

言語データベース統合管理システム

小倉健太郎 橋本一男 森元逞

ATR自動翻訳電話研究所

自然言語処理の研究を進めるためには、多様で複雑な言語現象を分析することが必要である。また、言語データから統計的なデータを抽出し、言語処理や音声認識へ応用できるようにするためには、言語データは大量なものである必要もある。本稿では、多様な言語情報を扱え、容易に言語分析を行うことができるように、高機能ワークステーションを利用したユーザインタフェースでは、言語データをオブジェクト指向で表現し、また大量な言語データを扱えるように、ホストコンピュータのRDBで言語データを一元的に管理する言語データベースシステムについてユーザインタフェースを中心に述べる。

An Integrated Linguistic Database Management System

Kentaro OGURA, Kazuo HASHIMOTO, Tsuyoshi MORIMOTO

ATR Interpreting Telephony Research Laboratories
Twin 21 Bldg. MID Tower, 2-1-61 Shiromi, Higashi-ku, Osaka 540, Japan

This paper describes a linguistic database management system which is able to treat a large amount of linguistic data and easily analyze the linguistic data from various aspects by connecting object-oriented linguistic data representation in user interface using a high performance workstation and a relational database(RDB) in host computer. In this paper, the functions which should be provided in the linguistic database management system are discussed and linguistic data representation in user interface and user interface functions are mainly presented.

1. はじめに

自然言語処理の研究にとって、言語データは言語現象を分析^[1]するための基礎データとして、また辞書^[4]作成のための基礎データとして重要である。複雑な言語現象を解明するためには、いろいろな側面から言語現象を眺めたデータである必要がある。また、言語データから統計的なデータを抽出し、言語解析や音声認識などに応用するためには^{[8][9][10]}、言語データは大量なものである必要がある。

ところが、現在ほとんどの言語データ^[20]は、量としては大量なものもあるが、そこに含まれる言語情報は品詞や活用形などの単語に関する情報のみであることが多い。言語データとして有名なブラウンコーパス^[21]もそのような英語の言語データである。日本語の言語データとしては国立国語研究所の言語データ^[19]があるが、これも単語情報を扱ったものである。構文情報を含む言語データとしては、CCPP(Computer Corpus Pilot Project)コーパス^[25]などがあるだけである。

そのような言語データを扱う言語データベース管理システム^[24]も、単純なもので十分であった。しかし大量でかつ多様な言語情報を扱わなければならない言語データベース管理システムでは、各種の言語情報を組み合わせて検索できることや統計処理など高度な機能を持つ必要がある。

一方、データベースシステムを考えた場合、データを関係モデルで表現する関係データベース(Relational Database(RDB))^[22]が現在の主流であるが、それは本研究で扱おうとしている言語データのような種々のものが複雑に関連しあったものを表現する能力が弱いことが指摘されている^[15]。また、最近注目を浴びつつあるオブジェクト指向データベース^[23]は、データの表現能力には優れているが、まだ試験的な段階で本格的に大量のデータを扱ったものはまだない。

そこで我々は、ワークステーションのユーザインタフェースでは、言語データをオブジェクト指向で表現することにより、多様な言語情報を扱え、容易に言語分析を行うことができるようにし、また大量な言語データを扱えるように、ホストコンピュータのRDBで言語データを一元的に管理する言語データベースシステムを開発している。このようなシステムでは、言語データへの高度な検索機能を含めて、広い意味でのユーザインタフェースが重要である。ここでは、言語データをオブジェクト指向で表現したユーザインタ

フェースを中心に言語データベース統合管理システム^{[7][16]}が持つべき機能について述べる。

2. 言語データ

まず、言語データとしてどのような言語情報を収集すべきかを考えてみると次のような性質を持つことが重要である。

(1)最も基本的な言語情報

(2)一般性を持った言語情報

(1)はそこから組み立てて、構文、意味、文脈構造抽出し分析することが可能であるための性質である。(2)はたくさんの人に言語データを利用できるための条件である。そのことを踏まえて、言語データベース統合管理システムでは、以下に示すような言語データを扱えるようにしている。

(1)言語の構成(全体部分関係)

(a)会話、(b)発話、(c)文、(d)文節、(e)単語

(2)単語に関する情報

(a)会話(文章)に現れた単語、

(b)その文脈での読み、

(c)標準表現、(d)品詞、

(e)活用型、(f)活用形、(g)音便

(3)単語間の関係^{[2][3]}

(a)係り受け関係、(b)格関係、(c)接続関係

(4)言語間の対応対応^[5]

(a)文、(b)格要素、(c)文節、(d)単語、

(e)その他

言語データは、一般性を持たせるため特殊な言語理論には依存しないようにしている。品詞体系は学校文法に準拠しており^[18]、意味関係も一般的なものを網羅したものにしている^[13]。品詞体系および意味関係の体系には、階層構造を持たせている。また、翻訳処理研究用として使えるように言語間の対応データを扱えるようにしている。

言語データは各種情報毎に収集支援システム^[6]を使って収集している。文の表現など言語表現を網羅し、言語の統計的なデータに実際上の有意性を持たせるため、言語データとしては、最低100万語集める必要があると考えて収集を進めている。^{[11][12][17]}

3. システム構成

言語データベース統合管理システムへの要求条件を満足するためのシステム構成について述べる。図1に言語データベース統合管理システムの構成を示す。システムは、言語データ収集システムにより得られたデータを格納管理する部分と、

言語データを利用するためのユーザインタフェース部^[16]からなる。大量の言語データを効率良く格納するため、ホストコンピュータに、RDB (Relational Database)を用いて言語データを格納している。複雑で、多様な検索を容易にするためワークステーション上では、言語データはオブジェクト指向の表現になっている。

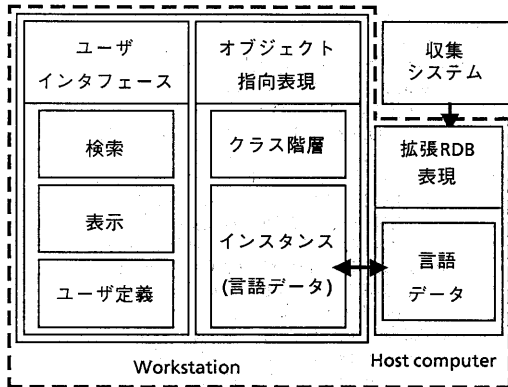


図1.言語データベース統合管理システムの構成

言語データベースには大きく分けて二つの利用の仕方がある。一つはあるテキストをいろいろな側面から詳しく分析するマイクロ分析であり、もう一つは言語現象を全体的な傾向から分析するマクロ分析である。

マイクロ分析は、一般的には分析するデータ量は少ないが、各種の情報を組み合わせた複雑な言語データの検索が必要になる。例えば、“の”で結び付けられた単語が、どのような意味関係で結び付けられている時に、英語でどのように表現されるかを検索するなどがマイクロ分析の例である。

またマクロ分析には、検索自体は比較的単純であるが、大量の言語データの検索が必要になる。マクロ分析は、言語データの統計的な分析である。例えば、電話会話のデータに対して、品詞や意味関係毎に頻度を求めることなどが考えられる。

このシステムは、マイクロ分析とマクロ分析用に、2つのモードを用意する。マイクロ分析モードでは、言語データは予めホストコンピュータからワークステーション上にロードされており、ワークステーションの中で実際の検索が行われる。マクロ分析モードでは、実際に言語データの検索要求があった時点で、ホストコンピュータの中で検索が実行され、結果がワークステーションに送られる。検索のコマンドは、形式的には、モードに関係なく同様の検索記述言語が使える。

データ格納に既存のデータベースシステムを改良したものを使い、主にRDBとユーザインタフェースのつながりと、ユーザインタフェースを開発するという方式により、上記の要求を満たすシステムを最初から開発するのに比べて効率的な開発が期待できる。

4. オブジェクト指向の言語データ表現

4.1 データ形式

上で述べたような言語データは、単純なものではなく、種々のものが複雑に関連しあったものである。上述したマイクロ分析では、これらの複雑な関係を直接的に検索や操作できる必要がある。このため、ワークステーション上では、言語データをオブジェクト指向で表現している。図2に言語データをオブジェクト指向で表現した例を示す。

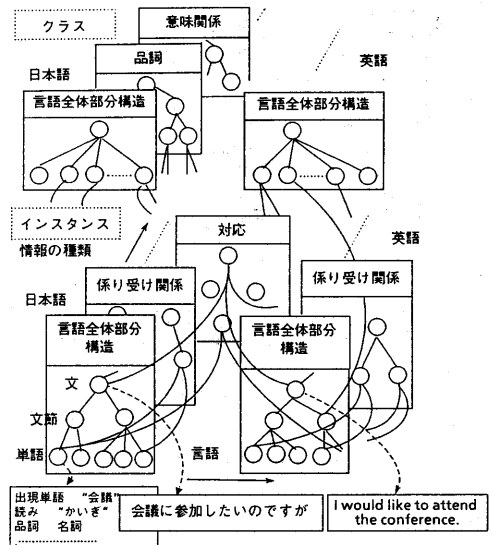


図2. 言語データをオブジェクト指向表現した例

言語データは二種類のデータからなる。一つは、同じ概念の集合を表すクラスであり、これは階層構造を持つ。もう一つは実際の言語データを表すインスタンスである。図3にオブジェクト指向表現言語データの記述例を示す。

クラスで表されるのは、言語の構成要素である会話、発話、文、文節、単語、各種の関係、品詞などである。品詞や意味関係をクラスで表すことにより、それらの階層構造を自然に表現できる。例えば、品詞は、動詞、名詞、助詞、助動詞などからなり、さらに助詞は格助詞、係助詞、副助詞、接続助詞、終助詞などに細分されるが、そ

クラスの例 (文クラス)	
(文	(superclass 言語構成要素)
	(has-slots
	(part-of
	:slot-value (instance-of 発話))
	(composed-of
	:slot-value (instance-of 文節))
	(priv-is
	:slot-value (instance-of 文))
	(next-is
	:slot-value (instance-of 文))
(対応	:slot-value (instance-of 対応関係))
インスタンスの例	
(文1	(instance-of 文)
	(part-of 発話1)
	(composed-of 文節1 文節2)
	(priv-is nil)
	(next-is 文2)
(対応	対応関係1)
(文節1	(instance-of 文節)
	(part-of 文1)
	(priv-is nil)
	(next-is 文節2)
	(composed-of 単語1 単語2)
.....	
(単語2	(instance-of 単語)
	(part-of 文節1)
	(priv-is 単語1)
	(next-is 単語3)
	(標準表現 会議)
	(読み かいぎ)
	(品詞 名詞)
.....	

図3. オブジェクト指向表現記述の例

れらをクラスで表し、superclassスロットにより上位概念を示すことにより、品詞の階層構造を簡単に表現できる。has-slotsによりそのクラスに含まれるインスタンスがどのようなスロット(属性、情報)を持っているのか、そのスロットの値はどのようなものでなければならないかを示す。

インスタンスの全体/部分関係はpart-ofとcomposed-ofにより示す。part-ofはそれを含む言語データの構成要素を表し、composed-ofはそれの部分要素を表す。文1の場合、この文は発話1の一要素になっており、また、文節1、文節2の二つの文節から構成されている。言語データの構成要素の順序関係は、priv-isとnext-isで表現する。これは前後の構成要素の順序関係を指定した検索やある構成要素の検索の後、前後の要素を調べに行くことが多いため付加した機能である。文1の場合、priv-isがnilなので、これが発話1の最初の文であり、次の文は文2である。

4.2 オブジェクト指向表現の利点

(1) いろいろな側面から見たデータを組み合わせた検索

例えば、「教えて欲しい」の言い方にはどのようなパターンがあり、また目的語にはどのような単語がくるか。さらに英語での表現の違いを検索したいとする。そのような要求を満足させる検索の一例を図4に示す。

```
(DEFSEARCH 検索1 "教えて欲しい"
(OBJECT-PATTERN (VAR 日文1)
(INSTANCE-OF 日文)
(COMPOSED-OF * 日文節1 *)
(スピーチアクト 依頼)
(対応関係 (VAR 文対応関係1)))
(OBJECT-PATTERN (VAR 日文節1)
(INSTANCE-OF 日文節)
(COMPOSED-OF * 日単語1 *))
(OBJECT-PATTERN (VAR 日単語1)
(INSTANCE-OF 日単語)
(標準表現 教える)))
(OBJECT-PATTERN (VAR 文対応関係1)
(INSTANCE-OF 文対応関係)
(日本語 (VAR 日文1))
(英語 (VAR 英文1)))
(RETURN-FORM (日文 英文) 0 0))
```

図4. 各種情報を組み合わせた検索の例
("教えて欲しい"の検索)

この例では、全体/部分関係と単語や文に関する情報と文の日英対応関係を使って検索を行っている。「教える」の目的語に関しては陽に指定していないが、検索結果に日本語を指定することによって、目的語が何であったかをユーザは知ることができる。もちろん、係り受け関係の情報を利用して、目的語となるものの条件を指定して検索するようなことも簡単にできる。

検索記述では、日文1、日文節1、日単語1、文対応関係1のようなオブジェクト変数を使って、条件を直接的に記述できる。このような検索条件の記述方法は非常に柔軟性のある記述方式なので、言語データのような複雑な構造を持ったデータの検索には有効である。検索のフレームワークについては、5.1(1)でもう少し詳しく説明する。

(2) 検索結果の環境を簡単に調べられる。

言語データを言語現象の分析に利用する時、単に検索ができるだけでなく、言語現象が生じた周りの状況を容易に調べられなくてはならない。このような機能は、言語データをオブジェクト指向で直接的に表現することにより周りの状況を調べる機能を容易に実現することができる。

例えば、図4のような検索の場合、検索結果は図5のようになる。

(1) 名前を教えてくださいか。	Give me your name, please.
(2) 取引銀行と口座番号を教えてくださいか。	Please tell me your bank and account number.
(3) 参加の手続きについてお教え願えますでしょうか。	How do I register?
(4) 申し込みの方法を教えてください。	Could you give me how I can apply?

図5. “教えて欲しい”の検索結果

教えるの目的語には、「名前」などの普通名詞の他に、「参加の手続き」、「申し込みの方法」などのような{サ変名詞+格助詞の+方法or手続き}などもある。前者の場合は、「tell me, give me+(目的語)」と訳せば良いが、後者の場合は、「tell me + how …」や「How do I…?」のような疑問文に訳す必要があることなどが分析できる。係り受け関係や文の前後関係の情報を調べることにより、より詳細な分析が容易に行える。

本システムで用意した具体的な周りの状況を調べる機能については、5.2で詳しく述べる。

(3) is-a階層構造を利用した検索

オブジェクト指向表現によるis-a階層構造を利用して、検索に実際の言語データに付けられているものの上位概念を指定することができる。図6の「動詞+*+助詞か」の検索を例にとると、“か”の実際の言語データには、品詞として終助詞もしくは副助詞が入っているが、終助詞や副助詞の上位概念である助詞を品詞に指定して検索することができる。

(4) 新たなクラスの定義が容易

クラスのスロットの値を特定化したり、新たな属性(スロット)を追加したような下位クラスを簡単に定義できる。それを利用して検索なども行える。

5.1 検索・表示機能

5.1 検索

(1) 検索のためのフレームワーク

マイクロ分析においては上述のように、複数の条件を組み合わせた検索など、多種多様な検索を可能としなければならない。このため以下に述べるような検索要求を記述するための一般的なフレームワークを用意している。このフレームワークを用いた検索パターンの例を図6に示す。検索パターンの記述は、次の5つの部分からなる。

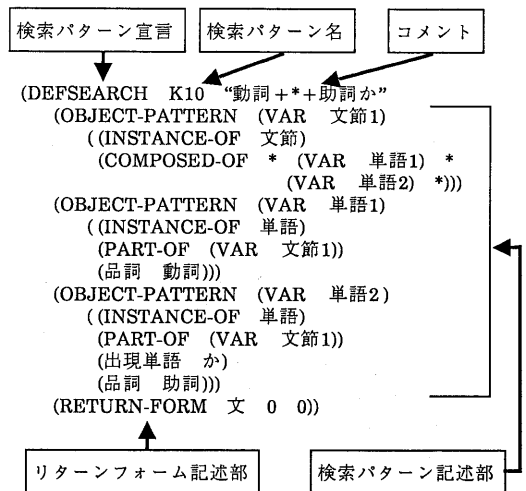


図6. 検索パターンの例
「動詞+*+助詞か」の検索パターン

- ① 検索パターン宣言(defsearch)
- ② 検索パターン名(任意、例ではK10)
- ③ コメント(“”で囲まれた部分)
- ④ 検索パターン記述部(object-pattern)
- ⑤ リターンフォーム記述部(return-form)

④は実質的な検索内容を記述する部分である。object-patternでは、個々の言語データ構成要素オブジェクトについての条件を記述する。⑤で検索結果の表示単位などを指定する。この例の場合には、条件を満足した文節を含む文を表示する。まわりの環境を見られるように、結果の表示単位で条件を満足した部分の前後を指定することができる。(RETURN-FORM 文 1 1)と指定すれば、条件を満足した文節を含む文とその前後1文の計3文を表示する。検索のためのフレームワークをBNF的な記述で表現すると以下ようになる。

```

(defsearch <<search-name>> <<comment>>
  <<matching object pattern list>>
  <<return form specification>>)

<<search-name>> ::= <atom>
<<comment>> ::= <string>
<<matching object pattern list>>
 ::= <<matching object pattern>> |
   <<matching object pattern>>
   <<matching object pattern list>>
<<matching object pattern>>
 ::= (object_pattern <<object-name>>
   <<slot condition>>*)
<<object-name>> ::= <<variable>>
<<slot condition>>
 ::= (<<slot-name>>
   <<value_restriction>>)

```

```

<<slot-name>> ::= <atom>
<<value_restriction>>
  ::= <<value>> |(not <<value>>)|
      (or <<value>> >$)
<<value>>
  ::= <atom> | <<variable>> | ? | * | %group
<<variable>> ::= (var <atom>)
<<return form specification>>
  ::= (return form <<level>>
      [- <<region>>] [+ <<region>>])
<<level>> ::= 会話 | 発話 | 文 | 文節 | 単語 | 文字 | 関係 |
              %count | (<<level>> <<level>> >$)
<<region>> ::= <number>
  ?は任意の一個の値とマッチング
  *は任意個の値とマッチング
  %は0回以上の繰り返し
  $は1回以上の繰り返し

```

検索は検索パターン名を指定することにより実行できる。検索パターンもオブジェクトになっており、実行時に、スロットの条件を追加したり、修正したりして検索できる。

(2) 検索パターンエディタ

検索のためのフレームワークはLISPのS式であり、必ずしもユーザにとって記述し易いものであるとはいえない。そこで、ユーザが検索パターンを会話的に容易に定義できるツールを用意した。これが、検索パターンエディタである。ユーザが言語データの全体/部分関係、順序関係、スロット条件等を対話的に記入すると、図6のような検索パターンの形式に変換する。

(3) 検索パターンの操作機能

検索パターンエディタで作成した検索パターンは、セーブしたり、ロードしたりして利用することができる(もちろん、検索パターンエディタを使わないで作成したフレームワークを用いた検索パターンも利用できる)。

本システムでは、この機能を利用して、初心者ように使用頻度の高いと思われる検索パターンをいくつか用意している。これにより、初心者から熟練者までいろいろなレベルのユーザがレベルに合わせて言語データを利用できる。

(4) 検索範囲の指定

会話単位で検索範囲の指定できる機能である。会話IDを指定するか、会話の属性(メディア、内容、領域、言語パターン)を指定することによって範囲の指定が行える。例えば、会話のメディアとして電話を指定すれば、電話会話の言語データを範囲として検索できる。会話の属性としてどのようなものかを以下に示す。

メディア: 電話、キーボード、手紙等

内容: 問合せ、依頼、感情表現等

領域: 国際会議、旅行案内等

言語パターン: 日本語-英語、
日本語-日本語、
英語-英語等

本機能については現在一部を除いてインプリメントを終了している。

5.2 表示

本システムでは、以下に示すような言語データの表示機能を持たせることにした。これらの機能はオブジェクト内の値を参照することにより、比較的容易に実現できる。また、これらの表示はマルチウィンドやアイコン化などにより、ユーザに分かり易く、使い易いものになっている。

(1) 検索結果の表示

図9左中央に検索結果の表示例を示す。反転表示によりマッチングに成功した部分を強調したユーザに分かりやすい形式で表示する。

(2) 検索マッチング表示

各々の検索結果がどのようにマッチングしたかを表示する。オブジェクトのスロットの値はマウスセンシティブになっており、マウスクリックにより内容をコマンドウィンドに表示するなどが簡単に行える。

(3) 言語データの全体部分関係の表示

検索結果の全体部分関係をツリー形式で表示する機能である。図9右下に例を示す。

(4) オブジェクトの内容の表示

検索結果の全体部分関係のツリー表示の各ノードは、マウスクリックすることにより、その内容を表示することができる。図9右下の単語0408や単語0411の内容表示がその例である。

(5) 単語情報の詳細表示

検索結果の内容表示の特殊なものである。単語レベルの情報は頻繁に参照されることが予想されるので、便利のために設けた。図9右中央に例を示す。

(6) 対訳表示

検索結果に対応する英語を対応させて表示する機能である。現在は文レベルの対応表示のみをインプリメントしている。図9左下に例を示す。

6. ユーザ定義

ユーザ定義機能とは、ユーザがシステムで用

意した品詞体系や意味体系など言語データの体系に満足できない場合に、体系の修正や追加をユーザごとに行え、そのユーザ定義に基づいて検索などが行えるようにする機能である。これにより、ある特定の言語理論に基づいて言語データを利用したいというようなユーザも、言語データを利用できる。

例えば、副助詞と係助詞を新しく副・係助詞として検索したい場合、品詞のクラス定義をユーザが本機能を用いて行えばよい。クラスの名称の変更、上位クラスの作成、クラスの削除などが簡単に行える。クラスの分割は、分割のための条件指定機能など考察すべき点が残っているので未だ実現は行っていない。

7. 統計機能

マクロ分析用の機能である。本機能の処理自体はホストコンピュータのRDB上で行われるが、検索インタフェースは原則的にマイクロ分析の検索機能を拡張したものとなっており、ユーザの使い易さを考慮している。図7に例を示す。%groupによりキーとして品詞を指定し、リターンフォーム記述部で%countを指定することにより、頻度のみを出力することを示す。

```
(DEFSEARCH COUNT1 "品詞ごとの頻度"
 (OBJECT-PATTERN (VAR 単語1)
 (INSTANCE-OF 単語
 (品詞 %group)))
 (RETURN-FORM %count))
```

図7. 統計機能の例

図8に品詞による統計の表示の例を示す。

品詞による統計	
名詞	1594
固有名詞	109
普通名詞	1265
サ変名詞	86
⋮	
数詞	122
⋮	
合計	3241

図8. 統計表示機能の例

8. データ格納

収集システムを使って集めた言語データはホストコンピュータのRDBに格納する。全体部分のような関係を効率良く扱えるように、値として多値を許すように拡張する。会話、発話、文、文節、単語、意味関係、構文関係、各レベルの対応関係を関係表(Relation)とし、実際のデータを組(Tuple)とし、発話者、品詞、意味関係名などを属性(Attribute)で表す。RDB表現とオブジェクト指向表現との対応を考えると、関係表が最下位のクラス、組がインスタンス、属性がインスタンス変数に対応している。言語データの修正、機密保持、一貫性、統計処理については既存のRDBに依存することによりシステム開発の手間を大幅に省くことができる。

9. おわりに

言語データベース統合管理システムを利用することにより、多様な言語情報を組み合わせることによって、いろいろな側面から、言語現象のマイクロ分析およびマクロ分析が容易に行える。また、言語データの維持・修正などの管理が容易になる。この言語データベース統合管理システムは、データ格納にVAX8810/ULTRIX、ユーザインタフェースにシンボリックス・リスプマシンを使用する。現在、ユーザインタフェース部のプロトタイプが完成し、RDB上の言語データをワークステーションに転送する機能の実現が終了したところである。RDB上の言語データの格納機能については、現在開発を進めている。

将来的には、当研究所の音声処理研究室で開発を進めている音声データベース^[4]と結び付けることによって、自然言語研究者と音声研究者に対して、より強力な研究環境を提供できる。

<謝辞>

本研究の機会を与えて下さったATR自動翻訳電話研究所樽松明社長に感謝する。また、熱心に議論して下さい下さったデータ処理研究室および応用技術(株)諸氏に感謝する。

<参考文献>

- [1] 有田・小暮・野垣内・前田・飯田、メディアに依存する会話の様式-電話会話とキーボード会話の比較-、情処学会NL研、NL-61-5、1987
- [2] 井ノ上・小倉・森元、係り受け意味関係の問題点とその考察、信学会NLC87-25、1988
- [3] 井ノ上・小倉・森元、言語データベース用単語間の関係データ、情処学会第37回全国大会5B-7、1988

- [4] 内田、電子化辞書について、文字言語・音声言語の知能的処理 第152委員会 第三回研究会資料 3-2、1987
- [5] 小倉、言語対比データの構築について、信学会創立70周年記念総合全国大会1642、1987
- [6] 小倉・篠崎・森元、言語データベース収集支援システム、情処学会第36回全国大会4U-4、1988
- [7] 小倉・橋本・森元、言語データベース統合管理システム、情報処理学会第37回全国大会5B-6、1988
- [8] 北・森元、テキストデータベースからの慣用表現の自動抽出、情処学会第37回全国大会、7B-6、1988
- [9] 北・川端・斎藤、HMM音韻認識と予測LRパーザを用いた文節認識、信学会 音声研究会、SP-88-88、1988
- [10] 鹿野、Trigram Modelによる単語音声認識結果の改善、信学会 音声研究会、SP-87-23、1987
- [11] 篠崎・小倉・森元、言語データベースの品質管理、情処学会第36回全国大会4U-3、1988
- [12] 篠崎・小倉・森元、言語データベース作成のためのシミュレーション会話;会話データのリアリティ改善、情処学会第37回全国大会5B-8、1988
- [13] 島津・内藤、日本語文意味構造の分類・名詞句構造を中心に、情処学会NL研、NL-47-5、1985
- [14] 武田・勾坂・片桐・桑原、音韻ラベルを持つ日本語音声データベースの構築、信学会 音声研究会、SP-87-19、1987
- [15] 田中、オブジェクト指向データベース・システム-その背景と概念、bit Vol.20 No.6 p.687-693
- [16] 橋本・小倉・森元、言語データベース統合管理システムのマンマシンインタフェース、情処学会第37回全国大会、2C-4、1988
- [17] 森元・小倉・飯田、自動翻訳電話研究用言語データベースの収集について、情処学会第36回全国大会4U-5、1988
- [18] 吉本、日本語品詞の分類、ATRテクニカル・レポート TR-I-0008、1987
- [19] 国立国語研究所、CL研究第1号、1987
- [20] J. Aarts and W. Meijs, Corpus Linguistics II: new studies in the analysis and exploitation of computer corpora, Rodopi Amsterdam, 1986
- [21] Brown University, Brown Corpus, Technical Report, Brown University, 1967
- [22] E. F. Codd, "A Relational Model of Data for Large Shared Data Banks", Comm. of ACM Vol. 13, No. 13, 1970
- [23] A. Ege, C. A. Ellis, "Design and Implementation of GORDION, an Object Base Management System", Proceedings Third International Conference on DATA ENGINEERING, 1987
- [24] L. F. Farina, "LDMS: A Linguistic Data Management System", Computers and the Humanities 17, 1983
- [25] University of Nijmegen, TOSCA The Nijmegen Research Group for Corpus Linguistics, 1987

The screenshot displays a complex graphical user interface for a Japanese language processing system. It features several overlapping windows and panels:

- Top Panel:** A menu with options like 'ユーザー定義' (User Definition), '検索範囲指定' (Search Range Specification), and '検索パターン定義' (Search Pattern Definition).
- Search Results Window:** Shows a search pattern 'K10-EHEC-1' with 4 items. The first item is 'はい、登録用紙は既に持っていますか?' (Do you already have the registration form?).
- Linguistic Analysis Window:** A tree diagram showing the analysis of the sentence 'はい、登録用紙は既に持っていますか?' into nodes like '文節0401' through '文節0404'. Below the tree, a table lists morphological and syntactic information for each node.
- Table of Morphological/Syntactic Information:**

正規表現	読み	正規表現	品詞	活用型	活用形
はい	はい	はい	感動詞		記号
登録用紙	どうろく	登録用紙	名詞		名詞
は	は	は	係助詞		
既に	すでに	既に	副詞		
お	お	お	接頭語		
持ち	もち	持ち	本動詞	連用	五段
でしょ	でしょ	でしょ	助動詞		未然
- Bottom Panel:** A translation window showing the English translation of the search results: '1. Do you already have the registration form? 2. Well, in that case, we will send you one. Could you give me your address, please? 3. How much will it cost ne? 4. Being an international conference, simultaneous interpretation will also be provided. There is no need to worry since it will be possible to listen to the Japanese version.'

図9. システムの画面イメージ