

# CFG とバイグラムの併用による話し言葉用文法の自動 学習と音声認識による評価実験

中川聖一 大谷耕嗣

豊橋技術科学大学 情報工学系

〒441 豊橋市天伯町字雲雀ヶ丘 1-1

Tel. (0532) 47-6777

E-mail {nakagawa, otani}@slp.tutics.tut.ac.jp

あらまし 一般に話し言葉の文法の構築は難しい。大量のテキスト文集合がある場合には、n-gram が適切である。しかし、比較的小規模なタスクやユーザが容易に設計したいタスクでは文脈自由文法 (CFG) が適切である。本稿では文のカバー率を改善するために、n-gram と CFG の利点を取り入れた文法規則の自動学習法について述べる。学習方法は、システムに入力された文が文法規則が不備なために受理できない時、システムがこの入力文を使って単語クラスペアやバイグラムの学習することによって実行する。

この単語クラスペアやバイグラムの適用方法を変えたいいくつかの実験をテキストと音声認識で行ない、文法学習による解析 (認識) 文数の向上及びパープレキシティの変化について調べた。

キーワード 文法の学習, CFG, バイグラム, カバー率, パープレキシティ, 音声認識

## An Automatic Learning of Grammar Rules for Spontaneous Speech by Combining CFG with Bigram, and Evaluation by Speech Recognition

Seiichi NAKAGAWA and Kouji OHTANI

Department of Information and Computer Sciences

Toyohashi University of Technology Tenpaku-cho, Toyohashi, 441, Japan

Tel. (0532) 44-6777

Email {nakagawa, otani}@slp.tutics.tut.ac.jp

**Abstract** In this paper, we describe an automatic learning method of the grammar rules for improving coverage of acceptable sentences. The learning method is carried out by a sentence which is not accepted by production rules. The system learns wordclass pairs or bigram using this input sentence.

We experimented on some strategies of applying wordclass pairs on text parsing and speech recognition level, and evaluated the coverage of acceptable sentences (or speech recognition rate) and perplexity using above methods.

**Keyword** learning of grammar, CFG, bigram, coverage, perplexity, speech recognition

### 1 序論

自然言語理解システムの目的の1つに人間・機械間のコミュニケーションの方法としての人間にやさしいインターフェイスの開発があげられる。我々の研究室では、“富士山周辺の観光案内システム”のタスクで自然発話を使った対話システムの開発を行なっている [1]。

対話システムはユーザが発話した言葉を受理するための文法を使って入力した文を認識する。近年においては、machine-readable な言語データの整備が進んだことと、計算機の進歩により大規模な言語データの取り扱いが可能になったことから、自然言語処理に使用する文法を自動的に獲得する様々な研究が盛んに行なわれてきた。このような大量の言語データを使って自動的に獲得した文法は、人的に作られたものに比べて、獲得した文法が文法設計者の主観に流されにくい、文法作成にかか

るコストが小さい、文法の適用範囲が広い、文法に何らかの統計情報を組み込むことができる、といった優れた特徴を持っているが、それでも完全な文法を作成することは難しい。ましてや、大量のデータが得られないような応用の開発などでは、このアプローチはとりにくい。その上、もしユーザがシステムの文法で受理できない文を話した時に対話システムはその文を認識することができない。ところが大量の言語データを用いて文法を構築してみても、あらかじめ完全な文法を構築するのは難しい上に、システム設計者があらかじめ全ての発話を考慮することは不可能である。そこで、文法の不備による認識間違いを減らすために、ユーザがシステムの文法で受理できない文を発話した時に、その文を使ってシステムの文法に登録されていない規則の学習を行ない、これらの文を受理できるようにするシステムの開発を行ってきた [2][3]。その結果、新しい入力文に対するカバー率の改善ができるようになった。

音声認識を前提とした文法規則学習の研究で、Wrightらは、ロバストなパーザのためにCFGとバイグラムの併用を提案している[4]。しかし、それらは統合されおらず並列に実行される。最近では竹澤らが、ポーズ情報で区切られた区間を部分木で表現し、部分木出力による音声認識実験に対して、前終端記号バイグラムを利用した再順序付けを行なう方法を提案している[5]。

我々の以前の研究で、生成規則の登録については、単語クラスペアをCFGの補間に使用する方法がカバー率とパープレキシティのバランスが一番良いことが分かった。また、CFGと単語(単語クラス)ペアに確率を付与してパープレキシティを抑えることを試み、その有効性を示した[3]。本稿では、これらのテキストレベルの結果に基づき、単語クラスペアやバイグラムを用いた文法を音声認識システムにインプリメンテーションして、実際の音声認識での各文法学習の効果について調べ、文法学習のための例文が少ない際にはCFGと単語クラスペアを併せて使用する本手法の有効性を実証した。

## 2 システムの認識エラーの原因

本研究で現在作成している富士山の観光案内をタスクとする対話システム[1]は、発話した音声解析する音声認識部と、効率良く認識を行なうためのユーザが発話するであろう文を記述している文法部からできている。

対話システムの認識間違いが起こる原因は大まかに言って次の2つがあげられる。

- 文法部の不備
- 認識部の不備

文法部の不備というのはユーザーがシステムに登録されている文法で受理できない文を発話したためにシステムが認識間違いを起こす場合のことである。つまり文法がまだ不完全な為に、認識間違いを起こす場合のことである。一方、認識部の不備というのは、ユーザーはシステムに登録されている文法に従って文を発話しているが、認識部分がまだ不完全な為にシステムが正しく理解できず認識間違いになる場合のことである。

認識間違いを減らすには上の2つの原因を解消してやれば良いが、本研究では、文法部の不備が原因による認識間違いを減らすために、文法で受理されなかった文を例文として使いシステムの文法に自動的に新しい規則を追加登録する方法について考える。

文法の不備には、単語の登録が不備な場合と規則の登録が不備な場合があるが、本稿では後者の場合のみ考察する。

## 3 生成規則の登録(ペアを使った規則の追加)

生成規則の登録には単語(単語クラス)ペアを使った規則の追加の他にCFGによる規則の追加登録を行なった[2]。しかし、あまりよい結果が得られなかったため、ここではペアを使った登録方法についてのみ述べる。

単語(単語クラス)ペアを使うことによる生成規則の不備により解析できない文のための規則の登録は、以下の手順で自動的にCFG規則の補助として単語(単語クラス)対制約の登録を行なう。

1. システムが入力文の解析に失敗した時に、ボトムアップのパーザを使って入力文の部分解析木を作る。

2. 入力文をカバーできる部分解析木の組合せで、木の組合せの数が最小となる組合せを見つける。
3. 部分解析木の組合せの各解析木に隣接する単語(単語クラス)のペアを登録する。つまり、図1の様最小の部分解析木の数が2個なら、図中の部分解析木 $P_1$ の最後の単語 $w_1$ と、 $P_2$ の最初の単語 $w_2$ の単語(単語クラス)のペア( $w_1, w_2$ )の登録を行なう。

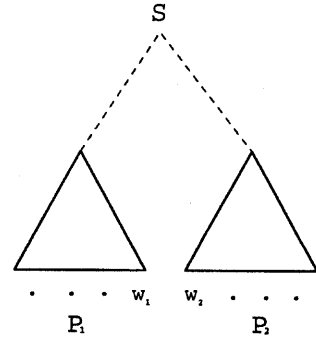


図1: 単語対制約の登録

この方法で登録した単語(単語クラス)ペアは部分解析木を繋ぐ制約として使用する。但し、実際には、CFGとペアを独立に用いて後続単語の予測を行なう。つまり、CFG規則と単語ペアは文脈に関係なくいつも使える。以上の方法で獲得した単語(単語クラス)ペアの獲得する節点とペアの使用方法との違いにより3つの解析方法が考えられる。

### 方法1 CFG + ペア

ペア: 上記の木の組合せの数が最小となる組合せの節点で獲得。

使用方法: 部分解析木を繋ぐ制約として使用。

### 方法2 ペアのみ

ペア: 学習文の全ての場所で獲得。

使用方法: ペアのみでの解析。

### 方法3 CFGのみ+ペアのみ

ペア: 学習文の全ての場所で獲得。

使用方法: ペアのみでの解析。但し、CFGでの解析も並列して行なう。

方法3については、CFGと単語ペアでの解析を並列して行なうが、完全に並列ではなく単語予測毎に各解析による予測単語をマージしながら解析を進める。

## 4 確率文法の導入

音声認識では文法の評価規準としてパープレキシティ(複雑度)がよく用いられる。パープレキシティは2のエントロピー乗で計算される言語の複雑度を示す規準の一つであり、予測分岐数に相当し文法の複雑さを示すのにより指標である。

音声認識ではパープレキシティが小さくカバー率が大きい文法が望ましい。そこでパープレキシティを抑え

るために確率の導入を単語（単語クラス）ペア、CFG各々に対して行なった。

単語（単語クラス）ペアの単語又は単語クラスのバイグラム（単語クラス）ペアの単語又は単語クラスのバイグラムの推定には学習セットとして使用している文を使用して任意の場所の単語のペア又は単語クラスのペアの出現頻度からバイグラムの確率を推定する。つまり入力文（単語列）に対して、単語  $W_i$  の出現頻度  $C(W_i)$  と単語の2つ組の出現頻度  $C(W_{i-1}, W_i)$  を数え上げる。この時のバイグラム確率は

$$P(W_i|W_{i-1}) = \frac{C(W_{i-1}, W_i)}{C(W_i)}$$

で求まる。単語クラスバイグラムについても同様な式を用いて求める。ただ本稿のように、1000文ぐらいの文を学習に用いた場合、単語ペアの数は十分でなく、データの不足を補うため単語ペア及び単語クラスペアは学習データの全ての場所を使って求めている。

次にCFGの各生成規則の確率化（SCFG:Stochastic CFG）について述べる。各生成規則に対する適用確率を求めるのに学習セット914文中でもともと解析が可能な文419文を使用し、これらの文に対する正しい解析木を調べ、使用されている生成規則の使用頻度をカウントする。そして同一の非終端記号を左辺に持つ生成規則の適用確率の和が1になるように正規化を行なう。

$$P(A \rightarrow \alpha | A) = \frac{\text{コーパス中の } A \rightarrow \alpha \text{ の使用回数}}{\text{コーパス中の } A \text{ の出現回数}}$$

この際、全ての生成規則が使用されているわけではないので、使用頻度が0回の規則については1の頻度によりフロアリングを行ない各規則に対する適用確率を求めた。また、バイグラムとCFGにより並列で解析する場合のパープレキシティは以下の方法で求めた。

該単語の予測確率を、CFGの場合には予測される確率をその単語での全ての解析木による予測単語数の逆数とし（該当単語が予測されない場合は確率=0）、SCFGの場合には全ての正しい解析木の予測確率（使った生成規則の生成確率の積）の和の逆数とし、これとバイグラムによる予測確率との平均値で近似した。つまり

$$P(W_i|CFG, \text{bigram}) = \lambda P(W_i|CFG) + (1 - \lambda)P(W_i|W_{i-1})$$

(実験では  $\lambda = 0.5$ )

この逆数をとることによって該単語の予測単語数（分岐数）とし、これらの相乗平均で求める。但し、バイグラムは任意の場所で適用する。

## 5 テキストでの評価

実験には“富士山周辺の観光案内”の対話システムの文法を使った。システム開発者が対話システムの入力文を想定して作成したオリジナルな初期文法の詳細は以下の通りである。なお、パープレキシティは解析できるようになった全評価文を用いて求めている。そのため、学習データ数や言語モデルによって多少文集が変わるが、評価データ（106文）のうちオリジナルな元の文法で解析できた39文を使って求めても大差はない[3]。（）内の数値はSCFGでのパープレキシティである。

- 終端記号数（単語数、語彙のサイズ） - 419
- 生成規則数 - 393
- 非終端記号数 - 137

- 前終端記号数（単語クラスのサイズ） - 110
- 前終端記号から終端記号への書き換え規則数 - 255
- パープレキシティ - 89.8(62.4)

学習と評価で使われたデータは53人の被験者から集めた1020文を使用した。実験では、53人の発話者にあらかじめ使える単語（名詞と動詞）を教え、一人当たり20文ぐらいの質問文を発話してもらった。この実験によって集まったデータの内106文を評価に、残りの914文を学習に使った。評価セットとして使用しているデータは初期文法を使って106文中39文が解析が可能である。

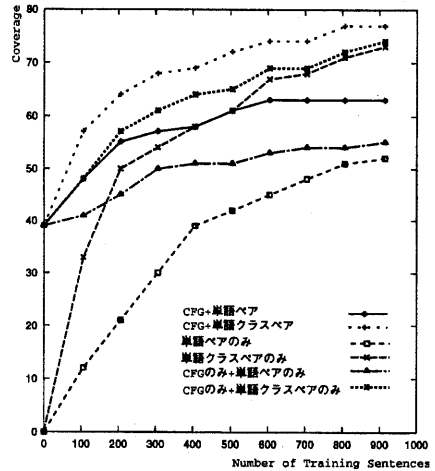


図2: 学習によるカバー率の変化

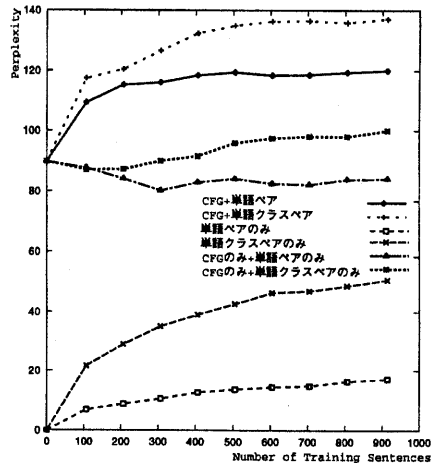


図3: ペアの登録によるパープレキシティの変化

単語（単語クラス）ペアの学習によるカバー率の変化をまとめたのが図2で、パープレキシティについてま

とめたのが図3である。図中の“CFG + ベア”はCFGでの解析とは別にベアを部分解析木の接続に使うことによる解析を並列に行なう方法で、“CFGのみ+ベアのみ”はCFGでの解析とは別に単語(単語クラス)ベアによる解析を並列して行なう方法のことである。単語ベアの学習より、単語クラスによる学習の方がカバー数が増えていることが分かる。また、カバー数は方法1のベア+CFGで77文(単語クラスベアの登録時)まで増えた。それに伴いパープレキシティはかなり増加している。パープレキシティだけを考えれば方法3のベアのみによる方法が一番良いが、初期文法のパープレキシティを考慮すればパープレキシティの増加は十分に抑えられていると言える。

図4はベアの代わりに確率を使って単語クラスバイグラムを学習した場合のパープレキシティの結果を示している(注:カバー率は単語クラスベアと同一である)。図より単語クラスベアによる確率の導入によって、パープレキシティがかなり改善されることが分かった。また、バイグラムだけで解析を行なうようにすればパープレキシティをかなり改善することができるので、単語クラスベアの確率化だけでなくCFG自体の確率化も行なうとより効果があるが結果については割愛する[6]。

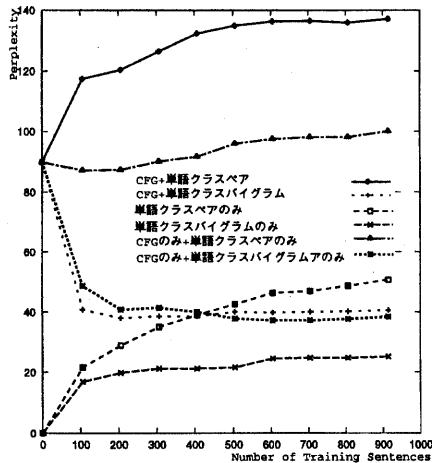


図4: バイグラムによるパープレキシティの変化

結局、学習文数が少ない間はCFGを中心とした解析を行ない、学習文数が増えるに従い単語クラスバイグラム中心とした解析を行なうようにするのがよいと言える。

## 6 音声認識実験による評価

本節では、今までテキストレベルで評価してきた各登録方法が実際の音声認識においてどれくらい有効であるかについて述べる。

### 6.1 実験条件

今回行なう音声認識実験は文法の学習効果を調べるのが目的である。従って、元の文法の解析可能文に対

表1: 音声の分析条件

サンプリング周波数	12kHz
解析窓	ハミング窓
解析窓幅	21.33ms (256 ポイント)
フレーム周期	8ms (96 ポイント)
LPC 分析次数	14 次

して高い認識率を得る音声認識モデルと音声データを使う方が文法の学習効果を調べる上では効果的である。そこで研究室のシステムに使いなれている本研究室の学生の発声した文を認識対象にして、50文で話者適応化を行なった話者適応モデルで認識実験を2話者について行なった。使用した文法はテキストでの実験と同様のものを使用した。終端記号(単語)は、テキストでの実験と同じものから間投詞を省いて実験を行なった。HMMは5状態4出力分布の連続出力分布型HMM(全共分散行列を使用)を使用した。評価データは、テキストでの実験と同様の106文を使用した。認識単位となる音節数は113である。

音声の分析条件は表1の通りである。音声特徴パラメータとして、10次元のLPCメルケプストラム係数と、動的特徴量としてその線形回帰係数を組み合わせ用いている。

### 6.2 規則の確率化を行なわない場合

今までいくつかの文法の学習法をテキストレベルでカバー率などを調べる実験をしてきたが、その中で基本となるCFGのみの文法以外にカバー率とパープレキシティのバランスがよかった単語クラスベアを用いた以下の学習法について音声認識実験を行なった。

#### 方法1 CFGのみ

従来のCFGだけを使って単語予測を行なう方法。

#### 方法2 CFG+単語クラスベア

従来のCFGに加えて、単語クラスベアをCFGの部分解析木の接続に使用して単語予測を並列に行なう方法。

#### 方法3 CFGのみ+単語クラスベアのみ

従来のCFGに加えて、単語クラスベアによる予測を並列に使用して単語予測を行なう方法。

#### 方法4 単語クラスベアのみ

単語クラスベアのみによって単語予測を行なう方法。

結果は文正解数(A)と助詞誤りを無視した文正解数(B)についてまとめた。文正解数は認識結果と正解文とが完全に一致した文の数をまとめたもので、助詞誤りを無視した文正解数は助詞については考慮せずに正解と一致した文の数についてまとめたものである(助詞誤りを無視した文正解数を示したのは、助詞の音声認識が難しいと、ほとんどの助詞は言語処理部で復元できるからである)。textの欄はテキスト入力の場合に解析された文数、括弧内の値は辞書から間投詞を除いた文法によるテストセットパープレキシティを表している。

方法2(CFG+単語クラスベアをCFGの補間に用いる方法)を除いて学習文数を増やすと認識率が上昇する傾向にあることが分かる(方法1については学習前

表 2: 406 文で文法を学習した場合の正解認識文数

A: 文正解数 (テスト文数: 106)

B: 助詞誤りを無視した文正解数

括弧内の値はテストセットパープレキシティを表す

話者	方法 1		方法 2		方法 3		方法 4	
	A	B	A	B	A	B	A	B
KO	21	33	31	50	37	58	36	54
AD	34	36	45	52	50	58	44	54
text	39(75.2)		72(121.9)		64(82.3)		58(38.9)	

表 3: 914 文で文法を学習した場合の正解認識文数

A: 文正解数 (テスト文数: 106)

B: 助詞誤りを無視した文正解数

括弧内の値はテストセットパープレキシティを表す

話者	方法 1		方法 2		方法 3		方法 4	
	A	B	A	B	A	B	A	B
KO	21	33	29	52	41	59	42	58
AD	34	36	44	50	53	60	53	60
text	39(75.2)		77(127.1)		74(91.8)		73(50.6)	

の CFG による結果なので、学習セットによって認識率は変わらない。方法 2 の認識率が上昇しなかった理由はパープレキシティの大きさによると考えられる。実際認識文数はあまり変わらない結果となっている。方法 3 (CFG+単語クラスペアで同時に予測を行なう方法) と方法 4 (ペアのみの方法) を比べると、学習データが少ない時は方法 3 の方が良く、学習データが増えれば方法 4 の方が良くなっている。単語クラスペアの学習数が少ない場合は CFG での予測で認識文数が上昇していたものが、単語クラスペアの学習数が増えると方法 4 と方法 3 は同程度の認識率が得られていることから、CFG で認識できていた文が単語クラスペアだけでも解析できるようになったためと考えられる。どの結果も元の文法での結果よりも良くなっているもの、単語クラスペアのみの方法でも良い結果が得られることから、学習データが多い場合には一番良い方法であると言える。また、学習データが少ない場合には単語クラスペアに CFG での予測を組み合わせた方が良い結果が得られることから、学習データが少ない場合は CFG を組み合わせた方が良く、学習が進むにつれて単語クラスペアのみの解析にした方が良いと言うテキストレベルの場合と同じ結論が得られた。

表 4: 38 文での正解認識文数 (確率なし)

(2 話者の認識文数の平均, 学習: 914 文)

	CFG	単語クラスペア
平均認識文数	27.5	26.5
パープレキシティ	74.7	54.7
SMR パープレキシティ	100.2	73.2

表 4 は、CFG、単語クラスペア両学習法で解析可能な 38 文での正解認識文数の違いについてまとめた結果である。これは、学習による認識率の上昇が単語クラスペアでの解析可能文数が増えたためとも考えられるので、両解析法で解析可能な文を使って評価を行なったものである。

SMR パープレキシティは中川らが提案している以下の式によって求まるパープレキシティに代わる評価規準である [7]。

$$F_{SMR}(P) = \left\{ \frac{1}{n} \left( \sqrt{\frac{1}{P(w_1|\cdot)}} + \sqrt{\frac{1}{P(w_2|w_1)}} + \dots + \sqrt{\frac{1}{P(\cdot|w_1 \dots w_n)}} \right) \right\}^2$$

パープレキシティと SMR パープレキシティは CFG と比べて単語クラスペアの方が小さいが、受理できる文に対しては、両方法とも文認識率は約 70% (助詞誤りを無視すれば約 80%) で、ほぼ同じであった。

### 6.3 規則の確率化を行なった場合

前節の確率を導入しない場合で認識率があまり上昇しなかった単語クラスペアを CFG の補間に使用する方法 (方法 2) を除いた 3 つの方法について、確率を導入した場合の認識率を調べる実験について述べる。前述した以下の 3 つの学習法についての認識実験を行なった。

方法 1 SCFG のみ

方法 3 SCFG のみ+単語クラスバイグラムのみ

方法 4 単語クラスバイグラムのみ

表 5: 確率法による正解認識文数

A: 文正解数 (テスト文数: 106)

B: 助詞誤りを無視した文正解数

括弧内はテストセットパープレキシティを表す

話者	方法 1		方法 3		方法 4	
	A	B	A	B	A	B
KO	20	32	50	64	50	64
AD	33	35	60	70	60	69
text	39(47.1)		74(34.3)		73(25.2)	

確率を使った言語モデルでもバイグラムのみの方が SCFG のみの結果よりも良く、解析できる文のうち約 75% の文が正しく認識できた (助詞誤りを無視すれば約 90%)。また表 4 と同様に両解析法で解析可能な 38 文で正解認識文数を比較した結果を表 6 に示す。

表 6: 38 文での正解認識文数 (確率あり)

(2 話者の認識文数の平均, 学習: 914 文)

	SCFG	単語クラスバイグラム
平均認識文数	26.5	33
パープレキシティ	46.4	19.7
SMR パープレキシティ	150.7	34.9

単語クラスバイグラムを使用することにより、パープレキシティ、SMR パープレキシティ共に SCFG の場合 に比べてかなり小さくなったことから認識率の向上 (68% → 85%) が見られた。この結果より単語クラスバイグラムの言語モデルが SCFG や確率を使わない言語モデルに比べて優れていることが分かる。

最後に、各学習法の学習文数が少ない場合 (506 文) の結果を表 7 に示す。解析の際の各パラメータは学習文数 914 の時の最も良かった結果に準じて使用した。

表 7: 学習文数の変化による認識文数の変動  
A: 文正解数 (テスト文数: 106)  
B: 助詞誤りを無視した文正解数

学習文数	話者	Bigram1		Bigram2		Bigram1-SCFG	
		A	B	A	B	A	B
406 文	KO	44	61	43	69	47	62
	AD	48	60	52	70	52	63
914 文	KO	50	64	44	70	50	64
	AD	60	69	57	75	60	69

表中の Bigram1 はスムージングを行わない場合のバイグラムのみでの解析結果、Bigram2 はユニグラムによってスムージングを行なった場合のバイグラムのみでの結果、Bigram1-SCFG は SCFG とバイグラムを併合した解析での結果を示している。学習文数を増やすことによって一様に認識率が良くなっており、学習の効果があることが分かる。しかし、Bigram1 (バイグラムのみ) の上昇に比べて、他の方法は認識率の上昇が少ない。これは、学習文数が少ない時に学習の不足をスムージングや SCFG によって補っていた効果が学習文数が増えたことにより効果が薄れてきたせいだと考えられる。

以上、テキストで調べてきた文法の学習方法と効果を実際の音声認識でどれぐらいの効果があるかを調べた。文法の学習効果については学習セットを増やした時に正解認識文数の向上を確認でき、効果があることが確認された。また学習方法については、学習セットがある程度ある場合には単語クラスバイグラムのみによる解析が最も効果があり、学習セットがあまりない場合には単語クラスバイグラムと SCFG を併せて解析する方法が優れているというテキストでの結果と同じ傾向となった。

## 7 結論

音声認識システムでユーザが受理できない文を入力した時に、新しい規則を登録し、入力文のカバー率を改善するシステムの開発を行ない、「富士山観光案内システム」のタスクを使うことにより評価を行なった。

テキストレベルでの評価においては、単語クラスベアを使った方法のうちベアを部分木の接続に使用する方法がカバー率とパープレキシティのバランスが一番良いということが分かった。

また、確率を導入することにより単語ベアのパープレキシティの増加を抑えることも試みた。CFG を使わずに単語 (単語クラス) バイグラムだけで解析する場合には効果があったが、カバー率は十分ではなかった。一方、CFG とバイグラムを組み合わせて使用する方法は、確率を学習するためのデータが不十分な場合にも有効であった。バイグラムと SCFG とを並列に使用する

場合はカバー率、パープレキシティの両面において非常に効果があった。

大規模なデータベースを利用できる場合はトライグラム等の確率モデルの方が CFG のような文法よりもパープレキシティの面で優れているが、前もって大量のデータが得られないようなアプリケーションでは CFG ベースの文法の方が利用価値がある。すなわち初期文法は CFG を使い、以後学習データの増加とともに n-gram への移行 (2 つの解析に対する重み  $\lambda$  を  $\lambda = 1$  から  $\lambda = 0$  に逐次変更することに相当) というのがアプリケーションによっては現実的であると思われるし、実際の実験結果でもそのような結果が得られた。本稿で提案した手法は、CFG と n-gram の両方の利点を活かした文法学習法と言える。

音声認識実験による評価においては、テキストレベルとは同等のことが言えた。言語モデルの確率を使わない場合には、学習セットが少ない間は単語クラスベアやバイグラムだけでなく CFG を組み合わせた文法を使った方が認識率が良く、学習が進むにつれて単語クラスベアやバイグラムのみで単語予測を行なった方が認識率からも文認識時間からも良い結果が得られた。

また、確率を用いても確率を用いない場合と同じ傾向であったが、パープレキシティが下がったことから認識率の面でさらに良い結果が得られた。テキストレベルから音声認識レベルまで一貫した評価を行なった結果によって、学習データが少ない場合の文法の構築において本稿で提案している CFG と単語クラスベアやバイグラムを組み合わせた文法が有効であることを示せた。

## 参考文献

- [1] 山本, 伊藤, 肥田野, 中川: 人間の理解手法を用いたロバストな音声対話システム, 情報処理学会論文誌, VOL.37, No.4, pp.471-482 (1996.4)
- [2] 大谷 耕嗣, 中川 聖一: CFG とバイグラムの結合による文法の半自動修正法, 情報処理学会 音声言語情報処理研究会, 95-SLP-9-15, pp.99-104 (1995.12)
- [3] 大谷 耕嗣, 中川 聖一: 話し言葉用文法の Bigram の追加による自動学習法, 電子情報通信学会, 言語理解とコミュニケーション技報, NLC96-12, (1996.7)
- [4] G.J.F. Jones, J.H.Wright, E.N. Wrigley: The HMM Interface with Hybrid Grammar-Bigram Language Models for Speech Recognition, Proc. ICSLP-92, pp.253-256(1992)
- [5] 竹澤 寿幸, 森元 暎: 部分木を単位とする構文規則と前終端記号のバイグラムを利用した連続音声認識, 情報処理学会音声言語情報処理研究会, 95-SLP-9-9, pp.55-62 (1995.12)
- [6] 大谷 耕嗣: 話し言葉用文法の自動学習, 豊橋技術科学大学修士論文 (1997.2)
- [7] 中川 聖一, 伊田 政樹: タスクの複雑さを表す新しい尺度 - SMR-Perplexity -, 電子情報通信学会, 音声技報, SP96-101, (1997.1)