自然言語処理 38－3
（1983. 7. 15）

日本語形態素解析の基本設計

坂本義行
（電子技術総合研究所）

概要

本研究は、日英機械翻訳システム[1]における科学技術文献を対象とした日本語文の形態素解析を行い、分析結果として日本語構文解析等に必要な情報を辞書システム[2]より検索し、出力することを目的としている。解析は、辞書駆動型による文字列間の接続関係を文の左端からチェックし、文を単位として、可能な解を全て出力するMultiple-path方式をとっている。本稿ではサンプル文としてJICST文献抄録文について実験を行っている。そこに出現した文の表現形式を細部にわたり分析し実用的な解析システムを構築するプロセスの開発をめざしている。

1. はじめに

日本語の形態素解析はその目的、用途によってその方法も自から異なったものとなるであろう。本研究では、日英機械翻訳システムとして必要な処理形態、言語的な興味といった目的を捨て実用という面から文献に現われる表記を可能な限り辞書に登録することにより解析の可能性を確めるとともに、処理ソフトウェアと辞書とを明確に区別し処理手順の簡略化、拡張の容易さをめざした。

日本語構文解析[3][4]に必要とされる情報は、すべてこの形態素解析レベルで辞書を引くことにより付与する方式をとっている。

処理手順は、大きく次の3つからなる。

1) 入力テキストの定形化処理
2) 形態素の分割
3) 形態素解析の出力処理

2. 入力テキストの定形化処理

翻訳の入力文であるJICSTの抄録文から日本語形態素解析の入力文となる標準形式をつくりだすために以下のようなフォーマッティングを行なう。

1) 抄録番号、記事番号といった書誌的事項の区分
2) 標題、本文、式、図などの区分
3) 印刷書式等の内容に直接関係しない部分の区分

2.1 システム中での処理の流れ（図1）

2.2 フォーマット情報

フォーマット情報の取扱いを以下の4種に分ける。

A1. 総合システムが管理するもの（翻訳フラグをNILとする。）
例. @PP@, @AR@

A2. 総合システム[5]がすべて管理するもの。フォーマット情報のうち首尾照応するもので、中間に翻訳対象を含むもの。

本研究は国の科学技術振興調整費による「日英技術文献の速報システムに関する研究」の1部として行ったものである。

研究協力者として、木村睦子（計量計画研究所）、山本稔（東洋情報システム）の諸氏、ならびに本研究の委託を受けて、その遂行のために、自然言語処理の専門家から成る言語処理システム作業分科会、さらに処理システム作業グループ（京大）、辞書作業グループ（JICST）を組織し、その審議、指導のもとに研究をすすめている。

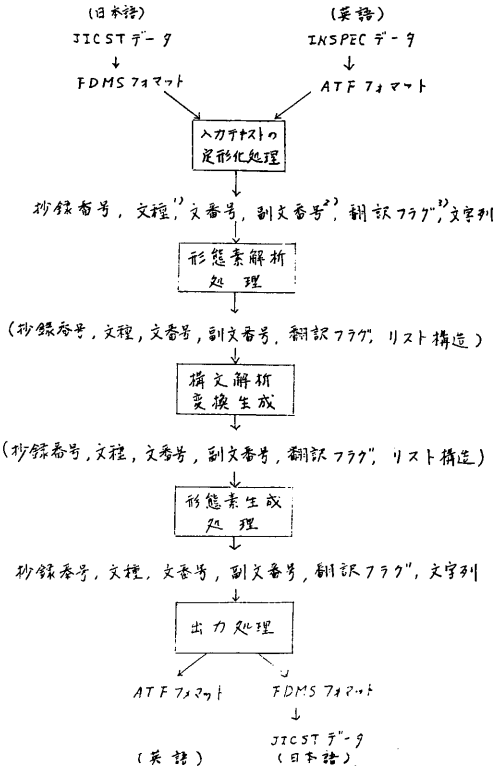


図1　翻訳システムのデータ形式

1) 文字列(文)の種類。例：標題，本文，レイアウト。

2) 同一文番号内での枝番号。一文字中にフォマット情報が付いて複数のレコードに分かれる場合

3) 翻訳を必要とするかどうかのフラグ
T：翻訳が必要（通常の文）
NIL：翻訳が不要（主としてページ様式に関するフォーマット情報など）

フォーマット情報と翻訳対象文を分けて別レコードとする。

例：　@HS=H12.3@抄録の機械翻訳システム@HE@

B.：形態素解析の時点で単語の性質に反映するもの（字種情報）

例：　@UL@システム@UE@
ただし、複数の単語にまたがって指定されている場合は「C」で扱う。

例：　@UL@ある種の単語@UE@
また、単語の一部分だけが指定されている場合は、その単語全体に対して指定したものとみなす。

C：句読点と同様に一語扱いとして、文法処理に引き渡すもの。

3. 形態素の分割

漢字かな混じりの日本文を文節、語あるいはさらに小さな単位に分割する処理方法として、その手順をプログラム化して行う方法と辞書の検索を中心としたプログラムにより行う、大きく分けて2種類の方法がある。前者は「小まわりのきく」方法であり、後者は「辞書」に依存するためプログラムは単純化されるといった特徴を有している。本研究では翻訳に適応する、すなわち英語に訳すに最適であり、改良、拡張が容易な後者の方法である辞書駆動型の解析手順を選んだ。

```
JEF FDMS DOCEDT V01/L07 モニタリスト          DATE:83.07.01    TIME:19.27.03    PAGE:  1
テキスト名 :                  作成日付 83.07.01

 項目番号      参照番号
 0. 0. 0    1.  1.  0   @PG=50,66,1,5,9,4@AR=5,58@HS=H12,3@
 0. 0. 0    1.  3.  0   抄録の機械翻訳システム@HE@このシステムは，ＪＩＣＳＴ
 0. 0. 0    1.  4.  0   の抄録を計算機により翻訳するものである．本システムの入力は，たとえば，富士通のＦＤＭＳのようなフォーマット機能を
 0. 0. 0    1.  5.  0   持ったシステムに対する入力であってもよい．@NL@このシステ
 0. 0. 0    1.  6.  0   ムの特徴は，次のとおりである．

 項目番号      参照番号
 0. 0. 0    1.  1.  0   E82060001, レイアウト, 0, 10, NIL, '@PG=50,66,1,5,9,4,@'
 0. 0. 0    1.  2.  0   E82060001, レイアウト, 0, 20, NIL, '@AR=5,58'
 0. 0. 0    1.  3.  0   E82060001, レイアウト, 0, 30, NIL, '@HS=H12,3@'
 0. 0. 0    1.  4.  0   E82060001, 標題, 0, 40, T, '抄録の機械翻訳システム'
 0. 0. 0    1.  5.  0   E82060001, レイアウト, 0, 50, NIL, '@HE@'
 0. 0. 0    1.  6.  0   E82060001, 本文, 1, 10, T, 'このシステムは，ＪＩＣＳＴの抄録を計算機により翻訳するものである．'
 0. 0. 0    1.  7.  0   E82060001, 本文, 2, 10, T, '本システムの入力は，たとえば，富士通のＦＤＭＳのようなフォーマット機能をもったシステムに対する入力であってもよい．'
 0. 0. 0    1.  8.  0   E82060001, レイアウト, 2, 20, NIL, '@NL@'
 0. 0. 0    1.  9.  0   E82060001, 本文, 3, 10, T, 'このシステムの特徴は，次のとおりである．'
```

図2．定形化処理の実例

3.1 分割の特徴

1) 構成要素 — プログラム、辞書、接続表等テーブル類

2) 処理方式 — Multiple-path 方式として、1)の言語情報により形態的に接続可能なあらゆる単位語列を作成する。

3) 処理単位 — 設定した「文」を単位にその中の文字列がすべて接続可能な場合のみを出力する。

3.2 辞書引きの方法

辞書には自立語（複合語を含む）、付属語、接辞等を見出しとし、形態素解析に必要な情報として、図1に示すような見出し情報、形態素情報が与えられる。

見出しは、用言については、終止形で表わされ語尾字数によって語幹を判定する。異形語はすべて見出しと同等に扱う。

分割の可否を判定するには、後接情報（後続語との接続関係を示す表3-1）と前接情報（先行語との接続関係を示す表3-2）との結合を接続表（表3-3）により判別する。すなわち、先行語の後接情報と後続語の前接情報とがうまく整合するか否かを、ブランク、1、2等の記号で示された表によって判別する。ブランク：接続不可

1：接続可だがそこで切ってはいけない。

2：接続が可で、かつ単語の切れ目となる。

表1 形態素辞書フォーマット

見出し情報	見出し語	
		語尾字数
		漢字部
		語基
		読み
	異形語	
	派生語	
	関連語	
形態素情報	形態	名 副名 動 形 形動 副 連体
	品詞	接 助 助動 接頭 接尾
	動詞	五 上一 下一 サ変 ザ変 カ変
	活用型	
	活用行	ア カ ガ サ ザ タ ダ
		ナ バ マ ラ ワ
	形動	ダナ ダナノ
	活用型	
	助動	形 形動 動 特
	活用型	
	助詞	格 接 副 並 終 準
	細分類	
	接尾	体 動 形 形動
	活用型	
	前接情報	後接情報

表3-1 後接情報

コード	内容	コード	内容	コード	内容
1	名詞	26	格助詞 まで、並助詞 と、か	50	五段動詞語幹 マ行
2	副詞的名詞（含数詞）	27	接続助詞 から	51	〃 〃 ラ行
3	副詞	28	接頭辞	52	〃 〃 ワ行
4	連体詞	29	五段動詞未然形（〜ない）	53	上一段動詞語幹 ア行
5	接続詞	30	〃 〃（〜う）	54	〃 〃 カ行
6	数字	31	〃 イ音便（清音）促音便	55	〃 〃 ガ行
7	サ五、上一、下一、カ変、サ変連用	32	〃 〃（濁音）撥音便	56	〃 〃 サ行
8	五段動詞（サ行以外）連用	33	動詞・形容詞仮定形	57	〃 〃 タ行
9	動詞終止形	34	上一、下一、カ変未然形	58	〃 〃 バ行
10	動詞・形容詞連体形	35	サ変未然形 さ	59	〃 〃 マ行
11	動詞命令形	36	〃 〃 せ	60	〃 〃 ラ行
12	上一、下一語幹未語尾	37	〃 〃 し	61	下一段動詞語幹 ア行
13	サ変語幹	38	形・形動未然形 く、で	62	〃 〃 カ行
14	〃 終止形 す	39	〃 連用形 かっ、だっ	63	〃 〃 ガ行
15	形容詞連用形 く	40	ザ変語幹	64	〃 〃 サ行
16	〃 終止形	41	形容詞語幹	65	〃 〃 ザ行
17	形容動詞連用形 で	42	形容動詞（ダナ型）語幹	66	〃 〃 タ行
18	〃 終止形	43	〃（ダノ型）〃	67	〃 〃 ダ行
19	〃 連体形			68	〃 〃 ナ行
20	て、ないで	44	五段動詞語幹 カ行	69	〃 〃 バ行
21	べき	45	〃 〃 ガ行	70	〃 〃 マ行
22	う、よう 終止形	46	〃 〃 サ行	71	〃 〃 ラ行
23	格助詞 と	47	〃 〃 タ行		
24	〃 〃 に	48	〃 〃 ナ行	0	終端状態
25	〃 と、で、から、より、へ	49	〃 〃 バ行		
		50	〃 〃 マ行		

表3-2 前接情報

コード	内容	コード	内容	コード	内容
1	名詞	26	のでの、のに	51	五段動詞語尾 ラ行
2	名詞以外の自立語、接頭辞	27	格助詞 が、と	52	〃 〃 ワ行
3	形式名詞	28	〃 に	53	上一段動詞語尾 ア行
4	補助動詞（てに接続形）	29	〃 から	54	〃 〃 カ行
5	助数詞	30	〃 まで、より	55	〃 〃 ガ行
6	助動詞 ない	31	〃 の	56	〃 〃 ザ行
7	〃 ぬ	32	格助 と、へ、で、並助 や	57	〃 〃 タ行
8	〃 れる、せる	33	と（引用）	58	〃 〃 バ行
9	〃 られる、させる	34	助詞 か	59	〃 〃 マ行
10	〃 しめる	35	係助詞 は、も、さえ、こそ	60	〃 〃 ラ行
11	〃 う	36	副助詞 など、ばかり、たち、くらい、のみ	61	下一段動詞語尾 ア行
12	〃 よう	37	接尾辞 さ	62	〃 〃 カ行
13	〃 まい、たい、つつ、ながら	38	〃（体言接続）	63	〃 〃 ガ行
14	〃 そうだ（推定）	39	サ変語幹接続助動詞	64	〃 〃 サ行
15	〃 た	40	サ変語尾	65	〃 〃 ザ行
16	〃 らしい、だろう、なら	41	ザ変語尾	66	〃 〃 タ行
17	〃 だ、であろ	42	形容詞語尾	67	〃 〃 ダ行
18	〃 べき	43	形動語尾 な	68	〃 〃 ナ行
19	〃 まい	44	五段動詞語尾 カ行	69	〃 〃 バ行
20	接続助詞 が、から、けれど、し	45	〃 〃 ガ行	70	〃 〃 マ行
21	助動詞 そうだ（伝聞）	46	〃 〃 サ行	71	〃 〃 ラ行
22	接続助詞 て	47	〃 〃 タ行		
23	〃 ば	48	〃 〃 ナ行	0	終端状態（下段点）
24	〃 と	49	〃 〃 バ行		
25	だ、で（「た」、「て」のにごり）	50	〃 〃 マ行		

活用語尾処理は、活用語尾表を用い、その前接情報により判定する。なお語幹を持たない語は、語尾字数により判定し活用処理を行う。(辞書台帳には特殊活用型として記載されており、辞書生成時に見出し等の情報が独立して記憶される。)

3.3 分割の手順

3.3.1 辞書による単語抽出

処理は辞書引きから始まる。文字列を左端から一字づつ辞書を引き、1回の操作で最長一致する文字列まで、すべてをスタックする。スタック中の最長一致の文字列に対して、そのつぎの文字列について辞書を引き接続チェックを行い可能な文字列をスタックする。活用する語には、活用語尾処理を施す。以上の操作をくり返す。接続不可となったら、以下の処理を実行する。

3.3.2 特殊記号処理

①特殊記号の判別方法の特徴

・文脈(前後の文字の字種や、文字の並びのパターン)によって判別される記号が多い。

・確定的な判別法はほとんどなく、実際の文例をもとにした経験的な手法が多い。

②特殊記号の取り扱いの特徴

・前の文字と同一の字種として扱う。

・後の文字と同一の字種として扱う。

・区切り符号、特殊符号や「つなぎ」の符号等のように独立の符号として取り扱う。

また、翻訳対象の広がりと共に、判別方法の変更の追加等が頻繁に行われ、その取り扱い方法も拡張されることが予想される。

以上の点から、ソフトウェアは機能の追加、変更に柔軟に対応できるようなテーブル駆動型の構造をとった。

③特殊記号処理の手順

特殊記号処理は、辞書で引けなくなった文字列に適用され、特殊記号処理テーブルにより判別を行う。特殊記号テーブルは、特殊記号、処理適用条件と実行処理からなる。

3.3.3 未知語の処理

辞書になく、特殊記号処理によっても単語抽出ができなかった場合で、かつ1つ前のスタックにも、文完成スタックにも隣がないとき、字種によるスキップ処理を行い、未知語としてスタックし、単語ポインタを変更する。

スキップ位置の決定方法として、未知語の先頭文字が、

漢字、
カタカナ、
英文、
ギリシャ文字

指定した特殊記号、であった場合にその字種が連続する最後の文字の次の文字までスキップする。

3.3.4 バックトラック探索

1) 1文完成内でのバックトラック

単語抽出に失敗したとき、それ以前にある単語スタックを探し、単語リストを書き換え、単語ポインタを変更し、3.3.1からの処理を繰り返す。

2) 文完成後のバックトラック探索

文末までの処理が完成した後、マルチパスを行うため、残余の単語スタックを探索し、スタックが空になるまで処理を続行する。

3.4 形態素解析の出力処理

3.4.1 辞書項目の付与

分割処理が終了し、文が完成した場合に辞書項目の付与を行う。

1) 見出し情報
見出しの形(活用語は、その見出しの形)、派生語、関連語等表に記述されている情報

2) 形態素情報
活用語には、その活用形を示す

3) 構文・意味情報

4) 格支配情報

5) 共起情報

6) 備考

以上のすべての情報をLISP-S式にて図3のように付与する。

3.4.2 日本語構文解析への出力

1) 複数の文の出力

分割が異なる解が得られたとき、文を単位として複数の文を出力する。

2) 複数の形態素情報をもつ(PARA)

単位として同一の分割が行われたが複数の形態素情報が得られ、形態素解析では判別不能の場合

例

「で」…並列要素(PARA) ― 格助詞
　　　　　　　　　　　― 形容動詞の語尾形

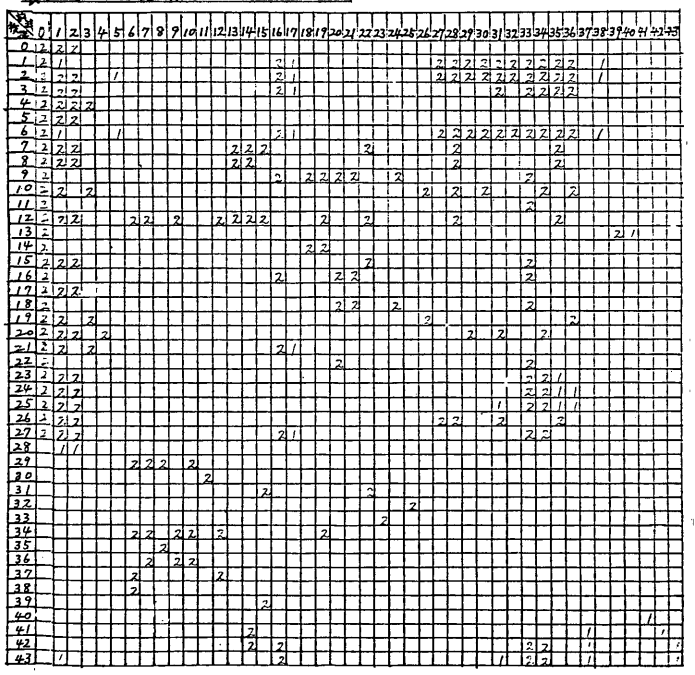
3) 連語

連接した文字列間に形態素的な結合関係がある。

例

「ヲ1回」…形態要素(MORCOMP) ― ヲ(接頭語)
　　　　　　　　　　　― 1(数)
　　　　　　　　　　　― 回(助数詞)

表3-3 形態素分析特統表(1)



注 ※ヘクト；対角線上が1

4) 複合語

辞書に見出しとして登録されているが語基情報をもっている

例

「機械翻訳」…語基要素(DICCOMP) ― 機械
　　　　　　　　　　　― 翻訳

4. おわりに

本稿では、5ク年度行った基本設計に基き、本年度開発に着手した段階での報告であって、完成されたものではない。今後、実験をすすめて、分割レベルの問題、構文解析へのインターフェイス等、多くの改良、拡張を必要とするとおもわれる。マルチパス方式の効率性の問題がある。それらについては、漸次報告する予定である。

参考文献

1) 長尾真、「科学技術庁機械翻訳プロジェクトの概要」。

2) 坂本義行、「格構造を中心とした用言と付属語」。

3) 辻井潤一、「日本語構文解析」。

4) 中村順一、「文法記述用ソフトウェアGRADE」。

5) 矢田光治、長尾真、「機械翻訳総合システムの基本設計」

以上のいずれも、本会で同時報告されたものである。

```
*** INPUT DATA ****  =>  (ＥＳ２０６０００３ 本文３１７ システム工学，情報工学，ソフトウエア工学など新しい技術の急
速な成長により，ＩＥＥおよび電気工学教育の全パターンでの重点が移動した。)

("ＥＳ２０６０００３" 本文３１７
 ($形態素分割
    ($辞書情報
       (($見出し情報 ($見出し語 "システム工学"))
        ($形態素情報 ($形態品詞 名))
        (レコード番号 51629))
       1)
    ($区切符号 "，" 0)
    ($辞書情報
       (($見出し情報 ($見出し語 "情報工学"))
        ($形態素情報 ($形態品詞 名))
        (レコード番号 52763))
       1)

    ($区切符号 "，" 0)
    ($辞書情報
       (($見出し情報 ($見出し語 "ソフトウエア工学"))
        ($形態素情報 ($形態品詞 名))
        (レコード番号 59726))
       1)
    ($辞書情報
       (($見出し情報 ($見出し語 "など") ($読み "など"))
        ($形態素情報
           ($形態品詞 助)
           ($助詞細分類 副)
           ($前接情報 36)
           ($後接情報 2))
        ($構文－意味情報 ($副助詞 (意味区分 "など") (格助詞との代替性 "１")))
        (レコード番号 394))
       2)
    ($単語情報
       ($辞書情報
          (($見出し情報
              ($見出し語 "新しい")
              ($語尾字数 1)
              ($音アク 1 1)
              ($読み "あたらしい"))
           ($形態素情報 ($形態品詞 形) ($後接情報 41))
           ($構文－意味情報
              ($構文品詞 "形容詞")
              ($品詞細分類 "性質・状態")
              ($格パターン "Ａ１")
              ($程度性 "尺度")
              ($名詞性程度 "－－"))
           ($格支配情報
              (格パターン "Ａ１")
              (表層格 "が" "深層格" "ＳＵＢ" "名詞意味コード" "必須性" "１"))
              (レコード番号 322)))
        ($語尾 い 連体形)
        10)
    ($辞書情報
       (($見出し情報 ($見出し語 "技術"))
        ($形態素情報 ($形態品詞 名))
        (レコード番号 50983))
       1)
    ($辞書情報
       (($見出し情報 ($見出し語 "の") ($読み "の"))
        ($形態素情報
           ($形態品詞 助)
           ($助詞細分類 格)
           ($前接情報 31)
           ($後接情報 5))
        ($構文－意味情報 ($格助詞 (表層格名称 "の") (修飾関係 "連体")))
        (レコード番号 441))
       5)
    ($単語情報
       ($辞書情報
          (($見出し情報
              ($見出し語 "急速だ")
              ($語尾字数 1)
              ($音アク語 1 2)
              ($読み "きゅうそくだ"))
           ($形態素情報
              ($形態品詞 形動)
              ($形動活用型 ダナノ)
              ($前接情報 21)
              ($後接情報 43))
           ($構文－意味情報
              ($構文品詞 "形容動詞")
              ($品詞細分類 "性質・状態")
              ($格パターン "Ａ１")
              ($程度性 "尺度")
              ($名詞性程度 "－－"))
           ($格支配情報
              (格パターン "Ａ１")
              (表層格 "が" "深層格" "ＳＵＢ" "名詞意味コード" "必須性" "１")
              (表層格 "より" "深層格" "ＣＯＲ" "名詞意味コード" "必須性" "０"))
              (レコード番号 203)))
        ($語尾 な 連体形)
        19)
    ($辞書情報
       (($見出し情報 ($見出し語 "成長"))
        ($形態素情報 ($形態品詞 名))
        (レコード番号 53183))
       1)
    ($辞書情報
       (($見出し情報 ($見出し語 "により") ($語基 １１２３) ($読み "により"))
        ($形態素情報
           ($形態品詞 助)
           ($助詞細分類 格)
           ($前接情報 27)
           ($後接情報 5))
        ($構文－意味情報
           ($格助詞
              (表層格名称 "によって")
              (修飾関係 "連用")
              (深層格 "原因－理由" "手段－道具" "条件")))
        (レコード番号 437))
       5)
    ($区切符号 "，" 0)
    ($未知語 "ＩＥＥ")
    ($未知語 "および")
    ($辞書情報
       (($見出し情報 ($見出し語 "電気工学教育"))
        ($形態素情報 ($形態品詞 名))
        (レコード番号 54378))
       1)
    ($辞書情報
       (($見出し情報 ($見出し語 "の") ($読み "の"))
        ($形態素情報
           ($形態品詞 助)
           ($助詞細分類 格)
           ($前接情報 31)
           ($後接情報 5))
        ($構文－意味情報 ($格助詞 (表層格名称 "の") (修飾関係 "連体")))
        (レコード番号 441))
       5)
    ($未知語 "全")
    ($辞書情報
       (($見出し情報 ($見出し語 "パターン"))
        ($形態素情報 ($形態品詞 名))
        (レコード番号 60954))
       1)
    ($辞書情報
       (($見出し情報 ($見出し語 "で") ($読み "で"))
        ($形態素情報
           ($形態品詞 助)
           ($助詞細分類 格)
           ($前接情報 32)
           ($後接情報 25))
        ($構文－意味情報
           ($格助詞
              (表層格名称 "で")
              (修飾関係 "連用")
              (深層格 "主体" "原因－理由" "手段－道具" "材料" "構成要素"
                      "方式" "条件" "観点" "その他")))
        (レコード番号 369))
       25)
    ($辞書情報
       (($見出し情報 ($見出し語 "の") ($読み "の"))
        ($形態素情報
           ($形態品詞 助)
           ($助詞細分類 格)
           ($前接情報 31)
           ($後接情報 5))
        ($構文－意味情報 ($格助詞 (表層格名称 "の") (修飾関係 "連体")))
        (レコード番号 441))
       5)
    ($辞書情報
       (($見出し情報 ($見出し語 "重点"))
        ($形態素情報 ($形態品詞 名))
        (レコード番号 52633))
       1)
    ($辞書情報
       (($見出し情報 ($見出し語 "が") ($読み "が"))
        ($形態素情報
           ($形態品詞 助)
           ($助詞細分類 格)
           ($前接情報 27)
           ($後接情報 5))
        ($構文－意味情報
           ($格助詞 (表層格名称 "が") (修飾関係 "連用") (深層格 "主体" "対象")))
        (レコード番号 351))
       5)
    ($単語情報
       ($辞書情報
          (($見出し情報
              ($見出し語 "移動する")
              ($語尾字数 2)
              ($音アク語 1 2)
              ($読み "いどうする"))
           ($形態素情報
              ($形態品詞 動)
              ($動詞活用型 サ変)
              ($前接情報 2)
              ($後接情報 13))
           ($構文－意味情報
              ($分野コード "電気")
              ($構文品詞 "動詞")
              ($格パターン "Ｖ１")
              ($アスペクト "瞬時")
              ($態 "使役")
              ($意志 "無"))
           ($格支配情報
              (格パターン "Ｖ１")
              (表層格 "が" "深層格" "ＳＵＢ" "名詞意味コード" "必須性" "１")
              (表層格 "に，へ" "深層格" "ＳＴＯ" "名詞意味コード" "必須性" "０")
              (表層格 "から" "深層格" "ＳＦＲ" "名詞意味コード" "必須性" "０"))
           ($構文－意味情報
              ($分野コード "電気")
              ($構文品詞 "動詞")
              ($格パターン "Ｖ２")
              ($アスペクト "瞬時")
              ($態 "受身" "可能")
              ($意志 "有"))
           ($格支配情報
              (格パターン "Ｖ２")
              (表層格 "が" "深層格" "ＳＵＢ" "名詞意味コード" "必須性" "１")
              (表層格 "を" "深層格" "ＯＢＪ" "名詞意味コード" "必須性" "１")
              (表層格 "に，へ" "深層格" "ＳＴＯ" "名詞意味コード" "必須性" "０")
              (表層格 "から" "深層格" "ＳＦＲ" "名詞意味コード" "必須性" "０")
              (表層格 "に" "深層格" "ＴＩＭ" "名詞意味コード" "必須性" "０"))
           (レコード番号 512)))
        ($語尾 し 連用形)
        7)
    ($単語情報
       ($辞書情報
          (($見出し情報 ($見出し語 "た") ($語尾字数 1) ($読み "た"))
           ($形態素情報
              ($形態品詞 助動)
              ($助動活用型 特)
              ($前接情報 15)
              ($後接情報 不定))
           ($特異語用形
              (未然形 (形 "たろ") (後接情報 30)))
              (終止形 (形 "た") (後接情報 16)))
              (連体形 (形 "た") (後接情報 10)))
              (仮定形 (形 "たら") (後接情報 5)) ((形 "たらば") (後接情報 5))))
           ($構文－意味情報
              ($構文品詞 "助動詞")
              ($品詞細分類 "相")
              ($相 "過去・完"))
              (レコード番号 100)))
        ($語尾 "た" 終止形)
        16)
    ($区切符号 "。" 0)))
```

図３　形態素解析の出力例