Internal Use Only (非公開)

002

TR-IT-0265

音声翻訳システムのための日本語音声認識言語モデル ('98 年 6 月版)

Japanese Language Models for Continuous Speech Recognition in Speech Translation System (June 1998 release)

塚田 元
Hajime Tsukada

中嶋 秀治
Hideharu Nakajima

伴 敏雄*
Toshio Ban

山本 博史
Hirofumi Yamamoto

1998.6.22

TDMT 形態素に基づく日本語言語モデルについて解説し，それを用いた連続音声認識実験について報告する．現在，音声翻訳通信研究所では，日英音声翻訳システム (日英 MATRIX) が稼働をはじめており，音声翻訳システム全体としての評価を進めつつある．また，音声認識部と言語翻訳部をより密接に結合することで，音声翻訳性能を高める研究も盛んになってきている．これらの要求に答えるため，TDMT 品詞体系に基づく言語モデルを作成・公開する．本モデルにより，はじめて本格的な音声翻訳実験が可能となる．本稿では，目的別に (1) 研究用言語モデル，および (2)MATRIX 用言語モデルの二種類を作成し，従来のモデルと認識率の観点から比較実験を行った．その結果，どちらのモデルとも，従来のモデルと同等の性能が得られることを確認した．

©ＡＴＲ音声翻訳通信研究所

©ATR Interpreting Telecommunications Research Laboratories

*エスイーティ株式会社

もくじ

1 はじめに 1

2 テストセットと学習セット 3

2.1 テストセット 3

2.2 学習セット 4

3 単語認識の評価方法 6

3.1 正解のゆれ 6

3.2 間投詞の扱い 9

3.3 音声データと言語データの不整合 10

(3.3.1) 間投詞挿入の問題点 10

(3.3.2) 言い淀み，言い直しの問題点 11

4 研究用言語モデル 12

4.1 言語モデルの作成 12

4.2 認識実験 12

(4.2.1) サーチのパラメータ選択 13

(4.2.2) 従来のモデルとの比較 15

4.3 今後の比較実験のための設定 16

5 MATRIX 用言語モデル 18

5.1 言語モデルの作成 18

5.2 認識実験 19

(5.2.1) サーチのパラメータ選択 19

(5.2.2) 従来のモデルとの比較 20

6 むすび 22

謝辞 23

参考文献 24

付録 A 言語モデルの在処 25

A.1 研究用言語モデル 25

A.2 MATRIX 用言語モデル 25

付録 B SDB テストセット 26

付録 C SLDB テストセット 27

付録 D MDB1 テストセット 28

付録 E 研究用言語モデルの認識実験 30

E.1 今後の比較用の認識結果 30

E.2 今後の比較用の設定ファイル 32

付録 F MATRIX 用言語モデルの認識実験 34

F.1 最適パラメータの認識結果 34

F.2 設定ファイル 37

1 はじめに

音声翻訳通信研究所では，日本語音声認識，日英翻訳，英語音声合成，そしてそれらを統合する会話管理からなる日英音声翻訳システム (日英 MATRIX) が稼働をはじめている．現在，第四研究室が中心となり，本格的にシステム全体としての評価も始められようとしている．また，第一研究室や第三研究室においては，単語グラフ形式を用いて，より密接に音声認識と言語翻訳をつなぐことを目指した研究も盛んになってきている．

日英 MATRIX では，音声認識部と言語翻訳部の間のインタフェースとして，翻訳に適した形態素体系 (TDMT 形態素体系) が用いられている．TDMT 形態素を出力するためには，音声認識の言語モデルとしても，この体系を採用する必要がある．しかし，TDMT 形態素に基づく言語モデルは，これまで語彙の限定されたもの [2][5] しか存在しておらず，多様な発話の音声翻訳実験ができないという問題があった．

この音声翻訳実験の要求に答えるため，多様な表現をカバーする言語モデルを目的別に二種類作成する (表 1)．一つは，自然発話音声データベース [7][8][9] を用いた音声認識 / 音声翻訳実験に用いるもので，**研究用言語モデル**と呼ぶことにする．研究用言語モデルは，旅行会話を認識対象とし，タスクに依存した先見的な知識を用いず，主にデータベースの統計情報に基づいて作成する．もう一つは，実時間で動作する音声翻訳システムで用いるもので，**MATRIX 用言語モデル**と呼ぶ．MATRIX 用言語モデルは，旅行会話の部分タスクであるホテル予約会話を認識対象とし，タスクに依存した知識を導入することで，タスク達成上重要な表現が偏り無く認識できることを目指している．両方の言語モデルとも，可変長 N-gram[3] として構成する．

まず，2節では，認識実験に用いたテストセットと学習セットについて解説する．次に 3節では，本稿の言語モデルを用いて得られる単語認識結果を評価する上で問題となる点を解説し，それらを解決するための評価方法について述べる．続いて，4節では，研究用言語モデルについて説明する．ここで作成する言語モデルは，TDMT 形態素に基づくが，これは従来の SLDB 形態素に基づく研究用言語モデル [1] と比べて，単語認識率で同等，音素認識率で若干上回る性能が得られた．最後の 5節では，MATRIX 用言語モデルについて説明する．今回作成する MATRIX 用言語モデルは，従来の MATRIX 用言語モデル [2][5] と比べて，より多様な表現を認識できるように，語彙も約 1.7 倍に増強する．この語彙の増加にもかかわらず，約 3 倍に増えた学習データのおかげで，デモ会話テストセットに対して，従来と同等の認識率が得られる．

	認識対象	タスク依存知識	N-gram の種類
研究用言語モデル	旅行会話	なし	可変長 N-gram
MATRIX 用言語モデル	ホテル予約会話	あり	可変長 N-gram

表 1: 二種類の言語モデル

2 テストセットと学習セット

2.1 テストセット

音声認識，音声翻訳のテストセットは，様々な観点で対象とするタスクの分布を表現していることが必要である．音響モデルの観点からは，音素環境や話者等に関してバランス良く分布していることが必要である．一方，翻訳を含む言語モデルの観点からは，申込者 / 担当者といった会話の役割がバランスしており，様々な言い回しが現れることが必要である．読み上げ音声データベースの場合，音響モデルと言語モデルの両方の要求を満たすように設計することは比較的容易である．しかし，SDB[8][9] や SLDB[7] のような模擬会話では，両者の要求を満たすことは著しく困難である．

そこで，研究用言語モデルを用いた認識実験 (4節) では，音響モデル側の要求を重視したテストセット (S1,S2,S4) および言語モデル側の要求を重視したテストセット (SLTA1) を併用する．どちらも人間どうしの模擬会話である．ただし，様々なパラメータの設定は，これまでの認識実験 [1] にならって，S1,S2 セットを用いることにする．

一方，MATRIX 用言語モデルの認識実験 (5節) では，模擬会話ではなく，デモで用いるような限定会話文を読み上げたテストセット (MDB1 の IS, DS, WS セット) を用いる．MDB1 は，言語モデル的には，デモで用いるような極端に限られた表現しか含まないものの，その範囲で音響と言語両モデルの要求を満たすよう設計されている．

[S1,S2,S4]

SDB の旅行会話のテストセット．特に S1, S2 は，ホテルの部屋の予約会話からなる[1]．S1,S2,S4 を合わせると，SDB の旅行会話で使われているプロットの分布が，大まかにではあるが，反映された会話セットとなっている．ただし，このテストセットの問題点として，発話が申込者側に著しく偏っていることがある．これは，かつて申込者側のみを認識対象としていたという経緯によるものである．

[SLTA1]

音声翻訳システムのオープンテスト用のホテル予約会話．SLDB 中のホテルの部屋を予約する会話 (TA*) から，次の基準で選んだ 23 片側会話からなる[2]．

- 最低限の音声の品質を確保
- 日英 / 英日がバランスしている

[1] S1 を言語モデル open, S2 を言語モデル closed として認識実験に用いられてきたという経緯がある

[2] 当初，SLTA1 は 24 片側会話からなっていたが，closed テスト用の一会話 (TAS12010.A) が含まれていたため，それを除外して，23 片側会話となった

- プロットが，SLDB の残りのホテル予約会話と大まかにバランスしている．
- TDMT の学習等に使われていない．(今後も使われない)
- 第四研究室で進められている音声認識用の文法の開発に使われていない．

SLTA1 には，ほどよく男女の話者が混じっている (男性 8 名，女性 16 名) ものの，同じ話者が何度も現れているため，音響モデルの観点からは妥当なテストセットとはいえない．そこで，現在，SLTA1 の内容を複数の話者により読み上げたデータベースを構築中である．発話様式は読み上げになるが，言語モデルと音響モデルの両方の要求を満たすものとして期待される．

[MDB1: IS, DS, WS]

日英 MATRIX デモのための日本語認識部テストセット．同一会話文を，複数の話者に読み上げてもらったもの．デモで用いる極端に限られた表現しか含まない．「金額表現」，「スペル」，カード番号などの「連続数字」などは含まれていないため，これらの表現に対する評価ができない．

IS(33 文): SDB や SLDB 中の重要単語を含む文を抜き出し，若干変更を加えたもの．

DS(19 文): '97 年 ATR 研究発表会のデモ用標準会話．

WS(21 文): アンケートで認識されにくいと言われた単語を含む文．

2.2 学習セット

研究用言語モデルの学習セットは，1998 年 5 月 7 日現在，収集と整備が完了している旅行会話 (SLDB, SDB: TRA1, TRA2, TRA3) から作成した．ただし，テストセット (S1, S2, S4, SLTA1) の片側会話を除くだけでなく，TAP␣␣␣␣␣ および S␣␣␣␣␣␣␣ についても除いてある．基本的に，テストセットの相手パートの片側会話は，学習セットに含むこととした．

一方，MATRIX 用言語モデルの学習セットは，研究用言語モデルの学習セットをホテル予約会話 (TA␣␣␣␣␣) に限定して作成した．これは，事前検討により，旅行会話全体を用いるより，データ量は減るものの，ホテル予約会話に限定した方が，良好な結果が得られることを確認したためである．

表 2に，テストセットと学習セットのデータ量を示す．表の単語数は，句点および括弧等の記号を含んでいないが，発声開始や終了を表す特殊単語を含んでいる．研究用言語モデルは，ReseachJ V2 および V3 で使われた SLDB 形態素言語モデル [1] と比べると，学習セットは一

致はしていないものの，ほぼ等しい．学習データ量は，片側会話数で 2% ほど減少している[3]が，これは主にTAP␣␣␣␣␣を除いたことによる．一方，MATRIX 用言語モデルでは，従来のモデル [2][5] と比べると，学習データは約 3 倍に増えている．

セット	片側会話数	のべ単語数	異なり単語数
SDB テストセット (S1,S2,S4)	42	6,342	843
SLDB テストセット (SLTA1)	23	4,665	591
MDB1 テストセット	-	30,255	239
研究用言語モデル学習セット	2,473	459,383	7,302
MATRIX 用言語モデル学習セット	683	140,899	2,434

表 2: テストセットと学習セット

[3]のべ単語数でみると，19% も減少しているが，TDMT 形態素の方が単位が長いため，のべ単語数での比較はその点を考慮する必要がある．

3 単語認識の評価方法

3.1 正解のゆれ

単語認識結果の正解ファイル (ANS ファイル) は，形態素ファイル (JMOR ファイル) から，「記号」を削除して自動的につくられる．単語は，七つ組 (表記形，振り仮名，標準形，品詞，活用型，活用形，音便またはコメント) からなり，マスター辞書によって，この七つ組に対して単語 ID が定義されている．ANS ファイルは，この単語 ID の列として記述される．図 1〜3に，これらのファイルの例を示す．

```
20|0020|30|70| もしもし | モシモシ | もしもし | 感動詞 ||||
20|0020|40|80| あたくし | アタクシ | あたくし | 代名詞 ||||
20|0020|50|90| 青山 | アオヤマ | 青山 | 普通名詞 ||||
20|0020|50|100| と | ト | と | 格助詞 ||||
20|0020|60|110| 申し | モウシ | 申す | 本動詞 | 五段サ | 連用 ||
20|0020|60|120| ます | マス | ます | 助動詞 | 特殊サ | 基本 ||
20|0020|60|130| が | ガ | が | 接続助詞 ||||
```

図 1: JMOR ファイル

```
20332    青山 | アオヤマ | 青山 | 普通名詞 ||||
10768    あたくし | アタクシ | あたくし | 代名詞 ||||
10327    申し | モウシ | 申す | 本動詞 | 五段サ | 連用 ||
10011    ます | マス | ます | 助動詞 | 特殊サ | 基本 ||
10009    と | ト | と | 格助詞 ||||
10005    もしもし | モシモシ | もしもし | 感動詞 ||||
5        UTT-START        #
6        UTT-END          #
```

図 2: マスター辞書

このように，書き起こしテキストをもとに単語を定義しているため，同一と考えられる単語に対して，表記のゆれによって複数の単語 ID が割り振られ (図 4)，正解がゆれてしまうとい

```
5
10005
10768
20332
10009
10327
10011
10012
6
```

図 3: ANS ファイル

う問題がある．したがって，音声認識としては正しく認識されたものが，そのような正解のゆれのため，不正解とみなされることがある．このように，正解のゆれを解消することは，厳密な評価をする上で重要であるが，ゆれの頻度はそれほど大きくないために，同一条件で認識率の相対的な比較をする分には，ほとんど支障のないことも多い．そこで，以降の認識実験では，正解のゆれを極力解消して評価する方法とそうでない方法の両方を，目的に応じて使い分けることにする．

正解のゆれを解消して評価するためには，二つの方法が考えられる．

学習時のマージ あらかじめ，認識用辞書，学習セット，テストセットの三者において，同一と考えられる単語を，一つの単語で代表させ，言語モデルの学習および単語認識を行う．以降，複数の単語を一つの単語で代表させることを**単語のマージ**と呼ぶことにする．

評価時のマージ 言語モデル学習時や単語認識時ではなく，評価時に単語をマージする．

従来の SLDB 形態素の言語モデルを用いた認識実験 [1][4] では，学習時のマージによる方法が用いられてきた．これは，SLDB 形態素の単語では，3 カラム目に標準形が定義されており，単語のマージが形態素情報だけで容易に行えるためである．しかし，すべての単語に対して標準形を定義することは困難であるため，TDMT 形態素では標準形を定義せず，必要に応じて単語のマージ情報を翻訳システムの中で管理するという仕様になっている．本稿の認識実験では，この考え方に習い，正解のゆれを解消して評価をする場合は，後者の評価時のマージによる方法を採用した．この際，TDMT 形態素におけるマージ情報は，人手で作成した．

評価時に単語のマージを行うためには，ATRresult に単語マージリストを与えればよい．単語マージリストとは，一行ごとに単語 ID をならべたもので，同一視する複数単語ごとに，

```
20001 [あ] a {|-}  # あ | ア | あ | 間投詞 ||||
20002 [あー] a a {|-}  # あー | アー | あー | 間投詞 ||||
20004 [ああ] a a {|-}  # ああ | アア | ああ | 間投詞 ||||

12607 [申込み] m o o sh i k o m i {|-} # 申込み | モウシコミ | 申込み | サ変名詞 ||||
14095 [申し込み] m o o sh i k o m i {|-} # 申し込み | モウシコミ | 申し込み | サ変名詞 ||||

20008 [あの] a n o {|-}  # あの | アノ | あの | 間投詞 ||||
20009 [あのー] a n o o {|-}  # あのー | アノー | あのー | 間投詞 ||||

11487 [翌日] j o k u zh i ts u {|-} # 翌日 | ヨクジツ | 翌日 | 普通名詞 ||||
13524 [よくじつ] j o k u zh i ts u {|-} # よくじつ | ヨクジツ | よくじつ | 普通名詞 ||||

10858 [ということは] t o i u k o t o w a {|-} # ということは | トイウコトハ | ということは | 接
続詞 ||||
11820 [とゆうことは] t o j u u k o t o w a {|-} # とゆうことは | トユウコトハ | とゆうことは |
接続詞 ||||
```

図 4: 表記のゆれ (単語マージリスト)

空行で区切られている．各行の単語 ID より右側は無視される．単語マージリストの書式は，先に示した図 4と同様である．

3.2 間投詞の扱い

現在，音声翻訳システムの翻訳部は，間投詞を無視して翻訳している．したがって，単語認識率を求める際も，間投詞に関する誤認識を除いて評価する方が，適当であると考えられる．このよな評価方法として，間投詞のマージと間投詞の無視という二つが考えられる．

間投詞のマージとは，間投詞どうしの置換はエラーとしないが，その他の誤認識パタンはエラーとするというものである．一方，間投詞の無視は，間投詞のマージをさらにゆるくしたもので，次のようなものである．

- 間投詞の脱落，挿入，間投詞どうしの置換は，エラーとしない．
- 間投詞がその他の単語に置換した場合，およびその他の単語が間投詞に置換した場合はエラー．

間投詞無視による評価は，間投詞のマージリストの先頭行に!NULL を挿入する (図 5) ことで行うことができるが， ATRSPREC r05r05 以前の ATRresult ではバグのため，!NULL を使った評価ができないといった問題もある．

```
!NULL
20001 [あ] a {|-}  # あ | ア | あ | 間投詞 ||||
20002 [あー] a a {|-}  # あー | アー | あー | 間投詞 ||||
20003 [あーっと] a a q t o {|-}  # あーっと | アーット | あーっと | 間投詞 ||||
20004 [ああ] a a {|-}  # ああ | アア | ああ | 間投詞 ||||
20015 [いー] i i {|-}  # いー | イー | いー | 間投詞 ||||
20016 [いや] i j a {|-}  # いや | イヤ | いや | 間投詞 ||||
20017 [いやー] i j a a {|-}  # いやー | イヤー | いやー | 間投詞 ||||
20018 [う] u {|-}  # う | ウ | う | 間投詞 ||||
20019 [うー] u u {|-}  # うー | ウー | うー | 間投詞 ||||
20020 [うーん] u u ng {|-}  # うーん | ウーン | うーん | 間投詞 ||||

12607 [申込み] m o o sh i k o m i {|-} # 申込み | モウシコミ | 申込み | サ変名詞 ||||
14095 [申し込み] m o o sh i k o m i {|-} # 申し込み | モウシコミ | 申し込み | サ変名詞 ||||
```

図 5: 間投詞の無視

4節の研究用言語モデルの認識実験では，間投詞に関する評価上の工夫は特に行わない．一

方，5節のMATRIX用言語モデルの認識実験では，より実質にあった評価をするために，間投詞のマージによる評価方法を用いる．

3.3 音声データと言語データの不整合

今回採用したTDMT形態素体系では，SLDB形態素体系と比べて，その仕様上，評価セットの音声データと正解ファイルを整合させることができないという問題がある．一つは，間投詞の扱いで，もう一つは言い淀み・言い直しの扱いに起因する．

正解ファイルは，形態素ファイルから作られるが，オリジナルのTDMT形態素ファイルでは，間投詞は除かれている．このため，間投詞を含む発話では必ず不整合が生じる．この問題を解決するために，間投詞を挿入したTDMT形態素ファイル(TDMTX JMOR)を作成し，それをもとに正解ファイルを作成した．間投詞の挿入に関しても，些細な問題は残っているが，それは(3.3.1)で補足する．

間投詞と同様に，TDMT形態素ファイルでは，言い淀みや言い直しも削除されている．しかし，間投詞の場合と異なり，言い淀みや言い直しを復元することは，TDMT形態素の仕様上，困難である．この問題については，(3.3.2)で補足する．そこで，言い淀みや言い直しを含む発声を，単語認識の評価に用いないという工夫で，この問題を避けることにした．このような評価方法を言い直し発声の無視と呼ぶことにする．従来の音声認識実験[1][4]では，テストセットを片側会話ごとに定義していた．今回の認識実験では，より細かい発声単位でテストセットを定義していると考えることができる．

(3.3.1) 間投詞挿入の問題点

TDMT形態素ファイルに対して，通常の間投詞は容易に復元できるが，単語内に挿入された間投詞は復元できないという問題がある．例えば，図6において，鍵括弧で括られた単語が間投詞であるが，これが単語“入れさ | イレサ | 入れる | 本動詞 | 一段 | せる ||”の中に挿入されている．このような場合，間投詞を復元することはできない．幸いなことに，このような例は極めて稀なため，今回の評価では無視することにした．

担当者：かしこまりました。三泊ご予約を入れ［え］させていただきます。

図 6: 間投詞が復元できない場合

(3.3.2) 言い淀み，言い直しの問題点

自然発話においては，言い淀みや言い直しによって，単語が崩れる場合がある．その例を，図7の書き起こしテキスト(以後，JTEXT)で示す．JTEXTでは，文として崩れている部分を円括弧で括って記述するが，その付近には崩れた単語が現れる．図7の円括弧で括られた「そ」や「いれ」は，「その」や「入れ」が崩れたものであると考えることができる．さらに「けれ（も）ども」の例では，「けれども」という一単語が「も」という言い淀みによって分断されて「けれ」と「ども」に分かれている．

SLDB形態素ファイルでは，このような崩れた単語を表すために，特殊な品詞「その他」が用意されている．一方，TDMT形態素ファイルでは，言い淀みや言い直しを扱うことをあきらめ，円括弧部分を取り除いて形態素ファイルを作成している．このように，TDMT形態素ファイルで，言い淀みや言い直しを扱うのは，仕様上不可能であるため，言い直し発声の無視による評価方法をとった．

申込者：じゃ、（そ）そのお部屋とって頂けますか。
担当者：かしこまりました。
　　それでは御予約、本田(honda)様で（いれ）［えー］入れておきます。

申込者：きょうそちらに泊まる者なんですけれ（も）ども、
　　時間より早くチェックインできませんか。

図 7: 言い淀み，言い直し

4 研究用言語モデル

4.1 言語モデルの作成

研究用言語モデルは可変長 N-gram (複合 N-gram[3]) として作成した．モデルの作成には，ATRSPREC の ATRLANGTOOLS を用い，従来のモデルとの比較のために，バージョンは r05r04 を用いた．

初期クラスには，品詞と活用型と活用形の組の異なる 230 クラス，および，発声開始，発声終了の 2 クラスをあわせた合計 232 クラスを用い，削除補完法を用いて 1000 回までクラス分離を行なった．学習時には，単語のマージおよびクラス内等確率化は行っていない．認識用辞書には，学習セットとテストセットの両方に含まれる単語をすべて登録した．

従来の言語モデル [2] の認識用辞書では，発音を人手によって修正し，多様な発音を認識できるようになっている．この貴重な情報を受け継ぐために，形態素としての整合性の高い名詞類について，従来の辞書の発音を用いた．ただし，SLDB 体系で品詞「人名」「住所」の単語，および，TDMT 体系で品詞「ローマ字」の単語については，旧辞書から発音を継承していない．

新しい研究用言語モデルの学習セットは，従来のものとほぼ一致しているため，違いは主に形態素体系にあると考えられる．そこで，SDB テストセット (S1, S2, S4) を用いて，テストセット・パープレキシティによって，両者のモデルを比較したものを，表 3に示す[4]．テストセット・パープレキシティは，通常，一単語あたりのクロス・エントロピーから求めるが，TDMT 形態素と SLDB 形態素では，単語の単位が異なっているため，単純な比較はできない．TDMT 形態素体系における新しい言語モデルのパープレキシティは，やや大きめの値になっているが，SLDB 形態素の単語数に換算してテストセット・パープレキシティを求めなおしたところ，従来のものとほぼ同一の値が得られた．このパープレキシティによる評価から，新しいモデルは従来の言語モデルと同等の性能が得られていることが確認できる．

4.2 認識実験

言語モデルや音響モデルが変更された場合，計算量との兼ね合いで，ビーム幅とともに，各々のモデルの尤度レンジを揃える言語重みの最適値を決定することが必要となる．まず最初にこれらのサーチパラメータの選択を行い，その後，テストセットに対する認識実験を行う．実験の設定を以下に示す．

- 音響モデル: “/dept1/work1/ResearchJ/amodel/19980121/AM.MF.bin”
  (ResearchJ V3 の実験 [4] と同一)

[4]SLDB 形態素に基づく言語モデルのテストセット・パープレキシティは，TR-IT-0235 の表 2 から抜粋した．

	新研究用モデル		旧研究用モデル
	TDMT 形態素	SLDB 形態素換算	SLDB 形態素
0 回分離	56.43	30.61	31.53
200 回分離	30.09	17.95	17.42
500 回分離	26.49	16.11	15.52
1000 回分離	26.20	15.97	15.07

表 3: SDB テストセット (S1,S2, S4) に対するテストセット・パープレキシティ

- 設定ファイルは，基本的には ResearchJ V3 の実験と同一．言語重みのみ修正．
- 探索および正解との照合: ATRSPREC (r05r05 版) の ATRlattres
- ランニングマシン: DEC Alpha Station

(4.2.1) サーチのパラメータ選択

TR-IT-0235[1] では，次の手順で最適なサーチのパラメータ (第一パス / 第一パス言語重み，ビーム幅) を求めている．

1. 第一パスと第二パスの言語重みを同じにして，言語重みとビーム幅を変えたときの認識率，計算時間などを調べる．
2. 1 の結果をふまえ，ビーム幅と第一パス言語重みを固定し，第二パス言語重みを求める．

一方，ここでは，従来の認識実験と同一ビーム幅での比較実験を目的に，最適パラメータを設定することとする．したがって，ビーム幅は，従来の認識実験 (ResearchJ V3) と同一の 85 に固定する．次に，テストセット S1 と S2 とを用いて，このビーム幅に対して，最適な言語重みを求める．従来から用いられている言語重みの選択方法の妥当性を確かめるために，ここでは，第一パスと第二パスの言語重みを両方変化させ，従来の手法で探索する周辺の値も求める．

第一パス言語重みと第二パス言語重みを同一にして，言語重みを変化させたときの，認識結果を表 4に示す．従来のパラメータ選択法によれば，第一パス言語重みは，最大の認識率の得られている 7 が最適であると予想される．

次に，第一パスだけでなく第二パス言語重みについても，変化させたときの認識結果を表 5に示す．この表から明らかなように，ビーム幅 **85** では，**第一パス言語重み 7**，**第二パス言語重み 15** で最大の認識率が得られる．この言語重みの組みは，従来の手法で選んでも同じものが得られる．このことから，従来のパラメータ選択法は妥当であると思われる．以後の認識実験は，このパラメータ設定で行う．

beam	lmscale		best word	net word	best word accuracy				speed	error	skip
	1st	2nd	acc/cor	acc/cor	total	Ins	Del	Sub	cpu/utt	utt	utt
85	6	6	75.54/80.84	92.03/94.60	1795	95	57	287	3.69	0	0
85	7	7	76.66/81.56	90.19/92.87	1795	88	64	267	1.93	0	0
85	8	8	70.75/76.55	81.56/85.74	1795	104	72	349	1.23	0	0
85	9	9	63.40/70.64	71.03/76.71	1795	130	86	441	0.97	0	0
85	10	10	58.61/66.85	65.07/72.26	1795	148	103	492	0.80	0	0

表 4: 第一パス言語重みを変化させたときの認識結果

		best word acc				cpu-time/utt-time			
		1st lmscale				1st lmscale			
		5	6	7	8	5	6	7	8
	7	77.23	77.05	76.66	-	10.21	3.56	1.88	-
	8	78.48	77.99	77.16	-	9.14	3.57	1.91	-
	9	78.44	78.38	77.83	71.09	10.11	3.55	1.90	1.22
	10	78.22	78.22	77.72	71.31	10.04	3.58	1.90	1.22
	11	78.50	78.38	78.16	71.64	10.03	3.59	1.91	1.22
	12	78.61	78.38	78.16	71.75	10.00	3.61	1.92	1.26
2nd	13	-	78.72	78.77	72.03	-	3.58	1.91	1.23
lmscale	14	-	78.72	78.72	71.92	-	3.61	1.91	1.23
	15	-	78.77	78.89	72.03	-	3.63	1.88	1.23
	16	-	78.38	78.27	71.75	-	3.59	1.89	1.23
	17	-	78.16	78.16	71.98	-	3.61	1.87	1.24
	18	-	78.11	78.11	71.92	-	3.60	1.89	1.23
	19	-	78.05	-	71.87	-	3.61	-	1.24
	20	-	78.33	-	72.09	-	3.61	-	1.23

表 5: ビーム幅 85 での言語重みの変化と認識結果

(4.2.2) 従来のモデルとの比較

(4.2.1)で決定したパラメータを使って得られる認識結果を，従来のモデルのもの (ResearchJ V3)[4] と比較する．テストセットには，S1, S2 だけでなく，S4 を含めた SDB テストセットを用いる．ただし，S4 中の TCS70004.B については，パワーが小さいため，異常に計算量がかかり，正当な認識結果が得られなかった．そのため，これを除く 41 片側会話での比較となっている．この比較では，言い直し発声を無視する評価方法をとることにする．実験条件をそろえるため，言語重み以外は，ResearchJ V3 の実験と同一の設定とする．その結果を表 6に示す．

		best	net	best accuracy				speed	error	skip
		acc/cor	acc/cor	total	Ins	Del	Sub	cpu/utt	utt	utt
旧モデル	音素	87.22/89.34	91.66/93.21	18607	395	977	1006	2.96	0	19
(正規形による	単語	74.02/79.30	82.92/86.93	5569	294	245	908	2.96	0	19
単語マージあり)	発声	42.17(218)	52.42(271)	517	-	-	-	2.96	0	19
新モデル	音素	89.53/91.40	93.73/95.13	18607	349	807	793	1.96	0	19
(単語マージなし)	単語	73.41/77.83	84.48/87.50	4664	206	222	812	1.96	0	19
	発声	36.17(187)	52.61(272)	517	-	-	-	1.96	0	19
新モデル	音素	89.53/91.40	93.73/95.13	18607	349	807	793	1.96	0	19
(単語マージあり)	単語	74.42/78.86	84.97/87.95	4664	207	223	763	1.96	0	19
	発声	38.30(198)	53.58(277)	517	-	-	-	1.96	0	19

表 6: SDB テストセット (41 片側会話) に対する認識結果

表 6によると，音素認識率では TDMT 形態素の言語モデル (新モデル) の方が，SLDB 形態素言語モデル (旧モデル) よりもやや高い認識性能が得られた．一方，単語認識率や発声認識率は，単語マージの条件を同一にできないため，直接の比較は難しいが，なるべく同じ条件になるような単語マージリスト[5]を用いて比較する．従来のものと比べて，発声認識率でやや劣るものの，一位単語認識率でほぼ同等という結果が得られた．特に，この同等の認識率を，より少ない処理速度で実現していることは特筆に値する．

TDMT 形態素に基づく言語モデルでは，明らかに音素認識率が向上している．したがって，SLDB 形態素に基づく認識用辞書レベルにまで発音辞書を整備すれば，さらに単語認識率，発声認識率を向上できる余地が残っていると考えられる．

[5]間投詞のマージなし．それ以外の単語で同一単語と考えられるものをマージ．

4.3 今後の比較実験のための設定

4.2節の認識実験では，ResearchJ V3 の実験と比較を行うため，言語重み以外は，ResearchJ V3 と同一設定でおこなった．しかし，この設定は探索の観点からは，必ずしも最適なものとはいえない．そこで本節では，今後の比較実験のため，推奨される設定について述べ，それに基づく認識結果を示す．

これまでの設定では，ATRlattice の二つのパラメータ，word_merge とword_boundary_skip は次のように設定されていた．

```
ATRlattice:word_merge=non
ATRlattice:word_boundary_skip=OFF            # default
```

しかし，これまでの探索の研究 ([10] など) によれば，この二つのパラメータは，次のように設定する方がよいことが確かめられている．

```
ATRlattice:word_merge=all
ATRlattice:word_boundary_skip=2              # r05r05 以降
## ATRlattice:word_boundary_skip=ON          # r05r04 以前
```

豊富な計算量が使える状況では，同じ計算量で比較したとき，両者の認識精度には差がない．しかし，計算量を減らした状況では，後者の設定の方が，より高い認識精度が得られる．ただし，単に上記の二つのパラメータをかえただけでは，計算量も減るかわり認識精度も落ちてしまう．そこで，この設定を生かすには，ビーム幅を広げるなどして，探索空間を広げる必要がある．

そこで，今後の比較実験のため，後者の設定を用いるとともに，ビーム幅を 95 に広げた認識結果を付録 E.1に，そのときの設定ファイルを付録 E.2に示す．この新しい設定では，従来のモデルとの比較実験の際に問題となった会話 (TCS70004.B) についても，問題なく処理することができた．

表 7は，新しい設定で求めた「SDB テストセット (S1, S2, S4)」に対する認識結果を，表 6と同条件 (41 片側会話に限定，言い直し発声を無視) で評価したものである．単語のマージを全く行っていないにもかかわらず，単語マージを行っている旧モデルの実験結果とほぼ等しい単語認識率が得られている．

		best	net	best accuracy				speed	error	skip
		acc/cor	acc/cor	total	Ins	Del	Sub	cpu/utt	utt	utt
新モデル	音素	89.68/91.48	94.39/95.73	18607	335	811	775	2.14	0	19
(単語マージ	単語	74.16/78.07	87.11/89.45	4664	182	229	794	2.14	0	19
なし)	発声	36.75(190)	57.45(297)	517	-	-	-	2.14	0	19

表 7: 新しい設定での SDB セット (41 会話) に対する認識結果

5 MATRIX 用言語モデル

従来の MATRIX 用言語モデル[6] [2][5] は，'97 年 ATR 研究発表会のデモ会話を認識するために作られた．したがって，音声翻訳システムとしての評価で用いるとなると，語彙数も小さく，多様な表現を認識できないという問題があった．例えば，テストセット SLTA1 や S1, S2 を使った未知語なしの音声認識の評価ができないばかりか，「～円」「～ドル」などの金額表現は，タスク達成上重要な表現にもかかわらず，認識されにくかった．そこで今回，従来の MATRIX 用言語モデルとくらべて，より多様な表現を認識することが可能な言語モデルを作成する．

5.1 言語モデルの作成

言語モデルは可変長 N-gram(複合 N-gram[3]) として，作成する．学習セットは，研究用言語モデルの旅行会話学習セットをホテル予約会話 (TA$_{\sqcup\sqcup\sqcup\sqcup\sqcup}$) に限定して用いる．これは，事前検討により，ホテル予約会話に限定した方が，学習データ量は減るが，良好な結果が得られることを確認したためである．認識用辞書は，研究用言語モデルの辞書を，学習セットとテストセット (S1 ～ 4， SLTA1， MDB1(DS, IS, WS)) の単語に限定し，それに従来の MATRIX 用言語モデルの認識用辞書の単語を加えて作成する．したがって，従来の MATRIX 用言語モデルの語彙はすべて含まれる．

言語モデルの作成にあたっては，次のような表現を，明示的に複合語として登録した．[新規] と記したものは，従来の MATRIX 用言語モデルにはなかった複合語である．

- 桁読み数字 (～万，～千，～百，～十) [新規]
- 金額表現 (～円，～ドル) [新規]
- 日時 (～月，～日，～時，～分)
- 宿泊 (～泊，～人，～名，～部屋)

初期クラスは，基本的に品詞，活用型，活用形の三つ組みで作成するが，それ以外にも，上記複合語のためのクラスや，普通名詞を細分化したクラス (棒読み数字，曜日，人名，ホテル名) を明示的に導入した．合計，発声開始，発声終了のクラスを含めて，205 クラスを初期クラスとして用いた．明示的に導入したクラスからは，単語が分離されないようにしながら，1000 回クラス分離して可変長 N-gram を作成した．さらに，研究用言語モデルの考え方とことなり，

[6]厳密には，'97 年 ATR 研究発表会のデモ使った言語モデル [2] と，音響学会で発表した認識実験 [5] の言語モデルは異なる．前者のモデルは，後者のモデルを改良したものである．本稿では，特に断らない限り，前者のモデル (/dept1/work1/MatrixJ/lmodel/19971113) を従来の MATRIX 用言語モデルと呼ぶことにする．

辞書に登録された単語はすべて認識できるようにするため，クラスからの生起確率 $P(w_i|C)$ を，すべての 単語 w_i について等確率とするクラス内等確率化を行った．

図 8 に，学習セットのデータ量，辞書サイズ，テストセット・パープレキシティをまとめる．学習データおよび辞書の単語数は，発声開始，発声終了を表す特殊な単語を含んでいる．パープレキシティは，テストセット MDB1(DS, IS, WS) により求めたものである．新モデルでは，語彙数は従来の 1.7 倍もあるものの，豊富な学習データのため，より小さなパープレキシティを実現している．

	学習データ量	認識辞書	testset perplexity
旧モデル (19971113)	210 片側会話　45,504 語	1,597 語	19.4
新モデル (19980605)	683 片側会話　140,899 語	2,736 語	18.4

表 8: MATRIX 用言語モデルの比較

5.2 認識実験

音声認識実験は，テストセット MDB1(IS, DS, WS) を用いて行った．音響モデル[7][6][5]，設定ファイル (付録 F.2) は， MatrixJ V2 の認識実験で用いたものを使う．ただし，設定ファイル中のビーム幅，言語重みについては，書き換える．単語認識率を求める際には，間投詞はすべてマージするとともに，表記のゆれの問題を解消するために，人手で作成したマージリストを用いる (3.1節参照)．

認識は， ATRSPREC r05r04 を用い，正解との照合は， r05r05 を用いる．認識時に古い版を用いる理由は， MatrixJ V2 の認識実験 [5] で用いた Python スクリプトを，流用するためである． r05r05 からは，イベント・ループの仕様が変り，古い Python スクリプトは使えない．

サーチパラメータの選択実験では， DEC Personal Workstation 600au (atrm03/atra50, CPU 600MHz) をランニングマシンとして用いた．さらに，従来の言語モデルとの比較には， DEC Personal Workstation 500au (pau04 ～ pau07, CPU 500MHz) を用いた．

(5.2.1) サーチのパラメータ選択

MATRIX の音声認識では，精度だけでなく処理時間が重要である．そこで，テストセットの CPU 時間の平均が実時間を切る条件で，認識精度が最大となるような第一パス・第二パス

[7] /dept1/work1/MatrixJ/amodel/19980106/{AM.M.803.5.bin,AM.F.803.5.bin}

の言語重み (lmscale)，およびビーム幅 (beam) の組み合わせを求めた．ビーム幅については，第一パス，第二パスで同じ値を用いることにする．

処理時間の大部分は，第一パスで費やされる．そこで，第一パスと第二パスの言語重みを同一にしつつ，言語重みとビーム幅を変化させて，目標の処理時間を達成する第一パス言語重みとビーム幅の組み合わせを調べたのが，表9である．表中の数字は“一位認識率 (処理時間)”を表す．処理時間は，かかったCPU時間が発声時間の何倍かで表されている[8]．

lmscale	beam		
	85	90	95
6	88.53 (<u>0.82</u>)	88.73 (1.12)	88.84 (1.39)
7	85.41 (0.73)	88.01 (<u>0.85</u>)	89.13 (1.02)
8	75.97 (0.61)	82.72 (0.71)	85.51 (<u>0.84</u>)

表 9: 認識率と処理時間 (第一パス言語重みと第二パス言語重みは同一)

当然のことながら，同じ言語重みであれば，ビーム幅が大きいほど，認識率は高くなるものの，処理時間もかかる．各々の言語重みにおいて，下線部のビーム幅で，実時間の約8割を達成している．次に，この組合わせで，第二パスの言語重みを変化させたのが，表10である．その結果，下線部 (第一パス言語重み 6, 第二パス言語重み 12，ビーム幅 85) で，一位認識率が最大となった．

(5.2.2) 従来のモデルとの比較

同一音響モデルを用い，同一の単語認識評価方法で，従来のMATRIX用言語モデル (19971113) と新しい言語モデル (19980605) を比較した．設定ファイルは，Matrix V2 の実験と同一のものを用い，言語重みとビーム幅のみ各々のモデルごとに設定する．従来の言語モデルの認識結果は，第一パス言語重みを 7，第二パス言語重み 10，ビーム幅 90 として求めた．このパラメータは，Matrix V2 の実験で求めた最適なものである．一方，新しい言語モデルについては，(5.2.1)で求めた最適パラメータを用いた．これらの認識結果を，表11に示す．認識率は，(5.2.1)で求めたものと，若干異なっている．これは，正解のゆれの条件を極力同一にするために，認識用辞書をソートして認識実験を行ったためである．新モデルは，旧モデルと比べると，格段に語彙数が向上して認識できる表現が増えているにもかかわらず，デモで用いるような非常に限られた表現に対して，同等の認識性能が得られることが示せた．今後の比較実験のために，新しい言語モデルのより詳細な認識結果と設定ファイルを，それぞれ付録F.1，F.2として載せておく．

[8]本来CPU時間は，計算機の混雑度に依存しない値であるが，ネットワークの負荷などによって，かなりのゆれがみられるようである．

lmscale		beam	accuracy (cputime/utt-time)
1st	2nd		
6	6	85	88.53 (0.82)
6	9	85	89.29 (0.92)
6	12	85	89.31 (0.92)
7	7	90	88.01 (0.85)
7	10	90	88.35 (0.86)
7	14	90	88.09 (0.88)
8	8	95	85.51 (0.84)
8	12	95	85.76 (0.84)
8	16	95	85.46 (0.86)

表 10: 認識率と処理時間

lmodel		best word	net word	best word accuracy				speed	error
version		acc/cor	acc/cor	total	Ins	Del	Sub	No/Wo(cpu/u)	utt.
1997	male	88.22/ 89.67	95.04/ 95.66	16848	244	435	1306	3.99(0.91)	0
1113	female	91.75/ 93.01	96.94/ 97.35	8157	103	127	443	3.19(0.78)	0
(old)	both	89.37/ 90.76	95.66/ 96.21	25005	347	562	1749	3.73(0.87)	0
1998	male	88.15/ 89.73	94.91/ 95.76	16848	267	436	1294	2.77(1.03)	0
0605	female	91.74/ 93.15	96.57/ 97.27	8157	115	99	460	2.32(0.87)	0
(new)	both	89.32/ 90.85	95.45/ 96.25	25005	382	535	1754	2.62(0.98)	0

表 11: 新旧 MATRIX 用言語モデルの比較

6 むすび

TDMT 形態素に基づく可変長 N-gram として，二種類の言語モデルを作成した．一つは，データベースの統計量を重視して作成した旅行会話認識のための研究用言語モデルである．もう一つは，先見的なタスク依存の知識を導入し，ホテル予約会話認識に用いる MATRIX 用言語モデルである．研究用言語モデルは，従来のモデルと比べて，単語認識率は同等ながら，若干ではあるが音素認識率を向上することができた．MATRIX 用言語モデルは，デモ会話に対する認識性能を落さず，多くの表現を認識可能とした．今後，これらの言語モデルを用いた音声翻訳実験により，音声認識部の性能だけでなく，翻訳部を含めた評価を進めて行く予定である．

謝辞

間投詞つきTDMT形態素データを作成するにあたっては，隅田英一郎氏，竹澤寿幸氏をはじめとする第三研究室，第四研究室の方々にご尽力頂いた．また，データ作成の実作業は，溝口淳子氏，森田千帆氏，伊藤いずみ氏らによって行われた．評価に用いた単語マージリストは手作業で作成したが，作成にあたっては中村 篤氏，政瀧浩和氏，谷垣宏一氏，西野敦士氏らに協力頂いた．さらに，MATRIX用モデルの定性的な評価は，菅谷 史昭氏にご協力頂いている．これらの皆様に感謝いたします．

参考文献

[1] 政瀧浩和ほか: “連続音声認識用ベースライン言語モデル”, ATR テクニカルレポート, TR-IT-0235, 1997

[2] 政瀧浩和: http://www.itl.atr.co.jp/~masataki/demosystem.html

[3] 政瀧浩和: “統計的手法による音声言語処理の研究”, ATR テクニカルレポート, TR-IT-0259, 1998

[4] 深田俊明ほか: “連続音声認識用音響モデル (version 2.0)”, ATR テクニカルレポート, TR-IT-0241, 1997

[5] 内藤正樹ほか: “日英音声翻訳システム ATR-MATRIX における音声認識用音響・言語モデル”, http://www.itl.atr.co.jp/~mnaito/Papers/ASJ9803SYS.ps, 音響学会春季講演論文集, 2-Q-20, 1998.3

[6] 内藤正樹: http://www.itl.atr.co.jp/~mnaito/matrix/amodel97/index.html

[7] Morimoto et. al., “Speech and language database for speech translation research,” Proc. of ICSLP'94, pp.1791–1794, 1994

[8] 塚田 元ほか: “研究用自然発話音声データベース解説書 ('97 年度公開版) – 旅行会話タスク –”, ATR テクニカルレポート, TR-IT-0222, 1997

[9] Nakamura et. al., “Japanese Speech Databases for Robust Speech Recognition,” Proc. of ICSLP'96, pp.2199–2202, 1996

[10] 清水 徹ほか: “大語い連続音声認識のための単語仮説数削減”, 信学論, Vol.J79-D-II, No.12, pp.2117–2124, 1996

付録 A　言語モデルの在処

A.1　研究用言語モデル

可変長 N-gram: /dept1/work1/ResearchJ/lmodel/19980521/LM.W

認識用辞書: /dept1/work1/ResearchJ/lmodel/19980521/LEX.W

学習セット: /dept1/work1/ResearchJ/ldata/19980507/ANS/TrainSet/*.ANS

テストセット:

/dept1/work1/ResearchJ/ldata/19980507/ANS/TestSet/{S1,S2,S4,SLTA1}/*.ANS

A.2　MATRIX 用言語モデル

可変長 N-gram: /dept1/work1/MatrixJ/lmodel/19980605/vngram.equal.bin

認識用辞書: /dept1/work1/MatrixJ/lmodel/19980605/RECOG_LEX.vn1000

学習セット: /dept1/work1/ResearchJ/ldata/19980507/ANS/TrainSet/TA*.ANS

テストセット:

/dept1/work1/MatrixJ/ldata/19971029/{*_DS,*_IS,*_WS}/*.ANS,

/dept1/work1/ResearchJ/ldata/19980507/ANS/TestSet/{S1,S2,SLTA1}/*.ANS

付録 B　SDB テストセット

S1

TAC70021	JTEXT	ENV	JMOR	WAV	TRS	LBL	MMAHA	64db-18db	1994.00.00	16	A	CUSTOMER
TAC70019	JTEXT	ENV	JMOR	WAV	TRS	LBL	FTOAR	64db-48db	1994.00.00	15	A	CUSTOMER
TAC70101	JTEXT	ENV	JMOR	WAV	TRS	LBL	FYOOO	64db-24db	1994.00.00	20	A	CUSTOMER
TAC70103	JTEXT	ENV	JMOR	WAV	TRS	LBL	MMAUC	64db-32db	1994.00.00	10	A	CUSTOMER
TAC70202	JTEXT	ENV	JMOR	WAV	TRS	LBL	MSAHA	64db-32db	1994.00.00	11	A	CUSTOMER
TAC70203	JTEXT	ENV	JMOR	WAV	TRS	LBL	FMASZ	64db-24db	1994.00.00	18	A	CUSTOMER
TAC70303	JTEXT	ENV	JMOR	WAV	TRS	LBL	FYUKI	64db-32db	1994.00.00	19	A	CUSTOMER
TAC70304	JTEXT	ENV	JMOR	WAV	TRS	LBL	MKEWA	64db-48db	1994.00.00	8	A	CUSTOMER

S2

TAC70015	JTEXT	ENV	JMOR	WAV	TRS	X	FTAAO	64db-48db	1994.00.00	14	A	CUSTOMER
TAC70016	JTEXT	ENV	JMOR	WAV	TRS	X	MNACH	64db-40db	1994.00.00	12	A	CUSTOMER
TAC70017	JTEXT	ENV	JMOR	WAV	TRS	X	MMINA	72db-40db	1994.00.00	11	A	CUSTOMER
TAC70022	JTEXT	ENV	JMOR	WAV	TRS	X	MJUHO	64db-40db	1994.00.00	13	A	CUSTOMER
TAC70023	JTEXT	ENV	JMOR	WAV	TRS	X	MHITA	64db-40db	1994.00.00	17	A	CUSTOMER
TAC70102	JTEXT	ENV	JMOR	WAV	TRS	X	FYOAZ	64db-40db	1994.00.00	18	A	CUSTOMER
TAC70201	JTEXT	ENV	JMOR	WAV	TRS	X	FKOTS	64db-40db	1994.00.00	15	A	CUSTOMER
TAC70301	JTEXT	ENV	JMOR	WAV	TRS	X	FTOHO	64db-40db	1994.00.00	11	A	CUSTOMER

S4

TCC70109	JTEXT	ENV	JMOR	WAV	TRS	X	MYUYO	72db-48db	1994.00.00	13	A	CUSTOMER
TCC70103	JTEXT	ENV	JMOR	WAV	TRS	X	FRIYU	72db-48db	1994.00.00	9	A	CUSTOMER
TCC71035	JTEXT	ENV	JMOR	WAV	TRS	X	FKAAR	66db-27db	1995.00.00	11	A	CUSTOMER
TCC71008	JTEXT	ENV	JMOR	WAV	TRS	X	MMAOX	61db-36db	1995.00.00	17	A	CUSTOMER
TCC71016	JTEXT	ENV	JMOR	WAV	TRS	X	FASSA	61db-38db	1995.00.00	16	A	CUSTOMER
TCC71001	JTEXT	ENV	JMOR	WAV	TRS	X	FASHI	58db-30db	1995.00.00	12	B	CLERK
TCS70070	JTEXT	ENV	JMOR	WAV	TRS	X	METYA	64db-32db	1994.00.00	8	A	CUSTOMER
TCC70201	JTEXT	ENV	JMOR	WAV	TRS	X	MSHSZ	64db-40db	1994.00.00	7	A	CUSTOMER
TCC70307	JTEXT	ENV	JMOR	WAV	TRS	X	MMAOK	64db-32db	1994.00.00	11	A	CUSTOMER
TCS70055	JTEXT	ENV	JMOR	WAV	TRS	X	MMAMU	64db-32db	1994.00.00	11	A	CUSTOMER
TCS70074	JTEXT	ENV	JMOR	WAV	TRS	X	MKENA	72db-40db	1994.00.00	12	A	CUSTOMER
TCC70212	JTEXT	ENV	JMOR	WAV	TRS	X	MKEOO	64db-32db	1994.00.00	18	A	CUSTOMER
TCS70034	JTEXT	ENV	JMOR	WAV	TRS	X	MKEWZ	64db-40db	1994.00.00	18	A	CUSTOMER
TCC71007	JTEXT	ENV	JMOR	WAV	TRS	X	FJUSA	56db-40db	1995.00.00	26	A	CUSTOMER
TCS70013	JTEXT	ENV	JMOR	WAV	TRS	X	FYUNZ	72db-40db	1994.00.00	7	A	CUSTOMER
TSC71005	JTEXT	ENV	JMOR	WAV	TRS	X	FAKMX	62db-33db	1995.00.00	15	B	CLERK
TCS70023	JTEXT	ENV	JMOR	WAV	TRS	X	FKYYA	64db-40db	1994.00.00	14	A	CUSTOMER
TCS70059	JTEXT	ENV	JMOR	WAV	TRS	X	FSAWA	64db-40db	1994.00.00	6	A	CUSTOMER
TCS70082	JTEXT	ENV	JMOR	WAV	TRS	X	FNOTA	72db-48db	1994.00.00	8	A	CUSTOMER
TCS70004	JTEXT	ENV	JMOR	WAV	TRS	X	FKENA	48db-16db	1994.00.00	15	B	CLERK
TCS70010	JTEXT	ENV	JMOR	WAV	TRS	X	FAYYA	64db-40db	1994.00.00	10	A	CUSTOMER
TCS70028	JTEXT	ENV	JMOR	WAV	TRS	X	FKITA	64db-40db	1994.00.00	13	A	CUSTOMER
TCS70025	JTEXT	ENV	JMOR	WAV	TRS	X	FYOSU	64db-48db	1994.00.00	9	A	CUSTOMER
TSC71013	JTEXT	ENV	JMOR	WAV	TRS	X	FHAOK	65db-33db	1995.00.00	14	A	CUSTOMER
TCS70020	JTEXT	ENV	JMOR	WAV	TRS	X	FYUNI	64db-24db	1994.00.00	11	A	CUSTOMER
TCS70047	JTEXT	ENV	JMOR	WAV	TRS	X	FYUKO	64db-48db	1994.00.00	12	A	CUSTOMER

付録 C　SLDB テストセット

SLTA1

TAC22013	JTEXT	ENV	JMOR	WAV	TRS	LBL	FHITA	72db-16db	1994.05.19	19	A	CUSTOMER
TAC23022	JTEXT	ENV	JMOR	WAV	TRS	X	FAKMO	58db-38db	1994.06.10	9	B	CLERK
TAS12002	JTEXT	ENV	JMOR	WAV	TRS	X	MMASE	64db-16db	1993.10.24	12	A	CUSTOMER
TAS12006	JTEXT	ENV	JMOR	WAV	TRS	X	FYOMA	64db-32db	1994.05.25	6	A	CUSTOMER
TAS12014	JTEXT	ENV	JMOR	WAV	TRS	X	FARSH	64db-24db	1994.07.28	13	A	CUSTOMER
TAS12016	JTEXT	ENV	JMOR	WAV	TRS	X	FARSH	58db-32db	1994.07.22	14	A	CUSTOMER
TAS12019	JTEXT	ENV	JMOR	WAV	TRS	X	FKAKI	59db-34db	1994.07.26	13	A	CUSTOMER
TAS12020	JTEXT	ENV	JMOR	WAV	TRS	X	FARSH	58db-27db	1994.07.28	15	A	CUSTOMER
TAS13002	JTEXT	ENV	JMOR	WAV	TRS	X	MMASE	56db-08db	1993.11.14	21	B	CLERK
TAS13014	JTEXT	ENV	JMOR	WAV	TRS	X	FARSH	62db-34db	1994.07.22	25	B	CLERK
TAS13018	JTEXT	ENV	JMOR	WAV	TRS	X	FARSH	58db-29db	1994.07.28	10	B	CLERK
TAS13019	JTEXT	ENV	JMOR	WAV	TRS	X	FKAKI	59db-34db	1994.07.26	18	B	CLERK
TAS13020	JTEXT	ENV	JMOR	WAV	TRS	X	FARSH	59db-29db	1994.07.28	16	B	CLERK
TAS13022	JTEXT	ENV	JMOR	WAV	TRS	X	FARSH	64db-32db	1994.07.22	18	B	CLERK
TAS13024	JTEXT	ENV	JMOR	WAV	TRS	X	FARSH	59db-28db	1994.07.28	19	B	CLERK
TAS32001	JTEXT	ENV	JMOR	WAV	TRS	X	FYOAS	64db-16db	1993.11.01	12	A	CUSTOMER
TAS32012	JTEXT	ENV	JMOR	WAV	TRS	X	MNOSA	63db-33db	1994.06.07	9	A	CUSTOMER
TAS32014	JTEXT	ENV	JMOR	WAV	TRS	X	MNOSA	61db-32db	1994.06.07	13	A	CUSTOMER
TAS32016	JTEXT	ENV	JMOR	WAV	TRS	LBL	MRYTA	64db-24db	1994.07.18	8	A	CUSTOMER
TAS33010	JTEXT	ENV	JMOR	WAV	TRS	X	MMATA	54db-28db	1994.05.25	16	B	CLERK
TAS33015	JTEXT	ENV	JMOR	WAV	TRS	X	MSHAK	57db-33db	1994.06.04	13	B	CLERK
TAS33020	JTEXT	ENV	JMOR	WAV	TRS	X	MRYTA	59db-32db	1994.07.18	16	B	CLERK
TAS33022	JTEXT	ENV	JMOR	WAV	TRS	X	FYOAS	57db-35db	1994.08.04	15	B	CLERK

付録 D　MDB1 テストセット

IS

FAH_IS	X	X	X	WAV	TRS	X	FAH	X	1997.10.03	33	A	NONE
FCM_IS	X	X	X	WAV	TRS	X	FCM	X	1997.10.03	33	A	NONE
FHK_IS	X	X	X	WAV	TRS	X	FHK	X	1997.10.03	33	A	NONE
FII_IS	X	X	X	WAV	TRS	X	FII	X	1997.10.03	33	A	NONE
FKO_IS	X	X	X	WAV	TRS	X	FKO	X	1997.10.03	33	A	NONE
FKT_IS	X	X	X	WAV	TRS	X	FKT	X	1997.10.03	33	A	NONE
FMN_IS	X	X	X	WAV	TRS	X	FMN	X	1997.10.03	33	A	NONE
FTM_IS	X	X	X	WAV	TRS	X	FTM	X	1997.10.03	33	A	NONE
FYD_IS	X	X	X	WAV	TRS	X	FYD	X	1997.10.03	33	A	NONE
FYO_IS	X	X	X	WAV	TRS	X	FYO	X	1997.10.03	33	A	NONE
FYS_IS	X	X	X	WAV	TRS	X	FYS	X	1997.10.03	33	A	NONE
FYZ_IS	X	X	X	WAV	TRS	X	FYZ	X	1997.10.03	33	A	NONE
MAN_IS	X	X	X	WAV	TRS	X	MAN	X	1997.10.03	33	A	NONE
MAY_IS	X	X	X	WAV	TRS	X	MAY	X	1997.10.03	33	A	NONE
MAZ_IS	X	X	X	WAV	TRS	X	MAZ	X	1997.10.03	33	A	NONE
MFS_IS	X	X	X	WAV	TRS	X	MFS	X	1997.10.03	33	A	NONE
MGT_IS	X	X	X	WAV	TRS	X	MGT	X	1997.10.03	33	A	NONE
MHI_IS	X	X	X	WAV	TRS	X	MHI	X	1997.10.03	33	A	NONE
MHM_IS	X	X	X	WAV	TRS	X	MHM	X	1997.10.03	33	A	NONE
MHN_IS	X	X	X	WAV	TRS	X	MHN	X	1997.10.03	33	A	NONE
MHT_IS	X	X	X	WAV	TRS	X	MHT	X	1997.10.03	33	A	NONE
MHX_IS	X	X	X	WAV	TRS	X	MHX	X	1997.10.03	33	A	NONE
MHY_IS	X	X	X	WAV	TRS	X	MHY	X	1997.10.03	33	A	NONE
MHZ_IS	X	X	X	WAV	TRS	X	MHZ	X	1997.10.03	33	A	NONE
MKT_IS	X	X	X	WAV	TRS	X	MKT	X	1997.10.03	33	A	NONE
MKZ_IS	X	X	X	WAV	TRS	X	MKZ	X	1997.10.03	33	A	NONE
MMN_IS	X	X	X	WAV	TRS	X	MMN	X	1997.10.03	33	A	NONE
MMS_IS	X	X	X	WAV	TRS	X	MMS	X	1997.10.03	33	A	NONE
MNA_IS	X	X	X	WAV	TRS	X	MNA	X	1997.10.03	33	A	NONE
MRI_IS	X	X	X	WAV	TRS	X	MRI	X	1997.10.03	33	A	NONE
MST_IS	X	X	X	WAV	TRS	X	MST	X	1997.10.03	33	A	NONE
MSY_IS	X	X	X	WAV	TRS	X	MSY	X	1997.10.03	33	A	NONE
MTB_IS	X	X	X	WAV	TRS	X	MTB	X	1997.10.03	33	A	NONE
MTM_IS	X	X	X	WAV	TRS	X	MTM	X	1997.10.03	33	A	NONE
MTT_IS	X	X	X	WAV	TRS	X	MTT	X	1997.10.03	33	A	NONE
MTZ_IS	X	X	X	WAV	TRS	X	MTZ	X	1997.10.03	33	A	NONE
MYS_IS	X	X	X	WAV	TRS	X	MYS	X	1997.10.03	33	A	NONE

DS

FAH_DS	X	X	X	WAV	TRS	X	FAH	X	1997.10.03	19	A	NONE
FCM_DS	X	X	X	WAV	TRS	X	FCM	X	1997.10.03	19	A	NONE
FHK_DS	X	X	X	WAV	TRS	X	FHK	X	1997.10.03	19	A	NONE
FII_DS	X	X	X	WAV	TRS	X	FII	X	1997.10.03	19	A	NONE
FKO_DS	X	X	X	WAV	TRS	X	FKO	X	1997.10.03	19	A	NONE
FKT_DS	X	X	X	WAV	TRS	X	FKT	X	1997.10.03	19	A	NONE
FMN_DS	X	X	X	WAV	TRS	X	FMN	X	1997.10.03	19	A	NONE
FTM_DS	X	X	X	WAV	TRS	X	FTM	X	1997.10.03	19	A	NONE
FYD_DS	X	X	X	WAV	TRS	X	FYD	X	1997.10.03	19	A	NONE
FYO_DS	X	X	X	WAV	TRS	X	FYO	X	1997.10.03	19	A	NONE
FYS_DS	X	X	X	WAV	TRS	X	FYS	X	1997.10.03	19	A	NONE
MAN_DS	X	X	X	WAV	TRS	X	MAN	X	1997.10.03	19	A	NONE
MAY_DS	X	X	X	WAV	TRS	X	MAY	X	1997.10.03	19	A	NONE
MAZ_DS	X	X	X	WAV	TRS	X	MAZ	X	1997.10.03	19	A	NONE
MFS_DS	X	X	X	WAV	TRS	X	MFS	X	1997.10.03	19	A	NONE
MGT_DS	X	X	X	WAV	TRS	X	MGT	X	1997.10.03	19	A	NONE

MHI_DS	X	X	X	WAV	TRS	X	MHI	X	1997.10.03	19	A	NONE
MHM_DS	X	X	X	WAV	TRS	X	MHM	X	1997.10.03	19	A	NONE
MHN_DS	X	X	X	WAV	TRS	X	MHN	X	1997.10.03	19	A	NONE
MHX_DS	X	X	X	WAV	TRS	X	MHX	X	1997.10.03	19	A	NONE
MHY_DS	X	X	X	WAV	TRS	X	MHY	X	1997.10.03	19	A	NONE
MHZ_DS	X	X	X	WAV	TRS	X	MHZ	X	1997.10.03	19	A	NONE
MKT_DS	X	X	X	WAV	TRS	X	MKT	X	1997.10.03	19	A	NONE
MKZ_DS	X	X	X	WAV	TRS	X	MKZ	X	1997.10.03	19	A	NONE
MMN_DS	X	X	X	WAV	TRS	X	MMN	X	1997.10.03	19	A	NONE
MMS_DS	X	X	X	WAV	TRS	X	MMS	X	1997.10.03	19	A	NONE
MNA_DS	X	X	X	WAV	TRS	X	MNA	X	1997.10.03	19	A	NONE
MRI_DS	X	X	X	WAV	TRS	X	MRI	X	1997.10.03	19	A	NONE
MST_DS	X	X	X	WAV	TRS	X	MST	X	1997.10.03	19	A	NONE
MTB_DS	X	X	X	WAV	TRS	X	MTB	X	1997.10.03	19	A	NONE
MTM_DS	X	X	X	WAV	TRS	X	MTM	X	1997.10.03	19	A	NONE
MTT_DS	X	X	X	WAV	TRS	X	MTT	X	1997.10.03	19	A	NONE
MTZ_DS	X	X	X	WAV	TRS	X	MTZ	X	1997.10.03	19	A	NONE

WS

FAH_WS	X	X	X	WAV	TRS	X	FAH	X	1997.10.03	21	A	NONE
FCM_WS	X	X	X	WAV	TRS	X	FCM	X	1997.10.03	21	A	NONE
FHK_WS	X	X	X	WAV	TRS	X	FHK	X	1997.10.03	21	A	NONE
FII_WS	X	X	X	WAV	TRS	X	FII	X	1997.10.03	21	A	NONE
FKO_WS	X	X	X	WAV	TRS	X	FKO	X	1997.10.03	21	A	NONE
FKT_WS	X	X	X	WAV	TRS	X	FKT	X	1997.10.03	21	A	NONE
FMN_WS	X	X	X	WAV	TRS	X	FMN	X	1997.10.03	21	A	NONE
FTM_WS	X	X	X	WAV	TRS	X	FTM	X	1997.10.03	21	A	NONE
FYD_WS	X	X	X	WAV	TRS	X	FYD	X	1997.10.03	21	A	NONE
FYO_WS	X	X	X	WAV	TRS	X	FYO	X	1997.10.03	21	A	NONE
FYS_WS	X	X	X	WAV	TRS	X	FYS	X	1997.10.03	21	A	NONE
FYZ_WS	X	X	X	WAV	TRS	X	FYZ	X	1997.10.03	21	A	NONE
MAN_WS	X	X	X	WAV	TRS	X	MAN	X	1997.10.03	21	A	NONE
MAY_WS	X	X	X	WAV	TRS	X	MAY	X	1997.10.03	21	A	NONE
MAZ_WS	X	X	X	WAV	TRS	X	MAZ	X	1997.10.03	21	A	NONE
MFS_WS	X	X	X	WAV	TRS	X	MFS	X	1997.10.03	21	A	NONE
MGT_WS	X	X	X	WAV	TRS	X	MGT	X	1997.10.03	21	A	NONE
MHI_WS	X	X	X	WAV	TRS	X	MHI	X	1997.10.03	21	A	NONE
MHM_WS	X	X	X	WAV	TRS	X	MHM	X	1997.10.03	21	A	NONE
MHN_WS	X	X	X	WAV	TRS	X	MHN	X	1997.10.03	21	A	NONE
MHT_WS	X	X	X	WAV	TRS	X	MHT	X	1997.10.03	21	A	NONE
MHX_WS	X	X	X	WAV	TRS	X	MHX	X	1997.10.03	21	A	NONE
MHY_WS	X	X	X	WAV	TRS	X	MHY	X	1997.10.03	21	A	NONE
MHZ_WS	X	X	X	WAV	TRS	X	MHZ	X	1997.10.03	21	A	NONE
MKT_WS	X	X	X	WAV	TRS	X	MKT	X	1997.10.03	21	A	NONE
MKZ_WS	X	X	X	WAV	TRS	X	MKZ	X	1997.10.03	21	A	NONE
MMN_WS	X	X	X	WAV	TRS	X	MMN	X	1997.10.03	21	A	NONE
MMS_WS	X	X	X	WAV	TRS	X	MMS	X	1997.10.03	21	A	NONE
MNA_WS	X	X	X	WAV	TRS	X	MNA	X	1997.10.03	21	A	NONE
MRI_WS	X	X	X	WAV	TRS	X	MRI	X	1997.10.03	21	A	NONE
MST_WS	X	X	X	WAV	TRS	X	MST	X	1997.10.03	21	A	NONE
MSY_WS	X	X	X	WAV	TRS	X	MSY	X	1997.10.03	21	A	NONE
MTB_WS	X	X	X	WAV	TRS	X	MTB	X	1997.10.03	21	A	NONE
MTM_WS	X	X	X	WAV	TRS	X	MTM	X	1997.10.03	21	A	NONE
MTT_WS	X	X	X	WAV	TRS	X	MTT	X	1997.10.03	21	A	NONE
MTZ_WS	X	X	X	WAV	TRS	X	MTZ	X	1997.10.03	21	A	NONE
MYS_WS	X	X	X	WAV	TRS	X	MYS	X	1997.10.03	21	A	NONE

付録 E　研究用言語モデルの認識実験

E.1　今後の比較用の認識結果

一番単純な単語認識評価方法に基づく結果を示す.

- 単語のマージなし.
- 言い直し発声をテストセットに含む.

SLDB テストセット (SLTA1)

/dept1/work1/ResearchJ/V4/result/b95_7_15_SLTA1/*.res

	best word acc/cor	net word acc/cor	best word accuracy total	Ins	Del	Sub	No/Wo (cpu/utt, Li/Wo)	error utt.
TAS12002.A	84.02/ 84.54	92.27/ 92.27	194	1	4	26	3.07(2.20, 4.35)	0
TAS13002.B	73.64/ 77.52	86.43/ 88.76	258	10	15	43	5.00(5.90, 7.42)	0
TAS32012.A	84.46/ 87.16	91.89/ 93.24	148	4	5	14	3.71(2.47, 5.51)	0
TAS32014.A	80.00/ 84.44	90.37/ 93.33	135	6	3	18	4.48(2.40, 6.26)	0
TAS32016.A	59.62/ 63.46	86.54/ 89.42	104	4	8	30	9.20(7.71, 15.30)	0
TAS33010.B	77.78/ 81.48	90.12/ 91.36	162	6	5	25	3.78(2.20, 5.22)	0
TAS33015.B	59.91/ 64.06	74.19/ 79.72	217	9	12	66	5.70(2.93, 9.68)	0
TAS33020.B	72.00/ 72.89	84.89/ 86.22	225	2	18	43	4.58(3.57, 6.98)	0
TAC22013.A	87.31/ 89.34	92.39/ 95.94	197	4	3	18	2.51(1.47, 3.10)	0
TAC23022.B	85.83/ 87.50	91.67/ 93.33	120	2	4	11	2.67(2.05, 3.62)	0
TAS12006.A	79.49/ 85.90	91.03/ 94.87	78	5	2	9	5.14(1.83, 8.12)	0
TAS12014.A	86.31/ 86.90	94.05/ 94.05	168	1	7	15	4.27(3.45, 6.34)	0
TAS12016.A	82.35/ 87.39	96.64/ 97.48	119	6	2	13	4.60(4.30, 6.16)	0
TAS12019.A	85.54/ 85.54	96.99/ 96.99	166	0	9	15	3.24(2.21, 4.59)	0
TAS12020.A	82.49/ 83.05	89.27/ 89.83	177	1	15	15	3.33(4.47, 4.50)	0
TAS13014.B	78.47/ 79.20	91.24/ 91.61	274	2	19	38	4.64(15.89, 6.68)	0
TAS13018.B	69.37/ 72.97	88.29/ 91.89	111	4	6	24	7.46(7.78, 11.60)	0
TAS13019.B	85.95/ 87.60	94.21/ 95.45	242	4	6	24	2.89(2.11, 3.99)	0
TAS13020.B	69.47/ 73.89	84.51/ 87.17	226	10	15	44	6.22(7.55, 10.45)	0
TAS13022.B	77.96/ 80.65	88.71/ 90.32	186	5	8	28	6.25(5.65, 9.35)	0
TAS13024.B	75.26/ 78.35	88.66/ 90.21	194	6	10	32	5.76(4.35, 8.78)	0
TAS32001.A	80.36/ 83.04	88.39/ 91.96	112	3	4	15	2.53(1.42, 3.18)	0
TAS33022.B	85.94/ 88.02	93.23/ 94.27	192	4	5	18	2.65(1.30, 3.41)	0
male	73.87/ 76.78	86.49/ 88.70	1443	42	70	265	4.76(3.54, 7.29)	0
female	80.84/ 83.06	91.22/ 92.78	2562	57	115	319	4.25(4.89, 6.22)	0
both	78.33/ 80.80	89.51/ 91.31	4005	99	185	584	4.43(4.41, 6.60)	0

SDB テストセット (S1, S2, S4)

/dept1/work1/ResearchJ/V4/result/b95_7_15_SDB/*.res

	best word acc/cor	net word acc/cor	best word accuracy total	Ins	Del	Sub	No/Wo (cpu/utt, Li/Wo)	error utt.
TAC70016.A	70.79/ 71.91	80.90/ 82.02	89	1	7	18	4.56(3.05, 6.48)	0
TAC70017.A	90.77/ 95.38	96.92/ 98.46	65	3	0	3	4.15(2.11, 4.69)	0
TAC70021.A	83.81/ 87.62	92.38/ 92.38	105	4	3	10	3.09(1.51, 3.51)	0
TAC70022.A	69.23/ 71.33	81.82/ 83.92	143	3	12	29	3.89(2.74, 5.99)	0
TAC70023.A	86.99/ 91.87	92.68/ 94.31	123	6	2	8	2.70(1.92, 3.28)	0
TAC70103.A	81.58/ 82.89	90.79/ 92.11	76	1	4	9	2.84(2.63, 3.33)	0
TAC70202.A	71.14/ 71.14	87.25/ 87.25	149	0	9	34	5.07(7.04, 7.87)	0
TAC70304.A	66.18/ 80.88	77.94/ 89.71	68	10	0	13	6.41(4.23, 9.71)	0
TCC70109.A	68.18/ 75.00	84.09/ 85.23	88	6	5	17	6.08(3.92, 8.10)	0
TCC70201.A	43.90/ 52.44	62.20/ 67.07	82	7	7	32	6.59(3.64, 9.85)	0
TCC70212.A	56.52/ 63.35	72.67/ 78.88	161	11	14	45	6.75(3.41, 10.21)	0
TCC70307.A	76.87/ 78.91	86.39/ 88.44	147	3	5	26	5.40(3.70, 8.00)	0
TCC71008.A	54.97/ 61.99	78.95/ 82.46	171	12	21	44	6.40(2.37, 10.11)	0
TCS70034.A	65.58/ 73.38	85.06/ 90.91	154	12	11	30	5.64(1.96, 7.87)	0
TCS70055.A	68.24/ 68.82	78.82/ 80.59	170	1	14	39	3.21(2.72, 4.84)	0
TCS70070.A	68.49/ 73.97	82.19/ 86.30	73	4	1	18	7.62(2.30, 11.36)	0
TCS70074.A	87.80/ 90.24	96.34/ 97.56	82	2	2	6	2.67(1.02, 2.90)	0
TAC70015.A	83.18/ 85.98	94.39/ 96.26	107	3	4	11	5.88(1.68, 9.24)	0
TAC70019.A	81.00/ 82.00	92.00/ 92.00	100	1	4	14	3.08(1.35, 3.72)	0
TAC70101.A	88.24/ 89.92	98.32/ 98.32	119	2	6	6	2.35(0.97, 2.48)	0
TAC70102.A	89.29/ 94.29	93.57/ 97.14	140	7	0	8	2.12(0.99, 2.44)	0
TAC70201.A	67.97/ 72.66	86.72/ 87.50	128	6	9	26	3.61(1.21, 5.46)	0
TAC70203.A	68.53/ 72.73	81.12/ 85.31	143	6	9	30	3.83(1.28, 5.12)	0
TAC70301.A	83.33/ 89.17	95.00/ 96.67	120	7	3	10	3.16(1.42, 4.34)	0
TAC70303.A	87.50/ 89.17	93.33/ 95.83	120	2	5	8	2.43(0.91, 2.72)	0
TCC70103.A	84.81/ 84.81	89.87/ 89.87	79	0	3	9	3.58(1.21, 4.81)	0
TCC71001.B	76.32/ 79.47	91.05/ 92.11	190	6	7	32	3.56(1.72, 5.34)	0
TCC71007.A	90.20/ 92.81	98.04/ 99.35	153	4	4	7	2.26(0.97, 2.35)	0
TCC71016.A	69.01/ 78.17	85.92/ 93.66	142	13	0	31	4.20(1.32, 5.98)	0
TCC71035.A	73.91/ 78.26	91.30/ 91.30	92	4	3	17	4.76(2.07, 6.47)	0
TCS70004.B	62.17/ 63.67	76.03/ 77.15	267	4	37	60	6.99(13.27, 11.85)	0
TCS70010.A	48.15/ 53.70	67.59/ 70.37	108	6	12	38	5.37(2.77, 8.06)	0
TCS70013.A	75.86/ 85.06	87.36/ 94.25	87	8	1	12	3.54(1.43, 4.66)	0
TCS70020.A	56.90/ 62.07	75.00/ 81.03	116	6	14	30	5.68(5.00, 8.28)	0
TCS70023.A	70.87/ 75.73	89.81/ 93.20	206	10	8	42	3.99(1.43, 5.91)	0
TCS70025.A	75.00/ 79.63	88.89/ 90.74	108	5	8	14	3.37(1.10, 4.56)	0
TCS70028.A	86.36/ 93.18	93.18/ 95.45	88	6	1	5	2.88(1.10, 3.21)	0
TCS70047.A	63.92/ 69.07	80.41/ 87.63	97	5	7	23	3.88(1.37, 4.96)	0
TCS70059.A	74.39/ 76.83	84.15/ 87.80	82	2	5	14	4.45(1.55, 7.29)	0
TCS70082.A	82.69/ 82.69	94.23/ 94.23	156	0	3	24	4.73(2.47, 7.33)	0
TSC71005.B	76.42/ 80.49	91.06/ 93.50	123	5	3	21	5.31(1.92, 7.25)	0
TSC71013.A	66.82/ 73.09	80.27/ 84.75	223	14	8	52	6.41(5.71, 10.44)	0
male	69.99/ 74.41	83.40/ 86.28	1946	86	117	381	4.90(2.93, 7.03)	0
female	74.50/ 78.51	87.46/ 90.10	3294	132	164	544	4.23(2.55, 6.05)	0
both	72.82/ 76.98	85.95/ 88.68	5240	218	281	925	4.48(2.69, 6.42)	0

E.2 今後の比較用の設定ファイル

/dept1/work1/ResearchJ/V4/config/config.W

```
#I/Ocontrol config
        I/Ocontrol:rpcTable=
        I/Ocontrol:rpcNumber=3
        I/Ocontrol:outputByteorder=BigEndian
        I/Ocontrol:outputFd=stdout
        I/Ocontrol:outputParamType=
        I/Ocontrol:outputParamSize=
        I/Ocontrol:outputFormat=Lattice
        I/Ocontrol:inputByteorder=BigEndian
        I/Ocontrol:inputFd=stdin
        I/Ocontrol:inputParamType=float
        I/Ocontrol:inputParamSize=26
        I/Ocontrol:inputFormat=FrameSync

#ATRresult config
        ATRresult:answer=
        ATRresult:dp_weight=1.0,1.0,1.0
        ATRresult:pause_symbol=-
        ATRresult:UTT_END=6
        ATRresult:UTT_START=5
        ATRresult:re_beam=
        ATRresult:N_best=1
        ATRresult:N_best_out=stdout
        ATRresult:lattice_out=stdout

#ATRlattice config file
# TDMT
        ATRlattice:lexicon=/dept1/work1/ResearchJ/lmodel/19980521/LEX.W
        ATRlattice:ngram=Class-2,/dept1/work1/ResearchJ/lmodel/19980521/LM.W
        ATRlattice:lmscale=7.000000,15.000000
        ATRlattice:beam=95.000000,95.000000
        ATRlattice:amname=/dept1/work1/ResearchJ/V4/amodel/AM.MF.bin
#
        ATRlattice:work_area=4000,200
        ATRlattice:wdpenalty=0,0
        ATRlattice:frame_shift=10
#
        ATRlattice:active_model=all
        ATRlattice:pause_symbol=-
        ATRlattice:phone_boundary=ON
        ATRlattice:word_boundary_skip=2          # r05r05 以降    default is off
        ## ATRlattice:word_boundary_skip=ON       # r05r04 以前
        ATRlattice:word_merge=all
        ATRlattice:UTT_START=5
        ATRlattice:UTT_END=6
```

```
        ATRlattice:backward_frame=-1
#
        ATRlattice:dimension=26
        ATRlattice:max_allophone=5000
        ATRlattice:amscale=1.000000

### EOF
```

付録 F　MATRIX 用言語モデルの認識実験

F.1　最適パラメータの認識結果

/dept1/work1/MatrixJ/V3/result/B85_LM6_12.res/*.res

DS

	best word acc/cor	net word acc/cor	best word accuracy total	Ins	Del	Sub	speed No/Wo(cpu/u)	error utt.
MAN_DS.A	90.57/ 90.57	98.11/ 98.74	159	0	4	11	1.84(0.90)	0
MAY_DS.A	91.19/ 91.82	98.11/ 99.37	159	1	3	10	2.14(1.02)	0
MAZ_DS.A	88.05/ 88.68	96.23/ 96.23	159	1	7	11	2.78(0.96)	0
MFS_DS.A	88.05/ 88.05	97.48/ 98.11	159	0	7	12	2.15(0.89)	0
MGT_DS.A	91.82/ 91.82	99.37/100.00	159	0	2	11	1.76(0.81)	0
MHI_DS.A	83.65/ 83.65	93.08/ 93.08	159	0	10	16	2.75(1.18)	0
MHM_DS.A	84.91/ 86.16	95.60/ 96.86	159	2	3	19	3.13(1.10)	0
MHN_DS.A	90.57/ 91.19	98.74/100.00	159	1	2	12	1.87(0.88)	0
MHX_DS.A	75.47/ 75.47	84.28/ 84.28	159	0	19	20	3.40(1.70)	0
MHY_DS.A	89.94/ 91.19	97.48/ 99.37	159	2	3	11	2.19(0.93)	0
MHZ_DS.A	93.08/ 93.71	99.37/100.00	159	1	2	8	1.79(0.86)	0
MKT_DS.A	89.31/ 89.31	98.11/ 98.74	159	0	3	14	2.26(0.96)	0
MKZ_DS.A	93.08/ 93.08	99.37/100.00	159	0	4	7	2.01(0.81)	0
MMN_DS.A	72.33/ 75.47	86.79/ 89.31	159	5	8	31	3.98(1.17)	0
MMS_DS.A	92.45/ 93.08	97.48/ 98.11	159	1	2	9	1.77(0.91)	0
MNA_DS.A	91.82/ 91.82	98.74/100.00	159	0	2	11	2.02(0.89)	0
MRI_DS.A	91.19/ 91.19	99.37/100.00	159	0	2	12	1.89(0.76)	0
MST_DS.A	91.82/ 91.82	98.74/ 99.37	159	0	3	10	1.87(0.88)	0
MTB_DS.A	84.91/ 84.91	97.48/ 98.11	159	0	5	19	2.96(1.22)	0
MTM_DS.A	81.13/ 82.39	91.82/ 93.08	159	2	14	14	3.45(1.01)	0
MTT_DS.A	93.71/ 93.71	98.74/ 99.37	159	0	2	8	1.72(0.74)	0
MTZ_DS.A	91.19/ 92.45	98.11/100.00	159	2	3	9	2.01(0.87)	0
FAH_DS.A	93.71/ 93.71	99.37/100.00	159	0	2	8	1.82(0.78)	0
FCM_DS.A	91.82/ 91.82	98.74/ 99.37	159	0	4	9	2.19(0.94)	0
FHK_DS.A	86.79/ 88.68	94.97/ 97.48	159	3	6	12	2.19(0.85)	0
FIL_DS.A	91.19/ 91.19	98.11/ 98.74	159	0	2	12	1.82(0.89)	0
FKO_DS.A	87.42/ 89.31	94.34/ 96.86	159	3	4	13	2.47(1.05)	0
FKT_DS.A	93.71/ 93.71	97.48/ 97.48	159	0	2	8	1.66(0.72)	0
FMN_DS.A	90.57/ 90.57	99.37/100.00	159	0	2	13	2.15(0.79)	0
FTM_DS.A	93.08/ 93.08	98.11/ 98.74	159	0	2	9	1.69(0.80)	0
FYD_DS.A	93.71/ 93.71	99.37/100.00	159	0	2	8	1.84(0.70)	0
FYO_DS.A	89.31/ 89.94	96.86/ 98.74	159	1	5	11	1.90(0.80)	0
FYS_DS.A	91.19/ 91.82	96.86/ 98.11	159	1	4	9	1.74(0.81)	0
male	88.19/ 88.71	96.48/ 97.37	3498	18	110	285	2.35(0.97)	0
female	91.14/ 91.60	97.60/ 98.68	1749	8	35	112	1.95(0.83)	0
both	89.17/ 89.67	96.86/ 97.81	5247	26	145	397	2.22(0.92)	0

IS

	best word acc/cor	net word acc/cor	best word accuracy total	Ins	Del	Sub	speed No/Wo(cpu/u)	error utt.
MAN_IS.A	95.56/ 96.94	97.78/ 98.06	360	5	0	11	1.69(0.78)	0
MAY_IS.A	94.17/ 95.28	98.89/ 98.89	360	4	5	12	2.42(1.19)	0
MAZ_IS.A	94.17/ 95.83	98.06/ 98.89	360	6	4	11	2.50(0.93)	0
MFS_IS.A	93.06/ 94.17	98.89/ 98.89	360	4	7	14	2.23(0.96)	0
MGT_IS.A	93.33/ 95.00	97.50/ 98.06	360	6	5	13	1.86(0.82)	0
MHI_IS.A	88.33/ 91.11	96.94/ 97.50	360	10	9	23	3.09(1.04)	0
MHM_IS.A	85.00/ 87.22	95.00/ 95.56	360	8	14	32	3.40(1.14)	0
MHN_IS.A	91.94/ 93.06	96.67/ 96.94	360	4	9	16	2.32(1.01)	0
MHT_IS.A	92.22/ 94.44	96.11/ 97.78	360	8	5	15	2.13(0.87)	0
MHX_IS.A	80.83/ 81.94	88.61/ 89.72	360	4	26	39	4.20(1.65)	0
MHY_IS.A	93.61/ 94.72	99.17/ 99.17	360	4	4	15	2.06(0.89)	0
MHZ_IS.A	95.00/ 96.39	98.61/ 98.89	360	5	3	10	2.17(0.90)	0
MKT_IS.A	94.17/ 95.28	98.89/ 98.89	360	4	3	14	2.36(1.01)	0
MKZ_IS.A	93.89/ 96.67	98.06/ 99.17	360	10	1	11	2.16(0.82)	0
MMN_IS.A	78.89/ 81.39	87.78/ 88.89	360	9	21	46	3.67(1.20)	0
MMS_IS.A	94.72/ 96.39	98.33/ 98.61	360	6	2	11	2.23(0.91)	0
MNA_IS.A	95.00/ 97.22	96.94/ 98.33	360	8	1	9	1.94(0.96)	0
MRI_IS.A	94.17/ 95.83	98.06/ 98.33	360	6	4	11	2.06(0.91)	0
MST_IS.A	96.39/ 98.33	99.72/ 99.72	360	7	0	6	1.75(0.92)	0
MSY_IS.A	85.28/ 87.22	92.50/ 93.06	360	7	14	32	3.12(1.06)	0
MTB_IS.A	84.72/ 88.06	93.89/ 96.39	360	12	7	36	4.03(1.73)	0
MTM_IS.A	83.33/ 85.28	90.83/ 91.67	360	7	18	35	3.46(1.06)	0
MTT_IS.A	96.67/ 98.06	98.61/ 98.89	360	5	1	6	1.93(0.80)	0
MTZ_IS.A	93.61/ 96.67	98.06/ 98.89	360	11	1	11	2.29(0.85)	0
MYS_IS.A	89.17/ 91.67	95.28/ 96.11	360	9	11	19	2.57(0.88)	0
FAH_IS.A	95.28/ 96.67	97.78/ 98.06	360	5	2	10	1.89(0.74)	0
FCM_IS.A	91.39/ 94.44	95.83/ 96.94	360	11	2	18	2.29(0.87)	0
FHK_IS.A	92.50/ 94.72	97.50/ 98.33	360	8	2	17	2.10(0.83)	0
FII_IS.A	96.11/ 97.50	99.44/ 99.44	360	5	2	7	2.10(0.93)	0
FKO_IS.A	95.28/ 96.94	99.72/ 99.72	360	6	1	10	2.02(0.82)	0
FKT_IS.A	97.50/ 98.61	98.33/ 99.17	360	4	0	5	1.99(0.78)	0
FMN_IS.A	90.56/ 93.61	95.00/ 96.39	360	11	3	20	3.30(0.98)	0
FTM_IS.A	96.94/ 98.33	99.17/ 99.17	360	5	0	6	1.97(0.80)	0
FYD_IS.A	96.67/ 97.78	99.17/ 99.44	360	4	1	7	1.79(0.74)	0
FYO_IS.A	95.56/ 97.22	98.06/ 98.61	360	6	2	8	1.62(0.74)	0
FYS_IS.A	95.00/ 96.67	97.78/ 98.61	360	6	2	10	2.12(0.79)	0
FYZ_IS.A	88.06/ 90.28	94.17/ 94.72	360	8	8	27	2.56(0.91)	0
male	91.09/ 92.97	96.37/ 97.01	9000	169	175	458	2.55(1.00)	0
female	94.24/ 96.06	97.66/ 98.22	4320	79	25	145	2.15(0.83)	0
both	92.11/ 93.97	96.79/ 97.40	13320	248	200	603	2.42(0.94)	0

WS

	best word acc/cor	net word acc/cor	best word accuracy total	Ins	Del	Sub	speed No/Wo(cpu/u)	error utt.
MAN_WS.A	88.51/ 90.80	93.10/ 94.25	174	4	3	13	2.36(0.81)	0
MAY_WS.A	85.06/ 87.36	93.68/ 94.25	174	4	4	18	3.77(1.19)	0
MAZ_WS.A	84.48/ 85.63	94.83/ 95.40	174	2	6	19	3.04(1.06)	0
MFS_WS.A	79.31/ 81.03	89.08/ 90.80	174	3	9	24	3.02(1.11)	0
MGT_WS.A	89.66/ 90.23	95.40/ 95.98	174	1	3	14	2.12(0.90)	0
MHI_WS.A	75.86/ 76.44	87.36/ 87.36	174	1	9	32	3.55(1.25)	0
MHM_WS.A	72.99/ 76.44	85.63/ 86.78	174	6	11	30	4.45(1.54)	0
MHN_WS.A	85.63/ 85.63	92.53/ 92.53	174	0	5	20	3.44(1.09)	0
MHT_WS.A	70.69/ 78.16	82.18/ 87.93	174	13	4	34	4.47(1.09)	0
MHX_WS.A	58.62/ 59.77	75.86/ 77.59	174	2	26	44	7.83(2.31)	0
MHY_WS.A	87.36/ 87.36	93.68/ 94.25	174	0	5	17	2.99(1.04)	0
MHZ_WS.A	89.08/ 90.23	94.25/ 94.83	174	2	2	15	3.52(1.12)	0
MKT_WS.A	83.91/ 86.78	91.95/ 93.68	174	5	5	18	3.32(1.11)	0
MKZ_WS.A	87.93/ 88.51	94.83/ 95.98	174	1	4	16	2.66(0.82)	0
MMN_WS.A	72.41/ 74.71	86.21/ 88.51	174	4	9	35	4.25(1.17)	0
MMS_WS.A	91.38/ 91.95	94.25/ 94.83	174	1	2	12	2.95(1.03)	0
MNA_WS.A	87.36/ 90.80	91.38/ 93.68	174	6	2	14	2.57(1.03)	0
MRI_WS.A	87.36/ 88.51	95.98/ 96.55	174	2	6	14	3.45(0.96)	0
MST_WS.A	89.66/ 90.23	94.83/ 95.40	174	1	2	15	2.61(0.99)	0
MSY_WS.A	83.33/ 83.91	90.80/ 91.95	174	1	8	20	3.80(1.18)	0
MTB_WS.A	78.16/ 80.46	88.51/ 89.66	174	4	2	32	7.11(2.11)	0
MTM_WS.A	62.07/ 67.82	79.89/ 82.76	174	10	15	41	5.08(1.33)	0
MTT_WS.A	86.21/ 86.21	93.10/ 93.10	174	0	3	21	2.29(0.90)	0
MTZ_WS.A	83.91/ 87.36	92.53/ 94.25	174	6	4	18	2.41(0.95)	0
MYS_WS.A	89.66/ 90.23	94.25/ 94.83	174	1	2	15	2.44(0.80)	0
FAH_WS.A	90.80/ 91.95	95.40/ 95.40	174	2	2	12	2.66(0.87)	0
FCM_WS.A	84.48/ 87.36	92.53/ 93.10	174	5	4	18	3.75(1.09)	0
FHK_WS.A	79.31/ 81.61	92.53/ 93.68	174	4	6	26	2.63(0.95)	0
FII_WS.A	86.21/ 87.93	90.80/ 91.38	174	3	2	19	2.75(1.07)	0
FKO_WS.A	86.78/ 87.93	95.40/ 95.98	174	2	5	16	2.80(0.92)	0
FKT_WS.A	90.23/ 92.53	95.98/ 96.55	174	4	1	12	3.03(1.05)	0
FMN_WS.A	87.93/ 88.51	93.10/ 94.25	174	1	4	16	4.05(1.06)	0
FTM_WS.A	89.08/ 89.08	94.25/ 94.83	174	0	2	17	2.59(0.91)	0
FYD_WS.A	88.51/ 89.66	94.25/ 95.40	174	2	2	16	2.58(0.87)	0
FYO_WS.A	88.51/ 89.08	91.95/ 92.53	174	1	4	15	2.56(0.91)	0
FYS_WS.A	90.80/ 92.53	94.83/ 95.40	174	3	1	12	2.80(0.98)	0
FYZ_WS.A	82.18/ 82.76	90.23/ 90.80	174	1	6	24	3.60(1.10)	0
male	82.02/ 83.86	90.64/ 91.89	4350	80	151	551	3.58(1.15)	0
female	87.07/ 88.41	93.44/ 94.11	2088	28	39	203	2.98(0.98)	0
both	83.66/ 85.34	91.55/ 92.61	6438	108	190	754	3.39(1.09)	0

DS+IS+WS

DS+IS+WS	best word acc/cor	net word acc/cor	best word accuracy total	Ins	Del	Sub	speed No/Wo(cpu/u)	error utt.
male	88.15/ 89.73	94.91/ 95.76	16848	267	436	1294	2.77(1.03)	0
female	91.74/ 93.15	96.57/ 97.27	8157	115	99	460	2.32(0.87)	0
both	89.32/ 90.85	95.45/ 96.25	25005	382	535	1754	2.62(0.98)	0

F.2 設定ファイル

/dept1/work1/MatrixJ/V3/config/config

```
# ATRSPREC r05r04 Config file for MATRIXJ V3 (Optimized to DEC Alpha 500)
        I/Ocontrol:inputFormat=NULL
        I/Ocontrol:inputParamSize=160
        I/Ocontrol:inputParamType=short
        I/Ocontrol:inputFd=stdin
        I/Ocontrol:inputByteorder=BigEndian
        I/Ocontrol:outputFormat=NULL
        I/Ocontrol:outputFd=stdout
        I/Ocontrol:outputByteorder=BigEndian
        I/Ocontrol:rpcNumber=2

        ATRwave2cep:CutoffHighFrequency=8000.0
        ATRwave2cep:CutoffLowFrequency=0.0
        ATRwave2cep:FilterBankOrder=16
        ATRwave2cep:FrequencyWarping=mel
        ATRwave2cep:inputParameter=waveRaw
        ATRwave2cep:AnalysisType=fft
        ATRwave2cep:CepstrumOrder=12
        ATRwave2cep:FilterBankOrder=16
        ATRwave2cep:LagWindowFactor=0.01
        ATRwave2cep:TimeWindow=hamming
        ATRwave2cep:SamplingFrequency=16000
        ATRwave2cep:FrameShift=10
        ATRwave2cep:FrameLength=20
        ATRwave2cep:Preemphasis=0.98
        ATRwave2cep:DebuggingLevel=0

        ATRcep2para:OutputParameter=pow+cep(12)+dpow+dcep(12)
        ATRcep2para:rho=1.0
        ATRcep2para:DDCepstrumPadding=zero
        ATRcep2para:deltaCepstrumPadding=zero
        ATRcep2para:DeltaCepstrumWindow=9
        ATRcep2para:CepstrumOrder=12
        ATRcep2para:DebuggingLevel=10

        ATRviterbi:amname=/dept1/work1/MatrixJ/V3/amodel/AM.M.803.5.bin,
                           /dept1/work1/MatrixJ/V3/amodel/AM.F.803.5.bin
        ATRviterbi:active_model=all
        ATRviterbi:beam=0
```

```
ATRviterbi:frame_shift=10
ATRviterbi:dimension=26
ATRviterbi:boundary=OFF

ATRresult:N_best=1
ATRresult:N_best_out=/dev/null
ATRresult:Lattice_out=/dev/null

##
## Be careful that 'ATRlattice:lexicon' and 'ATRlattice:ngram'
## will be overwrited by DoRecog.csh in MatrixJ V3 experiment
##
ATRlattice:active_amname=all
ATRlattice:lexicon=/dept1/work1/MatrixJ/V3/lmodel/RECOG_LEX.vn1000
ATRlattice:amname=/dept1/work1/MatrixJ/V3/amodel/AM.M.803.5.bin,
                  /dept1/work1/MatrixJ/V3/amodel/AM.F.803.5.bin
ATRlattice:lmscale=6.000000,12.000000
ATRlattice:wdpenalty=10000,10000
ATRlattice:ngram=Class-2,
           /dept1/work1/MatrixJ/V3/lmodel/vngram.sub.dat.equal.bin
ATRlattice:beam=85.0,85.0
ATRlattice:work_area=1000,60
ATRlattice:frame_shift=10
ATRlattice:pause_symbol=-
ATRlattice:dimension=26
ATRlattice:max_allophone=5000
ATRlattice:phone_boundary=ON
ATRlattice:word_merge=all
ATRlattice:UTT_START=5
ATRlattice:UTT_END=6
ATRlattice:backward_frame=-1
ATRlattice:amscale=1.000000
ATRlattice:FSA=
ATRlattice:null_trans=OFF
ATRlattice:UTT_END_delay=-1
ATRlattice:word_boundary_skip=ON          # r05r04 以前
##ATRlattice:word_boundary_skip=2          # r05r05 以降


#EOF
```