

日本語→欧語機械翻訳のための日本語の分析*

山 田 小 枝**

序 言

本報告の目的は、日本語を入力言語とし、中間言語を通して欧語へ機械翻訳するために、日本語を分析する方法に関し、次の諸点を明らかにすることである。

1) あらかじめ segmentation の操作をうけていない日本語を機械に入れ、辞書引き、分ち書き、構文分析をもっとも効果的に行なうため、この三つの操作を各々独立した段階とせず、交互に行なう方法。

2) 構文分析は、原理的に phrase-structure language を取り扱うことを目的としているが、phrase-structure grammar の文法規則の機械的運用が能率的に行なわれるように考えられている。文法規則を左から右へ逐次吟味して行くのは坂井方式¹⁾の応用であるが、さらに immediate constituents の検索を容易ならしめるため、Harvard Univ. で predictions をたくわえるために使っているプール²⁾に近いものの中に分析の結果そのものを入れ、問題とされる immediate constituents が何か、自動的にわかるようにする。

3) 自然言語そのものもつ、形態上、構文上の特徴を明らかにし、分析方法に利用する。

4) 書かれた日本語には、同綴異機能、同綴異義だけでなく、しばしば長さの ambiguity も存在するから、唯一の解答を出すだけでは、他の正解を逃す可能性がある。そのため、あらゆる可能性を列挙できるようにする。

5) 以上の方法の裏付けとなる文法規則と単語分数の方法を明らかにする。

以下、1…入力日本語の形式、2…分ち書き、3…構文分析、4…辞書、5…辞書単位の情報の順序で記述する。

1. 入力日本語の形式

現代日本語の graphemic レベル以上の言語的諸現象を機械翻訳の対象とする。

* Analysis of the Input Japanese Text in Mechanical Translation: by Sae Yamada, (Electrotechnical Laboratory, Tokyo)

** 電気試験所電子計算機部

1.1 文字、記号

書かれた日本語を構成する文字記号は 75 種類のひらがな、75 種類のかたかな、約 1,850 種類の漢字、アルファベットの大文字・小文字、アラビア数字、ローマ数字、並びにその他の符号である。

現在は入出力装置の制限により約 50 種類の記号しか使えないが、近い将来、すべての文字を扱う入出力装置を使用することを前提として実験している。たとえば graphemic レベルで、漢字 1 字はそれ以下に細分されない 1 単位であるから、実験段階でもそれを phonemic の単位にまで書き直す必要はない。またテキスト中のかたかなまたはアルファベットの 1 列は何かの単位となっているという情報も生かさなければならぬ。

1.2 日本語の外形

印欧語では文は

単語 # 単語 # …… # (# は空白)

で示されるが、日本語は次のようになる。

文字 φ 文字 φ 文字 …… # (φ は小空白)

日本語でいわゆる単語は、構文上、意味上から単位とされるものである。

1.3 辞書単位

日本語で、分ち書き、辞書要素および構文分析の最小単位として、いわゆる単語をとるのは適当でない。印欧語において、機械辞書作成のため form → lexes に切断し、各 lex を辞書に入れるのと同様の理由で、基本単位は単語以下の小さいものとなる。これを辞書単位とよぶ。すべての辞書単位は同等の手続きで処理され、辞書の見出し、分ち書きの単位、構文分析の最小単位となる。

2. 分ち書き (構文分析の一部としての)

日本語が入力言語の時、先ず文字の列を辞書単位の列に置き換える必要がある。

<文> ⇔ <単位> <単位> …… <・>

これが分ち書きであるが、分ち書きは印欧語で form → lexes に置き換える方法^{4),5)}すなわち segmentation,

<単語> ⇔ <接頭辞> <語幹> <接尾辞> <変化語

尾>

の形に似て来る。

ここで行なう分ち書きは、完結した一段階とならず、文頭から辞書単位を切り取りながら構文分析を行なうが、その「切り取り」のことをさす。

文頭から逐次文法規則を吟味しながら進んで行く方法は simple path または multiple path の predictive analysis の手順に似ている。

語の外限を示す空白は存在しないから、分ち書きにはある程度、形態的、構文的な情報を使う。そこで、同じような情報をまず分ち書きに使い、さらにもう一度構文分析に使うことは無駄である。また、単位の限界は文頭あるいは文末から逐次定めるが、文末から操作を始めることは技術的に困難であるから左→右への方法をとる。

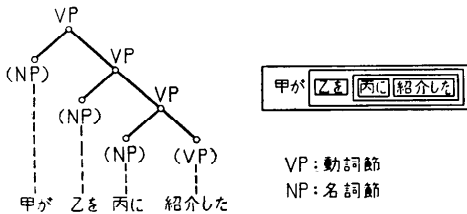
3. 構文分析

3.1 日本文の特徴

ここで次節に示す分析法の根拠となる日本文の構造の特徴について述べる。

1) 日本文の構造を木で示すと、局部的関係は left-recursive または self-embedding の形に、大局的關係(述語、補語関係)は right-recursive の形になる。つまり日本文の構成は大局的には Yngve のいう progressive construction の形をとる。いわゆる副文章は補語または修飾語のように他の文の中へはめこまれ、文全体として right-recursive の形をとることは変わらない。

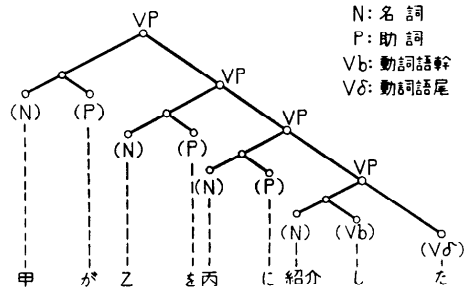
2) 上の特徴は、日本語では主格が絶対優位を示さず、動詞の任意的な補語であるという点からさらに助長される。



第1図

3) ある特定の単位間の関係は、全体的な文の構造とは無関係に局部的に定まる。

たとえば第2図の「甲+が」、「乙+を」、「丙+に」などの関係は全体的な条件に無関係に決めても矛盾は



第2図

生じない。

4) 辞書単位または節単位(辞書単位の結合したもの)に機能的 ambiguity があっても、その後にくる単位の品詞から判定が可能な場合がある(この場合、品詞とは辞書単位または節単位の機能名をいう)。

上記の特徴を構文分析に利用する。

3.2 構文分析の具体化

上の特徴を次のように形式化する。

日本文を構成する単位間に一種の operator が存在するものと考え、operator が強ければその両側の要素は局所的に結合し、弱ければ、強い operator が作用した後に作用する。この operators を6種類考える。

i) 5…非常に強い…動詞語幹と動詞語尾、形容詞語幹と形容詞語尾、本動詞語幹と補助動詞語幹との間、複合名詞となりうる2名詞の間、名詞と格助詞の間、その他これに類するもの間に存在。

ii) 4…かなり強い…属格として示された名詞と他の名詞またはその等価物の間、不変形容詞と名詞の間、その他、

iii) 3…やや強い…述語補語と述語との関係、ただし3で示される補語対述語の関係は間に他の補語・述語関係の存在を許さない。

iv) 2…やや弱い…述語補語と述語との関係、ただし2で示される補語・述語関係は間に他の補語述語関係の存在をゆるし、文全体に及ぶ。

v) 1…ごく弱い…通常、句点の前
なお、この他特別なものとして“\$”という結合関係を認める。\$はそれ以前にある2より大きい operators を全部作用させてから、4と同じ強さになる。

文中に仮定する operators の例を第3図に示す。

実際のテキスト中にはもちろんこれらの operators は存在しないから、文法規則は、二要素が結合する場

プロセス⁵ 計算機⁵ の⁴ 動作⁵ は^{5,2} 与え⁵ られ⁵ た⁵ プログラム⁵ の⁴ 量⁵ を³ ある⁴ 周期⁴ ごと⁵ に³ 繰り返え⁵ して³ 行なう⁵ こと³ である¹。

第3図

(格助詞は印欧語の変化語尾のように扱われている。)

合新たに何の機能をうるかの指示のほか、結合の強さを示すように作っておく。たとえば

$$X + Y \rightarrow Z/5$$

$X + W \rightarrow W/5$ (ただし X, Y, Z, W, は辞書単位または節単位の機能名)

$X + Y$ という組み合わせが不可能ならこの結合は ϕ で示され、Y は不適当であると判定される他のものとりかえられる。一方、各要素間の局部的あるいは大局的な関係を効果的にとらえるため push-down スタアの原理に近い記憶法を使う。すなわち文法表を使って得た分析の結果を「プール」と名付けられた一定の場所へ入れて行き、プールの入口または内部の2要素がその前後のものより強い結合を示していないかどうか吟味しながら進む。

分析の操作は、次のことの繰り返しとなる。

- 1) 辞書引きと単位→情報のおきかえ。
- 2) 結合関係 (operator の強弱) を吟味, concatenation を行なう。
- 3) 得られた結果の記憶。
- 2) の操作は6種類の operators を持つ数式の演算のようになる。

6種類の結合関係の意味は

- i) 5の時は両側の要素は無条件に結合し、より大きな節単位を構成する。
- ii) 結合関係が2~4の場合は、それより弱い関係が来るまで保留する。
- iii) 3, 4のあとに来る\$は3, 4より弱い。2は\$より弱い。
- iv) \$はそれ以前の3, 4の結合が完了したあとでは4として扱われる。

ここ⁵ に³ 用い⁵ て⁵ い⁵ る⁵ 解⁵ 法⁵ は² い² わ² ゆ² る⁴ 爆⁵ 発⁵ 法³ であ⁵ る¹。

上例の「ここに用いられている」が (vpft) という一つの節単位にまとまったのち、「解法」(n2)へ属性的にかかる。

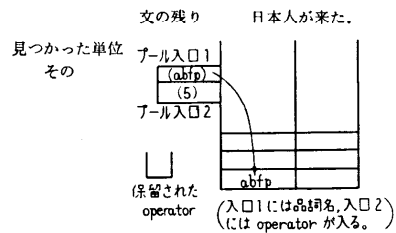
v) 結合関係が1の場合には、この結合は、それ以前のあらゆる要素が結合を完了して述語節になって初めて行なわれる。したがってその前に残っている要素がまとまらなければこの分析は正しくない。

構文分析が具体的にはどのように行なわれるかの例として「その日本人が来た」という文の分析過程を、以下に記述する。

まず辞書引きによって文頭の“その”という辞書単位が探し出され、その品詞名(機能名) abfp が情報辞書から得られる。第一の辞書引きの際“その”の他に“そ…動詞語幹”も見つかるが先ず長い方を吟味する。abfp はプールの第一の入口へ入れられる。分析の操作が始まる前に、プールは空にされているから、この abfp はプールの内容 # (ゼロ要素) と結合が可能かどうか調べられる。これは単位“その”が文頭に来りうるものかどうかの吟味である。#+abfp の組み合わせは文法表により、

$$\# + abfp \rightarrow abfp/5$$

と指定されるから abfp は文頭に来りうる品詞であると確認される。abfp を第一の要素としてプールの中へ入れ、単位自身を普通の記憶場所に記憶する(第4図)。



第4図

次に第2回目の辞書引きにより“日本人…n2”, “日本…n2”, “日…n2”, “日…n2”などがみつけれ、一番長いものが最初の候補として選ばれる。n2 がプールの第1の入口におかれ、プールの中の要素 abfp との関係が調べられ、見つかった operator が第2の入口へ入れられる。

$$abfp + n2 \rightarrow n2/4$$

であるから n2 をそのままプールに入れ、“日本人”を“その”の後に記憶し、結合関係4の保留を記録しておく。以下、句点に到るまでの操作を図式的に第5図に示す。

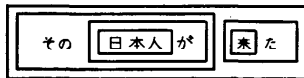
結局、文を構成する要素がすべてSにまとまり、正しい分析の行なわれたことがわかる。文を構成する単位の間には次のような operator が存在することが示され、

その⁴ 日本⁵ が³ 来⁵ た¹。

文の構造は次のように分析された。

	1	2	3	
可能な単位とその情報	その abfp そ vb1p	日本人 n2 日本 n2 日 n2 日 n4	が ps2 が pcon が n2	3までの操作が終了時のプールの状態
今採用するもの	その abfp	日本人 n2	が ps2	
使われる規則	# + abfp → abfp/5	abfp+n2 → n2/4	n2+ps2 → ns2/5	
操作の前後の string の状態	前 # — abfp 後 abfp ⑤ は operator 5	前及び後 abfp — n ²	前 abfp — n2 — ps2 後 abfp — ns2	
	4	5	6	6までの操作が終了時のプールの状態
可能な単位とその情報	来 vb2m	た vδ2ft た vb1p た abla	。 pF	6までの操作が終了時のプールの状態
今採用するもの	来 vb2m	た vδ2ft	。 pF	
使われる規則	ns2+vb2m → vb2m/3	vb2m+vδ2ft → vpft/5	abft+PF → S/1 ns2+vpft → s2vpft/3 s2vpft+PF → S/1	
操作の前後の string の状態	前 abfp — ns2 ③ — vb2m 後 ns2 — vb2m	前 ns2 — vb2m ⑤ — vδ2ft 後 ns2 — vpft	S2VPft+PF → s/1 前: ns2 — vpft — pF 前: s2vpft — pF 後 S	

第 5 図

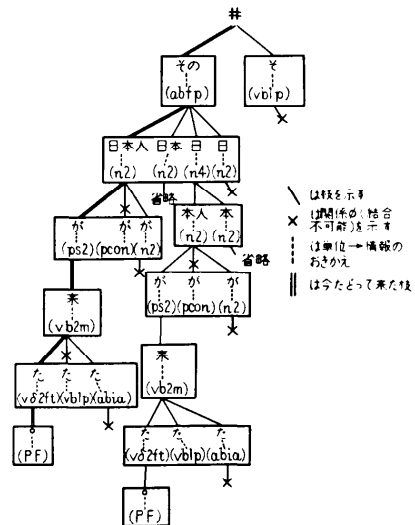


しかし、今までの経過の中に ambiguity の存在することが記憶されているから（たとえば、日本人という漢字列のところで最初は“日本人”をとったが、その他“日本”，“日”などの短い単位も可能であることが記憶されている）。第一の解答が出たあと、他の可能性との branching point にさかのぼって調べる。

今までの操作でたどって来た道および他の枝分れを第 6 図に示す。

今の場合、左最下端の句点の前の“た”のところで枝分れがある。つまり同綴異機能の単位が存在するから、それらを吟味する。

まず、はじめの“た…vδ2ft”までを消去し（分ち書き結果およびプールから）、vδ2ft の代わりに、第 2 の“た”の品詞名 vb1p をあらためてプールの入

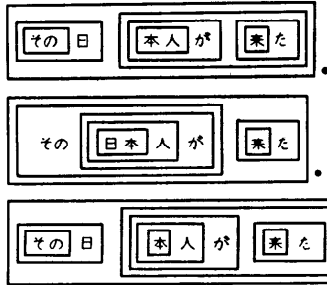


第 6 図

口に入れ、その前の要素 vb 2 m との結合関係を調べる。vb 2 m+vb 1 p は不可能であるから、これを消去し、もう一つの“た…abia”の吟味をする。規則

vb 2 m+abia → abip/5

により abia は前の要素 vb 2 m と結合し、abia とその結合の結果 abip がプールに入るが、次に来る単位は句点 pF で、abip+pF の結合は不可能である。この際、“た”のあとに続く単位は句点だけであるから abia 自身が不適当と判定される。同様にして、それ以前の枝分れも吟味される。その結果、他の正解



および

のほか

など構造から見て正しいがナンセンスな解答も出る。

これらの分析操作の大ざっぱな流れ図を第7図に示す。

4. 辞書

上に説明した分ち書きおよび構文分析に都合のよい辞書の形式ならびに内容を以下に示す。

先に述べたように、日本文では graphemic レベルにおける空白を示す単位の欠如のため、語を形態的には識別できない。したがって辞書単位の識別は印欧語で form を lexes に切断するのと同じになる。そこで辞書は切断の規準および切断された単位のもつ情報を与える。見出し辞書には辞書単位そのものが、情報辞書には辞書単位が自然言語の形の時有する情報が入られる。情報辞書内での情報列は、たとえば次のような形で与えられる。

“知”に対して vb | 1 | p | cd | cop | …

“る”に対して vd | 1 | ft | R | cd | cop | …

なお、vb は動詞語幹、vd は動詞語尾を示す。

上の情報列のうち、構文分析までに必要なものは、それぞれはじめの三つだけで、あとの意味分析のための情報、対応する出力言語へのつなぎの情報は構文分析の際にはふれられない。

前記のような分ち書き、構文分析を有効に行なうために特に必要となる辞書引きの条件には、

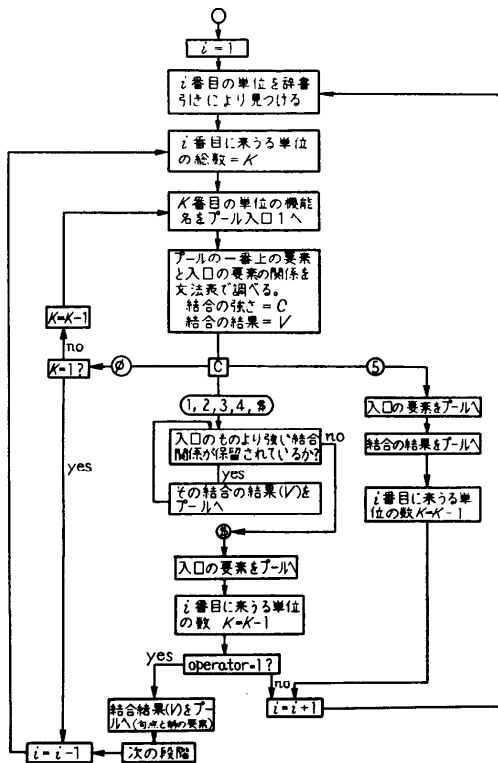
- 1) homographe, polysemie が存在する場合には全部をチェックする。
- 2) 与えられた string が数種類の長さに切断可能な場合、すべての長さの単位を見つけ出す。
- 3) この辞書引きの時間を節約する。

見出し辞書を木の形に作っておいて、辞書引きの際1文字を1回しか問題にしなれば、これらの条件がかなえられる。辞書引きは、テキストの与えられた箇所と見出し辞書をつき合わせ、一致する単位を短いものからいちいち記憶しながら、一番長い単位を見つけ出すまで進む。このことを図式的に示すと第8図のようになる。

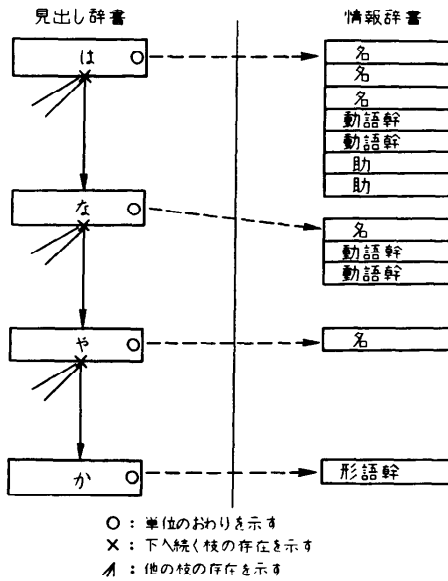
構文分析には、一番最後に見つかった単位、つまり一番長い単位から先に取り出される。

漢字、かなをすべて使用しても、見出し辞書の木の根の数は、語頭には来ることのない文字の存在するところから“漢字+かなの総数”以下になる。また枝の深さは、だいたい 2.5~3 程度のものが多くなる。

木形辞書の実現のためのテクニックとして、一般に

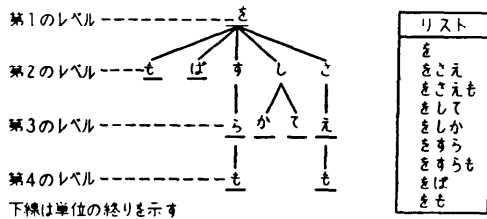


第7図

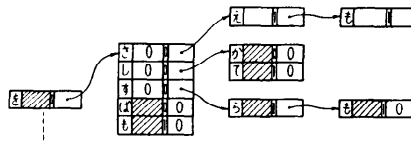


第8図

Lisp, IPL, FLPL などのリスト構造言語が用いられるが、検索の速さと経済性という点から考え、レベルが変わる時だけアドレスで指定し、同じレベルで次の文字に移る時にはインデックス・レジスタを使ってアドレスを一つずつ進める方法を採用する (Grenoble の Veillon 方式)。これによれば第9図の木は第10図のように表わされる。



第9図



第10図

なお、第9図、第10図に示されているものは“を”で始まる単位のリストである。

5. 自然言語の辞書単位とその情報

構文分析までに必要な情報によって日本語の辞書単位を分類すると第1表に示すクラス、サブクラスが分けられる。

第1表

クラス	サブクラス
1 名詞 (n)	1, 2, 2, 4
2 動詞語幹 (vb)	1p, 1a, 1m, 2p, 2a, 2m, 2s
3 動詞語尾 (vd)	1ft, 1ff, 1uf, 1uv, 2ft, 2uf, 2uv, 3ft, 3ff, 3uf, 3uv
4 形容詞語幹 (ab)	ip, ia, im, np, na, nm bp, fp
5 形容詞語尾 (ad)	if, ie, iuf, iuv, iav, ne, nav
6 コブラ (c)	ft, e, uf, uv
7 助詞 (p)	s2, s1, o, g, l, cj, nav, vav, or, ob, sb
8 副詞 (ad)	1, 2, 3
9 接続詞 (con)	1, 2
10 感嘆詞 (in)	
11 特殊接頭尾辞 (af)	n, a
12 雑詞 (vr)	

第2表

1 名詞 範囲: いわゆる名詞			
クラス	グループ		例
n	1; 複合名詞を作りうるもので、スル、デキルにも接続するもの	利用	n 1
	2; 複合名詞を作りうるが1に属さないもの	空気	n 2
	3; 複合名詞を作りえないもの	かれ	n 3
	4; いわゆる形式名詞	ため	n 4
2 動詞語幹 範囲: いわゆる動詞または動詞形助動詞の語幹			
クラス	グループ		
vb	1; いわゆる5段活用とそれに準ずるものの語幹	飛(ぶ)	vb 1 p
	2; その他の動詞およびそれに準ずるものの語幹	食(べる)	vb 2 p
	p: 助動詞とならないもの	た(がる)	vb 1 a
	a: 助動詞としてのみ使用されるもの	あげ(る)	vb 2 m
	m: 助動詞または本動詞として使用されるもの	す(る)	vb 2 s
	s: す(る), し(た), せ(よ), でき(る) など		

3 動詞語尾 範囲: 動詞語幹クラスに所属するものの語尾

クラス	グループ	カテゴリ

- vδ 1: 動詞クラス・グループ 1の後に来るもの
- 2: 動詞クラス・グループ 2の後に来るもの
- 3: 動詞クラス・グループ 1・2の後に来るもの
- ft: 終止体で属性的に用いられる語尾
- ff: 終止体で属性的に用いられない語尾
- uf: 保留体で第2義的主格をうけとめるもの
- uv: 保留体で第2義的主格をうけとめないもの

(泳) いた	vδ	1	ff
(来) たら	vδ	2	uf
(食べ) つつ	vδ	2	uv
(泳) ぐまい	vδ	1	ft

4 形容詞語幹 範囲: いわゆる形容詞, 形容動詞, 形容詞形助動詞, 形容動詞形助動詞の語幹および連体詞

クラス	グループ	カテゴリ

- ab i: いわゆる形容詞およびそれに準ずるものの語幹
- n: いわゆる形容動詞およびそれに準ずるものの語幹
- b: 上の i, n の両方にまたがるものの語幹
- f: いわゆる連体詞および不変形の形容詞
- p: 助形容詞とならないもの
- a: 助形容詞としてのみ用いられるもの
- m: 助形容詞または本形容詞となりうるもの

来 (い)	ab	i	p
らし (い)	ab	i	a
そう (だ)	ab	n	a
あらゆる	ab	f	p

5 形容詞語尾 範囲: 形容詞語幹クラスに属するものの語尾

クラス	グループ	カテゴリ

- aδ i: いわゆる形容詞の語尾にあたるもの
- n: いわゆる形容動詞の語尾にあたるもの
- f: 述語的および属性的に用いられるもの
- e: 属性的のみ用いられるもの
- uf: 保留体で第2義的主格をうけとめるもの
- uv: 保留体で第2義的主格をうけとめないもの
- av: 副詞的に用いられ

(赤) かった	aδ	i	f
(静か) な	aδ	n	e
(赤) かったり	aδ	i	uv
(早) く	aδ	i	av

るもの

6 コブラ 範囲: “だ”および“です”のすべての活用形

クラス	カテゴリ

- c f: 終止体で述語的に用いられるもの
- e: 属性的に用いられるもの
- uf: 保留体で第2義的主格をうけとめるもの
- uv: 保留体で第2義的主格をうけとめないもの

だ	c	ft
ならぬ	c	e
だったり	c	uv
だったら	c	uf

7 助詞 範囲: いわゆる助詞および助詞的に用いられているもの

クラス	カテゴリ

- p s1: 第一義的主格を示すもの
- s2: 第二義的主格を示すもの
- o: 目的格のみを示すもの
- g: 風格を示すもの
- l: その他の格を示すもの
- cj: 名詞間の接続関係を示すもの
- nav: 副詞的に使用されるもの(補充的に)
- vav: 副詞的に使用されるもの(動詞と)
- or: 並列接続詞の役目をするもの
- ob: 副詞的接続詞の役目をするもの
- sb: 名詞化の役目をするもの

は	p	s1
が	p	s2
を	p	o
の	p	g
へ	p	l
と	p	cj
くらい	p	nav
さえ	p	vav
し	p	or
が	p	ob
の	p	sb

第3表 文法規則の例

$X_1 + X_2 \rightarrow X_3/5$

$(X_1 = n_1, n_2)$
 $(X_2 = n_1, n_2)$

$vb1 X_4 + v\delta 1 X_5 \rightarrow vp X_6/5$

$vb2 X_4 + v\delta 2 X_5 \rightarrow vp X_6/5$

$vb2s + v\delta 2 X_5 \rightarrow vp X_6/5$

$(X_5 = ft, ff, uf, uv)$

$(X_6 = p, m)$

$vb2p + vb2 X_5 \rightarrow vb2 X_6/5$

$(X_5 = a, m)$

$vpuv + v X_6 X_7 \rightarrow v X_8 p/5$

$(X_6 = b1, b2)$

$vpuv + vb2s \rightarrow vb2s/5$

$n_1 + vb2s \rightarrow vb2p/5$

$abip + a\delta i Y_1 \rightarrow ap Y_1/5$

$abim + a\delta i Y_1 \rightarrow ap Y_1/5$

$abbm + a\delta i Y_1 \rightarrow ap Y_1/5$

$(Y_1 = f1, e, uf, uv)$

$abY_2p + abY_3a \rightarrow abY_4p/5$

$(Y_2 = i, n)$
 $(Y_3 = i, n)$

$vb2 Z_1 + abY_4a \rightarrow abY_4p/5$

$(Z_1 = p, m, s)$
 $(Y_4 = i, n)$

$vpuv + abY_4a \rightarrow abY_4p/5$

$vpuv + abY_4m \rightarrow abY_4p/5$

$apf + X_7 \rightarrow X_7/4$

$(X_7 = n_1, n_2, n_3, n_4)$

$vpft + X_7 \rightarrow X_7/4$

$apfp + X_7 \rightarrow X_7/4$

$n_1 + pM_1 \rightarrow nM_1/5$

$(M_1 = s_1, s_2, o, g, l, cj)$

$n_2 + pM_1 \rightarrow nM_1/5$

$ns1 + vpX_8 \rightarrow s1vX_8/2$

$no + vpX_8 \rightarrow ovX_8/3$

単位のクラス名およびサブクラス名が、構文分析に必要な情報のおもなものであるが、これが情報辞書内の情報列のはじめの部分となる。第2表にその形および例を図式的に示す。

文法規則の例を第3表に示す。

あとがき

上述の研究は、従来の phrase-structure grammar の規則を、特に日本語のために効果的に適用できるようにしたいという目的からなされたものである。

他の方法と比較して、どのくらい時間が短縮されるか、プログラムがどれだけ簡単になるか、記憶装置の使用部分をどのくらい節約できるか、また、余分な解答をどのくらい減らしうるかを、現在のところ、厳密に比較することはむずかしいが、これらの点に関して、かなり有利になることは確実である。

日本語で、第一義的な主格(“は”で示されるもの)を簡単にしかも正しく分析することはどの方法を採用するにしてもきわめて困難である。上記の方法では、operator 2 と 3 を区別し、“は”によって示される主格と動詞との関係を 2 で示しておいて、その主格がどの動詞にかかるかの可能性をもれなく捕えようとしているが、そのため意味的にナンセンスな解答も出て来る。たとえば“以下のことは諸兄のよく知るところと思う。”という文の“以下のことは”は“よく知る”の第一義的主格(諸兄のが第二義的主格)であるが

……ことは……よく知る……
 ↑

の他に

……ことは……よく知るところと思う。
 ↑

という解答も出る。しかし正解をもらさずにとらえて

あれば第2例のような解答は意味分析の過程でふいおとすことが可能であると思う。

また“日本人”と“日本人”のように全く同じものが、一つは辞書単位として、他の一つは辞書単位の組み合わせとして出て来る。このような重複はある程度やむをえないと考える。

なお、この報告は昭和38年10月28日、機械翻訳委員会幹事会(丹羽保次郎委員長)の席上で非公式に発表したものを発展させたものである。

終りに本研究のためいろいろと御指導、御助言をたまわった電気試験所の和田部長はじめ、諸氏に心から感謝の意を表します。

参考文献

- 1) I. Sakai: Syntax in Universal Translation, 1961, International Conference on MT, London, Her Majesty's stationery office. pp. 593~609
- 2) 小熊 均: 日本語における構文解析の方法, 機械翻訳委員会席上での発表, 1963年10月29日
- 3) S. Kuno & A.G. Oettinger: Multiple-path Syntactic Analyser, NSF-8, 1963, Harvard University.
- 4) S. Lamb & W. Jacobson: A high-speed-Capacity Dictionary System, Mechanical Translation, vol. 6, November 1961 pp. 76~107
- 5) G. Veillon: Consultation d'un Dictionnaire et Analyse syntaxique en Traduction Automatique, 1962 (Thèse présentée à la Faculté des Science à Grenoble)
- 6) 三上 章: 現代語法序説

(昭和39年3月12日受付)