

意味処理に基づいた機械翻訳システム

Machine Translation Based on Semantics

機械翻訳は実現困難なテーマと考えられた時期もあったが、言語学の進歩と計算機の性能向上により、人間の適切な援助を前提とすれば実用化が可能と考えられるようになってきた。日立製作所では、国際交流の活発化とともに急速に高まるニーズにこたえて、日英機械翻訳システムを製品化し、英日機械翻訳システムを研究開発中である。日英機械翻訳システムでは、言語独立性の高い意味表現「概念依存図式」を中間表現として両言語の構文ギャップを吸収する翻訳方式を開発した。英日機械翻訳システムでは、英語と日本語のそれぞれの特徴を考慮して設定した「疑似的句構造」と「格的句構造」の間で構文変換を行う翻訳方式を開発した。日英、英日両機械翻訳システムとも、意味情報を利用して構文的あいまいさを解消する技術を開発し、翻訳精度の向上を図った。また、前編集、後編集、ユーザー辞書作成などの支援機能を充実させ、実用性の高いシステムを実現した。

梶 博行* *Hiroyuki Kaji*
岡 島 惇* *Atsushi Okajima*
吉村紀久雄** *Kikuo Yoshimura*

1 緒 言

ある国の言語から他の国の言語への翻訳を、計算機に行わせることは実現不可能なテーマと考えられた時期もあった。現在でも言語理解のメカニズムはほとんど未解明であり、完全自動翻訳が不可能であることに変わりはない。しかし、計算言語学の分野での着実な技術の蓄積と最近の知識処理技術の進歩により、適切な対象分野で人間の援助を前提とすれば、実用システムを作ることができると考えられるようになった¹⁾。

一方、経済や科学技術分野の国際化とともに機械翻訳システムに対する期待が急速に高まってきた。特に我が国では、日本語が国際的に通用しないため、翻訳を必要とする文書が非常に多く、人手だけに頼れなくなりつつある。翻訳コストや翻訳に要する時間の問題だけにとどまらず、科学技術情報の輸入超過に対する諸外国の批判など、国家レベルの問題となってきている。

日立製作所ではこのような状況をいち早くとらえ、機械翻訳システムの開発に取り組んできた。現在までに日英機械翻訳システムHICATS/JE(Hitachi Computer Aided Translation System/Japanese to English)を製品化し、英日機械翻訳システムを研究開発中である。日英機械翻訳システムはマニュアル、特許抄録、論文などの科学技術文を主な対象とし、英日機械翻訳システムは科学技術文のほか経済記事を対象としている。いずれもHITAC Mシリーズ計算機でオペレーティングシステムVOS 3(Virtual-storage Operating System 3)の下で利用できる。日立機械翻訳システムの開発に当たっては、(1)意味処理の導入による翻訳精度の向上、(2)知識工学的手法による拡張性の高いシステムの実現、(3)実用性

を高める各種の支援機能の実現、に特に力を入れた。

2 機械翻訳の技術課題

機械翻訳システム実現上の主な技術課題と、それらに関して日立機械翻訳システムで開発した技術の概要を以下に述べる。

(1) 翻訳方式

機械翻訳は、文の構造や意味を形式的に表現する中間表現を介して行われる。翻訳方式は中間表現の設定の仕方により、トランスファ方式とピボット方式に大別される。トランスファ方式は、図1(a)に示すように、入力言語の中間表現と出力言語の中間表現を設定し、二つの中間表現の間に変換の過程を置く。一方、ピボット方式は、同図(b)に示すように、入力言語と出力言語に共通の中間表現を設定するもので、変換過程を含まない。どの言語にも依存しない中間表現を作るとは不可能との考えもあり、厳密な意味でのピボット方式の実現は難しい。しかし、入力言語と出力言語の両方に依存する変換の部分で、できるだけ小さくするという考え方は重要である。

翻訳方式は、入力言語と出力言語の特徴を考慮して決めなければならない。質の良い翻訳を実現するには、特に、入力言語の特徴を考慮することが重要である。入力文の持つ情報に基づいて翻訳するからである。日英機械翻訳システムでは、意味重視の中間表現によるピボット方式を基本としながら、両言語の発想の違いに関してトランスファ方式を採用した。一方、英日機械翻訳システムでは、構文重視の中間表現によるトランスファ方式を採用した。

* 日立製作所システム開発研究所 ** 日立製作所ソフトウェア工場

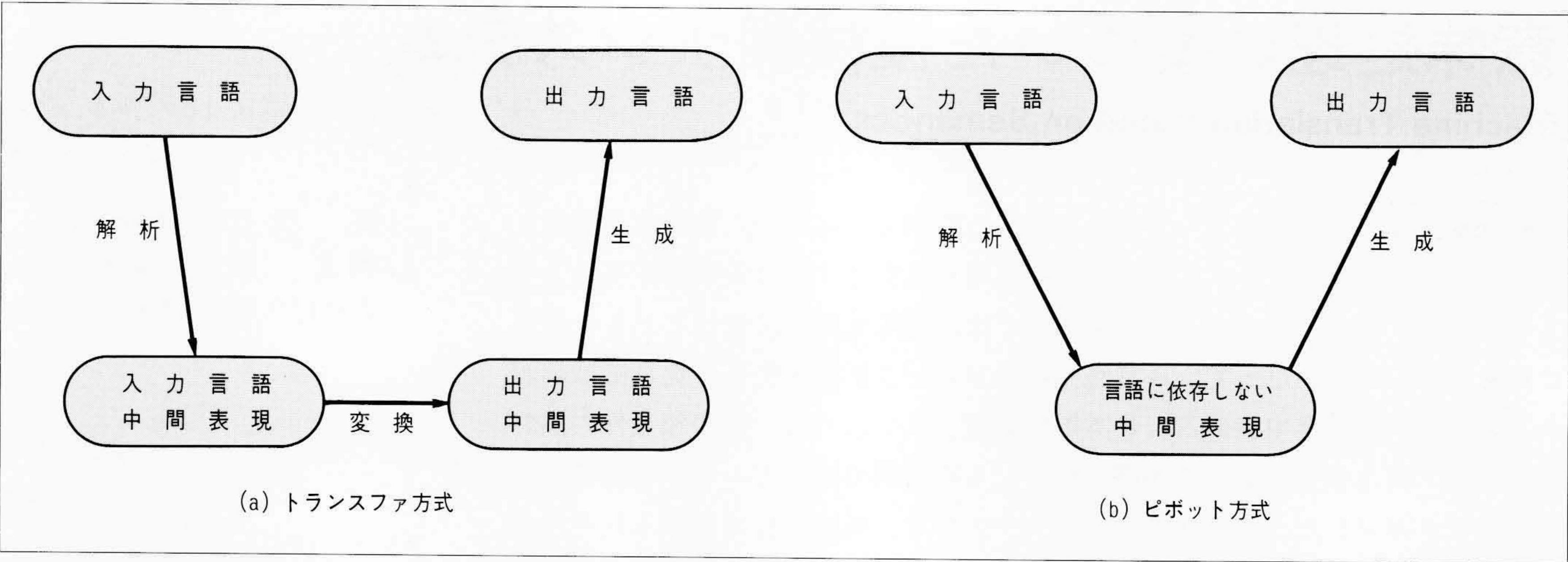


図1 2種類の翻訳方式 トランスファ方式は、二つの言語の中間表現の間に変換の過程を持つ。ピボット方式は、二つの言語に共通の中間表現を用いる。

(2) 構文・意味処理技術

入力文を中間表現に正しく写像し、適切な訳文を生成するには、構文、意味、更には文脈の処理が必要である²⁾。構文については言語学的にもかなり解明されている。しかし、意味や文脈についてはまだ分かっていないことが多く、計算機で形式的に扱うのが難しい。日立機械翻訳システムでは構文処理を基本としながら、構文情報だけでは解消できないあいまい

さを、意味情報を利用して解消する技術を開発することによって翻訳精度の向上を図った。

(3) 翻訳ソフトウェアの実現方式

翻訳には膨大な知識が必要であり、知識の追加・変更が容易にできる構造であることが実用システムの必要条件である。日立機械翻訳システムでは、図2に示すように、語彙(彙)に関する知識を蓄えた辞書だけでなく、文法的知識もデータと

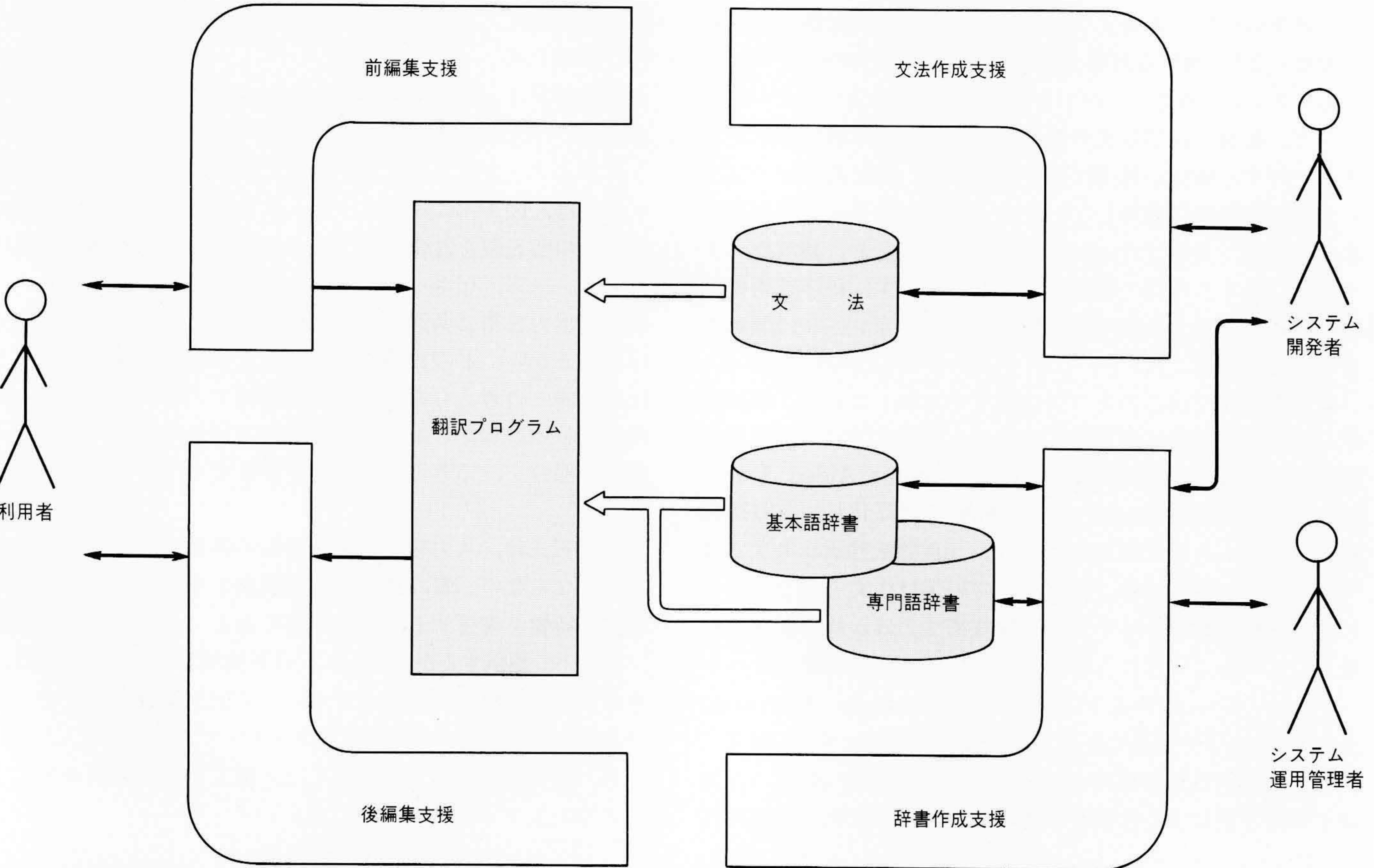


図2 日立機械翻訳システムの構成 翻訳プログラムから文法、辞書を分離することにより、拡張性・保守性の高いシステムを実現している。また、各種の支援機能により実用性を向上している。

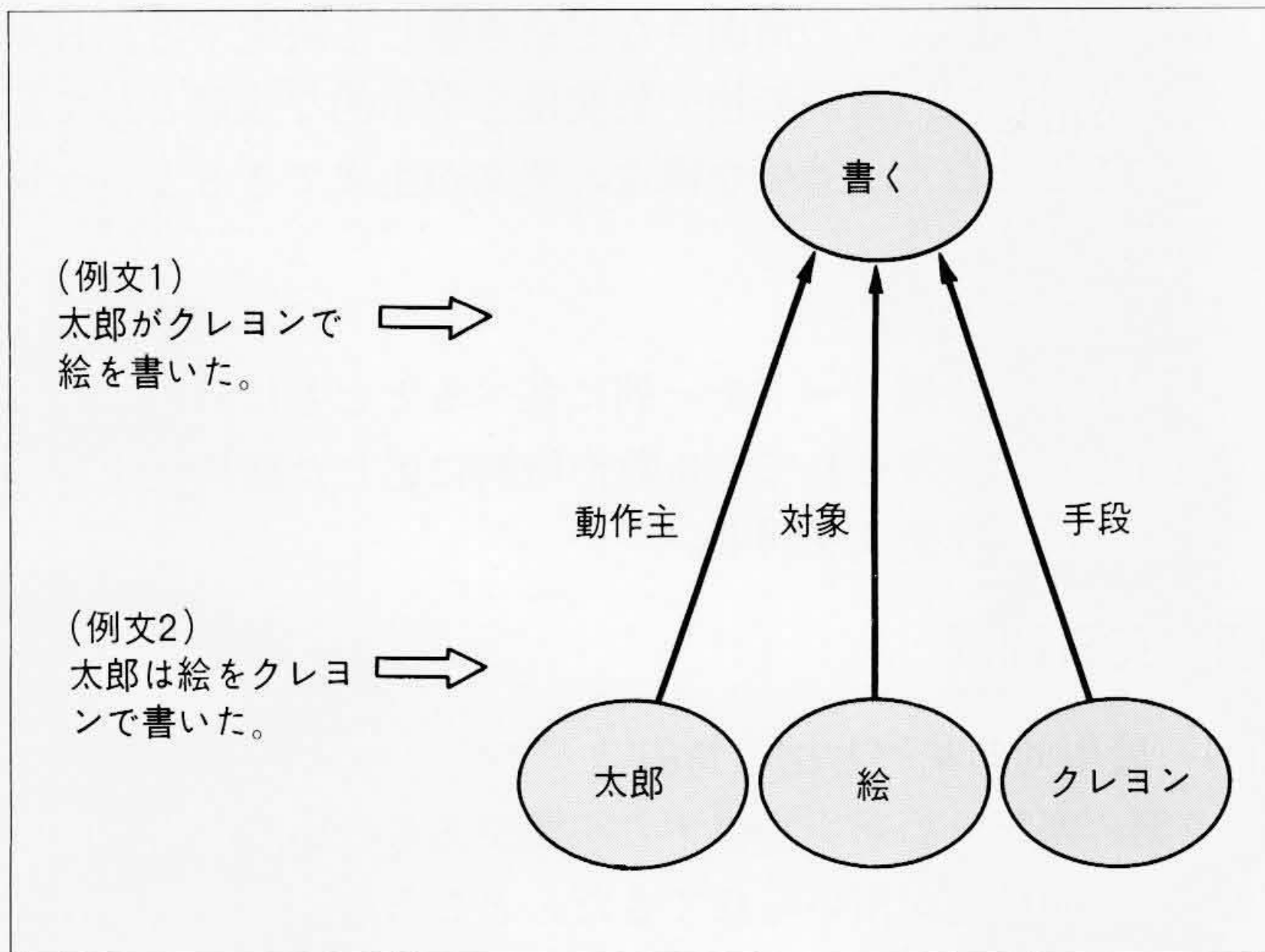


図3 概念依存図式 語順や語句が異なる文でも、文としての意味が同じであれば同一の概念依存図式に写像される。

して翻訳プログラムから分離し、拡張性・保守性の高いシステムを実現した。すなわち、翻訳プログラムは、翻訳処理に必要な木構造のパターンマッチングと変形操作の機能を持つプログラムであり、文法的知識は木構造の書換え規則として表現する^{3),4)}。翻訳精度の向上は、規則の追加・改良によって

行うことができる。

(4) 前編集・後編集支援機能

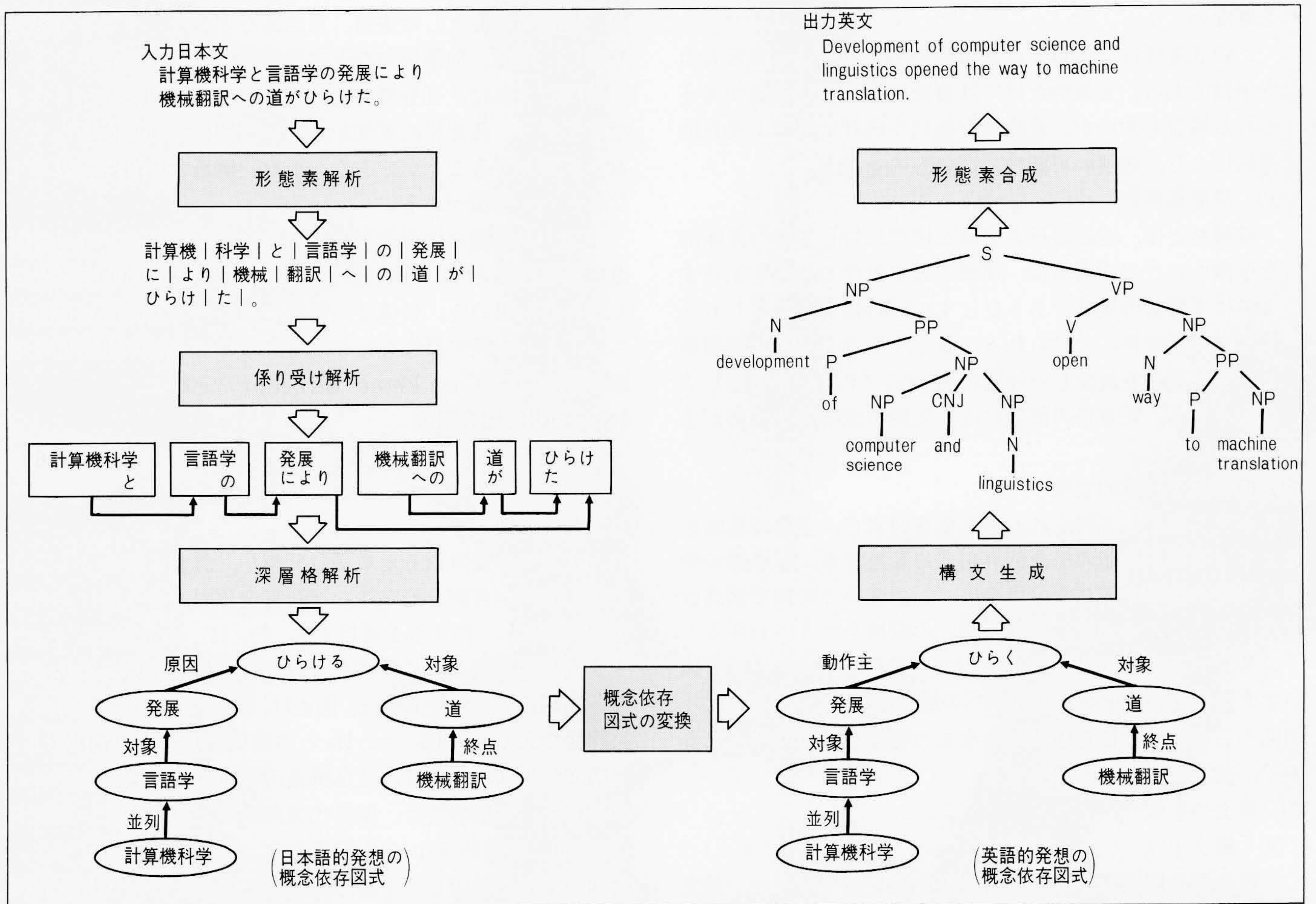
現在の技術レベルでは完全自動翻訳は不可能である。そのため、利用目的に応じて入力文を翻訳しやすいように修正する「前編集」と、訳文の不備を訂正する「後編集」が必要である。日立機械翻訳システムでは、前編集・後編集を効率よく行うための支援機能を提供し、実用性の向上を図った。

3 日英翻訳

3.1 概念依存図式

日本語と英語は異なる言語族に属し、構文ギャップが大きい。しかし、語句の表す概念の意味的な依存関係には共通点が多い。この点に着目し、「概念依存図式」と呼ぶ有向グラフで文の意味を表現することを試みた。概念依存図式では、ノードが事物や動作などの概念を表し、アークが動作主、対象、手段、場所といった深層格関係(述語に対する名詞の意味的な役割)を表す。文に含まれるほかの情報、例えば時制などはノードの属性として表現する。

概念依存図式の例を図3に示す。概念依存図式は、表層文中の語順や言い回しが違っていても、文としての意味が同じであれば同じ図式に写像されるという特徴がある。この特徴は日本語の特質に合致している。更に、日本語と英語のいずれにも偏



注：略語説明 S(Sentence), NP(Noun Phrase), N(Noun), VP(Verb Phrase), V(Verb), PP(Prepositional Phrase), P(Preposition), CNJ(Conjunction)

図4 日英翻訳の処理概要 日本語と英語で発想が異なる表現をとる場合、概念依存図式を用いた変換により、英語らしい文の生成を可能にする。なお、概念依存図式は言語独立性が高く、多くの文に関しては、日本文解析で得られる概念依存図式がそのまま英文生成の出発点になる。

らない中立的表現であるので、日本文解析部と英文生成部が共に相手言語に依存しない形で実現できるという利点がある。

3.2 翻訳処理の概要

日英翻訳は次のステップで実行する。一つの例を図4に示す。

(1) 日本文の解析⁵⁾

(a) 形態素解析

文の解析での最小単位である形態素(単語や接辞)を認識する処理である。日本文は単語と単語の間に空白がないので、単語の境界を認識するのに字種情報を利用したり、品詞の接続可否をチェックしたりしながら、辞書を検索する複雑な処理が必要である。単語への分割が終わると、辞書に記憶されている単語の文法的・意味的な性質を表す種々のコードが利用できるようになる。

(b) 係り受け解析

日本語は語順の自由度が高いため、文節間の修飾関係に注目した係り受け構造が、構文をとらえるのに適している。文節とは内容語(名詞、動詞、形容詞など)と、それに続く機能語(助詞、助動詞など)をひとまとめにしたものである。係り受け構造は次の3条件を満たす。

- (i) 文末以外の文節は必ず後方の一つの文節に係る。
- (ii) 係り文節から受け文節に矢印をひいたとき、矢印が交差することはない。
- (iii) 一つの文節が、同じ意味的役割の文節を複数受けることはない。

この3条件は、簡単ではあるが係り受けを決定する有力な手段である。3条件だけでは複数の係り受け構造が考えられる場合も多いが、意味的な制約や統計に基づく発見的規則により、正解の可能性の高いものを選ぶ。

(c) 深層格解析

深層格とは、述語に係る名詞が述語に対して持つ意味的な役割をいう。深層格は、述語の表す動作や状態を記述するのに不可欠の要素であるひっす(必須)格と、必ずしも必要でない任意格に分けられる。ひっす格は、助詞と深層格との対応関係を規定した述語の格支配パターンを利用して決定できる。任意格の決定には、名詞の意味的な性質が手掛かりになる。

(2) 概念依存図式の変換⁶⁾

日本文を解析して得られる概念依存図式を、全体の意味を変えずに英語的発想の概念依存図式に変換する。日本語の発想と英語の発想が同じ文の場合は、この変換の前後で図式は変化しない。この点で、構文や語いの変換を行ういわゆるトランスファ方式とは異なる。例えば、図4の例文が「計算機科学と言語学の発展が機械翻訳への道を開いた。」であれば、同図の右側の概念依存図式が日本文の解析から直接得られるので、変換の過程は不要になる。

(3) 英文の生成⁷⁾

(a) 構文生成

概念依存図式の各ノードに訳語を割り当てながら、英文の句構造を生成する。句構造は単語から句、更に文というように、文の構成要素のまとまりを階層的に表現するものである。句構造の生成で、文形や訳語の品詞は、英語の文

法的制約の下で文の簡潔さなどを考慮して決定する。日本語と英語の間の品詞転換や態変換を明示的な規則として記述することなく、適切な構造の英文が生成できるという特徴がある。

(b) 形態素合成

句構造の終端ノードを一行に並べるとともに、ノードの属性として表現されている数や時制に応じた語形変化を施すことにより訳文を得る。

4 英日翻訳

4.1 疑似的句構造と格的句構造

英語の解析は語順を重視する必要があるため、従来、解析木として句構造が用いられてきた。ところが、前置詞が幾つも使われた文の場合、句構造では木を生成する順に修飾関係を決めないとめんどろな処理が必要なことがある。例えば、図5(a)で、なんらかの理由によってwith a computerがstudiesを修飾することが分かっているが、studies machine translationの解析とin this laboratoryの修飾先の認識ができるまで、studies~computerを動詞句にまとめることができない。このような問題点を解決するため、修飾先のあいまいさが生じやすい任意格は、リンク[図5(b)の波線]で修飾先に直接つながるように拡張した表現方法を開発した。この表現方法を疑似的句構造QPS(Quasi Phrase Structure)⁸⁾と呼ぶ。QPSでは、動詞を中心とした主語、目的語などのひっす格を句構造で表し、副詞、前置詞句などの任意格をリンクで表現する。

QPSによれば、任意格の解析順序に関する制約がなくなるだけでなく、任意格的要素がどういう関係で何を修飾しているかを表示することができる。また、解析木の段数が少なくなるという特徴がある。

一方、日本語の中間表現としては、述語を中心とする格構造を基本単位とし、格構造の要素に格構造を再帰的に埋め込んだものを用いる。格構造の要素には、その深層格的意味やニュアンスを表す属性を付与する。この日本語中間表現を格的句構造CPS(Case Phrase Structure)⁹⁾と呼ぶ。

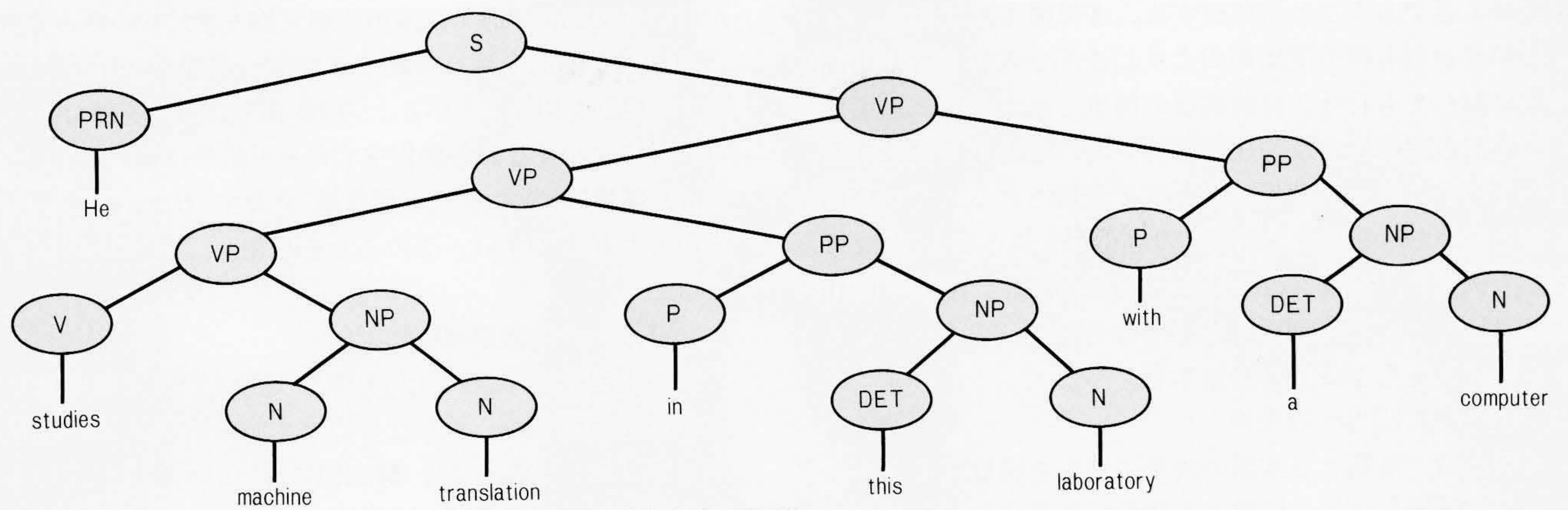
4.2 翻訳処理の概要

英日翻訳は、次のステップで実行する¹⁰⁾。一つの例を図6に示す。

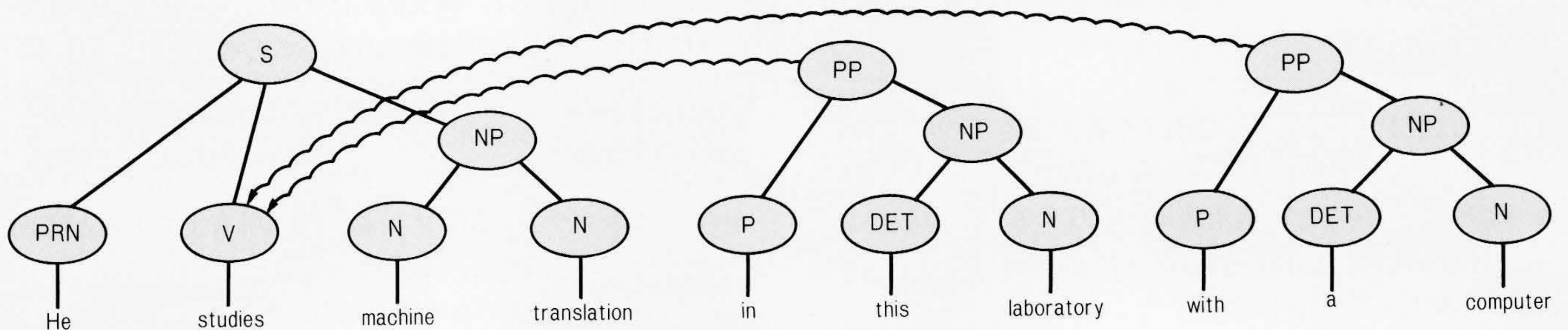
(1) 英文の解析

入力英文を受け取ると単語に分割し、辞書を検索する。この際、動詞、名詞、形容詞など語形変化のある単語はその形態素を解析し、標準形を選択する。例えば、studiesは「study + 3人称単数現在」に帰着させる。次に、品詞や属性の並びから多品詞語の品詞を決める。例えば、studyは名詞と動詞の多品詞語であるが、図5では、Heとの関係から動詞を選択する。

構文解析は、句解析、述部構造の解析、修飾関係の解析の順に行う。句解析では、関係代名詞節などを含まない単純句(名詞句、前置詞句など)を認識する。述部構造の解析では、動詞の文形を判定し、主語、目的語、補語といった構文的役割子を認識する。図6のIt~toのような構文も同時に認識する。例えば、It~toの構文では、isの真主語がto不定詞であること、更にIt~for~toの構文では、forの次の名詞句がto不



(a) 従来の句構造



(b) 疑似的句構造

注：略語説明 PRN(Pronoun), DET(Determiner)

図5 従来の句構造と疑似的句構造の違い
 従来の句構造は、疑似的句構造に比べて解析木の段数が少なく、修飾関係が見やすいという特徴がある。

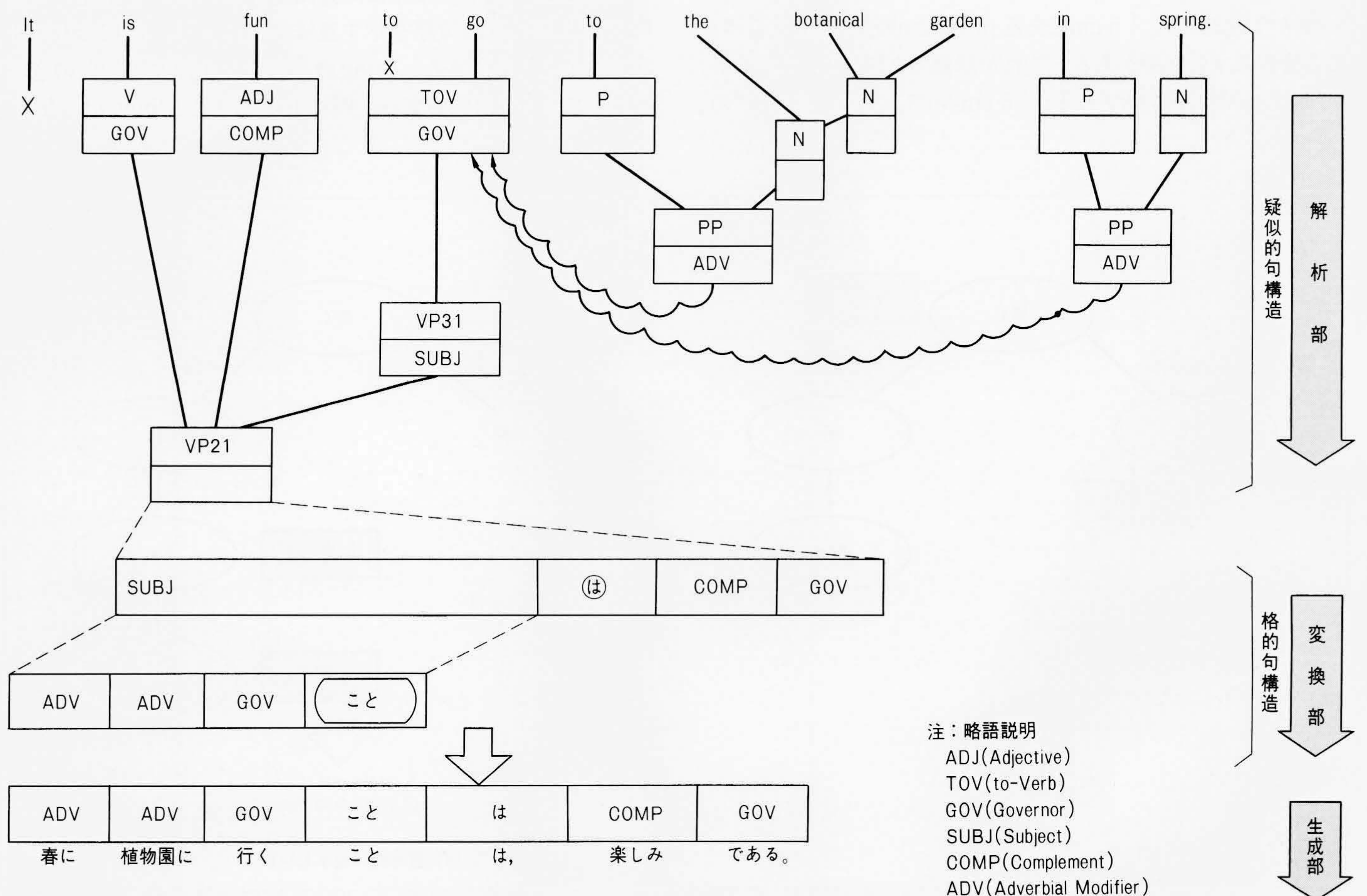


図6 英日翻訳の処理概要 疑似的句構造は、述部を軸として格的句構造に変換される。更に、助詞などの付加により日本語が生成される。

定詞の主語であることを認識する。述部構造の認識は、to不定詞のほか動名詞や関係節のように述部が述部を含む再帰的構造に対しても行う。修飾関係の解析では、副詞、前置詞句あるいは述部構造を持つ任意格的要素の修飾先を認識する。QPSでは、これらの関係はリンクで表現されるので、修飾する側と修飾される側の2項関係の解析が主体となる。例えば、**図5**の例文で、with a computerの修飾先は、「施設+with+物(道具)」と「行為動詞+with+物(道具)」のもっともらしさを判定することにより決定する。

(2) QPSからCPSへの変換

変換部は、英語と日本語の表現の違いを吸収するステップであり、語順を含めた格構造の変形で実現される。能動態を受動態に変換するとか、無生物主語を道具格的に解釈し直すなどの処理もこのステップで行う。

(3) 日本文の生成

日本文は、CPS中の構文的役割子や属性に従って生成する。主語、目的語といった構文的役割子に従い助詞を付与するとか、動詞、形容詞などの活用語尾を付与するなどの形態素処理が中心となる。語義やニュアンスによって異なる訳語の選択もこのステップで行う。

5 意味処理の例

日立機械翻訳システムでは、語の意味を考慮することにより、構文的あいまいさを解消したり、適切な訳語を選択する機能を実現している。そのうちの幾つかの例を以下に示す。

(1) 縮退表現の復元

図7の例文では、「5 cmである」が「5 cmの」に縮退していることを認識する必要がある。これが認識できないと、「半径が」が「描く」に係ると解析される。「5 cmの」のように、名詞に「の」

が付いた文節は、通常、連用修飾文節を受けないからである。このような縮退表現を復元するため、意味的な望ましさを判断する規則を用意している。「半径」は‘意志性’を持たないので、「描く」の主語すなわち動作主である可能性は低い。一方、「半径」は‘属性’を表すので、‘属性値’を表す「5 cm」のような数量表現と共起しやすい。このことを利用し、同図(b)の訳を選択する。

(2) 「は」の深層格解析

文の主題を表す助詞「は」を伴う文節は、構文的な係り先と意味的な修飾先が一致しないことがある。例えば、**図8**(a)の例文で、「本システム」は、構文的には「広い」に係っているが、意味的には「応用範囲」の所有主である。また、同図(b)の例文で、「文学作品」は、構文的には「難しい」に係っているが、意味的には「翻訳」の対象である。このような深層格関係の抽出は、構文的処理だけでは不可能であり、語の意味的な性質に基づく解析規則を用意している。同図(a)では、「応用範囲」が‘属性’を表す語で、「広い」が‘属性値’を表す語であることから、「本システム」が「応用範囲」の所有主であると解析する。一方、同図(b)では、「翻訳」が‘動作’を表す語で、「広い」が‘難易性’を表す語であることから、「文学作品」が「翻訳」の対象であると解析する。

(3) 英日翻訳での訳語選択

訳語選択は、(a) 入力言語内での多義と考えるべき場合と、(b) 出力言語の語の使い方の問題と考えるべき場合がある。いずれの場合も、ほかの語との共起関係が手掛かりとなる。訳語は**図9**に例示するように、意味的な共起制約を記述した意味文法と語の意味分類を示す意味辞書により、適切なものを選択することができる。bringの多義は、主語や目的語の意味コードによって解消できる。例文①では、主語‘boy’が「人」、

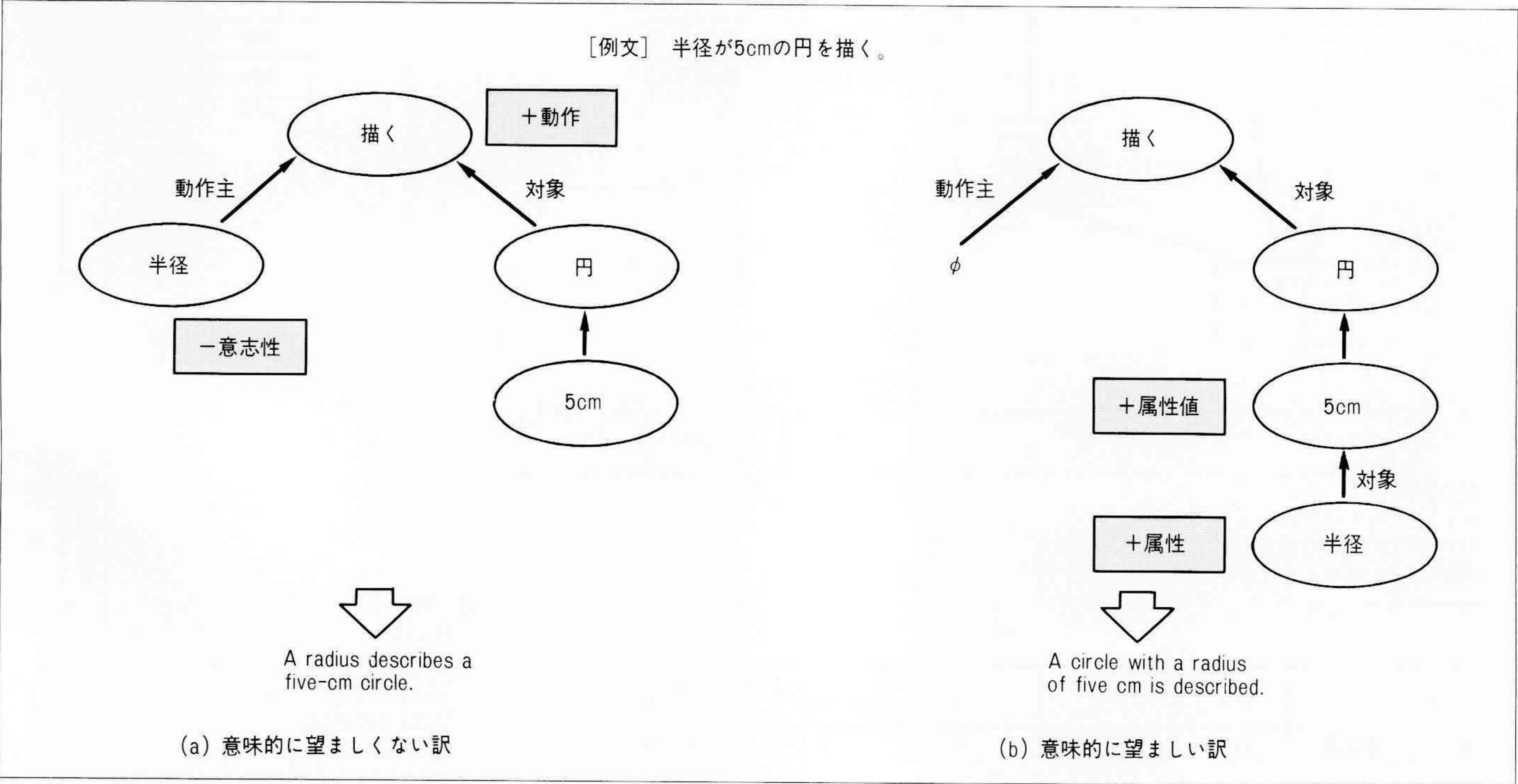


図7 縮退表現の復元 構文的には(a), (b)共に可能であるが、意味的な望ましきにより(b)を選択する。

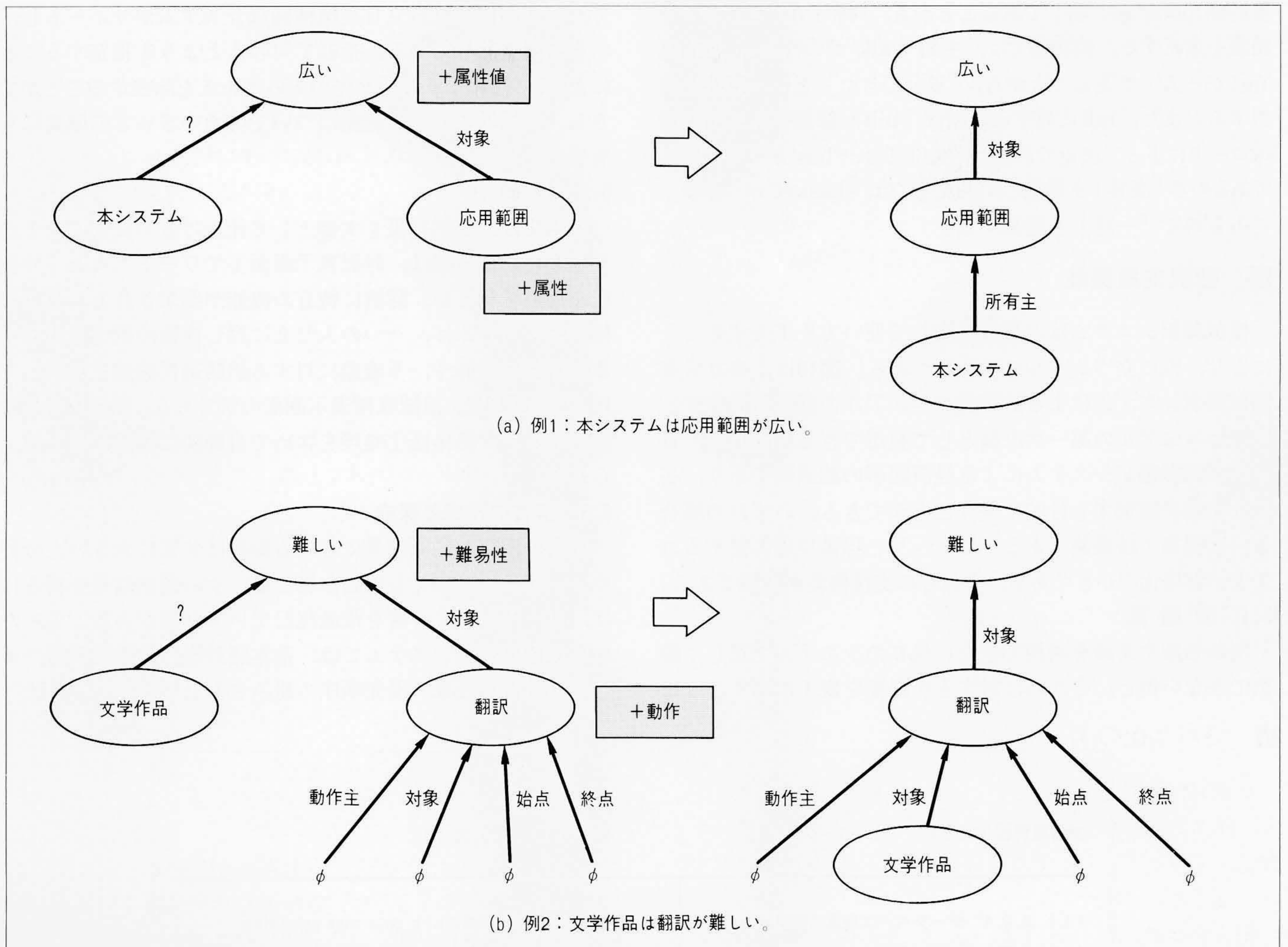


図8 語の意味的な性質に基づく深層格解析 構文的には関係のない「本システム」と「応用範囲」、あるいは「文学作品」と「翻訳」の間の意味的な関係を、語の意味的な性質に基づいて抽出する。

(1) 語義が異なる場合(主語・動詞・目的語の関係)

① The boy brought a book. 少年は本を持ってきた。

② An hour walk brought me here. 私は1時間歩いてここに来た。

(2) 訳語が異なる場合(主語・補語の関係)

③ The buses were all full. バスは満員だった。

④ The basket was full of oranges. バスケットはオレンジで一杯だった。

(a) 訳し分けの例

意味辞書				意味文法	
	内容	コード	訳		
boy	人(役割)	N211	少年	bring	S(人) & O(物) -> SがOを持ってくる
book	生産物	N533	本		S(行為) & O(人) -> OはSしてきた
walk	行為	NB	歩く		S(場所) & O(人) -> OはSを通ってきた
bus	乗物	N821	バス	full	ATR(乗り物) -> 満員
basket	容器	N811	バスケット		ATR(容器) -> 一杯

(b) 意味辞書と意味文法

注：略語説明 S(Subject), O(Object), ATR(Attribute)

図9 共起制約を利用した訳語選択 主語や目的語の意味コードを利用して、動詞や形容詞の訳語を選択する。

目的語‘book’が「物」であることから、「持ってくる」という語義を選択する。例文②では、主語‘walk’が「行為」、目的語‘me’が「人」であることから、「歩いてきた」という語義を選択する。また、fullに対する訳語は、fullが修飾する名詞の意味コードによって決定できる。例文③では、‘bus’が「乗り物」であるので「満員」を選択し、例文④では、‘basket’が「容器」であるので「一杯」を選択する。

6 翻訳支援機能

機械翻訳システムは、目的に応じて使い方を工夫することにより、役に立つレベルになりつつある。図10は、日立日英機械翻訳システムによる計算機マニュアルの翻訳例であり、英文マニュアルの第一次原稿として利用できる。図11は、日立英日機械翻訳システムによる経済記事の翻訳例であり、必要な記事を選別する目的に使うことができる。いずれの場合も、前編集・後編集、あるいはユーザー辞書作成に要するコストが実用化のかぎであり、それらの支援機能が重要である。

6.1 前編集

文脈知識や常識を利用できない現在のシステムで正しく翻訳できない例と、それらに対する前編集を表1に示す。こ

で、例2の括弧は日立日英機械翻訳システムがサポートしている前編集記号である。括弧で句のまとまりを指定することにより、「複雑な」の係り先のあいまいさを除去することができる。括弧は並列句の範囲についてのあいまいさの除去にも有効である。

6.2 後編集

後編集は、翻訳結果を文書として仕上げるのに欠くことができない。後編集は、対訳表示画面上でワードプロセッサ機能を用いて行うが、翻訳に特有の機能が要求される。日立機械翻訳システムは、一つの入力文に対し複数の訳文を出力する別解表示機能や、多義語に対する訳語選択機能を提供している¹¹⁾。図12は、訳語候補表示画面の例である。訳語を番号で選択すると、語尾変化処理を含めて自動的に訳語が置き換えられる。

6.3 辞書の作成と保守

辞書の内容が翻訳結果に与える影響は非常に大きい。分野やユーザーによっても用語が異なる。よい翻訳結果を得るには、ユーザー側で辞書を最適化していく必要がある。このため日立機械翻訳システムでは、基本語辞書、専門語辞書、ユーザー辞書を任意の優先順序で組み合わせて利用できるよう

文書識別名 (VOSO V) ページ (4)	段落 (1)	文 (1)
(1) 6.2.4 データベース作成の準備	(1) 6.2.4	Preparation for data base generation
(2) 必要なデータを蓄積し、データベースを作成する前に、ファイルを初期設定したり、物理ファイルや論理ファイルを定義したりしなければならない。	(2)	Before accumulating necessary data and generating a data base, files must be initialized, and physical files and logical files must be defined.
(3) (1) ディクショナリファイルとログファイルの生成	(3) (1)	Generation of dictionary files and log files
(4) RDBF ユティリティを用いて、ディクショナリファイルとログファイルを生成する。	(4)	Dictionary files and log files are generated by using RDBF utility.
(5) ディクショナリファイルはファイル、フィールドなどの定義情報やインデクスを格納するためのファイルであり、ログファイルは、障害が発生したときの回復に必要な情報を格納するためのファイルである。	(5)	The dictionary files are files to store definition information such as files and fields and indexes, and the log files are files to store information that is necessary for recovery at the time of an occurrence of a failure.
[E]		
PF : 1=ヘルプ 2=文校正 3=繰返し 4=文分割 5=文結合 9=上 10=下 11=終了 12=処理選択		

図10 日立日英機械翻訳システムによる計算機マニュアルの翻訳例 明解な日本語であれば、原文に忠実に理解容易な英文が得られる。

文書識別名 (SMPL1) ページ (1)		段落 (1)	文 (1)
(1) Personal income rose 0.5% in December, while consumer spending increased 1.2%.		(1)	個人所得は、12月に0.5%上がった、一方、消費者支出は1.2%増加した。
(2) Industrial output rose a strong 0.6% in December, providing further evidence that the economy accelerated by year-end.		(2)	鉱工業生産は、12月に0.6%上がった。これは、経済が年末まで加速したといういっそうの証拠を供給することになる。
(3) The better-than-expected reports show that the economy began 1987 on strong footing.		(3)	予想以上のレポートは、経済が強い足どりで1987年始まったということを表す。
(4) The President blamed the Fed for the length and severity of the recession.		(4)	大統領は連邦準備制度理事会を景気後退の長さ と 激しさについて非難した。
(5) The President's Council of Economic Advisors recommended some technical changes in Fed procedures for somewhat greater money supply growth.		(5)	大統領の経済諮問委員会は、幾分より大きなマネー・サプライの成長のために連邦準備制度理事会の手順のいくつかの技術的変更を勧告した。
[J]			
PF : 1=ヘルプ 2=文校正 3=繰返し 4=文分割 5=文結合 9=上 10=下 11=終了 12=処理選択			

図11 日立英日機械翻訳システムによる経済記事の翻訳例 実際の文書の翻訳には、専門語辞書の作成が重要である。

表1 前編集の例 文脈や非言語的知識を利用しないと正しく訳せない文もあるが、適切な前編集により訳せるようになる。

No.	問題点	原 文	前 編 集 後
1	品 詞	データを <u>読み込み</u> ファイルに格納する。 ⇒Data is stored in a reading file.	データを <u>読み込み</u> <u>,</u> ファイルに格納する。 ⇒Data is read and stored in a file.
2	係り受け	<u>複雑な</u> システムのメカニズム ⇒mechanism of a complicated system	複雑な <u>[</u> システムのメカニズム <u>]</u> ⇒complicated mechanism of a system
3	深 層 格	結果を出力する <u>装置</u> ⇒equipment which outputs the result	結果 <u>が</u> 出力 <u>される</u> 装置 ⇒equipment to which the result is output

にするとともに、次のような辞書作成・保守支援機能を提供している¹²⁾。

- (1) 辞書保守
画面誘導形の専用エディタにより、種々のコードを含む辞書レコードの内容を容易に更新できる。
- (2) 未知語検出
翻訳実行前に、辞書登録すべき語を一括抽出することができる。
- (3) 辞書データのソート出力
意味コードなど付与基準に揺れが生じやすいコードに対し、それをキーにソートしたリストを出力することにより、付与基準の統一を図ることができる。

7 結 言

計算機の出現以来、長年の夢であった機械翻訳もようやく実用化の第一歩を踏み出した。日立機械翻訳システムでは、意味処理により翻訳精度を向上するとともに、人間の援助を効果的に取り入れるための種々の支援機能を開発した。

日英翻訳では、語の表す概念の意味的な依存関係を表す「概念依存図式」を中間表現として、両言語の構文ギャップを吸

文書識別名 (CATALOG) ページ (1)		段落 (1)	文 (32)
機械翻訳 の 原理 を 簡単 に 説明する。		The principle of a machine translation is described briefly.	
辞書1	辞書2	辞書3	
1. 1 describe			
1. 2 illustrate			
1. 3 explain			
選択番号 []			
PF : 9 = 上 10 = 下 11 = 終了 送信 = 多義選択			
注 : プリンク表示			

図12 日立日英機械翻訳システムの訳語候補表示画面の例 訳語を番号で指定すれば、自動的に訳語が置き換えられる。

収する翻訳方式を開発した。概念依存図式は、言語独立性の高い意味表現であり、多言語翻訳への拡張性が高い方式である。英日翻訳では、英語と日本語のそれぞれの特質を考慮して設定した「疑似的句構造」と「格的句構造」の間で構文変換を行う翻訳方式を開発した。日英機械翻訳システム、英日機械翻訳システムとも、構文的あいまいさの解消や訳語の選択に意味情報を利用し、高品質の翻訳を可能にした。

文脈知識や常識などを用いなければ翻訳できないものも多く、これらに対しては長期的な基礎研究が必要である。しかし、現在のレベルでも、用途をうまく選び、人間が援助することによって実用化が可能である。今後、多様な言語現象に対応して辞書や文法を増補・改良するとともに、支援機能のいっそうの充実を図ることが必要である。

参考文献

1) 長尾：機械翻訳はどこまで可能か、岩波書店(1986)

2) 岡島, 外：英日機械翻訳におけるテキスト情報の利用について, 情報処理学会第33回全国大会論文集, pp.1767~1768(1986)

3) 梶, 外：日英機械翻訳システムATHENE/Nにおける文法記述言語, 情報処理学会第31回全国大会論文集, pp.1347~1348(1985)

4) 片桐, 外：英日機械翻訳システムにおける文法記述言語の検討, 情報処理学会第33回全国大会論文集, pp.1769~1770(1986)

5) 梶, 外：日英機械翻訳のための日本語文の依存構造解析, 情報処理学会第30回全国大会論文集, pp.1579~1580(1985)

6) 梶, 外：概念依存図式の変換による英文生成, 情報処理学会第27回全国大会論文集, pp.1085~1086(1983)

7) 梶, 外：概念依存図式からの英文生成, 情報処理学会第28回全国大会論文集, pp.907~908(1984)

8) 岡島, 外：自然言語処理における中間語表現と多義解消のしやすさとの関係, 情報処理学会, 自然言語処理技術シンポジウム(1983)

9) 山野, 外：英日機械翻訳システムATHENEにおける日本語文生成モデルについて, 情報処理学会第29回全国大会論文集, pp.1243~1244(1984)

10) 岡島, 外：機械翻訳の原理と実際, 日立評論, 65, 8, 579~584(昭58-8)

11) 臼井, 外：HICATS/JEのポストエディット支援環境, 情報処理学会第33回全国大会論文集, pp.1759~1760(1986)

12) 村上, 外：HICATS/JEの辞書作成支援環境, 情報処理学会第33回全国大会論文集, pp.1757~1758(1986)