

TR-S-0004

発話変動を考慮した音響・言語モデルの検討

Acoustic model and language model for  
speaking style variation

辻野 博紀  
Hiroki Tsujino

奥田 浩三  
Kozo Okuda

2000.9.8

現在の連続音声認識システムは、誤認識発声時の言い直し等、自然発話と異なる発話様式により発声された音声に対しては、認識率が劣化するという現象が生じている。本稿ではこのような問題に対し、音節発声に近い発話に対する現行音響モデルの性能評価と、このような発声に対し認識性能を向上するための方法について考察する。

©2000 A T R 音声言語通信研究所

©2000 by ATR Spoken Language Translation Research Laboratories

# 学外実習報告書

2000.09.08

当報告書作成者：龍谷大学 理工学部 電子情報学科  
t980206 辻野博紀

実習場所：(株)エイ・ティ・アール音声言語通信研究所  
ATR Spoken Language Translation  
Research Laboratories.

指導者：奥田浩三 研究員

テーマ：

発声変動を考慮した音響・言語モデルの検討

発声変動とは？

人が一番話しやすい話し方ではなく、多少の声の変動(強調、断片化など)のことをいう。

目的・背景：

人はいつも定常的な発声をするとは限らない。それを考慮した上で、音声認識の実験を行う。

目次

- 1 導入
- 2 実験概要
- 3 結果
- 4 考察
- 5 感想

# 1 導入

## 1.1 音声認識

音声認識とは人間が実際に発声した声、つまり音声の表層的内容(文字列情報、単語列情報)をコンピュータシステムによって同定することである。音声信号に対し信号処理を施し、事前に構築した音響モデル・言語モデルを用いて最も確からしい単語列を探索することにより認識を行う。

## 1.2 概要 - 音声認識のしくみ -

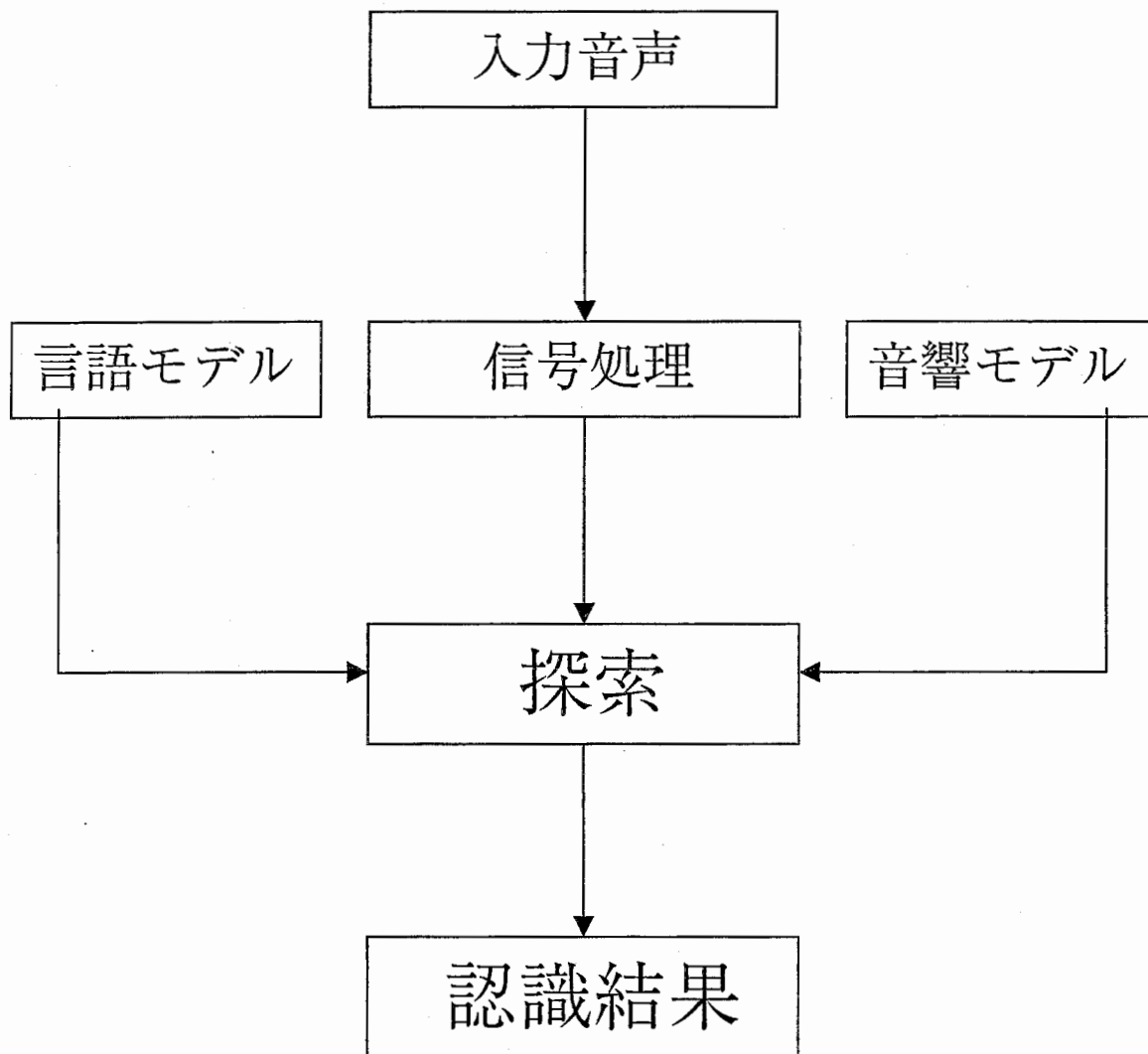


図1 音声認識のしくみ

### 音響モデル(Acoustic Model)

音声の構成単位(たとえばこの実験では 26 種類の音素)の音声認識に必要なスペクトル情報を、隠れマルコフモデル(HMM)の形式で確率的にモデル化する。

### 言語モデル

同定の対象となる単語、単語列を定める。また、単語と単語のつながりやすさを確率的にモデル化する。

### 信号処理

音声の波形信号を特徴ベクトルに変換する。

### 探索

音声の単語列情報を音響モデル、言語モデルを用いて同定する。

### 1.3 探索と音響モデル

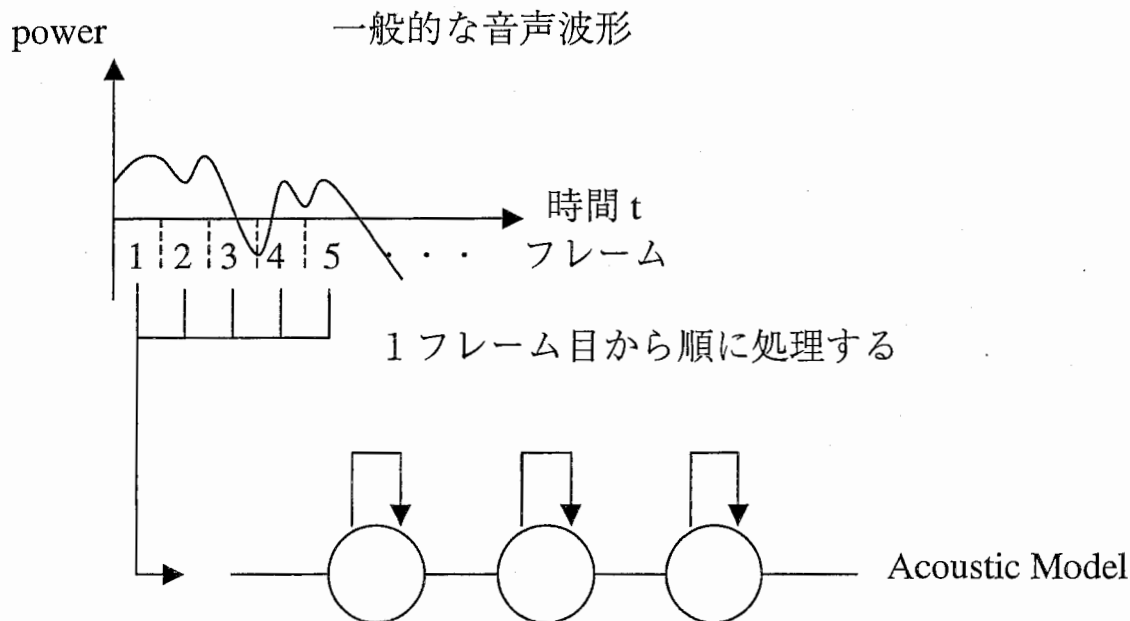
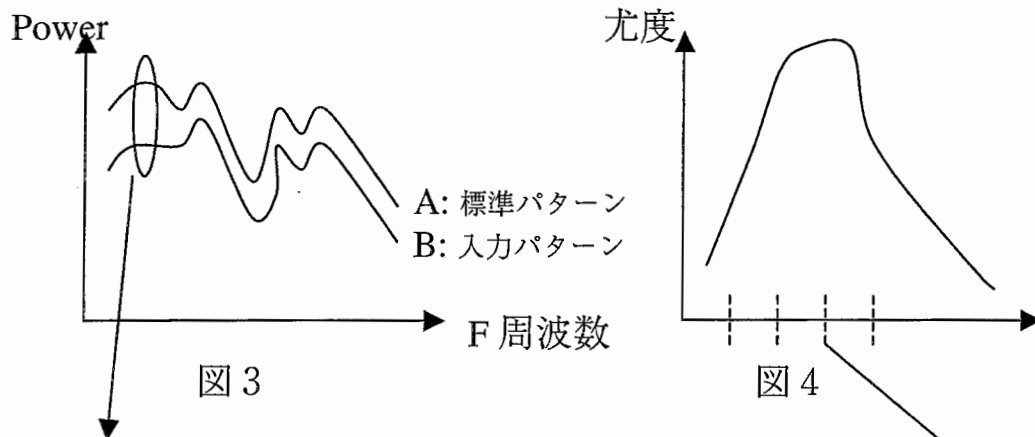


図 2 探索と音響モデル

ここでは探索(認識)にかけたときに、入力データがどのように音響モデルと比較されるかを概念的に簡単に解説する。

音響モデル(Acoustic Model)には一通りの Pass が存在する。そして 1 フレームごとに音声波形を読み込んでいくのだが、最終的に音声波形の読み込みが全て終了したとき、どのようなルートを通ってき

たのかということ考えると数パターン考えられる。そのそれぞれのパターンについて標準パターンと照らし合わせ、確率分布的な尤度を決定し、その尤度が一番高いものを一つ選択し、それを正解とする。



この A と B のズレを分布的に表現したものが図 4 である。

この尤度が一番高いものを一つ選出し、正解とする。

これが従来の音響モデルであるが、これを便宜上 Old Model と呼ぶことにする。今回はこの音響モデルに少し改良を施したものを用意した。それを New Model と呼ぶことにする。

#### 1.4 New Model の構造

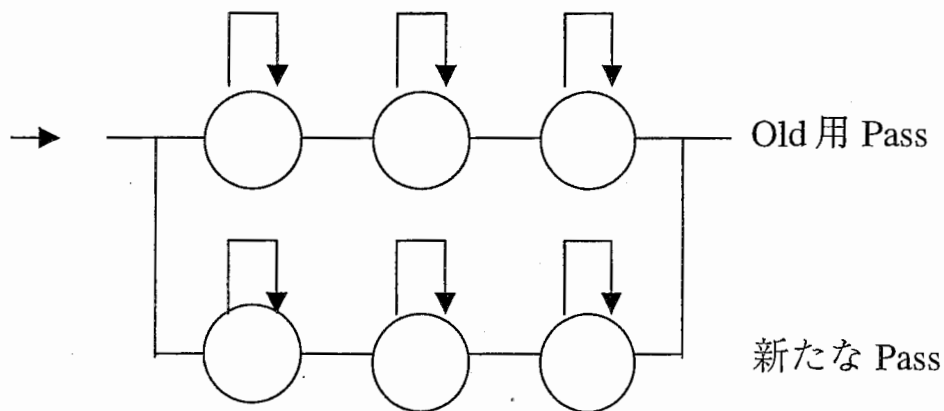


図 5 New Model の構造

Old Model に新たに Pass を追加したものを New Model とする。

そして、音声波形を読み込むときに2つの Pass に同時に読み込ませる。あとの処理(標準パターンとの比較、尤度の決定など)はすべて Old Model と同じである。

### 1.5 ケプストラム (cepstrum) とは？

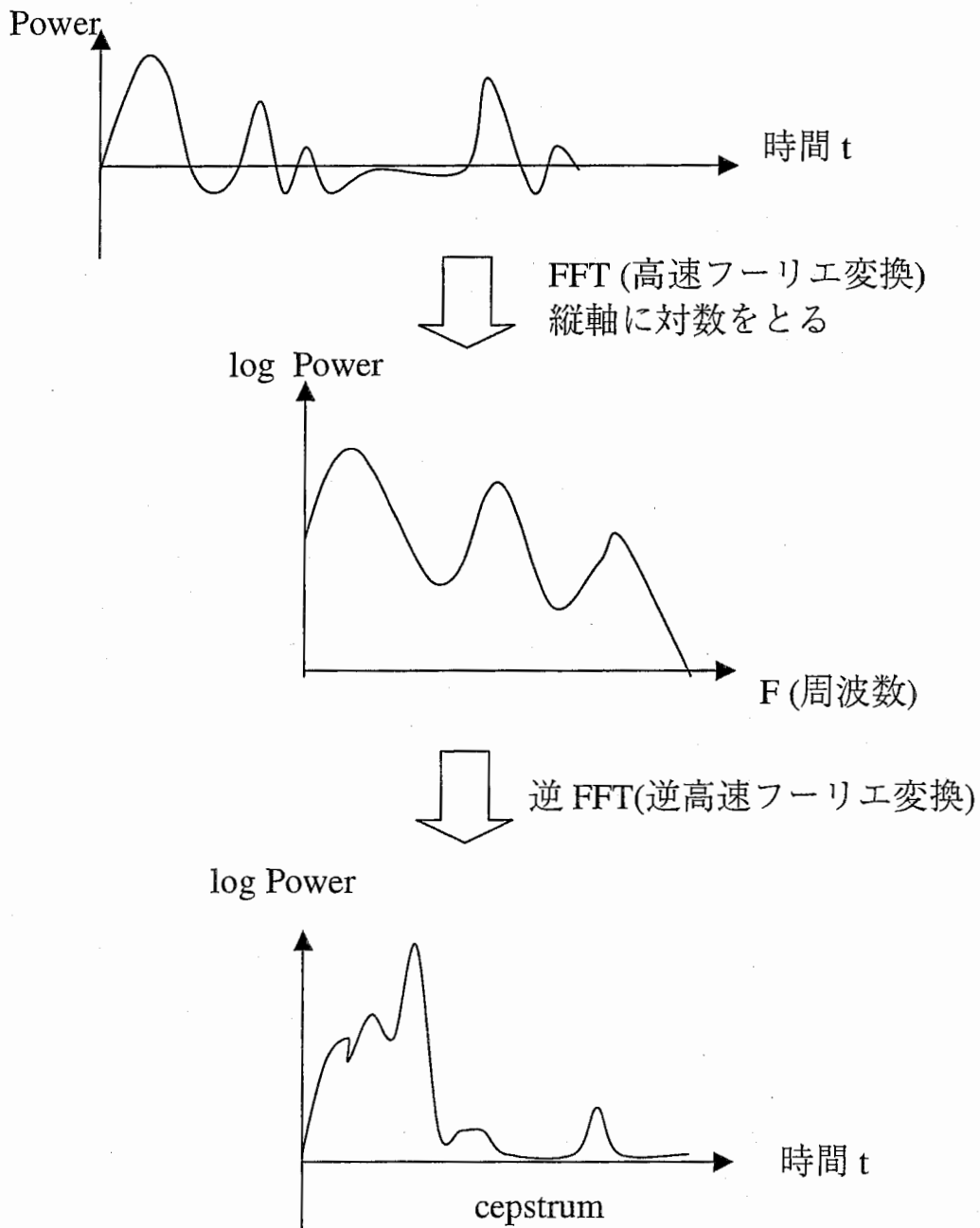


図 6 : ケプストラムと FFT

今回の実験では、ケプストラム 12次元と  $\Delta$ ケプストラム(時間的な変化量)が 12次元、さらに log Power が 1次元の計 25次元である音響モデルを使用した。

#### 1.6 Error の種類

- insertion error : 余計な文字が挿入される
- deletion error : 文字が消える
- substitution error : 文字が別の文字に変化する

#### 1.7 Word Accuracy と Word Correct の計算式

$$\text{Word Accuracy}(\%) = \frac{\text{total} - \text{ins} - \text{del} - \text{sub}}{\text{total}}$$

Word Correct は、Word Accuracy の ins を考慮しないものである。つまり ins = 0 としたものである。

注) total ins del sub は単語の個数

## 2 実験概要

### 2.1 音声収録

まず、被験者の方 8 人(男性 3 人、女性 5 人)にお願いした。そして 8 人を 4 つのグループ(group1~group4)に分割し、それぞれのグループの構成員を 2 人とする。それぞれのグループに共通の 10 文、それぞれのグループ内だけの 10 文、計 20 文を発声内容とし、3 パターンの発声方法で収録を行った。

その発声方法であるが、一つ目は、普通の声、つまりその人が一番ごく自然に発声することのできる喋り方(Normal)。次に、他の人に聞き取ってもらいやすいように、できるだけ明瞭に(Clear)。そして一番極端なのが、一文字一文字を区切って発声(Split)、の 3 パターンである。

収録風景としては、ディスプレイに次々と現れてくる文字を読んでもらう、ごく単純なものである。

### 2.2 実験

上のようにして得られた音声データを Old Model と New Model の 2 つの音響モデルに対して、探索(認識)を行う。

### 2.3 音響モデル学習条件

#### 音響分析

サンプリング周波数 16kHz, preemphaasis 0.98

フレーム周期 10ms, フレーム長 20ms (Hamming 窓)

$\Delta \log power$ , 12 次-MFCC(mel-frequency cepstrum coefficient),

12 次- $\Delta$ MFCC

ケプストラム平均、パワーを正規化

#### HMnet の構成

音声 1400 状態 5 混合、無音 3 状態 10 混合

自然発話学習用音声データセット(旅行対話)

男性：167 対話、総発話時間 約 2 時間

女性：240 対話、総発話時間 約 3 時間

朗読発話学習用音声データセット(音素バランス分)

男性：165 話者、総発話時間 約 9 時間

女性：235 話者、総発話時間 約 14 時間



## 2.4 発声リスト全 50 種類

### 様式

- ① TAC70015.0060.A
- ② 八月十五日の月曜日一泊です
- ③ ハチガツジュウゴニチノゲツヨウビイッパクデス
- ④ h,a,ch,i,g,a,ts,u,zh,j,u,u,g,o,n,i,ch,i,n,o,g,e,ts,u,j,o,o,b,i,i,q,p,a,k,u,d,e,s,u

①整理番号 ②発声内容(日本語) ③発声内容(カタカナ) ④音素列

---

### group 共通 10 文

TAC70015.0060.A  
八月十五日の月曜日一泊です  
ハチガツジュウゴニチノゲツヨウビイッパクデス  
h,a,ch,i,g,a,ts,u,zh,j,u,u,g,o,n,i,ch,i,n,o,g,e,ts,u,j,o,o,b,i,i,q,p,a,k,u,d,e,s,u  
TAC70015.0080.A  
ツインでお願いします  
ツインデオネガイシマス  
ts,u,i,ng,d,e,o,n,e,g,a,i,sh,i,m,a,s,u  
TAC70015.0100.A  
では予約をお願いします  
デハヨヤクヲオネガイシマス  
d,e,w,a,j,o,j,a,k,u,o,o,n,e,g,a,i,sh,i,m,a,s,u  
TAC70015.0120.A  
サービス料はどのようになっておりますか  
サービスリョウハドノヨウニナツテオリマスカ  
s,a,a,b,i,s,u,r,j,o,o,-,w,a,d,o,n,o,j,o,o,n,i,n,a,q,t,e,o,r,i,m,a,s,u,k,a  
TAC70015.0160.A  
分かりました  
ワカリマシタ  
w,a,k,a,r,i,m,a,sh,i,t,a  
TAC70015.0180.A  
ワシントンホリデーインホテルー〇二号室です  
ワシントンホリデーインホテルイチマルニゴウシツデス  
w,a,sh,i,ng,t,o,ng,h,o,r,i,d,e,e,i,ng,h,o,t,e,r,u,i,ch,i,m,a,r,u,n,i,g,o,o,sh,i,ts,u,d,e,s,u  
TAC70015.0200.A  
現金でお願いします  
ゲンキンデオネガイシマス  
g,e,ng,k,i,ng,d,e,o,n,e,g,a,i,sh,i,m,a,s,u  
TAC70015.0220.A  
承知しました  
ショウチシマシタ  
sh,j,o,o,ch,i,sh,i,m,a,sh,i,t,a  
TAC70015.0240.A  
PM六時の予定です  
ピーエムロクジノヨテイデス  
p,i,i,e,m,u,r,o,k,u,zh,i,n,o,j,o,t,e,e,d,e,s,u  
TAC70015.0260.A  
恐れ入りますがチェックアウトは何時ですか  
オソレイリマスガチェックアウトハナンジデスカ  
o,s,o,r,e,i,r,i,m,a,s,u,g,a,ch,e,q,k,u,a,u,t,o,w,a,-,n,a,ng,zh,i,d,e,s,u,k,a

## group1

TAC70015.0280.A

ではよろしく申し上げます

デハヨロシクオネガイシマス

d,e,w,a,j,o,r,o,sh,i,k,u,o,n,e,g,a,i,sh,i,m,a,s,u

TAC70016.0040.A

八月十三日から十五日までお願いしたいんですが

ハチガツジュウサンニチカラジュウゴニチマデオネガイシタイノデスガ

h,a,ch,i,g,a,t,s,u,zh,j,u,u,s,a,ng,n,i,ch,i,k,a,r,a,zh,j,u,u,g,o,n,i,ch,i,m,a,d,e,o,n,e,g,a,i,sh,i,t,a,i,ng,d,e,s,u,g,a

TAC70016.0080.A

シングルでお願いします

シングルデオネガイシマス

sh,i,ng,g,u,r,u,d,e,o,n,e,g,a,i,sh,i,m,a,s,u

TAC70016.0100.A

ダブルの方は空いてますでしょうか

ダブルノハウハアイテムスデショウカ

d,a,b,u,r,u,n,o,h,o,o,w,a,a,i,t,e,m,a,s,u,d,e,sh,j,o,o,k,a

TAC70016.0140.A

はい構いません

ハイカマイマセン

h,a,i,k,a,m,a,i,m,a,s,e,ng

TAC70016.0180.A

ワシントンのホリデーインホテル三〇五号室ですけども

ワシントンノホリデーインホテルサンマルゴゴウツツデスケドモ

w,a,sh,i,ng,t,o,ng,n,o,h,o,r,i,d,e,e,i,ng,h,o,t,e,r,u,-,s,a,ng,m,a,r,u,g,o,g,o,sh,i,t,s,u,d,e,s,u,k,e,d,o,m,o

TAC70021.0100.A

そうですかシングルは空いてませんか

ソウデスカシングルハアイテムセンカ

s,o,o,d,e,s,u,k,a,-sh,i,ng,g,u,r,u,w,a,a,i,t,e,m,a,s,e,ng,k,a

TAC70021.0120.A

ではツインで結構です

デハツインデケッコウデス

d,e,w,a,-,ts,u,i,ng,d,e,k,e,q,k,o.o.d,e,s,u

TAC70021.0140.A

ではそれをお願いいたします

デハソレデオネガイイタシマス

d,e,w,a,s,o,r,e,d,e,o,n,e,g,a,i,i,t,a,sh,i,m,a,s,u

TAC70021.0180.A

現金でお願いいたします

ゲンキンデオネガイイタシマス

g,e,ng,k,i,ng,d,e,-,o,n,e,g,a,i,i,t,a,sh,i,m,a,s,u

## group2

TAC70021.0200.A

六時頃お願いいたします

ロクジコロオネガイイタシマス

r,o,k,u,zh,i,g,o,r,o,o,n,e,g,a,i,i,t,a,sh,i,m,a,s,u

TAC70021.0300.A

今いる場所はワシントンシェラトンホテルです

イマイルバシヨハワシントンシェラトンホテルデス

i,m,a,i,r,u,b,a,sh,j,o,w,a,-,w,a,sh,i,ng,t,o,ng,sh,e,r,a,t,o,ng,h,o,t,e,r,u,-,d,e,s,u

TAC70021.0320.A

では八月十五日よろしくお願いいたします

デハハチガツジュウゴニチヨロシクオネガイイタシマス

d,e,w,a,-,h,a,ch,i,g,a,t,s,u,zh,j,u,u,g,o,n,i,ch,i,-,j,o,r,o,sh,i,k,u,o,n,e,g,a,i,i,t,a,sh,i,m,a,s,u

TAC70023.0320.A

はい分かりました

ハイワカリマシタ

h,a,i,-,w,a,k,a,r,i,m,a,sh,i,t,a

TAC70023.0340.A

よろしく申し上げます

ヨロシクオネガイシマス

j,o,r,o,sh,i,k,u,o,n,e,g,a,i,sh,i,m,a,s,u

TAC70201.0020.A

すみません一日滞在を延長したいんですけども

スミマセンイチニチタイザイラエンチヨウシタイノデスクレドモ  
s,u,m,i,m,a,s,e,ng,i,chi,i,n,i,chi,i,-,t,a,i,z,a,i,o,e,ng,chi,j,o,o,shi,t,a,i,ng,d,e,s,u,k,e,r,e,d,o,m,o  
TAC70201.0040.A  
二〇七号室です  
ニマルナナゴウシツデス  
n,i,m,a,r,u,n,a,n,a,g,o,o,sh,i,ts,u,d,e,s,u  
TAC70201.0160.A  
その場合はお値段はいくらぐらいになりますでしょうか  
ソノバアイハオネダンハイクラクライニナリマスデショウカ  
s,o,n,o,b,a,a,i,w,a,o,n,e,d,ang,w,a,i,k,u,r,a,g,u,r,a,i,n,i,n,a,r,i,m,a,s,u,d,e,sh,j,o,o,k,a  
TAC70201.0180.A  
そちらのお部屋は何階になりますでしょうか  
ソチラノオヘヤハナンカイニナリマスデショウカ  
s,o,chi,i,r,a,n,o,o,h,e,j,a,w,a,n,a,ng,k,a,i,n,i,n,a,r,i,m,a,s,u,d,e,sh,j,o,o,k,a  
TAC70201.0200.A  
それじゃそのようお願いします  
ソレジャソノヨウニオネガイシマス  
s,o,r,e,zh,j,a,s,o,n,o,j,o,o,n,i,o,n,e,g,a,i,sh,i,m,a,s,u

### group3

TAC70202.0040.A  
実は滞在の延長をお願いしたいんですけど  
ジツハタイザイノエンチヨウラオネガイシタイノデスクレドモ  
zh,i,ts,u,w,a,-,t,a,i,z,a,i,n,o,e,ng,chi,j,o,o,o,o,n,e,g,a,i,sh,i,t,a,i,ng,d,e,s,u,k,e,r,e,d,o  
TAC70202.0060.A  
一泊だけでいいんですけど  
イッパクダケデヨイノデスクレドモ  
i,q,p,a,k,u,d,a,k,e,d,e,i,ng,d,e,s,u,k,e,r,e,d,o  
TAC70202.0120.A  
構いません和室でお願いいたします  
カマイマセンワシツデオネガイイタシマス  
k,a,m,a,i,m,a,s,e,ng,-,w,a,sh,i,ts,u,d,e,o,n,e,g,a,i,i,t,a,sh,i,m,a,s,u  
TAC70202.0140.A  
こちらが御無理を言ってるんですから結構ですよ  
コチラガゴムリライッテルノデスクラケッコウデスヨ  
k,o,chi,i,r,a,g,a,g,o,m,u,r,i,o,i,q,t,e,r,u,ng,d,e,s,u,k,a,r,a,k,e,q,k,o,o,d,e,s,u,j,o  
TAC70202.0200.A  
出かける時にフロントに預ければよろしいんでしょうかね  
デカケルトキニフロントニアズケレバヨロシイノデショウカネ  
d,e,k,a,k,e,r,u,t,o,k,i,n,i,h,u,r,o,ng,t,o,n,i,a,z,u,k,e,r,e,b,a,j,o,r,o,sh,i,i,ng,d,e,sh,j,o,o,k,a,n,e  
TAC70301.0120.A  
もしツインのお部屋が空いてない場合はそれでも結構ですよ  
モシツインノオヘヤガイテナイバアイハソレデモケッコウデス  
m,o,sh,i,ts,u,j,ng,n,o,o,h,e,j,a,g,a,a,i,t,e,n,a,i,b,a,a,i,w,a,s,o,r,e,d,e,m,o,k,e,q,k,o,o,d,e,s,u  
TAC70301.0200.A  
はい御迷惑おかけします  
ハイゴメイワクオカケシマス  
h,a,i,-,g,o,m,e,e,w,a,k,u,o,k,a,k,e,sh,i,m,a,s,u  
TAC70303.0260.A  
十七日の十二時までです分かりました  
ジュウシチニチノジュウニジマデスネワカリマシタ  
zh,j,u,u,sh,i,chi,i,n,i,chi,i,n,o,zh,j,u,u,n,i,zh,i,m,a,d,e,d,e,s,u,n,e,-,w,a,k,a,r,i,m,a,sh,i,t,a  
TAC70303.0320.A  
そうですね  
ソウデスネ  
s,o,o,d,e,s,u,n,e  
TAC70304.0160.A  
はいそうして下さい  
ハイソウシテクダサイ  
h,a,i,-,s,o,o,sh,i,t,e,k,u,d,a,s,a,I

## group4

TCC70103.0080.A

今チェックインできませんか

いまチェックインできませんか

i,m,a,c,h,e,q,k,u,i,n,g,d,e,k,i,m,a,s,e,n,g,k,a

TCC70103.0100.A

分かりましたそれでは十二時にチェックインします

ワカリマシタソレデハジュウニジニチェックインシマス

w,a,k,a,r,i,m,a,s,h,i,t,a,-,s,o,r,e,d,e,w,a,zh,j,u,u,n,i,zh,i,n,i,c,h,e,q,k,u,i,n,g,s,h,i,m,a,s,u

TCC70103.0180.A

はいありがとうございますそれではよろしく願いいたします

ハイアリガトウゴザイマスソレデハヨロシクオネガイイタシマス

h,a,i,a,r,i,g,a,t,o,o,g,o,z,a,i,m,a,s,u,s,o,r,e,d,e,w,a,j,o,r,o,s,h,i,k,u,o,n,e,g,a,i,i,t,a,s,h,i,m,a,s,u

TCC70201.0100.A

じゃあまりひどくなると悪いから早くお願いします

ジャアマリヒドクナルトワルイカラハヤクオネガイシマス

zh,j,a,-,a,m,a,r,i,h,i,d,o,k,u,n,a,r,u,t,o,w,a,r,u,i,k,a,r,a,h,a,j,a,k,u,o,n,e,g,a,i,s,h,i,m,a,s,u

TCC70201.0140.A

じゃ大至急お願いします

ジャダイシキユウオネガイシマス

zh,j,a,d,a,i,s,h,i,k,j,u,u,o,n,e,g,a,i,s,h,i,m,a,s,u

TCS70020.0020.A

あしたのタクシーの手配をしてもらいたいのですが

アシタノタクシーノテハイラシテモライタイノデスガ

a,s,h,i,t,a,n,o,t,a,k,u,s,h,i,i,n,o,t,e,h,a,i,o,s,h,i,t,e,m,o,r,a,i,t,a,i,n,o,d,e,s,u,g,a

TCS70020.0080.A

いえ国際線です

イエコクサイセンデス

i,e,-,k,o,k,u,s,a,i,s,e,n,g,d,e,s,u

TCS70070.0100.A

じゃコンチネンタルブレックファーストでお願いします

ジャコンチネンタルブレックファーストデオネガイシマス

zh,j,a,k,o,n,g,c,h,i,n,e,n,g,t,a,r,u,b,u,r,e,q,k,u,f,a,a,s,u,t,o,d,e,o,n,e,g,a,i,s,h,i,m,a,s,u

TCS70070.0120.A

飲み物はジャトマトジュース頼めるかなあとホットコーヒーと

ノミモノハジャトマトジュースタノメルカナアトホットコーヒート

n,o,m,i,m,o,n,o,w,a,zh,j,a,t,o,m,a,t,o,zh,j,u,u,s,u,-,t,a,n,o,m,e,r,u,k,a,n,a,-a,t,o,h,o,q,t,o,k,o,o,h,i,i,t,o

TCS70070.0160.A

じゃよろしく

ジャヨロシク

zh,j,a,j,o,r,o,s,h,i,k,u

### 3 結果(全て New Model の Split のみリストアップ)

様式

@ 51.69 / 69.66 71.91 / 82.02 89 16 1 26 12.97 15/20  
① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩

- ① Word Accuracy(%) ② Word Correct(%)
- ③ 認識時において、正解だと思われる数個の候補が蓄積され、その中に正解が入っている度合いを示す数値  
(今回の実験では 20 個の候補を用いた)
- ④ ③から ins error を無視したもの
- ⑤ total の単語数 ⑥ ins error ⑦ del error ⑧ sub error
- ⑨ 1 秒あたりの音声を認識するのにかかった時間(sec)の割合
- ⑩ 読み込みに成功した文の数 / すべての文の数

#### /group1/speaker1/split

TAC70015.0080.A ツイン+で+お+願い+します  
TAC70015.0100.A では+予約を+お- 例外+します  
TAC70015.0160.A わかり+まし+た  
TAC70015.0180.A ワシントン工程+で行く+のホテルにイチマルで+一等 C S+です  
TAC70015.0200.A で+土+日+で+お+願い+します  
TAC70015.0220.A 承知+し+まし+た  
TAC70015.0240.A いいえ+え- ふろ口+の+予定+です  
TAC70015.0260.A 高層+と+言い+ます+が+チェックアウト+は+何&日+です+か  
TAC70015.0280.A では+よろしく+お+願い+します  
TAC70016.0080.A つい+にくると+お+願い+します  
TAC70016.0100.A ダブル+の+方+の+ばあい+です+で+しょう+か  
TAC70016.0140.A あまり+変わります+ね  
TAC70021.0120.A では+ツイン+で+え- 五+です  
TAC70021.0140.A では+それ+で+コーディネーター+に+いたし+ます  
TAC70021.0180.A では+ん- 自分+で+お+願い+いたし+ます  
@ 51.69 / 69.66 71.91 / 82.02 89 16 1 26 12.97 15/20

#### /group1/speaker2/split

TAC70015.0060.A あっち+行+か+た+す+う+十&五&日+の+え- 通用帯+に+バッグ+です  
TAC70015.0080.A ずいぶん+で+お+の+です+が+有り+します  
TAC70015.0100.A では+よ- 約五号+ので+絵+が+有り+します  
TAC70015.0120.A サービス料+五+は+どう+の+お+用具+に+な+つ+て+あり+ます+が  
TAC70015.0160.A では+帰+り+まし+た  
TAC70015.0180.A 和紙+の+日本+を+降+り+て+え- 一品+コップ類+って+いま+在+る+海+を+お+質+です  
TAC70015.0200.A 電気+言+う+と+思+う+の+で+は+い+し+ます  
TAC70015.0220.A 承知+し+まし+た  
TAC70015.0240.A いいえ+ブロック+安静+の+様+で+よい+です  
TAC70015.0260.A 高層+零+人+い+ます+が+チェックアウト+は+バン+氏+です+か  
TAC70015.0280.A で+一+和洋+ロー C U O+の+で+第一+し+ます  
TAC70016.0040.A あ+一&日+う+一十&三&日+から+十&五&日+まで+を+お+願+い+い  
TAC70016.0080.A C+か+くると+負+おう+の+で+か+に+し+ます  
TAC70016.0100.A ダブル+の+ご+会話+利用+税+札+で+え+正+午+か  
TAC70016.0140.A ない+まま+有+り+ませ+ん

TAC70016.0180.A 和紙イノモトあのコーリーで+えいまホテル差の+有る向こう五号+室です+けれども  
TAC70021.0100.A そう+です+か+SENT-START-END シーズン プールは+空いてません+か  
TAC70021.0120.A では+人+で+けっこう+です  
TAC70021.0140.A ではそう例でODがEEいたします  
TAC70021.0180.A で-なん+人で負おうので他にいたします  
@ 8.45 / 50.00 41.55 / 64.79 142 59 4 67 5.84 20/20

### /group2/speaker1/split

TAC70015.0060.A 八&月十&五&日の+月曜日-+泊です  
TAC70015.0080.A スキーで+お+願い+します  
TAC70015.0100.A で-和洋予約+を+お+願い+します  
TAC70015.0120.A サービス料は+どのようなディナーで+お降ります+か  
TAC70015.0160.A わかり+ました  
TAC70015.0200.A で+え-駅に+えで+お-え-お+願い+します  
TAC70015.0220.A S O O C がすいた  
TAC70015.0240.A T A N 六&時の+予定+です  
TAC70021.0200.A 六&時+頃+お+願い+いたします  
TAC70021.0320.A で-バーは四&月十&五&日よろしく+お+願い+いたします  
TAC70023.0320.A ワインわかり+ますに+行つた  
TAC70023.0340.A よろしく+お+願い+します  
TAC70201.0040.A 二+〇+七号+室+です  
TAC70201.0200.A ソウルで+え-じゃ騒音の+よう+にオーケーがいそいます  
@ 56.82 / 77.27 76.14 / 84.09 88 18 0 20 12.29 14/20

### /group2/speaker2/split

TAC70015.0060.A 八&月十&五&日の+十通曜日-+泊です  
TAC70015.0080.A ツイン+で+青を+お+願い+します  
TAC70015.0100.A では用意が国を+お-お+たり+します  
TAC70015.0120.A それ+は+良い州庁舎を同様に+なつて+お+ります+か  
TAC70015.0160.A わかり+ました  
TAC70015.0180.A ワシントン+を+降り+て+英文方+つていう-二+〇+二+号+室+です  
TAC70015.0200.A 現金+で+お-お+願い+します  
TAC70015.0220.A 承知+し+ました  
TAC70015.0240.A いいえ+ふる+コウジの+予定+です  
TAC70015.0260.A おお+そう+で+EBバス+が+チェックアウトは+何&時+です+か  
TAC70021.0200.A 同行+時+頃+を+お+願い+いたします  
TAC70021.0300.A イマイル+場所+は+ワシントン+制度+後+の+方+で+手+です  
TAC70021.0320.A では八&月十&五&日+洋式+空港+で+依頼+いたします  
TAC70023.0320.A はい+い+わかり+ました  
TAC70023.0340.A よ-ロス+胃痛+と+お-で+依頼+します  
TAC70201.0020.A すみません+SENT-START-END-+日以内+に+ツアー+以後+延長+し+たいの+です+けれども+え-ど  
うも  
TAC70201.0040.A 二+〇+七号+室+です  
TAC70201.0160.A そうの+ばあい+バスを+お+礼+だら+ま胃痛+が+有る+第+二+人+なり+ます+で+しょう+か  
TAC70201.0180.A 装置+だ+と+お+部屋+は+なに+変わ+り+に+なり+ます+けれども+ショー+を+売+つた  
TAC70201.0200.A そう+で+え+じゃ+そう+同様に+を+お+願い+します  
@ 38.62 / 72.41 59.31 / 77.24 145 49 0 40 16.86 20/20

### /group3/speaker1/split

TAC70015.0060.A はい+四&月+十五+五+二+の+合図+様+O D E R+です  
TAC70015.0080.A 水温計+お-を+絶対+バ+ンツ  
TAC70015.0100.A デンバー+よ-い+や-空港+を+お-で-じゃ+おいしい+いま+は+する+ので  
TAC70015.0120.A サービス+う+どん+はい+どうぞ+お-よう+に+なると+お-降り+います+か  
TAC70015.0160.A ワイバー+二+万+四+一+七  
TAC70015.0180.A ワシントン+と+なん+か+コー+ヒー+で+え-ディナー+の+方+で+でる+市+が+有る+か+二+号+室+です  
TAC70015.0200.A ゲーム+内+ので+多+で+ない+しい+いま+は+する+の  
TAC70015.0220.A 明日+の+大き+い+しい+いま+八+時間  
TAC70015.0240.A いいえ+え-もう+動く+内+の+O+か+予定+です+が+う-ん  
TAC70015.0260.A コース+を+お+礼+言+います+が+あ-チェック+は+ご+都合+は+なん+便+です+か  
TAC70202.0060.A はい+あ-コウダケイ+D+A D+二+で+S K+で+どう+ので  
TAC70202.0120.A 買+う+ば+あ+いに+バス+デー+も+和室+上+で+五+で+第+いい+や+し-いま+すぐ  
TAC70202.0140.A コー+チ+や+ま+あ+五+五+え+T+O Y+で+え+十+二+です+から+けっこう+です+よ+SENT-START-END+ん-

TAC70202.0200.A レイカーズでルートを+お+聞きに+もグローブの康に+は ツーデイでバーに+ご両親いま+の衣裳を+おを浮かれ+た  
TAC70301.0120.A もう惜しそうに胃炎の+方+へえーあーまああ+いいえ ないばあいバスを+入れでもけっこう+です  
TAC70301.0200.A 繁華街 要望 姪いま 空港は K C ます  
TAC70303.0320.A そう気+にケースで+え  
TAC70304.0160.A ハワイでショーを豪勢で+えすぐあー際+に  
@ -38.52 / 24.44 12.59 / 41.48 135 85 1 101 9.64 18/20

### /group3/speaker2/split

TAC70015.0060.A 八&月+十&二&日と必要具合一+泊です  
TAC70015.0080.A すー インディアンを+おーで借りすれ+ば明日  
TAC70015.0100.A でー 和洋いー額を大勢で+は 良い市バス  
TAC70015.0160.A わかり+ました  
TAC70015.0180.A 和紙のトム コーリーでと+言う行程でうー+〇二+五号+室+です  
TAC70015.0200.A J R+の二+二で五で依頼+します  
TAC70015.0220.A ショウ M C C バーすら  
TAC70015.0240.A キーケーブルと+おー九&時の+ご用で E です  
TAC70202.0040.A じつはワインさえ移動は日にちおー五+五でえ+では E C I E の+です+けれども  
TAC70202.0060.A キーは下ってで+えーいの+です+けれども  
TAC70202.0120.A あまり男性で和室+で手を+お+願い+いた+します  
TAC70202.0140.A こちら+が後部に合計ていうの+です+がラケットを+おーです+よ  
TAC70301.0120.A もしスリー分の+おーお+部屋が+空いて+ないばあい+に+はそう例で+もけっこう+です  
TAC70301.0200.A はい英語+で祝うお+掛け+します  
TAC70303.0260.A ジュース せいぜい二+千+円の+画集を意味するまでです+のでわかり+ました  
TAC70303.0320.A 総合+です+ので  
TAC70304.0160.A はい+SENT-START-END いいえ 総合し+てくださいいま  
@ 21.05 / 51.88 45.11 / 63.91 133 41 1 63 21.87 17/20

### /group4/speaker1/split

TAC70015.0060.A 八&月 十&五&日の+えー 通用を定員は区+です  
TAC70015.0080.A そういうでこうね側+に+します  
TAC70015.0100.A D は+予約を+おー ねえ 変わり+します  
TAC70015.0120.A 沢ドイツ料理はどう思うよう+にディナーで+お降ります+か  
TAC70015.0160.A わくわく+しました  
TAC70015.0180.A わあ C デートを+えーここを+降り+で有効で L E Z うまく日光を+し+すです  
TAC70015.0200.A D+A K+A D O D で+はに+します  
TAC70015.0220.A 承知+しました  
TAC70015.0240.A はい+えーもう六&時の+え O K+A です  
TAC70015.0260.A もう遅でいて+います+がチェックアウトは+何&日+です+か  
TCC70103.0080.A いいまチェックインえ定期いま製+です+か  
TCC70103.0100.A わかり+ました+SENT-START-END そう 零+零は十&二&時に+チェックイン+します  
TCC70103.0180.A あー イースターより+はとう ございます SENT-START-END+それでは用下ろし 空港をねが+有り+いた+します  
TCC70201.0100.A ではあまり移動区がルートは+どう行っからはい+やー 空港+で絵が+有り+します  
TCC70201.0140.A ジャグリ C Q O D Y +します  
TCS70020.0020.A あ明日+のタクシー+のセーター様を+してもらい+たい+の+です+が  
TCS70020.0080.A それでこうコース+は二+千+円別送  
TCS70070.0100.A ジャ五&日+で L R それと和室 お+部屋を+お+願い+します  
TCS70070.0120.A 飲物を和室+がトマトジュースか飲めるか+な SENT-START-END+後 ホットコーヒーと  
TCS70070.0160.A ではよーローズ地区で  
@ 2.86 / 56.43 42.14 / 69.29 140 75 1 60 8.02 20/20

### /group4/speaker2/split

TAC70015.0060.A 八&月 十&五&日の+月曜日一+泊+です  
TAC70015.0080.A 発ウイーンで+お+願い+いた+します  
TAC70015.0100.A では料理屋区おーお+願い+します  
TAC70015.0120.A サービス料は+どのよう+になっており+ます+が  
TAC70015.0160.A わかり+ました  
TAC70015.0180.A ワシントンのホリデイインホテルー五+〇二+号+室+です  
TAC70015.0200.A 現金+で+お+願い+します  
TAC70015.0220.A 承知+しました

TAC70015.0240.A B A M 六&時の+予定+です  
TAC70015.0260.A あお恐れいります+がチェックアウトは+何&時+です+か  
TCC70103.0080.A いまチェックインでき+ません+か  
TCC70103.0100.A わかり+まし+た+SENT-START-END+それでは+十&二&時に+チェックイン+します  
TCC70103.0180.A はい+ありがとうございます+SENT-START-END+それではよろしく+お+願ひ+いたし+ます  
TCC70201.0100.A じゃあまりひどくなると悪いから早くを+お+願ひ+します  
TCC70201.0140.A じゃワインで+は E C U O で借り+します  
TCS70020.0080.A いいえここ山陰線にです  
TCS70070.0100.A じゃ五&日イベント有るブレイクのストで+お+願ひ+します  
TCS70070.0160.A じゃ+よろしく  
@ 71.19 / 89.83 81.36 / 92.37 118 22 0 12 17.82 18/20

			OLD MODEL	NEW MODEL
group1	A	Nor	95.77	95.77
		Cle	95.07	95.05
		Spl	-50	51.69
	B	Nor	97.18	97.18
		Cle	93.66	97.18
		Spl	-16.9	8.45
group2	C	Nor	93.1	92.41
		Cle	95.86	93.1
		Spl	-80.69	56.82
	D	Nor	93.1	92.41
		Cle	95.17	95.17
		Spl	13.1	38.62
group3	E	Nor	88.75	88.12
		Cle	87.5	85
		Spl	-60	-38.52
	F	Nor	88.13	86.25
		Cle	94.38	94.38
		Spl	-15.63	21.05
group4	G	Nor	95.71	96.43
		Cle	93.57	92.86
		Spl	-38.57	2.86
	H	Nor	90	89.29
		Cle	94.29	95
		Spl	38.57	71.19

図7 全実験認識結果

注) 数字はすべて Word Accuracy(%)

図7をグラフ化した図8、図9は最終ページへ



## 4 考察

無作為に結果を選んでみた。(すべて new model の split)

1. 「では予約をお願いします」
  - 「でー和洋予約をお願いします」 (group2,C)
  - 「では予約をおー例外します」 (g1,A)
  - 「Dは予約をおーねえ変わりします」 (g4,G)
  - 「でー和洋いー額を大勢では良い市バス」 (g3,F)
2. 「PM6時の予定です」
  - 「いいえブロック安静の様でよいです」 (g1,B)
  - 「いいええーふろ口の予定です」 (g1,A)
  - 「TAN六時の予定です」 (g2,C)
  - 「いいえふろコウジの予定です」 (g2,D)

以下は、図7のNew ModelのSplitに焦点を絞って考察を行う。  
誤りの傾向として多いものは、

1つ目として、子音(かあ、きい、くう、けえ、こお)をのぼして発声したことによって、子音の前の部分(か、き、く、け、こ)や、後の部分の母音(あ、い、う、え、お)までも新たな単語として認識されているというミス。

2つ目として、どこからどこまでをその単語とするのかという区切りのミスである。

前者は、ある意味では「か」を「か、あ」の二つのものとして区切っているといえるので、結局、どこからどこまでが単語なのかという見切りが悪いということがいえ、両者とも区切りに関係してることがわかる。だが、両者は少し意味が違う。完全に同じ種類の誤りとは言い切れない。

このような現象がおこる原因の一つとして、直感的には、人それぞれ喋る速さが違っているからと考えられる。ひとつの区切る文字の発声に時間をかけた喋り方をするとこの種の誤りがよく起こると考えられる。

問題解決方法を考えるにあたって、大まかに二通りのアプローチができる。

1. ユーザ負担
  2. ソフトウェア負担
- の二つである。

1について、今回行ったすべての実験では音響モデル中のトレーニングデータ(多数の人の音声波形を集めて構築された確率的な分布のモデル)は同一のものをを用いて行った。つまり New Model を使用しての実験であるにもかかわらず、Model 中で使われているトレーニングデータは Old Model のものをを用いているのである。この理由は実習期間が短かったため、独自に音響モデルのトレーニングデータを構築していくことが不可能だったというのが最大の要因である。

よって、Split 専用の音響モデルのトレーニングデータを構築していけば単純に認識率は上昇することが期待できる。

しかし、問題は音響モデル構築時においてデータ取りをお願いする話者に負担がかかるということである。だが、これはどのような音響モデル構築時においても言えることなのでここではこれ以上言及しない。

次に2について述べる。

先ほど挙げた誤りの傾向について、どこまでが単語なのかという単語間の結びつきの力(この単語の後には確率的にこの単語がよくつながってくるなど)を強化する必要がある。これを強化するということは、つまり、言語モデルの強化と等価である。すると単にサンプル数(標準パターン)をふやせば認識率は上昇は期待できるが、ある程度のジャンル(たとえば科学、芸能、文学など)ごとに分けて言語モデルを構築していけば言語モデルの肥大化を少しはとどめられるのではないか。そしてそのジャンルを認識実行時に音響モデルにさらにプラスして、extra、option で選択できるようにしておけば少しは効果があると考えることができる。

この方法における問題として、ジャンル分けの定義をどうするかというのが挙げられる。きれいにすべての単語をそれぞれのジャンルに分けるのは難しい(専門語、固有名詞などならよいが)。そのようなときには単純にその単語を複数のジャンルに登録させるとよいかもしれない。以上。

## 5 感想

3週間という短い間であったが、実際に自分が8人の被験者をお願いして音声データを収録した。データ取りの時間が大半を占めていたように思われた。しかし、データがなくては実験開始すらできないので、データ取りの重要性を知るにはいい体験ができた。

データの数は、20文×3変動×8人=480文(50種類)である。それらを Acoustic Model と New Model の2種類の音響モデルで認識実験を行うのだが、本人の技術力(Linux)が乏しく Directory などの整理に時間がかかり、あまりスピーディーに行えず、少し苦痛を感じた。だが、苦勞した分だけ達成感もあった。

今回行った実験は比較的研究が進んでいない分野であり、この実験結果が少しでも何らかの足がかりになれば光栄である。

### 発声変動を含んだ音声認識実験の結果

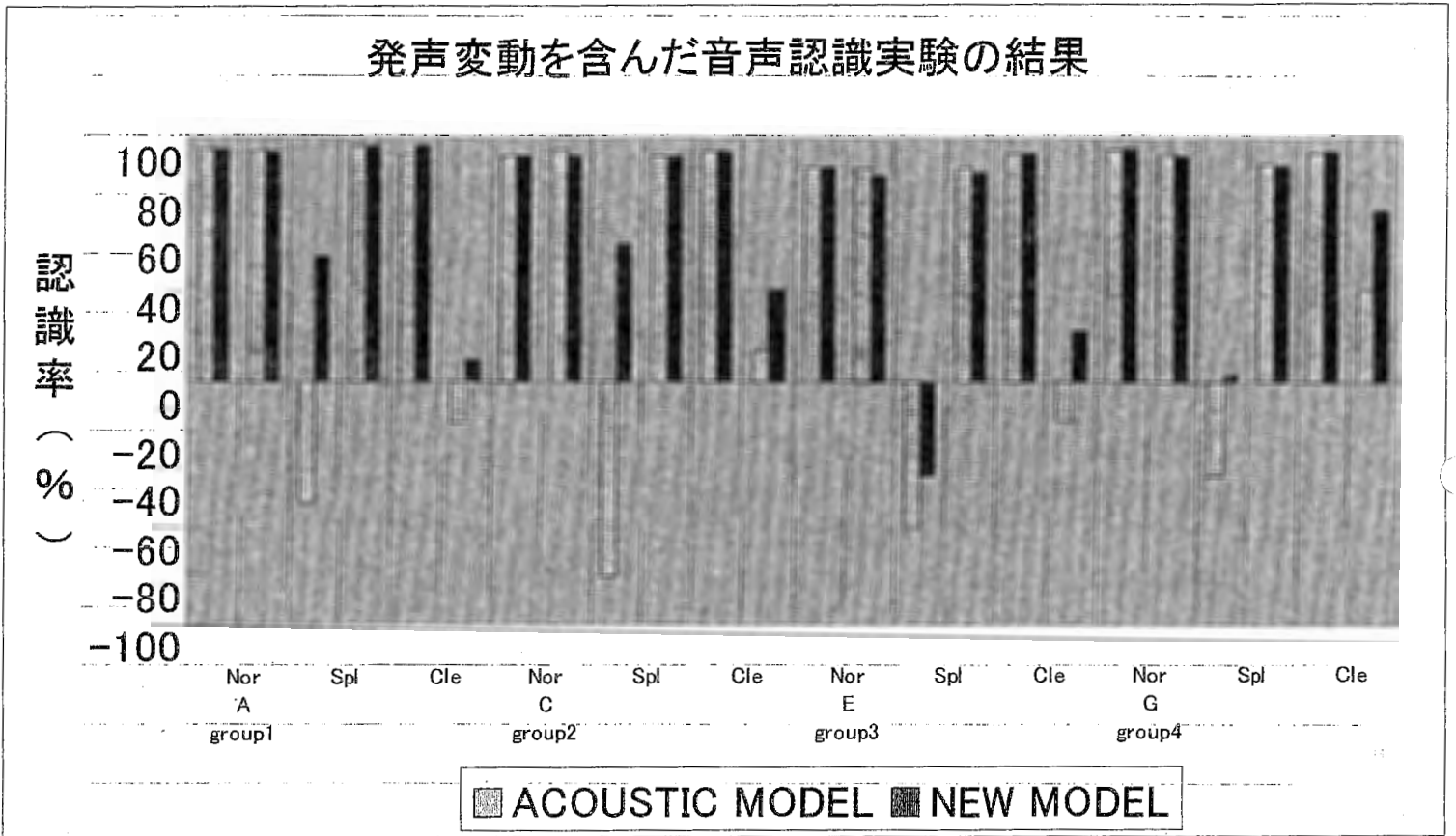


図 8

### 発声変動を含んだ認識実験の結果(splitのみ)

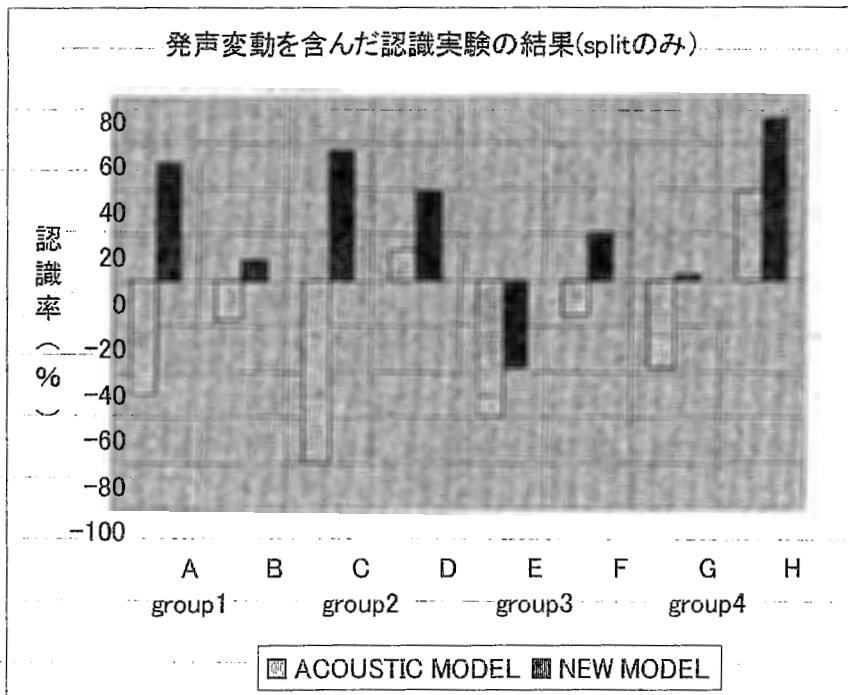


図 9