

# 確率文脈自由文法を用いた 言語モデルの構築と音声認識実験による評価

堀 智織<sup>†</sup> 加藤正治<sup>‡</sup> 伊藤彰則<sup>‡</sup> 好田正紀<sup>‡</sup>

<sup>†</sup> 東京工業大学 情報理工学研究所

〒 152-8550 目黒区大岡山 2-12-1

chiori@cs.titech.ac.jp

<sup>‡</sup> 山形大学工学部

〒 992 米沢市城南 4-3-16

あらまし 確率文脈自由文法 (Stochastic Context Free Grammar : SCFG) は,  $N$ -gram のような局所的な制約だけでなく, 文全体に渡る大局的な制約をも記述できることから, 非常に表現力の高い言語モデルとして知られている. しかし, SCFG のパラメータ推定には Inside-Outside アルゴリズムを用いる必要があり, 非終端記号数の 3 乗, かつ入力系列長の 3 乗に比例する莫大な計算量を要する. そのため, これまで SCFG が音声認識用の言語モデルとして利用される事はほとんどなかった. そこで本研究では, Inside-Outside アルゴリズムの計算量を削減するために, 文節単位の係受け文法を適用した SCFG を提案する. EDR コーパスを用いた実験では, 提案法を含む各種 SCFG の性能とパラメータ推定に要する処理量を比較し, 提案法の計算量削減の効果を示す. そして, 毎日新聞コーパスを用いて大規模な文節単位の係受け SCFG を構築し, 大語彙連続音声認識システムに実装して, Trigram との認識性能の比較を行った結果について報告する.

キーワード 音声認識, 言語モデル, 確率文脈自由文法, 係受け文法, Inside-Outside アルゴリズム

## Construction of Language Models based on Stochastic Context Free Grammar and Evaluation in Speech Recognition

Chiori HORI<sup>†</sup>, Masaharu KATOH<sup>‡</sup>, Akinori ITO<sup>‡</sup>, and Masaki KOHDA<sup>‡</sup>

<sup>†</sup>Tokyo Institute of Technology

2-12-1 Ookayama, Meguro-ku, 152-8550 Japan

<sup>‡</sup>Faculty of Engineering, Yamagata University

4-3-16 Johnan, Yonezawa-shi, 992-8510 Japan

**Abstract** It is shown that Stochastic Context Free Grammar:SCFG is a very expressive language model since it enables to express not only local constraints like  $N$ -gram, but also global constraints over a whole sentence. However, to estimate parameters of SCFG, Inside-Outside algorithm has to be used, which needs huge computation in proportion to cube of the number of the non-terminal symbols and the length of the input sequences. Therefore, SCFG has rarely been used for speech recognition. In this paper, we proposed a new SCFG to which applied phrase-based dependency grammar to decrease the huge computation. In the test using EDR corpus, we compared the proposed model with the other types of SCFG in terms of perplexity and computation amount. And then, we constructed large-scale SCFG using Mainichi news corpus, and compared it with trigram on a 5,000-word Japanese newspaper reading task.

**key words** speech recognition, language model, Stochastic Context Free Grammar, dependency grammar, Inside-Outside algorithm

## 1 はじめに

近年、高精度な大語彙連続音声認識 (Large Vocabulary Continuous Speech Recognition : LVCSR) を実現しようという試みがますます盛んになり、新聞記事読み上げ音声や放送ニュース音声の認識を目的とした多くの検討が行われている。新聞記事や放送ニュースなどの文声を認識する場合、音響情報だけでは曖昧なため各単語を一意に同定することが難しい。そこで、言語モデルを用いて言語的な制約を加え、文法を逸脱した単語の連鎖を避け、音響的な曖昧性を補うことが重要である。

現在の連続音声認識システムでは、言語モデルとして  $N$ -gram が用いられる。 $N$ -gram は、単純であるが認識性能の改善が大きく、モデルのパラメータの推定が容易であるため、広く一般に用いられている。しかし、 $N$ -gram は局所的な単語間の制約を表現するモデルであるため、文全体に渡る文の構造的な文法を示すような制約は全く表現できない。

一方、他の言語モデルとして、自然言語の構文解析に用いられる文脈自由文法に基づいた確率文脈自由文法 (Stochastic Context Free Grammar : SCFG) がある。SCFG は  $N$ -gram に対し、文全体に渡る制約をも表現できる言語モデルであるが、自然言語には文法的な曖昧さがあるため、その学習には Inside-Outside アルゴリズム [1][2] を用いる必要がある。その結果、モデルのパラメータ推定において非終端記号数の 3 乗、かつ入力系列長の 3 乗のオーダーの莫大な計算量を要することから、これまで SCFG が音声認識に適用される事はほとんどなかった。

そこで、我々は日本語の持つ言語的特性である係受け文法を SCFG に適用する事により、非終端記号数の 3 乗から 2 乗のオーダーまで計算量を削減する方法を提案した [3][4]。パラメータの初期値および非終端記号数の適正化をはかることで、SCFG は Trigram と同等の性能を確保できることを示した。しかしながら、SCFG の性能を Trigram と同程度にするには非終端記号を多数用いる必要があり、大量のコーパスからの大規模な SCFG を構築できるほどの処理量削減には至らなかった。

本報告では、更なる計算量削減のために、文節単位の係受け文法を適用した SCFG を提案する。EDR コーパスを用いた実験では、提案法を含む各種 SCFG の性能とパラメータ推定に要する処理量を比較し、提案法の計算量削減の効果を示す。そして、毎日新聞コーパスを用いて大規模な文節単位の係受け SCFG を構築し、大語彙連続音声認識システムに実装して、Trigram との認識性能の比較を行った結果について報告する。

## 2 確率文脈自由文法 (SCFG)

SCFG は文脈自由文法 (Context Free Grammar: CFG) の生成規則に次式のように確率を付与したものである。

$$\alpha \rightarrow \beta, \quad P(\alpha \rightarrow \beta) \in [0, 1]$$

$$\alpha \in N, \quad \beta \in (N \cup T)^*$$

$$\sum_{\beta} P(\alpha \rightarrow \beta) = 1$$

( $N$ :非終端記号の集合,  $T$ :終端記号の集合)

SCFG で生成される単語列には生成確率が定義され、その生成確率は導出に用いられた全ての生成規則の確率の積で表される。導出木が一意に定まらない場合は、全ての導出木の生成確率の総和を単語列の生成確率とする。

SCFG の生成規則の確率は、言語コーパス中の各文章に対する導出木を求めて生成規則の適用回数をカウントすれば推定可能である。しかし、一般に SCFG における導出木は一意に定まらないため、生成規則の確率は Inside-Outside アルゴリズムを用いて推定する必要がある。その際、生成規則として言語学的な見地で作成された文法を用い、体言、名詞句、名詞などの全ての文法を作成した場合、非終端記号数が莫大となり、生成規則の確率を自動推定するには計算量が多く困難である。そこで、本研究では既存の文法によらず非終端記号数のみを決定し、可能な全ての生成規則が存在すると仮定して確率推定を行う。

### 2.1 Inside-Outside アルゴリズム

Inside-Outside アルゴリズムは、SCFG の最尤推定に基づいた学習アルゴリズムである。ただし、文法は次式で示される Chomsky 標準形で表現される。

$$\alpha \rightarrow \beta\gamma$$

$$\alpha \rightarrow w$$

$$\alpha, \beta, \gamma \in N, w \in T$$

$\alpha \rightarrow \beta\gamma$  で表される非終端記号間の書き換え規則の適用確率を  $a(\alpha \rightarrow \beta\gamma)$ 、 $\alpha \rightarrow w$  で表される終端記号 (単語) の生成確率を  $b(\alpha \rightarrow w)$  とする。

#### (1) Inside 確率

入力テキスト系列を  $w_1, w_2, w_3, \dots, w_L$  とした場合、非終端記号  $\alpha$  が  $i$  番目から  $j$  番目の単語列に置き換わる確率が Inside 確率であり、図 1 に示す導出木の影のつ

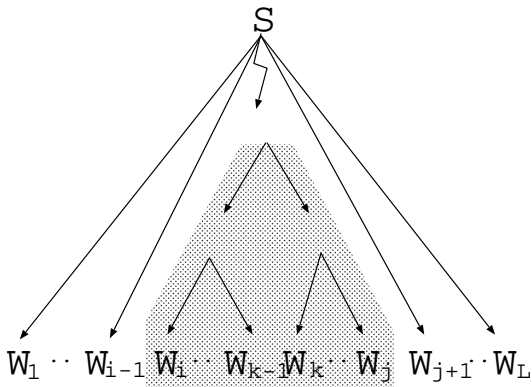


図 1: Inside 確率

いた領域の生成確率である．次式のように  $e(i, j|\alpha)$  で表現される．

$$\begin{aligned}
 e(i, j|\alpha) &= P(\alpha \rightarrow w_i \dots w_j) \\
 &= \begin{cases} \sum_{k=i}^{j-1} \sum_{\beta\gamma} a(\alpha \rightarrow \beta\gamma) e(i, k|\beta) e(k+1, j|\gamma) & \text{if } i < j \\ b(\alpha \rightarrow w_i) & \text{if } i = j \end{cases}
 \end{aligned} \tag{1}$$

### (2) Outside 確率

一方，図 2 に示す導出木の影のついた領域の生成確率が Outside 確率である．次式のように  $f(i, j|\alpha)$  で表現される．

$$\begin{aligned}
 f(i, j|\alpha) &= P(S \rightarrow w_1 \dots w_{i-1} \alpha w_{j+1} \dots w_L) \\
 &= \sum_{k=1}^{i-1} \sum_{\beta\gamma} a(\beta \rightarrow \gamma\alpha) e(k, i-1|\gamma) f(k, j|\beta) \\
 &\quad + \sum_{k=j+1}^L \sum_{\beta\gamma} a(\beta \rightarrow \alpha\gamma) e(j+1, k|\gamma) f(i, k|\beta)
 \end{aligned} \tag{2}$$

### (3) パラメータ再推定

単語列  $w_1, w_2, \dots, w_L$  の生成過程で，導出木の任意の位置で  $\alpha \rightarrow \beta\gamma$  の規則が適用される確率は，次式で計

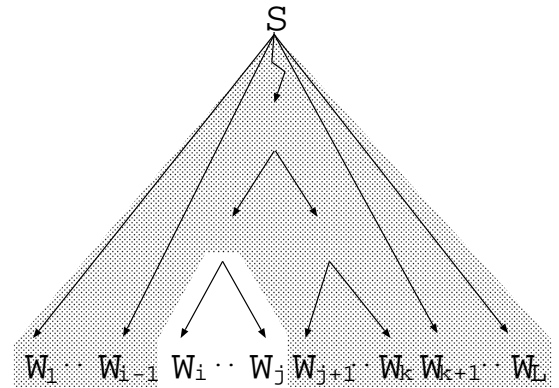
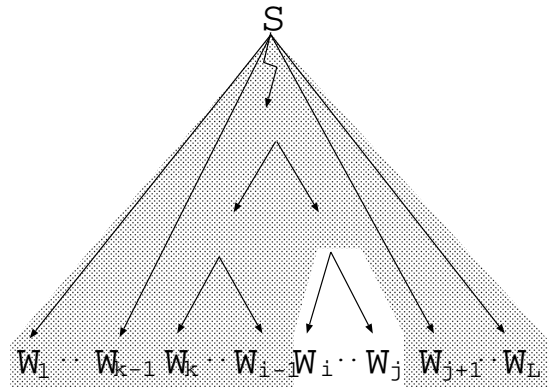


図 2: Outside 確率

算される．

$$\hat{a}(\alpha \rightarrow \beta\gamma) = \frac{\sum_{i=1}^{L-1} \sum_{j=i+1}^L \sum_{k=i}^{j-1} g(i, k, j; \alpha \rightarrow \beta\gamma)}{e(1, L|S)} \tag{3}$$

一方，単語  $w$  が生成される確率は，次式で計算される．

$$\hat{b}(\alpha \rightarrow w) = \frac{\sum_{i; w_i=w} b(\alpha \rightarrow w) f(i, i|\alpha)}{e(1, L|S)} \tag{4}$$

ここで， $g(i, k, j; \alpha \rightarrow \beta\gamma)$  は， $\alpha$  から  $\beta\gamma$  が生じ，さらに  $\beta$  から単語  $w_i \dots w_k$  が生じ， $\gamma$  から単語  $w_{k+1} \dots w_j$  が生じる確率で，Inside-Outside 確率を用いて次のように求められる．

$$\begin{aligned}
 g(i, k, j; \alpha \rightarrow \beta\gamma) &= e(i, k|\beta) e(k+1, j|\gamma) a(\alpha \rightarrow \beta\gamma) f(i, j|\alpha)
 \end{aligned} \tag{5}$$

学習は，各生成規則の初期値を設定，Inside-Outside 確率を計算，生成規則の確率を再推定，再推定された確率を新たな初期値として再学習，という手順で行われる．しかし，SCFG を Inside-Outside アルゴリズムによって推定するには， $O(N^3 L^3)$  の膨大な処理時間を要する．

### 3 文節単位の係受け SCFG を用いた言語モデルの構築

本研究では, SCFG の学習に要する計算量を削減するために, 係受け文法に加えて文節単位の SCFG を導入した新しいタイプの SCFG を提案する.

係受け文法は, 日本語の特徴である前方の文節が後方の文節を修飾する句構造文法であり, 次式の規則で表現できる.

$$\alpha \rightarrow \beta\alpha \quad (6)$$

$$\alpha \rightarrow w \quad (7)$$

係受け文法を適用する事で, 非終端記号間の書き換え規則が  $\alpha \rightarrow \beta\gamma$  から  $\alpha \rightarrow \beta\alpha$  に単純化され, 非終端記号数を  $N$  とした場合, Inside-Outside アルゴリズムの計算量は  $O(N^3)$  から  $O(N^2)$  に削減される.

また, 以前我々は, 次式に示すような実質語と機能語の区別を組み込み, 確率正規文法を併用する係受け SCFG を提案した [3].

$$\alpha \rightarrow \beta\alpha \quad (8)$$

$$\alpha \rightarrow w_c \quad (9)$$

$$\alpha \rightarrow \beta w_f \quad (10)$$

ここで,  $w_c$  と  $w_f$  は実質語と機能語を表す. そして, 提案した係受け SCFG が Trigram と同等の性能を有することを示した. しかし, Inside-Outside アルゴリズムが入力系列長の 3 乗に比例するという問題は解決されていないかった.

そこで, 本研究では計算量を削減する新たな手法として, SCFG の制約を文節単位で適用する文節単位の係受け SCFG を提案する. この SCFG は, 文節の区切りを明示的に与えることによって, 式 (8) の規則を文節間文法, 式 (9)(10) の規則を文節内文法とした, 式 (6)(7) の係受け SCFG の特殊なケースと言える.

文節単位の係受け SCFG では, Inside-Outside アルゴリズムにおける計算領域は図 3 のように削減される. 従って, 1 文中の単語数を  $L$ , 1 文中の文節数を  $M (L \geq M)$  とした場合, 計算量は  $O(L^3)$  から  $O(M^3)$  になる. つまり, 1 文節平均 2 単語と仮定すると, 計算量は単語単位の SCFG に比べて  $(\frac{1}{2})^3$  になる.

各種 SCFG の計算量の関係を表 1 にまとめる.

表 1: 各種 SCFG の計算量の比較

		SCFG	
		単語単位	文節単位
係受けの制約	なし	$\alpha \rightarrow \beta\gamma$	文節間 $\alpha \rightarrow \beta\gamma$
		$\alpha \rightarrow w$	文節内 $\alpha \rightarrow w_c$ $\alpha \rightarrow \beta w_f$
	計算量 $O(N^3 L^3)$		計算量 $O(N^3 M^3)$
	あり	$\alpha \rightarrow \beta\alpha$	文節間 $\alpha \rightarrow \beta\alpha$
$\alpha \rightarrow w$		文節内 $\alpha \rightarrow w_c$ $\alpha \rightarrow \beta w_f$	
		計算量 $O(N^2 L^3)$	計算量 $O(N^2 M^3)$

$N$ : 非終端記号数

$L$ : 1 文中の単語数

$M$ : 1 文中の文節数 ( $\approx \frac{1}{2}L$ )

### 4 文節単位の係受け SCFG の確率推定アルゴリズム

アルゴリズムを説明するためにまず以下の記号を定義する.

$M$ : 1 文中の文節数

$w_{mc}$ :  $m$  番目の文節の実質語

$w_{mf,i}$ :  $m$  番目の文節の  $i$  番目の機能語

$K_m$ :  $m$  番目の文節の機能語の数

但し, 文節とは実質語 + 0 個以上の機能語からなる単語列と定義する. つまり, 1 つの実質語から次に現れる実質語の前までの単語列を 1 文節と見なす.

以下に文節単位の係受け SCFG の確率推定アルゴリズムを示す.

#### (1) 初期値設定

初期値として,  $a(\alpha \rightarrow \beta\alpha)$  は等確率,  $b(\alpha \rightarrow w_c)$ ,  $c(\alpha \rightarrow \beta w_f)$  は乱数に設定する.

但し, 品詞列テキストや品詞 + 高頻度単語のテキストを用いた初期値設定法 [4] を利用することもできる.

#### (2) 文節内 Forward 確率

非終端記号  $\alpha$  から  $m$  番目の文節の単語列  $w_{mc} \dots w_{mf,i}$  が生成される確率を文節内 Forward 確率として次式のように求める.

例文：この／国／は／日本／です／か

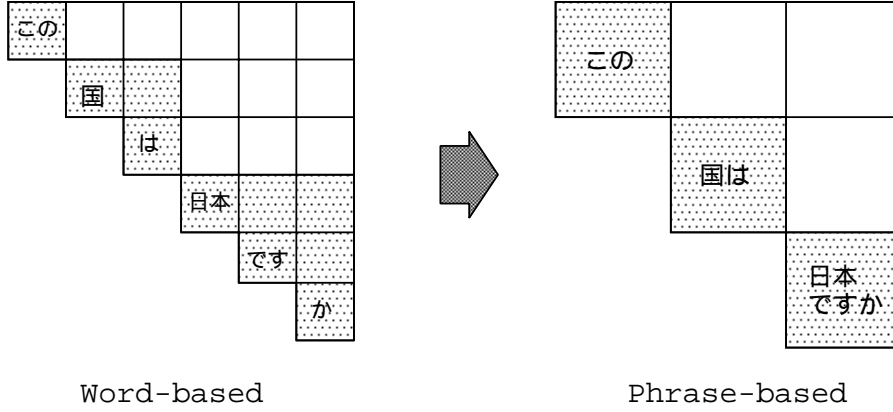


図 3: Inside-Outside アルゴリズムの計算領域

$$h(m, i, \alpha) = \begin{cases} b(\alpha \rightarrow w_{mc}) & \text{if } i = 0 \\ \sum_{\beta} h(m, i-1, \beta) c(\alpha \rightarrow \beta w_{mf, i}) & \text{if } i > 0 \end{cases} \quad (11)$$

### (3) 文節間 Inside 確率

文節内 Forward 確率を用いて，文節間 Inside 確率を次式のように計算する．

$$e(m, n | \alpha) = \begin{cases} h(m, K_m, \alpha) & \text{if } m = n \\ \sum_{l=m}^{n-1} \sum_{\beta} a(\alpha \rightarrow \beta \alpha) e(m, l | \beta) e(l+1, n | \alpha) & \text{if } m \neq n \end{cases} \quad (12)$$

### (4) 文節間 Outside 確率

文節間 Inside 確率を用いて，文節間 Outside 確率を次式のように計算する．

$$\begin{aligned} f(m, n | \alpha) &= \sum_{l=1}^{m-1} \sum_{\beta} a(\alpha \rightarrow \beta \alpha) e(l, m-1 | \beta) f(l, n | \alpha) \\ &+ \sum_{l=n+1}^M \sum_{\beta} a(\beta \rightarrow \alpha \beta) e(n+1, l | \beta) f(m, l | \beta) \end{aligned} \quad (13)$$

### (5) 文節内 Backward 確率

文節間 Outside 確率を用いて，文節内 Backward 確率を次式のように計算する．

$$r(m, i, \alpha) = \begin{cases} f(m, m, \alpha) & \text{if } i = K_m \\ \sum_{\beta} c(\beta \rightarrow \alpha w_{mf, i+1}) r(m, i+1, \beta) & \text{if } i < K_m \end{cases} \quad (14)$$

### (6) パラメータ再推定式

(2)～(5) で求めた確率を用いて，次の再推定式に従ってパラメータを更新する．

$$\hat{a}(\alpha \rightarrow \beta \alpha) = \frac{\sum_{m=1}^{M-1} \sum_{n=m+1}^M \sum_{l=m}^{n-1} g(m, l, n; \alpha \rightarrow \beta \alpha)}{e(1, M | S)} \quad (15)$$

$$\hat{b}(\alpha \rightarrow w_c) = \frac{\sum_{m=1, w_{mc}=w_c}^M b(\alpha \rightarrow w_c) r(m, 0, \alpha)}{e(1, M | S)} \quad (16)$$

$$\begin{aligned} \hat{c}(\alpha \rightarrow \beta w_f) &= \frac{\sum_{m=1}^M \sum_{i=1, w_{mf, i}=w_f}^{K_m} h(m, i-1, \beta) c(\alpha \rightarrow \beta w_f) r(m, i, \alpha)}{e(1, M | S)} \end{aligned} \quad (17)$$

ここで，

$$g(m, l, n; \alpha \rightarrow \beta \alpha)$$

$$= e(m, l | \beta) e(l + 1, n | \alpha) a(\alpha \rightarrow \beta) f(m, n | \alpha) \quad (18)$$

である。

(7) 収束するまで (2) ~ (6) を繰り返す

## 5 評価実験

### 5.1 実験条件

EDR コーパスを用いて、表 1 に示す 5 種類の SCFG を構築し、Perplexity による評価を行った。実験条件を表 2 に示す。

初期値の設定は等確率と乱数によって行い、学習回数 0, 5, 10, 15, 20, 25, 30 のモデルを 5 種類の SCFG それぞれについて作成した。

### 5.2 実験結果と考察

テスト文 100 文に対する Perplexity を図 4、学習時間を図 5 に示す。但し、図中の SCFG、P-SCFG、K-SCFG、K-SCFG2、PK-SCFG はそれぞれ単語単位の SCFG、文節単位の SCFG、実質語と機能語の区別がない係受け SCFG、実質語と機能語の区別がある係受け SCFG、文節単位の係受け SCFG (提案法) を表す。

図 4 より、文節単位の係受け SCFG は他の SCFG と同等の Perplexity を示しており、文節単位の導入がモデルの性能にほとんど影響しないことが分かる。また、単語単位の SCFG に対し、文節単位の SCFG は学習の収束が早いので、モデル構築の時間を短縮できるものと考えられる。

一方、図 5 より、1 回の学習にかかる処理時間は、係受けを適用する事で 1/20 に削減され、更に文節単位で SCFG を適用する事により 1/8 まで削減されることが分かる。

以上の結果より、文節単位の係受け SCFG は、従来 SCFG の性能を保ちつつ大幅に計算量を削減すること

表 2: SCFG の評価実験条件

非終端記号数	20	
語彙サイズ (終端記号数)	3034 単語 (学習テキスト出現頻度 2 回以上)	
使用テキスト	EDR コーパス	
	学習テキスト	評価テキスト
文章数	2000 文	100 文
総単語数	53910 単語	2782 単語
未知語の割合	10.3%	22.0%

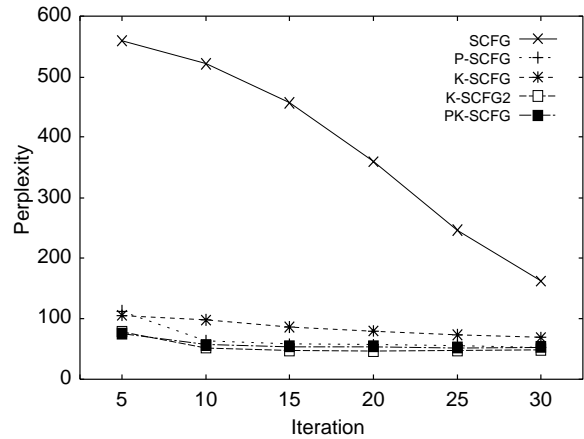


図 4: SCFG の性能比較

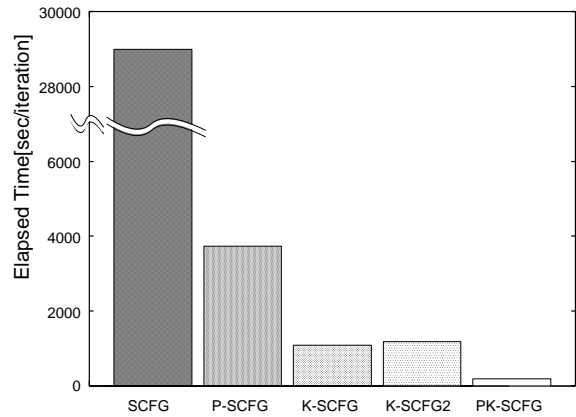


図 5: SCFG の学習時間

が可能な有効なモデルであることが示された。

## 6 大語彙連続音声認識実験

### 6.1 認識タスク

認識タスクは、新聞記事の読み上げ音声を対象とした不特定話者連続音声認識である。認識対象語彙は毎日新聞 1991 年 1 月から 1994 年 9 月までの記事に含まれる頻出上位 5000 単語 (形態素) とし、未知語は含まれないものとする。

### 6.2 音声資料・音声分析

音声資料は、音響モデル学習用として、日本音響学会新聞記事読み上げ音声コーパスの男性 102 名が発声した新聞記事と音素バランス文、計 15732 文を用い、評

表 3: 音声分析条件

標本化周波数	16 kHz
量子化	16 bit
フレーム長	32 msec
フレーム周期	8 msec
分析窓	ハミング窓
高域強調	$1 - z^{-1}$
特徴ベクトル	1~12 次の LPC メルケプストラム係数 と対数パワー, およびその一次と二次の 回帰係数 (計 39 次元)
正規化	発話毎のケプストラム平均正規化

備用として, 同コーパスの学習話者とは異なる男性 10 名がそれぞれ 10 文発声した計 100 文の新聞記事を用いる. 評価用の 100 文は毎日新聞の記事 94 年 10 月~12 月から選ばれた未知語を含まない 5000 語で閉じた文章である.

音声分析条件を表 3 に示す.

### 6.3 音響モデル

音響モデルとして状態クラスタリングによる HM-Net を用いる [5]. 音素カテゴリーは 34 音素 + 無音 (sil) とする. 初期 HM-Net として, 各音素 3 状態の音素環境独立 HMM を並列に接続した 81 状態の HM-Net を用い, 状態分割は 2000 状態まで行う. 構造決定後に各状態の出力確率分布を 16 混合に再構成して連結学習を行う.

### 6.4 言語モデル

学習テキストとして, 毎日新聞 1994 年 1 月~1994 年 9 月までの記事から 5000 語彙で閉じた 1 文 3 単語以上で構成される 46301 文を選択した. 言語モデルとして, Bigram, Trigram, 文節単位の係受け SCFG (非終端記号数 100 と 120, 初期値の設定に品詞 + 高頻度単語のテキストを利用 [4]) を作成した.

但し, 第一パスに用いる Bigram は, 毎日新聞 1991 年 1 月~1994 年 9 月の記事より作成している.

### 6.5 100-Best Rescoring

第一パスより単語グラフを生成し, その単語グラフより求めた上位 100 文の候補を以下の式に従ってリスコアする.

$$Score(W|O)$$

$$= w_{scfg} \log P_{scfg}(W) + w_{ng} \log P_{ng}(W)$$

$$+ \text{acoust}(O|W) + \text{len}(W) \times \text{penalty} \quad (19)$$

ここで,  $P_{scfg}(W)$ ,  $P_{ng}(W)$  は, SCFG および  $N$ -gram によって計算される言語尤度,  $w_{scfg}$ ,  $w_{ng}$  はそれぞれの言語尤度に対する重み,  $\text{acoust}(O|W)$  は音響尤度,  $\text{len}(W)$  は文候補の長さ (単語数),  $\text{penalty}$  は挿入ペナルティを表す. 但し, 上式は SCFG と  $N$ -gram を併用する場合のスコアであり, それぞれ単独で用いる場合は一方の言語モデルに対する重みを 0 とおく. そして, スコアが最大となる文を認識結果として選択し, 単語誤り率によって各種言語モデルを評価する.

単語誤り率は以下のように計算する.

$$WER = \frac{S + I + D}{L} \times 100 \quad [\%] \quad (20)$$

ここで,  $S, I, D$  はそれぞれ単語の置換誤り, 挿入誤り, 脱落誤りの数,  $L$  は正解文の単語数を表す.

## 6.6 実験結果と考察

各種言語モデルの正解テキスト (評価用 100 文の発話内容) に対する Perplexity を図 6, 大語彙連続音声認識に適用した場合の単語誤り率を図 7 に示す. 但し, 各モデルの単語誤り率は, リスコアリングの際に言語モデルの重みと挿入ペナルティを様々に変化させた場合の最適値であり, そのときの重みとペナルティは表 4 のようになっている.

正解テキストに対する Perplexity では, Trigram や Bigram の方が SCFG に比べより低い値を示している. これに対し, 実際の音声認識による評価では, SCFG は Trigram と Bigram の間の誤り率を示し, 更に, SCFG と  $N$ -gram を併用する事で, Trigram 単独よりも低い誤り率を達成した. これは, Trigram によって表現され

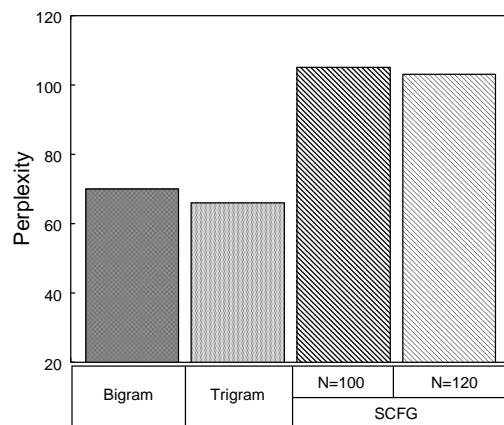


図 6: 正解テキストに対する Perplexity

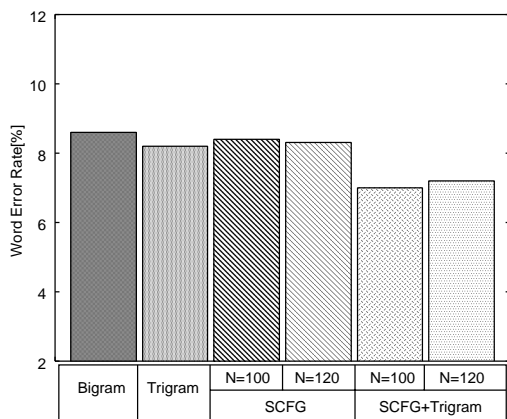


図 7: 音声認識実験における単語誤り率

表 4: 言語モデルの重みと挿入ペナルティの最適値

	$w_{ng}$	$w_{scfg}$	penalty
Bigram	16	0	-24
Trigram	17	0	-20
SCFG 100	0	20	-20
SCFG 120	0	18	-20
Trigram+SCFG 100	10	10	0
Trigram+SCFG 120	10	10	0

ていなかった文全体にわたる制約が SCFG によって補完された事が大きな理由と考えられる。

以下に, Trigram 単独では誤りであったが, SCFG と組み合わせる事により改善された認識結果の例を示す。

正解文章 (1) : それを 自衛隊 が やっ て  
いいわけが ない  
Trigram : それを 自衛隊 が あっ て  
いいわけが ない  
SCFG+Trigram : 誤りなし

正解文章 (2) : ほかに も ライバル が 多い  
Trigram : ほかに も ライバル 多い  
SCFG+Trigram : 誤りなし

いずれの結果も構文的, 意味的に不適当であった認識結果が, SCFG によって意味の通る文に改善されている。特に助詞が補われるケースは上記の例以外にもいくつか見られた。

一方, SCFG によって誤りが増えるケースもある。以下は, 文の終りに助詞が挿入されてしまった例である。

正解文章 (3) : 石油 業界 で は 少ない 規制

緩和 推進 派

Trigram : 誤りなし

SCFG+Trigram : 石油 業界 で は 少ない 規制  
緩和 推進 派 は

## 7 まとめ

確率文脈自由文法を用いた言語モデルの構築と音声認識実験による評価について検討した。SCFG の学習に要する計算量を大幅に削減する文節単位の係受け SCFG を提案し, これまで困難であった大量コーパスからの SCFG のパラメータ推定を可能にした。提案した文節単位の係受け SCFG は, 語彙サイズ 5000 の大語彙連続音声認識において, Trigram と組み合わせることにより, Trigram 単独よりも高い認識性能を有することを示した。

## 参考文献

- [1] K.Lari, S.J.Young : “The estimation of stochastic context free grammars using the Inside-Outside algorithm”, *Computer Speech and Language*,4, pp.35-56 (1990).
- [2] K.Lari, S.J.Young : “Application of stochastic context free grammars using the Inside-Outside algorithm”, *Computer Speech and Language*,5, pp.237-257 (1991).
- [3] 柳沼, 加藤, 伊藤, 好田 : “係り受け文法による確率文脈自由文法を用いた言語モデルの検討”, 情報処理学会第 54 回全国大会, 2-217 (1997).
- [4] 柳沼, 加藤, 伊藤, 好田 : “文脈自由文法を用いた言語モデルにおける初期値の検討”, 平 9 秋音講論集, 2-1-2 (1997).
- [5] 堀, 加藤, 伊藤, 好田 : “状態クラスタリングによる HM-Net の構造決定法の検討”, 信学論 (D-II), J81-D-II, No.10, pp.2239-2248 (1998).
- [6] 堀, 岡, 加藤, 伊藤, 好田 : “大語彙連続音声認識のための音素グラフに基づく仮説制限法の検討”, 情報処理学会論文誌, Vol.40, No.4, 掲載予定 (1999).
- [7] 河原, 他 : “大語彙日本語連続音声認識研究基盤の整備 — 評価用連続音声認識プログラムの開発 —”, 情報処理学会研究報告, 97-SLP-18-1, pp.1-6 (1997).