音声認識を目指した確率GLR法を用いた言語モデルの構築

今井宏樹† 田中穂積† 徳永健伸†

本論文では，音声認識処理に確率GLR法を利用した言語モデルを導入する試みについて述べる．確率GLR法は，通常のGLR法で生成される構文木に対してその生成確率を与える確率モデルの総称を指す．その中でも，特にbigram LR表やPGLRモデルは，文脈依存性を扱うことができる，終端記号の連接情報（bigram）を同時に扱うことができる，などの利点がある．一方，音声認識においては，精度の良い確率言語モデルを構築することが重要な課題となっている．本研究では，異音を終端記号とするCFGを用いてbigram LR表，PGLRモデルの2種類の言語モデルを構築し，異音のbigramやtrigramのみの確率モデルとの比較実験を行った．その結果，PGLRモデルはtrigramモデルのテストセットパープレキシティの値を約15%下げることができた．

Construction of Language Models Using Probabilistic GLR Methods toward Speech Recognition

HIROKI IMAI,† HOZUMI TANAKA† and TAKENOBU TOKUNAGA†

We describe an attempt to apply language models using Probabilistic GLR methods to speech recognition. Probabilistic GLR methods are stochastic models that assign a probability to each generated parse tree. In particular, the PGLR model and the bigram LR table have the advantages that they are mildly context sensitive and at the same time readily interfaces with information about the connectability of terminal symbol pairs (so called bigrams). On the other hand, the construction of a good stochastic language model is one important task in speech recognition. To evaluate the relative merits of the given methods, we built two language models, the PGLR model and the bigram LR table, using a CFG with allophone symbols, and compared our model with both allophone-level bigram and trigram models. Experiments showed that the PGLR model reduced the test-set perplexity of the trigram model by about 15%.

1. はじめに

音声認識の分野において，音響モデルだけでなく，精密な言語モデルの作成が重要な課題の1つとなっている．最近では，単語n-gramモデルを基本とした言語モデルが主流であり，単語のbigramとtrigramを併用して，認識精度を落とさずに計算量を削減する手法が提案されている[1]．

しかしながら，音声は言語と深く結び付いており，単純に音声から単語列を取り出すだけでなく，構文処理や意味処理などのより深い言語処理と関連付ける必要があると我々は考えている．このような視点から，我々の研究グループでは，構文解析アルゴリズムの1つであるGLR法[2]を基本として，形態素解析や音声認識との統合を目指した研究を進めてきた[3)~5)]．

† 東京工業大学大学院情報理工学研究科
Graduate School of Information Science and Engineering, Tokyo Institute of Technology

現在は，図1に示すような，HMM-LR音声認識システム[6]を拡張した認識システムの構築を目指している．GLR法では，LR表を利用した先読み記号の予測が可能なため，音声認識に応用すれば音素の予測に役立つとされている．このシステムでは，

- n-gramモデルよりも精密な言語モデルの提供
- 音声認識・形態素解析・構文解析の統合的な解析

が期待できる．

本研究では，確率LR法を用いた言語モデルを作成し，n-gramを基本とした統計的言語モデルとの比較を行うことを目的とする．2章では，本研究で対象とするPGLRモデルおよびbigram LR表の概要を述べ，3章で異音モデルへの適用について説明する．4章では，音声対話コーパスを用いた予備実験の結果を報告する．

2. 確率GLR法を用いた言語モデル

確率GLR法は，GLR法[2]を拡張して，生成された

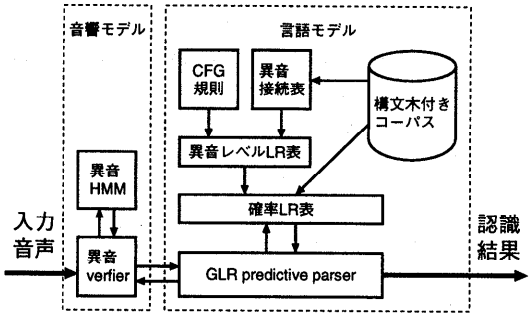


図 1 確率 GLR 法を組み込んだ HMM-LR システム
Fig. 1 An HMM-LR system with a Probabislitic GLR method.

構文解析木に対してその生成確率を付与する確率モデルの総称である．今までに確率モデルの与え方の提案がいくつかなされており，PCFG[7]，Briscoe，Caroll のモデル[8]，bigram LR 表[4]，PGLR モデル[9]，がその主な手法としてあげられる．

本研究では，これらの手法の中で，bigram LR 表と PGLR モデルを対象とする．以下，2.1 節，2.2 節では，bigram LR 表および PGLR モデルの概要をそれぞれ説明し，2.3 節で各モデルの特徴を述べる．

また，複数の連接制約を LR 表に組み込むことで異音レベル，形態素レベル，構文レベルの制約を統合して扱う枠組が提案されている[5]．2.4 節では，この手法の確率 GLR 法への応用について述べる．

2.1 bigram LR 表

bigram LR 表[4] は，終端記号列の bigram 確率を LR 表に統合するための手法である．この手法は，終端記号の bigram 情報が確率付きの接続表として表現できることに注目して，LR 表中の各動作に対しその生起確率を割り当てる点に特徴がある．以下に bigram LR 表の構成方法の概略を示す．

(1) 訓練コーパスから終端記号の bigram 情報を学習する．

(2) 与えられた CFG から通常の LR 表生成アルゴリズムにより LR 表を作成する．

(3) ある終端記号 a を shift して状態 s へ遷移する動作（shift s）を実行した直後の状態 s において，先読み記号 b の欄に定義されている各動作 A_j に対して以下のように確率値を割り当てる．

$$P(A_j) = \frac{P(b|a)}{\sum_{i=1}^{N}\{P(b_i|a)\}\cdot n} \qquad (1)$$

ここで，$\{b_i : i = 1,\ldots,N\}$ は状態 s で定義されているすべての先読み記号を，n は状態 i，先読み記号 b の欄に定義されている動作の数（競合を起こしている動作の数）を，$P(y|x)$ は終端記号列 x，y の bigram 確率を，それぞれ表す．つまり，式 (1) は，動作 A_j が行われる確率が状態 s で定義されるすべての動作について正規化されていることを意味する．

(4) それ以外の状態 t（reduce 動作を実行した直後の状態）に対しては，状態 t で定義されている動作が実行される確率はすべて等しいと見なす．

$$P(A_j) = 1/n \qquad (2)$$

ここで n は式 (1) と同様，同じ欄で競合を起こしている動作の数を表す．

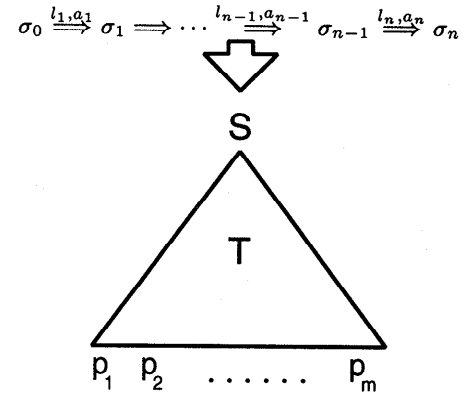


図 2 解析過程と構文木の関係
Fig. 2 Relation between a parsing process and a parse tree.

2.2 PGLR モデル

PGLR モデルは，LR 表中の各 shift 動作，reduce 動作に対してその生起確率を与える形で定義される確率モデルである．GLR 法は，構文的曖昧性を含む一般の CFG も扱えるように LR 法を拡張したものであり，与えられた CFG からあらかじめ LR 表と呼ばれるプッシュダウンオートマトンを作成し，LR 表に記述された動作に従って解析を行う．GLR 法では，初期状態から解析成功までに実行された一連の動作により生成される状態遷移系列が 1 つの解析木に相当する．したがって，状態遷移系列の生起確率が構文木の生成確率と等価になる．

このような考え方に基づいて GLR 法における状態遷移系列に確率を与える手法は，Briscoe ら[8] により最初に提案された．しかし，彼らの手法では正規化に問題があり，計算された値が構文木の生成確率とはなっていなかった．Inui ら[9] はその問題点を指摘し，構文木の生成確率を正しく与える PGLR モデルを提案した．以下に PGLR モデルの構成方法の概略を示す．

図 2 において，σ_i，l_i，a_i はそれぞれパーザの状態，先読み記号，実行された動作を表している．初期状態

σ_0 から受理状態 σ_n までのパーザの状態遷移の結果として木 T が生成されるから，その生成確率 $P(T)$ は

$$P(T) = P(\sigma_0, l_1, a_1, \sigma_1, \ldots, l_n, a_n, \sigma_n) \quad (3)$$

と表せる．PGLR モデルでは，各解析ステップの実行される確率は，直前のパーザの状態のみに依存するという仮定を導入し，式 (3) を以下のように近似する．

$$P(T) \approx P(\sigma_0) \cdot \prod_{i=1}^{n} P(l_i, a_i, \sigma_i | \sigma_{i-1}) \quad (4)$$

さらに，GLR 法においては，

- l_i と a_i が決まれば σ_i は必ず一意に決まる．
- reduce 動作では先読み記号を消費せず，直前の動作と同じ先読み記号が用いられる．よって，reduce 動作では，先読み記号を予測する必要がない．

という 2 つの特徴があるため，式 (4) の各条件付き確率の推定は式 (5)，(6) で行うことができる．

$$P(l_i, a_i, \sigma_i | \sigma_{i-1}) \approx P(l_i, a_i | s_{i-1}) \quad (s_{i-1} \in \boldsymbol{S_s}) \quad (5)$$

$$P(l_i, a_i, \sigma_i | \sigma_{i-1}) \approx P(a_i | s_{i-1}, l_i) \quad (s_{i-1} \in \boldsymbol{S_r}) \quad (6)$$

ただし，$\boldsymbol{S_s}$ は shift 動作直後に遷移する状態の集合，$\boldsymbol{S_r}$ は reduce 動作直後に遷移する状態の集合をそれぞれ表す．

式 (5)，(6) の右辺のそれぞれの条件付き確率は，訓練コーパスを GLR パーザで解析して，式 (5) に対しては状態 s_{i-1} において l_i, a_i が発生する回数，式 (6) に対しては状態 s_{i-1}，先読み記号 l_i において a_i が発生する回数から，それぞれ学習することができる．そして，LR 表に定義されている各動作に対してこの条件付き確率を割り当てることにより，PGLR モデルが構築される．

2.3 2 つの言語モデルの特徴

bigram LR 表と PGLR モデルには以下のような特徴がある．

(1) 文脈依存性が考慮される
bigram LR 表や PGLR モデルでは，その構築過程において LR パーザの状態が考慮される．パーザの状態はその時点までに解析されたスタックの状態に依存するため，左文脈にある程度依存していると考えることができる．
また，bigram LR 表では shift 動作後に遷移する状態 $\boldsymbol{S_s}$ のみに対して確率値を割り当てているのに対し，PGLR モデルでは reduce 動作後に遷移する状態 $\boldsymbol{S_r}$ に対しても確率値を割り当てている．

(2) 終端記号の連接情報が考慮される
bigram LR 表では，その定義より明らかである．また，PGLR モデルにおいても，式 (5) は，スタックの状態 σ_{i-1} から次の先読み記号 l_i を予測するモデルとなっている．一方，スタックの状態はそれまで解析済みの入力列 $l_1, \ldots, l_{i-1}$ に関する構文情報を表しているので，終端記号の連接確率 $P(l_i | l_{i-1})$ の前件部分をスタックの状態 σ_{i-1} によってさらに細分化していると考えることができる．したがって，PGLR モデルは終端記号の bigram モデルと同様な終端記号の連接情報が反映される．bigram LR 表，PGLR モデルは終端記号の bigram 情報に構文的制約が加わっていると考えられるため，終端記号の bigram モデルに比べて良い性能を持つことが予想される．

(3) 比較的容易にモデルを学習できる
bigram LR 表の場合は，モデルの学習方法は単なる bigram モデルの学習方法と同一である．PGLR モデルの学習は，コーパスを解析して使用された各動作の回数を数え上げることにより行われる．与えられた各例文に対して正解となる構文構造が与えられていれば，学習のコストは bigram や trigram の学習と比較してもそれほど大きくはならない．

2.4 複数の連接制約の LR 表への組込み手法の利用

bigram LR 表や PGLR モデルを 3 章に示す手法で異音モデルに組み込んだ場合，異音レベルの連接関係しか考慮されない．そのため，以下の要因で曖昧性が組合せ的に増大する可能性がある．

- 多品詞語が連続して現れる
たとえば，「お/尋ね/し/た/い/ん/で/す/けど」という発話を 4 章の実験で用いた文法で解析する場合，「し」「た」「い」「ん」の各単語には 6 種，4 種，9 種，6 種の品詞カテゴリがそれぞれ割り当てられており，少なくとも $6 \times 4 \times 9 \times 6 = 1296$ 種類の解析木が生成される．
- 不適当な語切りが発生する
たとえば，「何名/様/で」という句が異音列に展開されると，形態素情報が失われるために，"nanmei/sa/made" という誤った形態素列も生成される．

この問題を解決するために，綾部ら[5] は複数の連接制約を LR 表に組み込む手法を提案し，異音の連接関係だけでなく，細品詞の連接関係も同時に LR 表に反映させる手法を示した．ここで細品詞とは，**表 1** に示さ

表 1 本研究で用いた名詞に関する細品詞の体系
Table 1 A list of morphological categories of nouns used in our experiments.

細品詞ラベル	分類の説明	単語事例
hutu-meisi	普通名詞	さしみ，乗車券，空港
hutu-meisi-post	名詞句の先頭に来ない普通名詞	以上，客
n-keisiki	形式名詞	こと，ところ，もの
hutu-meisi	副詞的名詞	いかほど，ため
zikan-meisi	時間名詞	さ来週，翌日
zikan-meisi-ippai	「いっぱい」が後続する時間名詞	あした，今日
suuryousi	数量詞	何日，何人，全員
koyuu-meisi	固有名詞	ニューヨーク，京都駅
myoji-first	日本人の名字	林，田中，鈴木，西郷
namae-last	日本人の名前	あきこ，弘子，和夫
myoji-last	外国人の名字	フィリップス，レノン
namae-first	外国人の名前	フィリップ，メアリー，ジョン
daimeisi	代名詞	あれ，僕
daimeisi-domo	「ども」が後続する代名詞	私
daimeisi-exp	代名詞の特殊表現	あたし，わたくし，こちら
wh-daimeisi	疑問代名詞	いつ，誰，どちら
wh-daimeisi-nan	疑問代名詞「何」	何

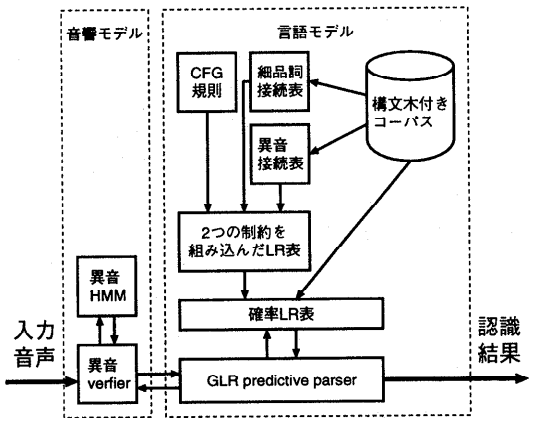


図 3 細品詞の連接制約を組み込んだ拡張 HMM-LR システム
Fig. 3 An extended HMM-LR system with constraints on the connectability of morphological categories.

れるような品詞を細分類した言語単位を指す．

図 1 のシステムは，この拡張によって図 3 のようになる．細品詞のレベルで連接制約が加わることによって，より精密な言語モデルの提供が可能になると期待される．

しかし，綾部らは手法の提案のみ行っており，確率 GLR 法に対しては応用可能であることを示唆したにとどまっている．そこで，我々は，bigram LR 表について，異音の連接情報のみを利用した場合，異音と細品詞の 2 つの連接情報を同時に利用した場合の比較を行い，この手法の有効性を評価する．なお，我々は PGLR モデルに 2 つの連接情報を同時に組み込んだ場合の確率の与え方は 2.2 節の手法とほぼ同じでよいことをすでに確認しているが，現在実験による評価を行っている段階であり，その手法と結果は機会を改めて発表する予定である．

3. 確率 GLR モデルの異音モデルへの適用

この章では，まず，音声認識で使用される認識単位と異音モデルの関係について述べ，次に異音モデルを終端記号とする CFG を作成し確率 GLR モデルを構築する手順を示す．

3.1 異音モデル

音声認識の分野においては，音声の認識単位として音素を使用するのが一般的である．しかしながら，同じ音素であっても，前後の音素との連接関係により音響的特徴も変化することが知られている．そこで，認識単位をモデル化する際に，音素の文脈依存性が反映されることが望ましい．

異音モデルは，このような音素文脈依存モデルの 1 つである．音素文脈依存モデルでは，すべての音素文脈パターンを別々に扱うと学習や認識のコストが増大するため，音響的に似ている音素文脈パターンをクラスタリングするのが一般的である．

異音モデルは，すべての音素文脈パターンを 1 つの状態で表すところから始めて音響的特徴の差異の大きい音素文脈から順に状態分割を行う SSS (Successive State Splitting)[10] と呼ばれる手法を用いて音素文脈のクラスタリングを行っている．トップダウンなクラスタリングのため，すべての音素文脈パターンは必ずいずれかのクラスタに属し，学習用音声データが多くなくても比較的良いクラスタを作成することが可能であるという利点がある．

我々は，ATR で学習された 401 状態の HMnet から生成された 1547 個の異音モデルを使用した．基に

```
元の辞書規則：  <koyuu-meisi> -> ニューヨーク
異音展開後の辞書規則：  <koyuu-meisi> -> <_n> j5 u25 u26 j7 o244 o402 k38 <_u>
異音導出規則：  <_n> -> n1
                <_n> -> n2
                     :
                <_n> -> n25
                <_u> -> u1
                     :
                <_u> -> u48
```

注：非終端記号は便宜上 "<>" で囲んである．また，"_" で始まる記号は音素を表している．

図 4 辞書規則の異音列への展開
Fig. 4 Expansion of a word to a sequence of allophones in a dictionary rule.

表 2 基になる音素モデル一覧
Table 2 A list of basic phone models.

基になる音素体系（計 26 音素）
- a i k j o zh z u d m g ch
ng r sh ts s e b q t w n p h

表 3 音素 /b/ の異音モデル
Table 3 Allophone models for a phone /b/.

音素 /b/ の異音		
	音素文脈	ラベル
左	- i j zh u r e q w	b1
中	b	
右	i k j zh z u d m ch ng r sh e b q t n p h	
左	- i j zh u r e q w	b2
中	b	
右	- a o g ts s w	
左	a k o z d m g ch ng sh ts s b t n p h	b3
中	b	
右	i k j zh z u d m ch ng r sh e b q t n p h	
左	a k o z d m g ch ng sh ts s b t n p h	b4
中	b	
右	- a o g ts s w	

なる音素モデルの一覧を**表 2** に，異音モデルの一部を**表 3** にそれぞれ示す．

3.2 異音を終端記号とする CFG の作成

異音を終端記号とする CFG を作成するためには，異音列に展開された単語辞書規則を用意する必要がある．具体的には，次のような手順で辞書規則を作成する．

(1) 辞書規則の右辺の単語を音素列に展開する．
(2) 音素列を前後の音素文脈を見ながら異音列に変換する．ただし，先頭と末尾の音素は文脈が決定できないのでそのままにしておく．
(3) 各音素ごとに，
音素 → 異音
なる展開規則を追加する．

図 4 に，「固有名詞 → ニューヨーク」という辞書規則を異音列に展開した例を示す．

3.3 確率 GLR 法による言語モデルの生成

2.1，2.2 節でそれぞれ示した確率 GLR 法の定義に基づいて確率 LR 表を生成する手順は，以下のとおりである．

(1) 学習用例文を異音列に展開し，異音の連接情報を抽出する．
(2) CFG と (1) で得られた連接情報から制約伝播アルゴリズム[11] を用いて LR 表を作成する．
(3) (2) で得られた LR 表を基に，bigram LR 表，PGLR モデルの各生成手順に従って確率 LR 表を作成する．

ここで問題となるのが，学習用データの不足によるデータスパースネス問題である．文法の規模が大きくなると LR 表中に定義される動作の数も増えるため，学習用データの解析では使用されない動作が多数現れる．そこで，我々は，bigram LR 表に対しては bigram 確率の計算においてバックオフスムージング[12] を，PGLR モデルに対してはあらかじめすべての動作に対して一定の低い頻度を与えるフロアリングを，それぞれ行っている．

4. 対話コーパスを用いた評価実験

4.1 コーパスと文法

本研究では，音声対話を書き起こした対話コーパスを使用して，連続音声認識実験を行うための予備実験を行った．実験用の対話コーパスとして，ATR で収録された，618 対話，約 21000 文の対話コーパス[13] を使用した．このコーパスには，形態素情報と構文情報がそれぞれ付与されている．文法は，衛藤らがこのコーパスの解析用に作成した細品詞列を導出する日本語句構造 CFG[14]（細品詞数 441）に，コーパスから自動的に作成した細品詞から異音列を導出する辞書規則，音素から異音を導出する規則をそれぞれ付加したものを用いる．この文法の詳細を**表 4** に示す．

今回は，コーパス全体のうち，衛藤の文法体系で解

表 4 実験に使用した文法
Table 4 The grammar for experiments.

規則の内訳	細品詞導出	辞書規則	音素 → 異音	全体
規則数	859	5002	1547	7408
終端記号数	—	1547	1547	1547
平均規則長	1.39	7.21	1.00	5.23

析可能で，かつ半自動的に正解の構文木を付与できた文を実験に用いた☆．その中で，「はい」などの形態素数 1 の文を除いた 9794 文から評価データとして 500 文をランダム抽出し，残りを学習用データとした．実験データ 9794 文の長さ分布を，形態素について**図 5** に，異音について**図 6** にそれぞれ示す．

4.2 評価方法

本研究では，言語モデルの評価尺度としてテストセットパープレキシティ[15] を用いた．パープレキシティは，解析のある時点で次に予測されるシンボル（ここでは異音となる）の候補数の平均値を表し，その値が小さいほど良いとされる．音声認識の分野においては，与えられたタスクの難しさの評価や言語モデルの性能比較などにおいて比較的よく用いられる．

bigram LR 表および PGLR モデルと比較するための言語モデルとして，現在音声認識で一般的に用いられている bigram モデルと trigram モデルを選んだ．2.3 節で述べたように，bigram LR 表と PGLR モデルは bigram モデルよりは良い性能を持つことが予想できるが，trigram との性能差については明らかではない．

また，bigram LR 表については，2.4 節で述べた複数の連接制約を組み込む手法の有効性を検証するため，細品詞間の連接制約を組み込んだモデル☆☆と組み込まないモデルの 2 種類を作成し，実験を行った．

本研究では，4.1 節の評価データに対して各言語モデルで文生成確率を求め，テストセットパープレキシティを計算し，その値を比較した．PGLR モデルでは，フロアリングの値を変化させて最も良い結果を得た値を採用した．ただし，bigram LR 表と PGLR モデルの文生成確率は，確率値の上位 10 位までの解析木の生成確率の和で近似した．その理由として以下の 2 点があげられる．

- 曖昧性の多い文に対してすべての解析木の生成確率値を求めるのが困難である．**表 5** に示すように，異音のみの連接制約を加えた LR 表で生成される解析木の数は膨大である．
- 下位の解析木の生成確率は上位のそれと比較して無視できる程度に小さくなる．**図 7**☆☆☆より，異

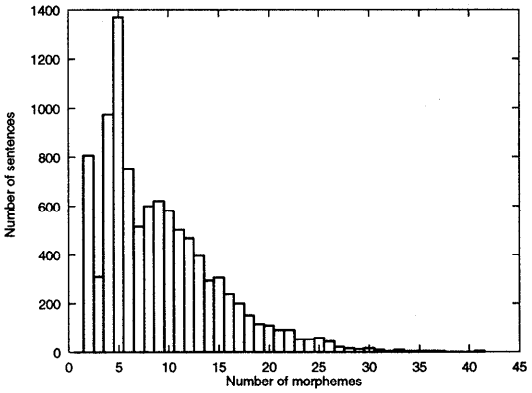


図 5 実験データの 1 文あたりの形態素数分布
Fig. 5 Distribution of the number of morphemes per sentence in the ATR corpus.

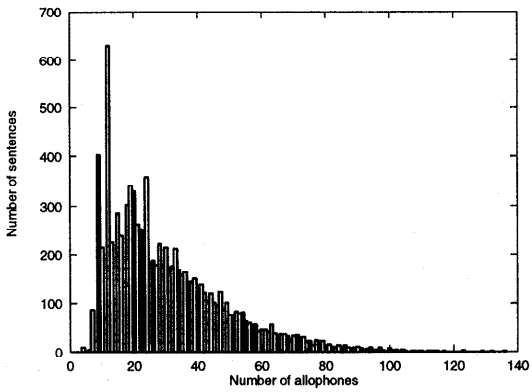


図 6 実験データの 1 文あたりの異音数分布
Fig. 6 Distribution of the number of allophones per sentence in the ATR corpus.

表 5 評価データを解析して得られる 1 文あたりの解析木の数
Table 5 The number of mamimum/average parse trees per sentence in the test-set.

制約の種類	最大値	平均値
異音のみの連接制約	$1.7 \times 10^{9\dagger}$	$4.7 \times 10^{7\dagger}$
異音＋細品詞の連接制約	7.9×10^{4}	5.9×10^{2}

† 解析木数が 4 バイト整数型で扱える最大値を超えた 106 文を除いて計算したため実際にはこの値より大きい．

☆ 衛藤文法は，元のコーパスに比べて品詞が細かく分類されている，動詞と後置詞句の係受け関係が文法に記述されている，などの理由により，コーパスの構文木をそのまま使用することができない．

☆☆ 異音の bigram 確率は異音の連接制約を内包しているので，異音レベル，細品詞レベルの 2 つの連接制約が組み込まれていることになる．

☆☆☆ 異音の連接制約のみ考慮された 2 つの言語モデルでは，1 文あたりの解析木の数が多く，計算機のメモリ不足で 40 位以下の解析木を取り出すことができなかった．

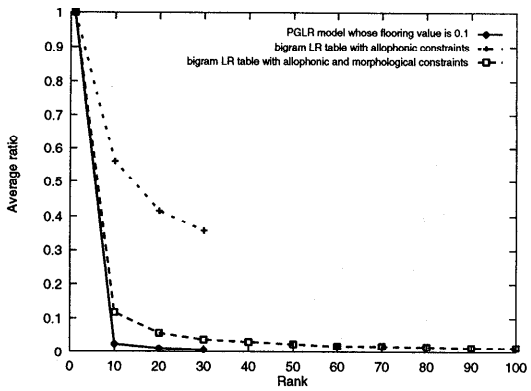


図 7 各言語モデルごとの 1 位の解析木の生成確率に対する比率の平均

Fig. 7 The average ratio of the probability of the best parse tree to that of the n-th parse tree for each language model.

表 6 各言語モデルの異音テストセットパープレキシティ

Table 6 Test-set perplexity of allophones for each language model.

言語モデル	test-set perplexity
異音 bigram	4.41
異音 bigram LR 表	4.20
細品詞接続表+異音 bigram LR 表	3.71
異音 trigram	3.19
異音 PGLR モデル	2.69 (floor 0.1)

音の連接制約のみの bigram LR 表以外の言語モデルでは，下位の値は 1 位の値と比較して十分小さいことが分かる．

4.3 結果と考察

評価データ 500 文のうち，計算機のメモリ不足で解析に失敗した 18 文を除いた 482 文で異音のパープレキシティを計算した結果を**表 6** に示す．5 つのモデルの中で，異音 PGLR モデルが最も小さい値を示している．

Imai ら[16] は，bigram LR 表はパープレキシティの値において bigram モデルよりは良い結果を示したが，trigram モデルほどは良い結果を得られなかった，と報告している．PGLR モデルは，式 (5)，(6) に示されるように，shift 動作直後の状態 S_s と reduce 動作直後の状態 S_r の両方に確率値を割り当てる．一方，bigram LR 表では，式 (6) に相当する，状態 S_r に対しての確率値の割当てが行われない．この S_r への確率値の割当ての有無が，trigram モデルに対するパープレキシティの相対的な差の要因となっていると考えることができる．

また，同じ bigram LR 表でも，異音の連接関係のみが考慮されているモデルよりも，細品詞の連接制約も同時に考慮されたモデルの方が良い結果を示している．さらに，細品詞の連接制約を統合したモデルのパープレキシティは，異音 trigram モデルと比べてそれほど大きな差はない．このことから，複数の連接制約の LR 表に組み込むことが精密な言語モデルの構築に大きく寄与していると考えられる．具体的には，細品詞の連接制約を考慮することで，誤った単語列の生成が抑制され，正しい構文木を生成する動作に対する確率値を大きくすると思われる．

以上の考察から，確率 GLR 法を用いた言語モデルは，音声認識用言語モデルとして有効であると期待される．しかしながら，パープレキシティはあくまでも言語モデルの評価尺度の 1 つにすぎない．今後，実際の音声データを使用した認識実験を行い，文認識率や単語認識率による評価を行う必要がある．

5. ま と め

本論文では，確率 GLR 法を音声認識の言語モデルに応用する試みについて報告した．ATR の音声対話コーパスを用いた比較実験では，PGLR モデル，trigram モデル，bigram LR 表，bigram モデルの順に小さいパープレキシティの値が得られた．したがって，PGLR モデルは音声認識の言語モデルとして有効であると期待される．また，細品詞の連接制約を組み込んだ異音 bigram LR 表は，異音 trigram に対してそれほど大きな差がなかったことから，複数の連接制約を LR 表へ組み込むことが言語モデルの精度向上に貢献することが確認できた．

今後の課題としては，以下の点があげられる．

(1) 大規模な評価実験
今回は対話コーパスの一部を使用しての実験だったため，各言語モデルの学習が不十分であった．十分な学習が行える程度の規模で実験を行う必要がある．

(2) 複数の連接制約の PGLR モデルへの組込み
PGLR モデルに対しても細品詞の連接制約を組み込むことが可能になれば，より高精度な言語モデルを構築できると期待できる．

(3) 音声認識実験による評価
現在使用している確率 GLR パーザはテキスト解析用であり，音声認識部との統合が行われていない．今後，実際の音声データを使用した認識実験を行い，どの程度の認識精度があるかを確認したい．現在，IPA のディクテーションソフトウェア[17] に収録されている音響モデルを利用して，拡張 HMM-LR システムによる認識

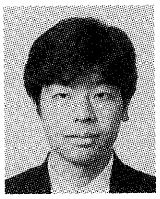
実験を行う予定である.

謝辞 本研究を行うにあたり，対話コーパスおよび異音モデルを提供いただいた ATR の竹沢寿幸氏，林輝昭氏，対話コーパス用日本語 CFG を提供いただいたランゲージウェアの衛藤純司氏に感謝いたします．また，論文の内容に関して適切な助言をくださった査読者の方々に感謝いたします．

参 考 文 献

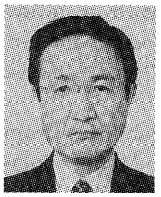
1) Lee, A., Kawahara, T. and Doshita, S.: An Efficient Two-pass Search Algorithm Using Word Trellis Index, *ICSLP98* (1998).
2) Tomita, M.: *Efficient Parsing for Natural Language: A Fast Algorithm for Practical Systems*, Kluwer Academic Publishers (1986).
3) 植木正裕，徳永健伸，田中穂積：EDR 辞書を用いて日本語文の形態素解析と統語解析を行なうシステム，EDR 電子化辞書利用シンポジウム論文集，pp.33–39 (1995).
4) Li, H.: Integrating Connection Constraints into a GLR Parser and Its Applications in a Continuous Speech Recognition System, Technical Report TR96-0003, Department of Computer Science, Tokyo Institute of Technology (1996).
5) 綾部寿樹，徳永健伸，田中穂積：複数の接続表の制約の LR 表への組み込み—LR 表工学 (2)，情報処理学会研究報告，NL, No.117-10, pp.67–74 (1997).
6) Kita, K., Kawabata, T. and Saito, H.: HMM continuous speech recognition using predictive LR parsing, *ICASSP89*, pp.703–706 (1989).
7) Wright, J.: LR Parsing of Probabilistic Grammars with Input Uncertainty for Speech Recognition, *Computer Speech and Language*, Vol.4, No.4, pp.297–323 (1990).
8) Briscoe, T. and Carroll, J.: Generalized probabilistic LR parsing of natural language (corpora) with unification-based grammars, *Computational Linguistics*, Vol.19, No.1, pp.25–59 (1993).
9) Inui, K., Sornlertlamvanich, V., Tanaka, H. and Tokunaga, T.: A New Formalization of Probabilistic GLR Parsing, *International Workshop on Parsing Technologies*, pp.123–134 (1997).
10) Takami, J. and Sagayama, S.: A Successive State Splitting Algorithm for Efficient Allophone Modeling, *ICASSP92*, pp.I-573–576 (1992).
11) Tanaka, H., Li, H. and Tokunaga, T.: Incorporation of Phoneme-Context-Dependence into LR Table through Constraint Propagation Method, *Journal of Japanese Society for Artificial Intelligence*, Vol.11, No.2, pp.246–254 (1996).
12) Katz, S.: Estimation of Probabilities from Sparse Data for the Language Model Component of a Speech Recognizer, *IEEE Trans. ASSP*, Vol.35, No.3, pp.400–401 (1987).
13) Morimoto, T., Uratani, N., Takezawa, T., Furuse, O., Sobashima, O., Iida, H., Nakamura, A., Sagisaka, Y., Higuchi, N. and Yamazaki, Y.: A Speech and Language Database for Speech Translation Research, *ICSLP94*, pp.1791–1794 (1994).
14) 田中穂積，竹澤寿幸，衛藤純司：MSLR 法を考慮した音声認識用日本語文法—LR 表工学 (3)，情報処理学会研究報告，SLP, No.15-25, pp.145–150 (1997).
15) Jelinek, F.: Self-Organized Language Modeling for Speech Recognition, *Readings in Speech Recognition*, Waibel, A. and Lee, K. (Eds.), pp.450–506, Morgan Kaufmann (1990).
16) Imai, H. and Tanaka, H.: A Method of Incorporating Bigram Constraints into an LR Table and Its Effectiveness in Natural Language Processing, *New Method in Language Processing and Computational Natural Language Learning*, pp.225–233 (1998).
17) Kawahara, T., Kobayashi, T., Takeda, K., Minematsu, N., Itou, K., Yamamoto, M., Yamada, A., Utsuro, T. and Shikano, K.: Sharable Software Repository for Japanese Large Vocablary Continuous Speech Recognition, *ICSLP98* (1998).

(平成 10 年 9 月 30 日受付)
(平成 11 年 2 月 8 日採録)



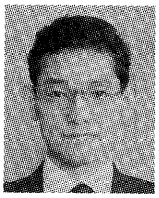
今井 宏樹

1971 年生．1994 年東京工業大学工学部情報工学科卒業．1996 年北陸先端科学技術大学院大学情報科学研究科修士課程修了．同年東京工業大学大学院情報理工学研究科博士課程入学．現在，同大学院博士課程に在学．自然言語処理，音声認識に関する研究に従事．人工知能学会，言語処理学会各学生会員．



田中　穂積（正会員）

1964年東京工業大学理工学部制御工学科卒業．1966年同大学院修士課程修了．同年電気試験所（現，電子技術総合研究所）入所．1983年より東京工業大学工学部助教授．現在，同大学院情報理工学研究科教授．自然言語処理，人工知能に関する研究に従事．工学博士．電子情報通信学会，認知科学会，人工知能学会，計量国語学会，言語処理学会，Association for Computational Linguistics 各会員．



徳永　健伸（正会員）

1983年東京工業大学工学部情報工学科卒業．1985年同大学院理工学研究科修士課程修了．同年（株）三菱総合研究所入社．1986年東京工業大学大学院博士課程入学．現在，同大学院情報理工学研究科助教授．自然言語処理，計算言語学の研究に従事．工学博士．認知科学会，人工知能学会，言語処理学会，計量国語学会，Association for Computational Linguistics 各会員．