

# 単語 bigram とクラス bigram の融合による話題変化に頑健な 統計的言語モデル

肥田木 康明 小林 哲則

早稲田大学 理工学部

あらし:

統計的言語モデルは大量の学習データを用いて推定される。しかし学習データにおける単語連鎖の統計的上の性質(以後言語的特徴と呼ぶ)は、学習データの収集された時期や話題に大きく依存している。このため汎用の音声認識システム構築の際には、言語的特徴変化に頑健な言語モデルの作成が不可欠である。本研究では、より精密な言語的特徴を表現できる単語 bigram と、あまり精密ではないが言語的特徴の変化に頑健なクラス bigram を融合することにより性能の安定した言語モデルを構築することを試みた。また、単語 bigram、クラス bigram、融合 bigram を用いた認識実験により、クラス bigram の言語的特徴の変化に対する頑健性と融合 bigram による性能向上について評価した。

## Combination of Class Bigram and Word Bigram for Topic-Robust Language Modeling

Yasuaki HIDAKI, Tetsunori KOBAYASHI

School of Science and Engineering, Waseda University

**Abstract:**

Linguistic feature (statistical characteristics of the word sequence) of the corpus extremely depends on the topics or the time when the corpus was collected. Therefore, it is important to generate language model which is robust against the change of linguistic feature when we develop speech recognition system for general purpose. This paper proposes some new approaches for generate language models which are robust against the change of linguistic feature by combining class bigram and word bigram. And we evaluate the robustness of the class bigram and advantage of the hybrid bigram through the comparison test.

### 1 はじめに

近年、新聞記事読み上げデータベースのような大規模な音声言語データベースが利用できるようになり、言語モデルの高精度な初期学習が可能になった。このことにより、タスクを限定すればかなり高精度な連続音声認識が可能になっている。

言語モデルには一般的に単語 N-gram モデルが使用されることが多い。これは大量の学習データを用いて推定されるが、次のような問題点がある。

- 学習データに現れなかった単語列は、出現頻度が 0 と推定され、そのような単語列が実際に発話されても認識できない。

連絡先: 〒169 東京都新宿区大久保 3-4-1 早稲田大学理工学部 電気電子情報工学科 小林 哲則  
TEL: (03)5286-3379 FAX: (03)3205-9381 E-mail: koba@tk.elec.waseda.ac.jp

- 学習データ（あるいは認識対象）の言語的「くせ」により、ある単語列の出現確率が不当に低く、または高く推定されてしまうことがある。

これらの問題を解決するために、従来からさまざまな手法が用いられている。主に前者に対するものとして、確率推定値0のものに対し、一定の小さな値を割り当てる flooring 法、N-gram 確率が低い（または0）のものに対し(N-1)-gram 確率を用いて推定値を補正する back-off 平滑化法 [1][3] などがあり、後者に対するものとして、大量の学習データで学習した一般的な言語モデルと認識タスクに特化した少量の学習データで作成した適応モデルを一定の重みでたし合わせる方法 [5] や、性質の良く似た単語で単語クラスを作り、クラスどうしの接続確率を用いるクラス N-gram 法 [6] などがある。

本研究では言語モデルとして、より精密である単語 bigram、およびあまり精密でないが言語的特徴変化に頑健なクラス bigram を採用し、これらを融合することによって話題やデータの収集時期による言語的特徴の変化に対して頑健な言語モデルの構築を試みる。

## 2 bigram 言語モデルの学習

学習データが無限にあるなら（対象とする言語の言語的特徴をすべて含み、すべての単語連鎖について出現確率を推定するのに十分な標本数を備えているなら）、単語 bigram に対するクラス bigram の優位性は見当たらない。

しかし、実際には学習データは有限であり、上記の前提は成り立たない。このためクラス bigram は単語 bigram に対して次のような優位性があると考えられる。

1. いくつかの単語をひとまとめにした単語クラスの持つ言語的特徴を扱うため、話題や時期などによる言語的「くせ」に対して頑健である。
2. 単語 bigram よりクラス bigram の方が組合せの数が少ないために単語 bigram 確率を推定する時よりも多くの標本が得られ、推定された bigram 確率の信頼性が高い。

たとえば次のような単語列が発話された場合、

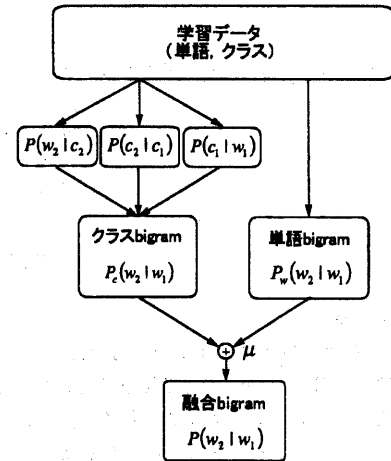


図 1: 単語 bigram とクラス bigram の融合

橋本 総理 大臣  
[固有名詞:人名] [名詞] [名詞]

単語 bigram を用いた場合、学習データが橋本氏が総理在任中の新聞記事などの場合は問題ないが、そうでない場合は、 $P(\text{総理} | \text{橋本})$  はほとんど0になってしまい認識できないことが予想される。一方これに対し、クラス bigram を用いた場合には、 $P([\text{名詞}] | [\text{固有名詞} : \text{人名}])$  が適用されるため、こうした問題は起きにくいことが期待できる。

また、

従業員 43 名  
[名詞] [名詞:数] [名詞:接尾:助数詞]

のような単語列も、単語 bigram よりもクラス bigram を用いた方が有利であると考えられる。

以上の仮定を基に、本研究では単語 bigram とクラス bigram をいくつかのアルゴリズムで重み付けし、融合することによってこれらの問題を解決することを試みる（図1）。

### 2.1 bigram 言語モデル学習条件

bigram 言語モデルの学習にあたり学習データとして、RWC テキストデータベース [7] および、CD-

毎日新聞データ集 [8] のうち 1994 年の新聞記事 1 年分 (943,870 文) を使用した。

単語としては形態素を採用し、単語クラスとしては基本的に RWC テキストデータベースで使用されている品詞体系 (THiMCO95) を採用した。ただし、付属語 (助詞, 助動詞) についてはクラスを作らない (それぞれ一つの単語で一つのクラスを作る) こととした。これは付属語は決まり文句的な言い回しが多く種類数も少ないため [9] 複数の付属語でクラスを作ることの優位性があまりないと予想したためである。

この結果、単語クラス数は 557 となった。

さらに、読み上げ時に発音されない”、。「」▽などの記号、および難読漢字の読みがなや捕捉説明に用いられる” ( ) ” で囲まれた部分を取り除いたものを学習データとして使用した。

語彙は上記学習データに含まれる単語の内、出現頻度上位 5000 語 + <silB> (文頭無音区間) + <silE> (文末無音区間) の 5002 語とした。ただし、これらの単語には複数の発音が考えられるもの (例: 「通っ」 = かよっ / とおっ) があるので、それらを発音ごとに展開し、実際には 5314 語とした。

## 2.2 単語 bigram モデル

単語 bigram 確率は以下の式を用いて推定する。

$$P_w(w_2|w_1) = \frac{c(w_1, w_2)}{c(w_1)} \quad (1)$$

ここで  $c(W)$  は引数で指定された単語または単語列が学習データ中に出現した回数を表す。

## 2.3 クラス bigram モデル

クラス bigram そのものは、単語 bigram と同様、以下のようにして推定する。

$$P(c_2|c_1) = \frac{c(c_1, c_2)}{c(c_1)} \quad (2)$$

ここで  $c_k$  は単語  $w_k$  がそのコンテキストで属していた単語クラスである。

しかしこのままでは単語 bigram と融合する際に不都合であるので、以下のようにして単語 bigram の形式 (単語から単語への遷移確率の形式) に拡張し、以後これをクラス bigram と呼ぶことにする。

$$P_c(w_2|w_1) = \sum_{i,j} P(w_2|c_2^i)P(c_2^i|c_1^j)P(c_1^j|w_1) \quad (3)$$

ここで  $P(c_k|w_k)$  は単語  $w_k$  がクラス  $c_k$  に属する単語として出現する確率

$$P(c_k|w_k) = \frac{c(w_k : c_k)}{c(w_k)} \quad (4)$$

であり、 $P(w_k|c_k)$  はクラス  $c_k$  を仮定した時、単語  $w_k$  を出力する確率

$$P(w_k|c_k) = \frac{c(w_k : c_k)}{c(c_k)} \quad (5)$$

である (ただし、 $c(w_k : c_k)$  は単語  $w_k$  がクラス  $c_k$  として出現した回数を表す)。一つの単語がコンテキストによって異なるクラスに属することがあるので式 (3) のように  $w_1, w_2$  が属するクラスすべての組合せで和をとる。

## 2.4 単語 bigram とクラス bigram の融合

ここで、単語 bigram とクラス bigram の性質の違いとして予想されることを表 1 にまとめてみる。

表 1: 単語 bigram とクラス bigram の性質

	単語 bigram	クラス bigram
認識単位種類数	多い	少ない
表現力	高い	低い
話題への頑健性	低い	高い
確率推定精度	低い	高い

単語 bigram の方が表現力が高いと考えられるので、基本的には単語 bigram の方を優先し、タスクによって出現確率が大きく変わってしまうような単語列や、サンプル数が少なくて推定精度が落ちてしまうような単語列についてはクラス bigram を優先したい。そこで以下の 2 つの基準で bigram 融合の重み関数  $\mu$  を定義することにする。

### 2.4.1 単語 unigram 頻度を基準とする方法

単語 bigram 確率は式 1 により推定される。ここで真の単語 bigram 確率を  $P_w(w_2|w_1) = p$  とすると、単語 bigram 頻度  $c(w_1, w_2)$  の分布は 2 項分布  $B(c(w_1), p)$  に従い、

$$P(X = c(w_1, w_2))$$

$$= \binom{c(w_1)}{c(w_1, w_2)} p^{c(w_1, w_2)} (1-p)^{c(w_1) - c(w_1, w_2)} \quad (6)$$

となり、分散は、

$$\sigma^2(X) = c(w_1)p(1-p) \quad (7)$$

である。分散に関する公式より、

$$\sigma^2(aX + b) = a^2\sigma^2(X) \quad (8)$$

が成り立つので、単語 bigram 確率の推定値

$P_w(w_2|w_1) = \frac{c(w_1, w_2)}{c(w_1)}$  の分散は、

$$\sigma^2\left(\frac{X}{c(w_1)}\right) = \frac{p(1-p)}{c(w_1)} \quad (9)$$

となる。

これは単語 unigram 頻度  $c(w_1)$  が  $\alpha$  倍になると、同じ信頼区間での誤差範囲が  $\frac{1}{\sqrt{\alpha}}$  になることを意味する。このことから、単語 bigram の重みを  $c(w_1)$  の関数としておき、 $c(w_1)$  が大きくなるほど重くなるように与えることが理にかなうと考えられる。そこで、重み関数として以下の5種類を定義する。

$$\begin{aligned} \text{hybrid1: } & \mu(w_1) = 1 - c(w_1)^{-0.5} \\ \text{hybrid2: } & \mu(w_1) = 1 - c(w_1)^{-0.3} \\ \text{hybrid3: } & \mu(w_1) = 1 - c(w_1)^{-0.1} \\ \text{hybrid4: } & \mu(w_1) = 1 - (1 + 0.03c(w_1))^{-0.5} \\ \text{hybrid5: } & \mu(w_1) = 1 - (1 + 0.003c(w_1))^{-0.5} \end{aligned} \quad (10)$$

この方法の基本的な考え方は、「 $w_1$  の出現頻度の低いものは、単語 bigram が信頼できないのでクラス bigram の重みを大きくする」というものである。

以上の関数と単語 unigram 頻度の関係を図2に示す。

これらの重み関数を用いて融合 bigram を計算する(式11)。

$$\begin{aligned} P(w_2|w_1) = & \mu(w_1)P_w(w_2|w_1) \\ & + (1 - \mu(w_1))P_c(w_2|w_1) \end{aligned} \quad (11)$$

#### 2.4.2 単語 bigram 確率を基準とする方法

N-gram の back-off 平滑化の基本的な考え方は、「N-gram 確率推定値が小さいものは信頼できないのでカットオフしてより低次の推定値で補間する」

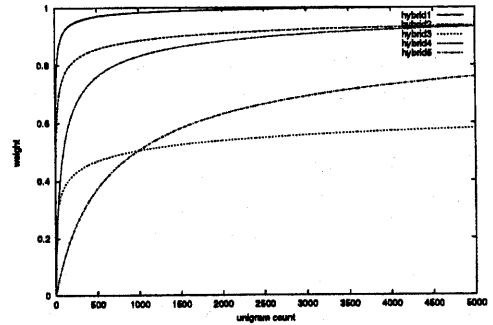


図2: 単語 bigram の重みと単語 unigram 頻度

というものである。この考え方は、先の例で示したような、不当に低く見積もられてしまった単語 bigram 確率推定値を補間するという意味でも有効であると考えられる。そこで、単語 bigram 確率を基準とする重み関数を考える。

単語 bigram 確率推定値は0から1の間に分布するが、実際の分布を調べると、1回以上出現した単語列のうち、約99.7%が0.01以下であった。そこで、この確率のlogをとったものの分布を調べた(図3)。

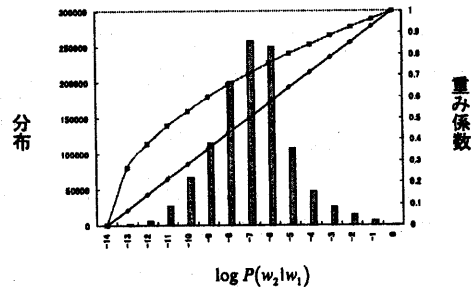


図3: log 単語 bigram 確率の分布と重み関数

これを基に重み関数を以下のように定義する。

$$\text{hybrid6: } \mu(w_1, w_2)$$

$$\text{hybrid7: } \mu(w_1, w_2) = \begin{cases} 0 & (\log P_w(w_2|w_1) < -14) \\ \frac{1}{14} \log P_w(w_2|w_1) + 1 & (-14 \leq \log P_w(w_2|w_1)) \end{cases}$$

$$= \begin{cases} 0 & (\log P_w(w_2|w_1) < -14) \\ \sqrt{\frac{1}{14} \log P_w(w_2|w_1) + 1} & (-14 \leq \log P_w(w_2|w_1)) \end{cases} \quad (12)$$

これらの重み関数を用いて融合 bigram を計算する (式 13)。

$$P'(w_2|w_1) = \mu(w_1)P_w(w_2|w_1) + (1 - \mu(w_1))P_c(w_2|w_1)$$

$$P(w_2|w_1) = \frac{P'(w_2|w_1)}{\sum_{w_1, w_2} P'(w_2|w_1)} \quad (13)$$

### 3 認識実験

言語モデル学習データと同一タスクのもの、異なるタスクのもの 2 種類の評価データについて、前章で作成した各 bigram 言語モデル、および比較のための単語 bigram モデル (onlyword)、クラス bigram モデル (onlyclass) を用いて認識実験を行ない、言語モデルの言語的特徴変化に対する頑健性について評価した。

実験諸条件、および評価データの詳細を以下に示す。

#### 3.1 実験条件

音響モデルは、文脈独立の音素 HMM (5 状態 3 ループ、連続型、混合数 4) 36 種および文頭無音区間 (silB)、文末無音区間 (silE) の合計 36 種とし、特徴量は mfcc 12 次元 + log パワーおよびその  $\Delta$  の 26 次元とした。また、音響モデルの学習には日本音響学会の新聞記事読み上げ音声コーパス [10] から男性話者 100 名分 10,448 文を使用した。

認識アルゴリズムとしては、Onepass Viterbi アルゴリズムを使用し、言語モデルと音響モデルの重みは 7:1 とした。

#### 3.2 評価データ

言語モデルの学習データと同一タスク (言語的特徴が同じ) の評価データとして、日本音響学会の毎日新聞記事読み上げデータから言語モデルの語彙で閉じているもの 94 文 (男性話者 10 名, 1992 ~ 1994 年のもの) を使用した (newspaper94)。

言語モデルの学習データと異なるタスク (言語的特徴が異なる) の評価データとして、NHK ニュース原稿 [11](1996 年のもの) から、言語モデルの語彙で閉じていて、1 文が 40 単語以内のもの 120 文 (男性話者 1 名) を使用した (NHK120)。

#### 3.3 結果

認識結果の単語認識率を表 2 に示す。

ただし、正解率

$$\text{Correct} = \frac{H}{N} \times 100[\%] \quad (14)$$

正解精度

$$\text{Accuracy} = \frac{H - I}{N} \times 100[\%] \quad (15)$$

である。

正解精度で見ると、言語モデル学習データと異なるタスク (NHK120) の時は hybrid3 のモデルが最も性能が良く、単語 bigram モデルと比べて 1.35%、同一タスク (newspaper94) の場合は hybrid4 のモデルが最も性能が良く単語 bigram モデルと比べて 1.32% 向上していることがわかる。

### 4 short pause の導入による認識実験

表 2 を見ると、どのモデルでも NHK120 タスクでは、認識精度が 10% 程度下がってしまっていることがわかる。

ここでそれぞれの評価データの平均文長を見てみると、newspaper94 タスクでは 9.63 単語なのに対し、NHK120 タスクでは 17.26 単語となり、NHK120 タスクの方が平均文長が 2 倍近く長いことがわかる。今回の認識システムでは short pause (息継ぎなどによる無音部) を考慮していないが、NHK120 タスクでは newspaper94 タスクよりも多くの short pause が含まれていると考えられる。こ

のため、NHK120 タスク時における性能低下原因の一つとして short pause が考えられる。

そこで、short pause に相当する音素モデル sp および一つの sp からなる単語<sp>を導入し、同様の認識実験を行なう。

#### 4.1 実験条件

音響モデルの学習時には、発話内容を示すラベルに short pause の情報が含まれていないので、直接音素 sp を学習することはあきらめ、silB のモデルで代用する。

単語 bigram モデルの学習時には、従来無意味語として無視していた記号のうち、通常一旦区切って読まれる”、[ ]・”の4種類を単語<sp>として学習する。またクラス bigram モデル学習時には単語<sp>のみからなるクラスを設けて学習を行なう。

#### 4.2 結果

short pause を導入したモデルで行なった認識結果を表3に示す。

各モデルともおおむね short pause なしの場合より良好な結果が得られているが、newspaper94 タスクの方が改善の度合いが顕著であることがわかる。また、各タスクにおいて単語 bigram と融合 bigram を比較すると、short pause 有りのモデルでは融合 bigram の優位性が薄れているものが多い。

今回の実験では、単語 bigram に対して認識精度が最も向上しているものは、NHK120 タスクでは hybrid5 モデルで 1.54%、newspaper94 タスクでは hybrid2 モデルで 0.11%であった。

### 5 考察

#### 5.1 short pause について

short pause を導入することにより、いずれのタスクでもおおむね性能が向上していることから short pause 導入の有効性が示されたが、newspaper94 タスクでは認識精度が単語 bigram モデルで 5.30%向上しているのに対して、NHK120 タスクでは 0.77%にとどまっております、NHK120 タスクは文長が長く、short pause が多いために性能が低下するという仮定は立証されなかった。この原因

を調べるために単語 bigram 確率の推定値が 0 になっていない単語列の種類数(学習データに含まれていた単語列の種類数)を調べた。この結果 short pause なしのモデルでは、全組合せ  $5,314 \times 5,314 = 28,238,596$  通りのうち 1,125,815 通り (3.99%) だったのに対し、short pause ありのモデルでは全組合せ 28,249,225 通りのうち 887,562 通り (3.14%) であった。これは short pause なしのモデルでは観測されていた単語列が、<sp>の導入により多く断ち切られたためである。このことにより、newspaper94 タスクではより強い言語モデルの制約で有利に正解単語列を導けたのに対し、NHK120 タスクでは出現した単語列の bigram 確率推定値が低くなり認識されにくくなった(言語モデルが言語的特徴の変化に敏感になった)と考えられる。

#### 5.2 クラス bigram モデルの言語的特徴に対する頑健性

クラス bigram モデルは、言語的特徴の変化に対して頑健であるという仮定で bigram の融合を行っていたが、このことを検証する(表4)。表4

表4: タスクによる認識精度の変化

<sp>	モデル	(a)	(b)	(c)
なし	onlyword	45.15	57.13	20.97
	onlyclass	39.50	46.41	14.89
あり	onlyword	45.92	62.43	26.44
	onlyclass	40.61	51.82	21.63

(a) NHK120 [%]

(b) newspaper94 [%]

(c)  $\frac{(b)-(a)}{(b)} \times 100$  [%]

より、newspaper94 タスクから NHK120 タスクに変更した時の認識精度の減少率が<sp>ありの場合でもなしの場合でも、単語 bigram モデルの時よりクラス bigram モデルの方が少なかった。このことから、相対的にはクラス bigram モデルの方が単語 bigram モデルよりも言語的特徴変化に対して頑健なことがうかがえるが、融合 bigram モデルにおいてドラスティックに性能を改善するには至らなかった。

### 5.3 bigram 融合の効果

今回作成した融合 bigram のうち hybrid1 から hybrid5 までは重み関数を  $w_1$  の出現頻度で決めたものであり、hybrid6, hybrid7 は単語 bigram 確率で決めたものであった。これらのうち、hybrid1 から hybrid5 までのモデルの方が、いずれのタスクでもおおむね良い性能が得られた。

また、hybrid1 から hybrid5 は  $w_1$  の頻度の変化に対し、それぞれ次のような特徴があった。

- hybrid1 から hybrid3 は、 $w_1$  の頻度が 500 程度までは急激に重みが増加し、その後は緩やかに変化するがその時の重みがことなる。
- hybrid4, hybrid5 は hybrid1 から hybrid3 より重みの変化のしかたが緩やかである。

しかし、認識結果を見ると、際だった差は見られなかった。

また、short pause なしのモデルではいずれのタスクでも bigram 融合の効果を得られたが、short pause ありのモデルでは newspaper94 タスクでの効果がほとんどなくなってしまった。これは、short pause ありのモデルでは同一タスクに対して十分な精度の単語 bigram が得られたためと考えられる。

## 6 まとめと課題

言語モデルとしてより精密である単語 bigram、およびあまり精密でないが言語的特徴変化に頑健なクラス bigram を融合することによって話題や年月による言語的特徴の変化に対して頑健な言語モデルを構築した。これにより、言語モデル学習時と同一タスク（言語的特徴が同じ）の認識実験で最大 1.32% (<sp>なし)、0.11% (<sp>あり)、言語モデル学習時と異なるタスク（言語的特徴が異なる）の認識実験で最大 1.35% (<sp>なし)、1.54% (<sp>あり) の性能向上を得た。

short pase を単語として導入することによって、認識性能は向上するが、bigram モデルは言語的特徴変化に敏感になることが確認された。

今回の方法で作成したクラス bigram モデルは、単語 bigram モデルよりも言語的特徴変化に頑健なことが確認されたが、bigram 融合による性能向上はわずかであった。

今後は bigram 融合により、性能向上が得られたケースと得られなかったケースについて、その特徴を調査するとともに、別のクラス bigram 作成法や bigram 融合アルゴリズムを検討したい。

## 7 謝辞

本研究には、RWC テキストデータベース、毎日新聞 CD-ROM、新聞記事読み上げ音声コーパス (JNAS)、NHK ニュース音声データベースを使用させていただいた。

## 参考文献

- [1] S.M.Katz "Estimation of Probabilities from Sparse Data for the Language Model Component of a Speech Recognizer" IEEE Trans. Acoust., Speech Signal Processing, vol. ASSP-35 no.3, pp.400-401, Mar 1987
- [2] A.Nádas "On Turing's Formula for Word Probabilities" IEEE Trans. Acoust., Speech and Signal Processing, vol. ASSP-33, no.6, pp.1414-1416 Dec 1985
- [3] 川端, 田本: "二項事後分布に基づく N-gram 言語モデルの Back-off 平滑化" 信学技報, NLC95-58, SP95-93(1995-12)
- [4] 伊藤, 好田: "対話音声認識のための事前タスク適応の検討" 情処研報 96-SLP-14, pp91-98(1996-12)
- [5] 伊藤, 代島, 丸山, 加藤, 好田 "大語彙言語データベースからの N-gram 構築とタスク適応の検討" 情処研報 96-SLP-11, pp25-30(1996-5)
- [6] W.Ward, S.Issar: "A Class Based Language Model for Speech Recognition" IEEE Proc. ICASSP'96, vol.1, pp416-418
- [7] "RWC テキストデータベース報告書" 新情報処理開発機構, テキストグループ, データベースワークショップ (1996)
- [8] "毎日新聞 CD-ROM 版 1991 年版~1994 年版" 毎日新聞社, (1994-1996)
- [9] 小林, 中野, 肥田木, 小林: "統計的言語モデルにおける付属語の扱いに関する一考察" 日本音響学会平成 9 年度秋期研究発表会講演論文集, J, 2-1-6, pp59-60
- [10] "新聞記事読み上げ音声コーパス" JNAS Vols.1-16, 日本音響学会
- [11] "NHK ニュース音声データベース" 日本放送協会放送技術研究所

表 2: 単語認識率

testset	model	H	D	S	I	N	Correct[%]	Accuracy[%]
NHK120	onlyword	1061	205	805	126	2071	51.23	45.15
	onlyclass	910	283	878	92	2071	43.94	39.50
	hybrid1	1064	206	801	123	2071	51.38	45.44
	hybrid2	1057	211	803	113	2071	51.04	45.58
	hybrid3	1060	237	774	97	2071	51.18	46.50
	hybrid4	1061	213	797	112	2071	51.23	45.82
	hybrid5	1056	226	789	100	2071	50.99	46.16
	hybrid6	1037	245	789	87	2071	50.07	45.87
	hybrid7	1049	243	779	93	2071	50.65	46.16
newspaper94	onlyword	562	98	245	45	905	62.10	57.13
	onlyclass	463	143	299	43	905	51.16	46.41
	hybrid1	565	98	242	48	905	62.43	57.13
	hybrid2	570	101	234	45	905	62.98	58.01
	hybrid3	559	107	239	37	905	61.77	57.68
	hybrid4	575	100	230	46	905	63.54	58.45
	hybrid5	567	105	233	39	905	62.65	58.34
	hybrid6	544	115	246	39	905	60.11	55.80
	hybrid7	543	116	246	39	905	60.00	55.69

表 3: short pause 有りのモデルによる単語認識率 (かっこ内は前回比)

testset	model	H	D	S	I	N	Correct[%]	Accuracy[%]
NHK120	onlyword	1060	200	811	109	2071	51.18(-0.05)	45.92( 0.77)
	onlyclass	919	286	866	78	2071	44.37( 0.43)	40.61( 1.11)
	hybrid1	1064	199	808	109	2071	51.38( 0.00)	46.11( 0.67)
	hybrid2	1061	208	802	98	2071	51.23( 0.19)	46.50( 0.92)
	hybrid3	1049	241	781	80	2071	50.65(-0.53)	46.79( 0.29)
	hybrid4	1065	210	796	99	2071	51.42( 0.19)	46.64( 0.82)
	hybrid5	1070	225	776	87	2071	51.67( 0.68)	47.46( 1.30)
	hybrid6	1013	261	797	77	2071	48.91(-1.16)	45.20(-0.67)
	hybrid7	1030	255	786	82	2071	49.73(-0.92)	45.77(-0.39)
newspaper94	onlyword	587	103	215	22	905	64.86( 2.76)	62.43( 5.30)
	onlyclass	493	144	268	24	905	54.48( 3.32)	51.82( 5.41)
	hybrid1	585	106	214	23	905	64.64( 2.21)	62.10( 4.97)
	hybrid2	587	110	208	21	905	64.86( 1.88)	62.54( 4.53)
	hybrid3	580	110	215	23	905	64.09( 2.32)	61.55( 3.87)
	hybrid4	583	112	210	22	905	64.42( 0.88)	61.99( 3.54)
	hybrid5	578	114	213	21	905	63.87( 1.22)	61.55( 3.21)
	hybrid6	561	123	221	23	905	61.99( 1.88)	59.45( 3.65)
	hybrid7	565	124	216	24	905	62.43( 2.43)	59.78( 4.09)