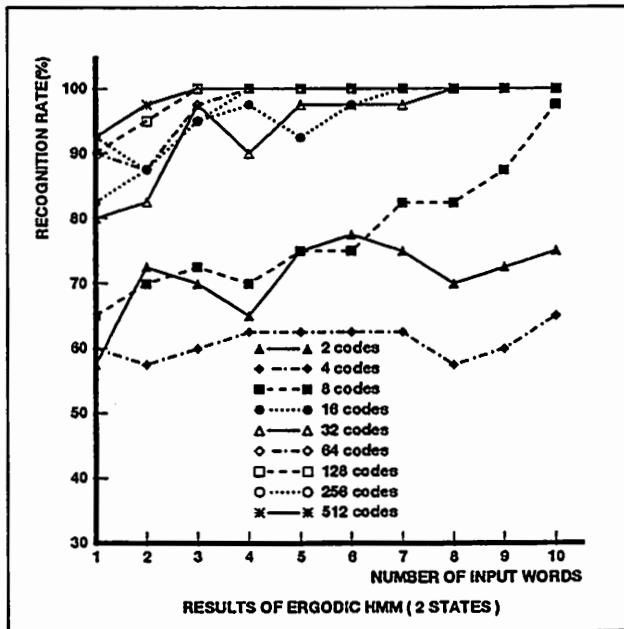


図 9: ergodic HMM (2 states) による識別結果

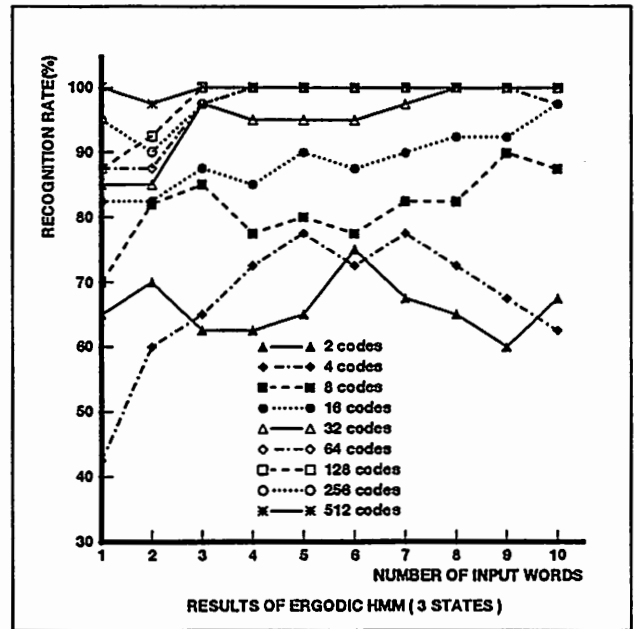


図 10: ergodic HMM (3 states) による識別結果

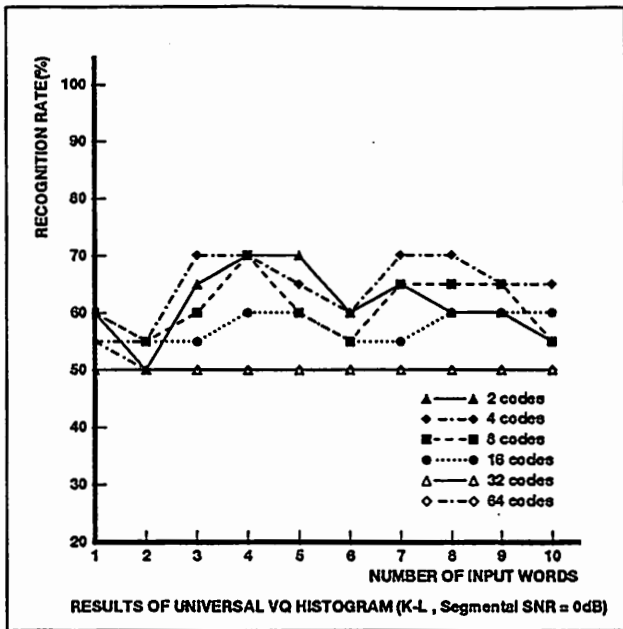


図 11: フレーム毎に雑音が重畳した場合の識別結果 (ヒストグラム法 (K-L 情報量))

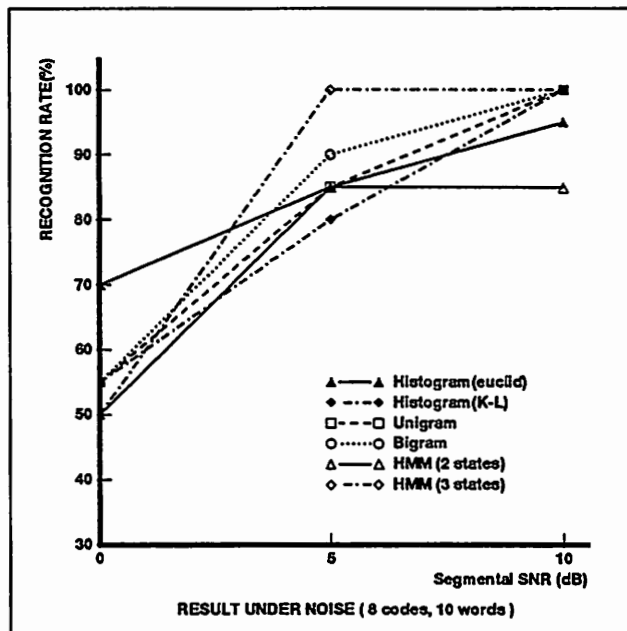


図 13: フレーム毎に雑音が重畳した場合の識別結果

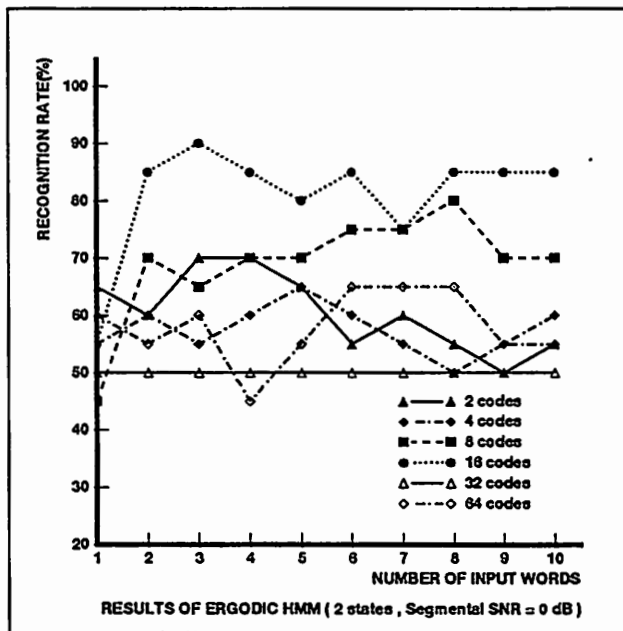


図 12: フレーム毎に雑音が重畳した場合の識別結果 (ergodic HMM (2 states))

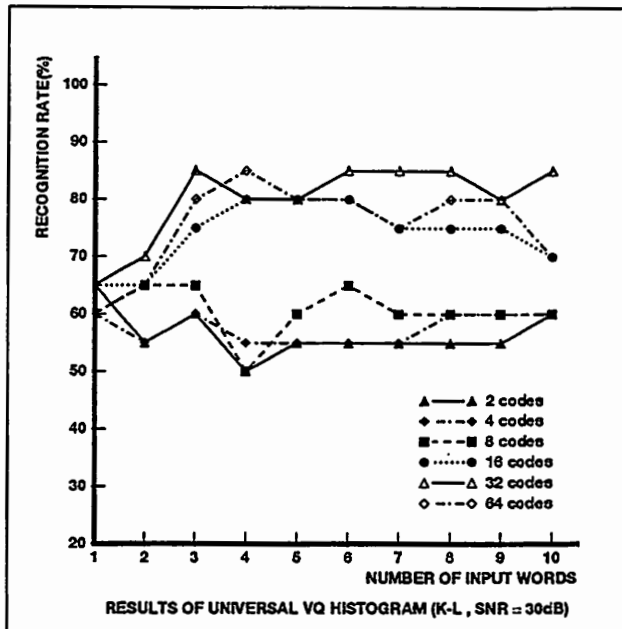


図 14: 単語全体に雑音が重畳した場合の識別結果 (ヒストグラム法 (K-L 情報量))

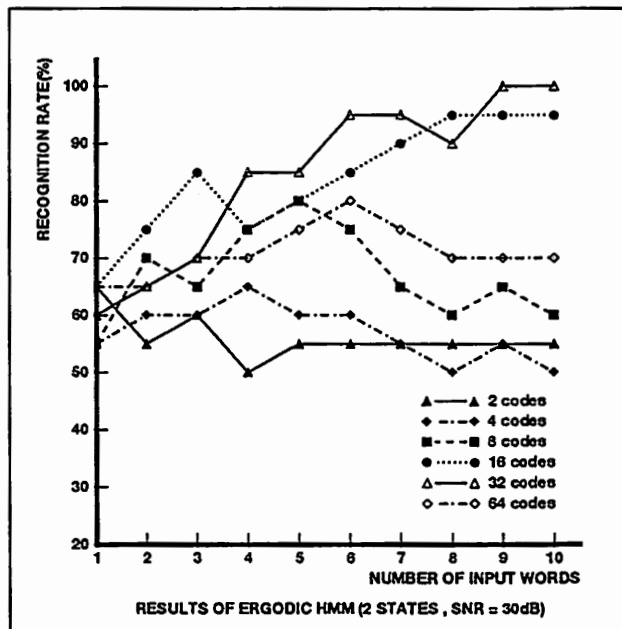


図 15: 単語全体に雑音が重畳した場合の識別結果 (ergodic HMM (2 states))

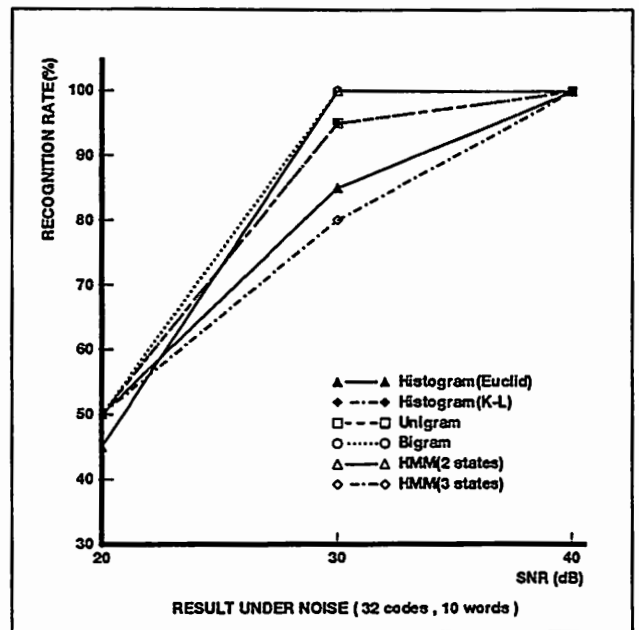


図 16: 単語全体に雑音が重畳した場合の識別結果

A 日・英語の識別率

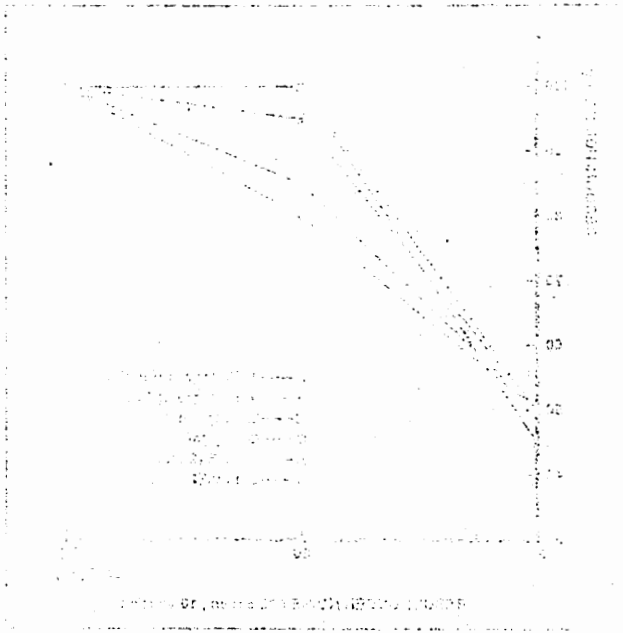


図1 日・英語の識別率

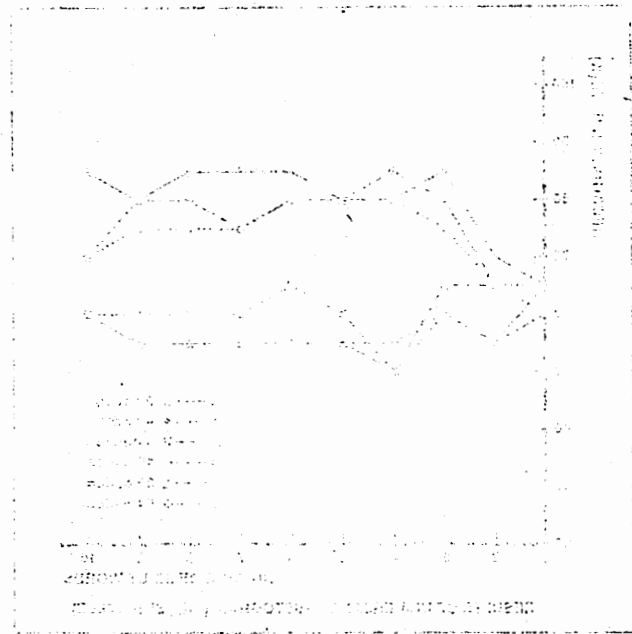


図2 日・英語の識別率

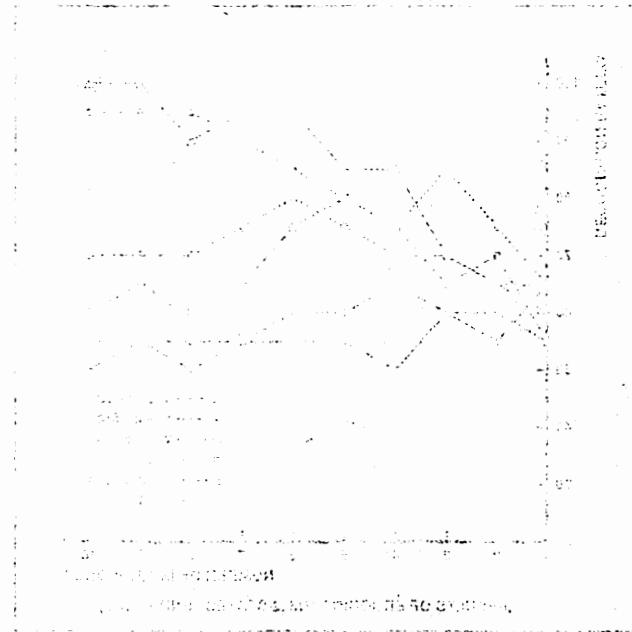


図3 日・英語の識別率

表 4: Universal VQ 符号帳 (Euclid 距離) による日・英の識別率 (話者: 男性 1 名 (MAO))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	70.0	65.0	70.0	75.0	60.0	75.0	75.0	75.0	75.0
4	55.0	75.0	70.0	70.0	80.0	85.0	75.0	80.0	90.0	85.0
8	60.0	85.0	85.0	80.0	85.0	85.0	90.0	90.0	90.0	95.0
16	80.0	85.0	80.0	85.0	90.0	90.0	95.0	95.0	95.0	100.0
32	80.0	90.0	95.0	90.0	90.0	95.0	95.0	95.0	95.0	100.0
64	75.0	90.0	100.0	95.0	95.0	100.0	95.0	100.0	100.0	100.0
128	75.0	95.0	95.0	90.0	95.0	100.0	95.0	100.0	100.0	100.0
256	80.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	95.0	100.0
512	80.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	95.0	100.0

表 5: Universal VQ 符号帳 (K-L 情報量) による日・英の識別率 (話者: 男性 1 名 (MAO))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	70.0	65.0	70.0	75.0	60.0	75.0	75.0	75.0	80.0
4	55.0	75.0	80.0	65.0	75.0	80.0	75.0	85.0	80.0	85.0
8	60.0	80.0	90.0	75.0	80.0	85.0	90.0	90.0	90.0	100.0
16	80.0	85.0	90.0	80.0	95.0	100.0	100.0	100.0	100.0	100.0
32	80.0	80.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	85.0	90.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	80.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	90.0	85.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 6: 0-gram ( unigram ) による日・英の識別率 (話者: 男性 1 名 (MAO))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	70.0	65.0	70.0	75.0	60.0	75.0	75.0	75.0	80.0
4	55.0	75.0	80.0	65.0	75.0	75.0	75.0	85.0	85.0	85.0
8	60.0	80.0	90.0	90.0	80.0	90.0	90.0	90.0	100.0	100.0
16	80.0	85.0	90.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0
32	80.0	80.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	85.0	90.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	80.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	90.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	95.0	100.0
512	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	95.0	100.0

表 7: 1-gram ( bigram ) による日・英の識別率 (話者: 男性 1 名 (MAO))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	70.0	65.0	70.0	75.0	60.0	75.0	75.0	75.0	80.0
4	60.0	75.0	80.0	70.0	80.0	85.0	75.0	85.0	90.0	90.0
8	65.0	80.0	95.0	80.0	85.0	90.0	100.0	90.0	95.0	100.0
16	80.0	85.0	95.0	100.0	95.0	100.0	100.0	100.0	100.0	100.0
32	85.0	85.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	80.0	85.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	80.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	85.0	95.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0
512	85.0	100.0	85.0	90.0	95.0	95.0	95.0	95.0	95.0	95.0

表 8: ergodic HMM ( 2 states ) による日・英の識別率 (話者: 男性 1 名 (MAO))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	70.0	60.0	60.0	65.0	70.0	75.0	70.0	70.0	65.0
4	55.0	60.0	60.0	65.0	65.0	65.0	70.0	55.0	55.0	70.0
8	60.0	70.0	70.0	60.0	75.0	75.0	80.0	75.0	85.0	90.0
16	75.0	85.0	95.0	95.0	85.0	95.0	100.0	100.0	100.0	100.0
32	75.0	80.0	100.0	85.0	100.0	100.0	100.0	100.0	100.0	100.0
64	85.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	85.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	90.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 9: ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	70.0	70.0	45.0	60.0	50.0	70.0	60.0	65.0	50.0	55.0
4	30.0	45.0	50.0	60.0	60.0	60.0	65.0	60.0	55.0	50.0
8	70.0	85.0	95.0	70.0	80.0	75.0	75.0	75.0	85.0	75.0
16	80.0	85.0	95.0	90.0	95.0	95.0	95.0	95.0	95.0	100.0
32	80.0	80.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	80.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	80.0	85.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	90.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	100.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0



表 10: Universal VQ 符号帳 (Euclid 距離) による日・英の識別率 (話者: 女性 1 名 (FCF))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	60.0	75.0	55.0	65.0	65.0	75.0	70.0	65.0	75.0
4	65.0	60.0	70.0	60.0	70.0	65.0	65.0	70.0	65.0	70.0
8	80.0	70.0	75.0	85.0	80.0	80.0	85.0	90.0	85.0	90.0
16	85.0	80.0	95.0	90.0	95.0	95.0	95.0	100.0	100.0	100.0
32	85.0	85.0	85.0	90.0	90.0	90.0	95.0	95.0	95.0	100.0
64	95.0	85.0	95.0	95.0	95.0	95.0	95.0	100.0	100.0	100.0
128	95.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	95.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 11: Universal VQ 符号帳 (K-L 情報量) による日・英の識別率 (話者: 女性 1 名 (FCF))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	60.0	75.0	55.0	65.0	65.0	75.0	70.0	65.0	75.0
4	60.0	55.0	60.0	55.0	65.0	60.0	70.0	70.0	60.0	65.0
8	75.0	70.0	80.0	80.0	75.0	80.0	90.0	90.0	90.0	95.0
16	90.0	90.0	95.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0
32	90.0	90.0	95.0	95.0	95.0	95.0	95.0	100.0	100.0	100.0
64	95.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	100.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 12: 0-gram ( unigram ) による日・英の識別率 (話者: 女性 1 名 (FCF))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	60.0	75.0	55.0	65.0	65.0	75.0	70.0	65.0	75.0
4	60.0	55.0	65.0	55.0	65.0	60.0	70.0	70.0	60.0	65.0
8	75.0	70.0	90.0	80.0	75.0	80.0	90.0	90.0	90.0	95.0
16	90.0	90.0	95.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0
32	90.0	90.0	100.0	95.0	95.0	95.0	95.0	100.0	100.0	100.0
64	95.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	100.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 13: 1-gram ( bigram ) による日・英の識別率 (話者: 女性 1 名 (FCF))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	60.0	80.0	55.0	60.0	60.0	70.0	75.0	70.0	75.0
4	60.0	65.0	70.0	55.0	60.0	70.0	70.0	75.0	65.0	70.0
8	70.0	70.0	75.0	85.0	80.0	75.0	85.0	90.0	90.0	95.0
16	90.0	90.0	95.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0
32	90.0	85.0	95.0	95.0	95.0	95.0	95.0	95.0	95.0	100.0
64	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	95.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	90.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	85.0	85.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0

表 14: ergodic HMM ( 2 states ) による日・英の識別率 (話者: 女性 1 名 (FCF))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	75.0	80.0	70.0	85.0	85.0	75.0	70.0	75.0	85.0
4	65.0	55.0	60.0	60.0	60.0	60.0	55.0	60.0	65.0	60.0
8	70.0	70.0	75.0	80.0	75.0	75.0	85.0	90.0	90.0	100.0
16	90.0	90.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
32	85.0	85.0	95.0	95.0	95.0	95.0	95.0	100.0	100.0	100.0
64	95.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
128	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	95.0	90.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 15: ergodic HMM ( 3 states ) による日・英の識別率 (話者: 女性 1 名 (FCF))

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	70.0	80.0	65.0	80.0	80.0	75.0	65.0	70.0	80.0
4	55.0	75.0	80.0	85.0	95.0	85.0	90.0	85.0	80.0	75.0
8	70.0	80.0	75.0	85.0	80.0	80.0	90.0	90.0	95.0	100.0
16	85.0	80.0	80.0	80.0	85.0	80.0	85.0	90.0	90.0	95.0
32	90.0	90.0	95.0	90.0	90.0	90.0	95.0	100.0	100.0	100.0
64	95.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	95.0
128	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
256	100.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
512	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 16: Universal VQ 符号帳 (Euclid 距離) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 0dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	50.0	65.0	70.0	70.0	60.0	65.0	65.0	60.0	55.0
4	55.0	55.0	65.0	70.0	70.0	70.0	75.0	75.0	65.0	65.0
8	55.0	50.0	60.0	60.0	60.0	60.0	60.0	60.0	60.0	50.0
16	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	00.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 17: Universal VQ 符号帳 (Euclid 距離) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 5dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	70.0	75.0	75.0	80.0	85.0	85.0	80.0	85.0	90.0
4	55.0	70.0	70.0	75.0	75.0	85.0	75.0	75.0	85.0	85.0
8	65.0	80.0	80.0	70.0	75.0	75.0	80.0	85.0	85.0	85.0
16	75.0	75.0	80.0	90.0	80.0	95.0	90.0	95.0	90.0	100.0
32	75.0	75.0	80.0	85.0	80.0	95.0	95.0	100.0	100.0	100.0
64	70.0	55.0	65.0	65.0	70.0	70.0	70.0	70.0	60.0	55.0

表 18: Universal VQ 符号帳 (Euclid 距離) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 10dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	65.0	65.0	65.0	80.0	75.0	85.0	85.0	85.0	75.0
4	50.0	70.0	70.0	80.0	80.0	80.0	75.0	80.0	85.0	90.0
8	60.0	85.0	85.0	75.0	85.0	85.0	75.0	95.0	90.0	95.0
16	80.0	80.0	85.0	85.0	85.0	90.0	95.0	100.0	100.0	100.0
32	80.0	80.0	85.0	85.0	90.0	100.0	95.0	100.0	100.0	100.0
64	80.0	85.0	95.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0

表 19: Universal VQ 符号帳 (K-L 情報量) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 0dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	50.0	65.0	70.0	70.0	60.0	65.0	60.0	60.0	55.0
4	55.0	55.0	70.0	70.0	65.0	60.0	70.0	70.0	65.0	65.0
8	60.0	55.0	60.0	70.0	60.0	55.0	65.0	65.0	65.0	55.0
16	60.0	55.0	55.0	60.0	60.0	55.0	55.0	60.0	60.0	60.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	55.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 20: Universal VQ 符号帳 (K-L 情報量) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 5dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	70.0	75.0	75.0	80.0	80.0	85.0	80.0	85.0	90.0
4	60.0	60.0	70.0	75.0	75.0	80.0	80.0	80.0	90.0	80.0
8	55.0	80.0	85.0	75.0	75.0	85.0	75.0	80.0	75.0	80.0
16	80.0	90.0	85.0	90.0	80.0	85.0	90.0	90.0	95.0	100.0
32	70.0	75.0	90.0	95.0	95.0	100.0	95.0	100.0	100.0	100.0
64	80.0	65.0	65.0	80.0	85.0	90.0	85.0	90.0	80.0	75.0

表 21: Universal VQ 符号帳 (K-L 情報量) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 10dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	65.0	65.0	65.0	80.0	75.0	85.0	85.0	85.0	75.0
4	50.0	70.0	75.0	75.0	85.0	80.0	75.0	85.0	90.0	85.0
8	60.0	80.0	90.0	85.0	75.0	80.0	80.0	100.0	100.0	100.0
16	80.0	85.0	90.0	90.0	90.0	95.0	100.0	100.0	100.0	100.0
32	80.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	85.0	85.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0

表 22: 0-gram ( unigram ) による日・英の識別率 (話者 : 男性 1 名 (MAO),segmental SNR 0dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	50.0	65.0	60.0	70.0	60.0	60.0	60.0	60.0	55.0
4	60.0	50.0	60.0	60.0	65.0	60.0	65.0	65.0	65.0	65.0
8	65.0	55.0	65.0	70.0	60.0	55.0	65.0	65.0	65.0	55.0
16	60.0	60.0	60.0	60.0	60.0	55.0	55.0	55.0	60.0	60.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	55.0	55.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 23: 0-gram ( unigram ) による日・英の識別率 (話者 : 男性 1 名 (MAO),segmental SNR 5dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	65.0	80.0	75.0	75.0	75.0	75.0	70.0	75.0	85.0
4	70.0	60.0	75.0	80.0	70.0	70.0	80.0	85.0	80.0	80.0
8	60.0	75.0	90.0	80.0	75.0	75.0	80.0	95.0	85.0	85.0
16	85.0	85.0	90.0	85.0	75.0	85.0	90.0	95.0	100.0	100.0
32	85.0	75.0	90.0	95.0	95.0	100.0	95.0	100.0	100.0	100.0
64	75.0	70.0	70.0	80.0	80.0	90.0	80.0	85.0	85.0	80.0

表 24: 0-gram ( unigram ) による日・英の識別率 (話者 : 男性 1 名 (MAO),segmental SNR 10dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	65.0	70.0	65.0	75.0	75.0	85.0	80.0	85.0	80.0
4	65.0	60.0	75.0	75.0	80.0	85.0	75.0	90.0	90.0	90.0
8	60.0	80.0	85.0	80.0	75.0	85.0	85.0	100.0	100.0	100.0
16	80.0	85.0	90.0	90.0	90.0	95.0	100.0	100.0	100.0	100.0
32	80.0	80.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	85.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 25: 1-gram ( bigram ) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 0dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	50.0	65.0	60.0	70.0	60.0	60.0	60.0	60.0	55.0
4	60.0	50.0	60.0	55.0	60.0	60.0	65.0	65.0	65.0	65.0
8	65.0	55.0	65.0	65.0	60.0	55.0	65.0	65.0	65.0	55.0
16	60.0	50.0	55.0	55.0	55.0	55.0	55.0	55.0	55.0	55.0
32	65.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	50.0	55.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 26: 1-gram ( bigram ) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 5dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	65.0	80.0	75.0	75.0	75.0	75.0	70.0	75.0	85.0
4	70.0	60.0	75.0	80.0	85.0	75.0	85.0	85.0	60.0	85.0
8	65.0	80.0	90.0	80.0	75.0	90.0	85.0	95.0	85.0	90.0
16	85.0	90.0	90.0	85.0	80.0	100.0	90.0	95.0	100.0	100.0
32	85.0	80.0	95.0	95.0	90.0	90.0	100.0	100.0	100.0	100.0
64	70.0	80.0	85.0	90.0	90.0	90.0	90.0	100.0	95.0	95.0

表 27: 1-gram ( bigram ) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 10dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	65.0	70.0	65.0	75.0	80.0	85.0	80.0	85.0	80.0
4	70.0	55.0	75.0	75.0	80.0	85.0	75.0	90.0	90.0	90.0
8	65.0	80.0	95.0	75.0	75.0	85.0	85.0	100.0	100.0	100.0
16	85.0	85.0	95.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0
32	80.0	85.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	80.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 28: ergodic HMM ( 2 states ) による日・英の識別率 (話者 : 男性 1 名 (MAO),segmental SNR 0dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	60.0	70.0	70.0	65.0	55.0	60.0	55.0	50.0	55.0
4	55.0	60.0	55.0	60.0	65.0	60.0	55.0	50.0	55.0	60.0
8	45.0	70.0	65.0	70.0	70.0	75.0	75.0	80.0	70.0	70.0
16	55.0	85.0	90.0	85.0	80.0	85.0	75.0	85.0	85.0	85.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	60.0	55.0	60.0	45.0	55.0	65.0	65.0	65.0	55.0	55.0

表 29: ergodic HMM ( 2 states ) による日・英の識別率 (話者 : 男性 1 名 (MAO),segmental SNR 5dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	70.0	70.0	70.0	70.0	75.0	85.0	75.0	75.0	75.0	70.0
4	55.0	65.0	55.0	70.0	70.0	65.0	55.0	65.0	65.0	65.0
8	65.0	75.0	75.0	85.0	80.0	85.0	80.0	90.0	85.0	85.0
16	85.0	90.0	85.0	90.0	85.0	90.0	90.0	90.0	100.0	100.0
32	85.0	80.0	90.0	90.0	95.0	95.0	95.0	100.0	100.0	100.0
64	80.0	75.0	85.0	85.0	90.0	90.0	90.0	95.0	100.0	100.0

表 30: ergodic HMM ( 2 states ) による日・英の識別率 (話者 : 男性 1 名 (MAO),segmental SNR 10dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	70.0	65.0	70.0	65.0	80.0	75.0	70.0	70.0	70.0
4	55.0	65.0	65.0	70.0	60.0	75.0	60.0	55.0	70.0	70.0
8	65.0	75.0	75.0	80.0	75.0	80.0	75.0	85.0	80.0	85.0
16	85.0	85.0	90.0	90.0	85.0	95.0	100.0	100.0	100.0	100.0
32	75.0	80.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	90.0	80.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 31: ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 0dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	55.0	65.0	65.0	65.0	50.0	65.0	50.0	55.0	60.0
4	40.0	55.0	55.0	55.0	55.0	60.0	65.0	65.0	55.0	50.0
8	50.0	50.0	55.0	55.0	55.0	55.0	55.0	50.0	50.0	50.0
16	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	55.0	50.0	55.0	50.0	55.0	50.0	50.0	60.0	50.0	50.0

表 32: ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 5dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	70.0	65.0	65.0	65.0	60.0	65.0	55.0	60.0	55.0	55.0
4	35.0	50.0	60.0	60.0	60.0	65.0	70.0	65.0	55.0	65.0
8	65.0	70.0	85.0	80.0	80.0	90.0	90.0	95.0	95.0	100.0
16	80.0	80.0	75.0	80.0	85.0	90.0	95.0	95.0	95.0	95.0
32	85.0	75.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0
64	65.0	65.0	70.0	85.0	90.0	90.0	80.0	80.0	80.0	75.0

表 33: ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 10dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	90.0	65.0	70.0	65.0	65.0	55.0	60.0	60.0	65.0	65.0
4	35.0	50.0	50.0	65.0	60.0	60.0	55.0	60.0	55.0	55.0
8	65.0	80.0	85.0	75.0	85.0	90.0	90.0	90.0	95.0	100.0
16	80.0	85.0	90.0	95.0	90.0	95.0	90.0	100.0	100.0	100.0
32	80.0	80.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	80.0	85.0	95.0	95.0	100.0	100.0	95.0	100.0	100.0	100.0



表 34: Universal VQ 符号帳 (Euclid 距離) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 20dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
4	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
8	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
16	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 35: Universal VQ 符号帳 (Euclid 距離) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 30dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	55.0	60.0	50.0	55.0	55.0	55.0	55.0	55.0	60.0
4	60.0	55.0	60.0	55.0	55.0	55.0	55.0	60.0	60.0	60.0
8	60.0	65.0	65.0	50.0	60.0	65.0	60.0	60.0	60.0	60.0
16	65.0	65.0	75.0	80.0	80.0	80.0	75.0	75.0	75.0	70.0
32	65.0	70.0	85.0	80.0	80.0	85.0	85.0	85.0	80.0	85.0
64	60.0	65.0	80.0	85.0	80.0	80.0	75.0	75.0	80.0	70.0

表 36: Universal VQ 符号帳 (Euclid 距離) による日・英の識別率 (話者: 男性 1 名 (MAO), segmental SNR 40dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	70.0	65.0	70.0	80.0	70.0	80.0	85.0	85.0	75.0
4	55.0	70.0	75.0	70.0	80.0	80.0	75.0	80.0	85.0	85.0
8	65.0	85.0	85.0	80.0	85.0	85.0	90.0	90.0	90.0	95.0
16	75.0	85.0	80.0	85.0	90.0	90.0	95.0	95.0	95.0	100.0
32	80.0	90.0	85.0	90.0	90.0	95.0	95.0	95.0	100.0	100.0
64	85.0	85.0	100.0	90.0	95.0	100.0	95.0	100.0	100.0	100.0

表 37: Universal VQ 符号帳 (K-L 情報量) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 20dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
4	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
8	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
16	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 38: Universal VQ 符号帳 (K-L 情報量) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 30dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	55.0	60.0	50.0	55.0	55.0	55.0	55.0	55.0	60.0
4	55.0	55.0	55.0	55.0	55.0	55.0	60.0	60.0	60.0	60.0
8	60.0	60.0	70.0	65.0	65.0	65.0	55.0	60.0	60.0	60.0
16	65.0	60.0	85.0	80.0	80.0	75.0	80.0	80.0	75.0	80.0
32	65.0	65.0	90.0	90.0	95.0	95.0	95.0	95.0	95.0	95.0
64	65.0	60.0	65.0	65.0	75.0	75.0	70.0	70.0	65.0	70.0

表 39: Universal VQ 符号帳 (K-L 情報量) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 40dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	70.0	65.0	65.0	80.0	70.0	80.0	85.0	85.0	80.0
4	55.0	70.0	80.0	70.0	80.0	80.0	75.0	85.0	90.0	85.0
8	60.0	80.0	90.0	75.0	80.0	85.0	90.0	100.0	90.0	100.0
16	75.0	85.0	90.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0
32	80.0	80.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	85.0	90.0	90.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 40: 0-gram ( unigram ) による日・英の識別率 (話者: 男性 1 名 (MAO),SNR 20dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
4	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
8	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
16	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 41: 0-gram ( unigram ) による日・英の識別率 (話者: 男性 1 名 (MAO),SNR 30dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	55.0	60.0	50.0	55.0	55.0	55.0	55.0	70.0	60.0
4	55.0	55.0	55.0	55.0	55.0	55.0	60.0	60.0	60.0	60.0
8	60.0	60.0	70.0	65.0	65.0	65.0	55.0	60.0	60.0	60.0
16	65.0	60.0	85.0	80.0	80.0	75.0	80.0	80.0	75.0	80.0
32	65.0	65.0	90.0	90.0	95.0	95.0	95.0	95.0	95.0	95.0
64	65.0	60.0	65.0	65.0	75.0	75.0	70.0	70.0	65.0	70.0

表 42: 0-gram ( unigram ) による日・英の識別率 (話者: 男性 1 名 (MAO),SNR 40dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	70.0	65.0	65.0	80.0	70.0	80.0	85.0	85.0	80.0
4	55.0	70.0	80.0	70.0	80.0	80.0	75.0	85.0	90.0	85.0
8	60.0	80.0	90.0	75.0	80.0	85.0	90.0	95.0	90.0	100.0
16	75.0	85.0	90.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0
32	80.0	80.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	85.0	90.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 43: 1-gram ( bigram ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 20dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
4	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
8	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
16	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 44: 1-gram ( bigram ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 30dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	55.0	60.0	50.0	55.0	55.0	55.0	55.0	55.0	60.0
4	60.0	55.0	60.0	55.0	55.0	55.0	60.0	60.0	60.0	60.0
8	60.0	65.0	70.0	60.0	65.0	70.0	60.0	60.0	60.0	60.0
16	65.0	60.0	90.0	80.0	85.0	75.0	85.0	80.0	85.0	85.0
32	65.0	70.0	90.0	95.0	95.0	95.0	95.0	95.0	95.0	100.0
64	70.0	65.0	70.0	70.0	80.0	80.0	70.0	70.0	70.0	80.0

表 45: 1-gram ( bigram ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 40dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	60.0	70.0	70.0	65.0	80.0	70.0	80.0	85.0	85.0	80.0
4	60.0	75.0	80.0	75.0	80.0	80.0	75.0	85.0	90.0	85.0
8	65.0	80.0	95.0	80.0	85.0	100.0	90.0	95.0	95.0	100.0
16	80.0	85.0	95.0	95.0	95.0	100.0	100.0	100.0	100.0	100.0
32	85.0	85.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	80.0	85.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 46: ergodic HMM ( 2 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 20dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
4	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
8	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
16	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	60.0	60.0
32	50.0	50.0	45.0	45.0	45.0	45.0	45.0	45.0	45.0	45.0
64	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 47: ergodic HMM ( 2 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 30dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	55.0	60.0	50.0	55.0	55.0	55.0	55.0	55.0	55.0
4	55.0	60.0	60.0	65.0	60.0	60.0	55.0	50.0	55.0	50.0
8	55.0	70.0	65.0	75.0	80.0	75.0	65.0	60.0	65.0	60.0
16	65.0	75.0	85.0	75.0	80.0	85.0	90.0	95.0	95.0	95.0
32	60.0	65.0	70.0	85.0	85.0	95.0	95.0	90.0	100.0	100.0
64	65.0	65.0	70.0	70.0	75.0	80.0	75.0	70.0	70.0	70.0

表 48: ergodic HMM ( 2 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 40dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	55.0	70.0	65.0	75.0	75.0	75.0	70.0	75.0	70.0	65.0
4	55.0	60.0	65.0	65.0	65.0	65.0	70.0	55.0	60.0	70.0
8	60.0	70.0	70.0	80.0	75.0	80.0	80.0	75.0	85.0	90.0
16	75.0	80.0	95.0	85.0	85.0	100.0	100.0	100.0	100.0	100.0
32	75.0	80.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	80.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

表 49: ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 20dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
4	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
8	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
16	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
32	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
64	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0

表 50: ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 30dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	65.0	50.0	50.0	55.0	55.0	55.0	55.0	55.0	55.0	50.0
4	60.0	55.0	60.0	55.0	60.0	60.0	60.0	60.0	60.0	60.0
8	65.0	60.0	60.0	50.0	55.0	60.0	60.0	60.0	60.0	60.0
16	65.0	55.0	65.0	60.0	55.0	60.0	60.0	60.0	60.0	60.0
32	65.0	55.0	85.0	85.0	90.0	80.0	85.0	90.0	90.0	80.0
64	60.0	55.0	65.0	65.0	70.0	70.0	65.0	65.0	65.0	65.0

表 51: ergodic HMM ( 3 states ) による日・英の識別率 (話者: 男性 1 名 (MAO), SNR 40dB)

符号帳の 大きさ	入力単語数									
	1	2	3	4	5	6	7	8	9	10
2	75.0	65.0	65.0	70.0	70.0	70.0	75.0	70.0	70.0	65.0
4	30.0	45.0	55.0	55.0	60.0	60.0	60.0	60.0	55.0	55.0
8	70.0	85.0	75.0	70.0	80.0	75.0	80.0	75.0	90.0	85.0
16	75.0	85.0	95.0	90.0	95.0	95.0	95.0	95.0	95.0	100.0
32	80.0	80.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
64	80.0	85.0	95.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

## B 識別実験に用いたソフトウェア

~/ の下にあるディレクトリ及びファイルの説明

~/BIGRAM/ : 1-gram (bigram) の学習と認識を行なう。

main : 二つ組の符号の出現確率を求める (学習時) 実行ファイル  
main.run : main を動かすシェル  
REC/ : 認識用のディレクトリ  
main : 認識実行ファイル  
main.run : main を動かすシェル

~/Cor/ : 波形データから自己相関分析をするディレクトリ

WaveCorr : 分析実行ファイル  
run.WaveCorr : WaveCorr を動かすシェル

~/EUC/ : Euclid 距離で確率密度関数による識別を行なうディレクトリ

main : 実行ファイル  
main.run : main を動かすシェル

~/HMM/ : ergodic HMM に関するディレクトリ

MK\_3S\_HMM/ : 3 状態の ergodic HMM に関するディレクトリ

REC/ : 認識用ディレクトリ  
main : 認識実行ファイル  
main.run : main を動かすシェル  
init/ : 3 状態の ergodic HMM の初期値を求めるディレクトリ  
mk\_init : 初期値を求める実行ファイル  
mk\_init.run : mk\_init を動かすシェル

main : ergodic HMM (3 states) の学習を行なう実行ファイル

main.run : main を動かすシェル

PRG\_S/ : 2 状態の ergodic HMM の学習を行なうディレクトリ

main : ergodic HMM (2 states) の学習を行なう実行ファイル  
main.run : main を動かすシェル

REC\_L/ : 2 状態の ergodic HMM の認識を行なうディレクトリ

main : 認識実行ファイル  
main.run : main を動かすシェル

init/ : 2 状態の ergodic HMM の初期値を求めるディレクトリ

mk\_init : 初期値を求める実行ファイル  
mk\_init.run : mk\_init を動かすシェル

\*sym/ : VQ された符号系列の格納ディレクトリ

~/KL\_HIST/ : K-L 情報量で確率密度関数による識別を行なうディレクトリ

main : 実行ファイル  
main.run : main を動かすシェル

~/LABEL/ : 音声データのラベルファイルの格納ディレクトリ

~/SEQ/ : 分析対象のファイル名を列記してあるファイルの格納ディレクトリ

~/UNIGRAM/ : 0-gram (unigram) の学習と認識を行なうディレクトリ

- REC/ : 認識用のディレクトリ
- main : 認識実行ファイル
- main.run : main を動かすシェル

~/VQ/ : ベクトル量子化に関するディレクトリ

- HIST\_DIC/ : 符号の確率密度関数の格納ディレクトリ
- MK\_HIST : 符号の確率密度関数を求めるディレクトリ
  - main : 関数を求める実行ファイル
  - main.run : main を動かすシェル
- vq/ : ベクトル量子化するディレクトリ
  - vq : VQ 実行ファイル (無雑音, フレーム毎の雑音)
  - vq.run : vq を動かすシェル
- word\_noise\_vq/ : ベクトル量子化するディレクトリ
  - vq : VQ 実行ファイル (単語全体の雑音)
  - vq.run : vq を動かすシェル

~/cdbk : コードブックを作るディレクトリ

- make\_cdbook : 実行ファイル
- run.make\_cdbook.test: make\_cdbook を動かすシェル



実験手順について

1. ~/Cor/WaveCor により LPC 相関係数を求める。
2. ~/cdbk/make\_cdbook によりコードブックを作成する。
3. ~/MK\_HIST/main により確率密度関数を求める。  
(確率密度関数 (Euclid, K-L), unigram による認識に使用する。)
4. ~/VQ/vq/vq によりベクトル量子化し、符号系列を求める。  
(学習単語からの符号系列は、bigram , ergodic HMM の学習に使用する。)
5. ~/BIGRAM/main により二つ組の符号の出現確率を推定する。
6. ~/HMM/PRG\_S/main により ergodic HMM (2 states) の学習を行なう。
7. ~/HMM/MK\_3S\_HMM/main により ergodic HMM (3 states) の学習を行なう。

以上で学習を終了する。

8. ~/EUC\_HIST/main により確率密度関数による認識 (Euclid 距離) を行なう。
9. ~/KL\_HIST/main により確率密度関数による認識 (K-L 情報量) を行なう。
10. ~/UNIGRAM/REC/main により unigram による認識を行なう。
11. ~/BIGRAM/REC/main により bigram による認識を行なう。
12. ~/HMM/REC\_L/main により ergodic HMM (2 states) による認識を行なう。
12. ~/HMM/MK\_3S\_HMM/REC/main により ergodic HMM (3 states) による認識を行なう。

以上で認識を終了する。