論文

言語における話者の認識と多段翻訳方式

Speaker's Recognition and

Multi-Level-Translating Method

based on it

Speaker's Recognition and Multi-Level-Translating Method based on it

by

SATORU IKEHARA,

MASAHIRO MIYAZAKI,

SATOSHI SHIRAI and

YOSHIHIKO HAYASHI

(NTT Communications and Information Processing Laboratories)

〈Abstruct〉

Taking notice the process construction of a natural language, two theme, namely, ①the analysis about speaker's recognition both to the subject and the object, and ②simultaneous analysis about a syntax and an meaning, are pointed out to be fundamental in a natural language processing. Based on this result, Multi-Level-Translating-Method is proposed. This method is constructed by two sub-methods, such as the separation and unite method for subjective expressions and the multi-level transfer method for objective expressions. The former analyzes the subject

iveexpression, extracts subjective emotions or intentions, and combines them into a object language. The latter transfers the objective expression remained into a object language by three steps based on an abstruction level of a sentence structure. In these three steps, a special recognition structure, an individual recognition structure are extracted and transferred in order, and the expressions remained are transferred by general rules.

〈著者名〉

＊池　原　　悟　　：会員６９０２８９０

(Satoru IKEHARA)

宮　崎　正　弘　：会員７３００９１２

(Masahiro MIYAZAKI)

白　井　　諭　　：会員８１１３７４５

(Satoshi SHIRAI)

林　　良　彦　　：会員８３０５２１６

(Yoshihiko HAYASHI)

〈所　属〉

ＮＴＴ情報通信処理研究所

(NTT Communications and Information

Processing Laboratories)

〈連絡先〉〒238横須賀市武一丁目2356

NTT情報通信処理研究所

自然部 池 原　悟

Tel. 0468-59-2515

〈梗　概〉

言語過程説の立場から,「対象」,「認識」,「表現」の関係に着目して，①主体と客体に対する話者の認識の分析と，②統語構造と意味の統一的扱い，の２点が自然言語処理の重要な課題である事を示し，これらの課題を実現する機械翻訳の方式として「多段翻訳方式」を提案する。本方式は上記の課題に対応する２つの部分的な方式，即ち，主体的表現／客体的表現分離融合方式と多段変換方式から構成される。前者は原言語における主体的表現を解析し，話者の主観的感情や意志を抽出して目的言語に組み入れるものである。また後者は主体的表現情報抽出後の原文（客体的表現）を客体認識の構造的抽象性のレベルに応じて特殊的認識構造，個

別的認識構造，汎用的認識構造の３段階の構造的枠組みで捉え，それぞれに対応する３つの変換方式，即ち，慣用表現変換，意味的結合価パターン変換，汎用パターン変換によって目的言語に変換するものである。

本方式は目的言語への変換の過程ではもちろん，原言語の解析の過程においても文構造の持つ意味を掬い取るものであるため，言語解析の基本的課題である多義解釈においても優れた効果が期待できる。また，変換規則は相互独立性が高く，相互矛盾の検証範囲が極小化されるため，システムの成長が容易であると期待される。

１．まえがき

自然言語の研究は自然科学と異り，人間の知的産物である自然言語が研究対象であるため，人間の精神活動の捉え方の違いによって種々の言語の説明が行なわれてきた。例えばソシュールの構造言語学[(1)(2)(3)]では人間の精神活動を先天的観念実体ラングの中に位置づけ，社会的で有限な言語規範をその内容と説明したが，このような形式に着目した構造言語学や言語の形式と機能を組み合わせた従来文法[(4)]では，同形式異内容の言語現象の説明ができない。そこで，チョムスキー[(5)(6)]はもっと抽象的な性質を持つ構造を意味とすべきだとし，万人に共通の思考の能力を想定して深層構造を設定した。これは表現の内容に目を向けた点で

は評価されるが，内容を対象と切り離して考えたため，対象の反映論が欠落し，形式と内容を対立的に捉えた二元論的な説明(7)(8)となっている。+しかし，言語の形式（表現の構造）は対象のあり方とそれに対する話者の認識のあり方が反映したものであるため，形式と内容は相互に支え合う構造をもっている。従って，表層と深層を分け表現（表層）と離れたところに深層構造のような意味構造を仮定するのではなく，表現に結びつけられた対象と話者の認識の関係に意味を見る必要があると考えられる。

+ ﾁｮﾑｽｷｰは当初，標準理論（文献9）において深層構造を統語構造のみで説明しようとしたが，ｶｯﾂ等の批判（文献10）に会って同形式異内容の説明が困難となり，訂正拡大標準理論（文献11）では変形による意味の変化をも認めるに至った。このため，深層にあるとした意味が今度は表層にもあることになり，形式と内容の二元論から内容自体の二元論的性格をもつようになっている。

最近の機械翻訳の研究[12][13]においても生成変形文法の流れをくむものが多く,「言語間で内容は共通」とする立場から，言語に共通した意味構造を仮定した翻訳方式が志向されている。しかし，言語が「対象」,「認識」,「表現」からなる過程的構造[14][15]をもつことに着目すれば，言語に共通するのは「対象」だけであり，それに対する認識の仕方は個人で異ると同様，言語でも異ることが指摘できる。従って，言語に共通する深層構造を表現の意味として仮定する事には困難さがあり,+ 質の良い翻訳を実現するには原言語と目的言語の認識構造の違い[16][17][18][19]を考慮した翻訳方式

+ 生成変形文法の「深層構造=話者と読者の共通の認識」の設定を変えて「深層構造=対象のあり方」とすると，今度は話者の認識が欠落してしまい，言語の違いによる認識の枠組みの違いが扱われなくなるため，やはり問題は解決しない。

を考えることが必要と思われる。

本論文では言語過程説[14][15]の考えに従い，言語における話者の認識の枠組みの違いを考慮する立場から，①主体的表現と客体的表現で表される認識の内容，②文構造と意味の一体性，の２点の扱いが重要であることを示し，これらを考慮した多段翻訳方式を提案する。また，本方式を実現した日英翻訳実験システムを取り上げ，本方式による翻訳の過程を示す。

２．言語の過程的構造と翻訳

２．１　従来の翻訳方式の問題点

チョムスキーに始まる生成変形流の言語学説と時枝学説の言語の捉え方の違いを示すと図１のようになる。生成変形文法では言語を表層構造と深層構造の二者の

関係で捉え，深層構造を対象のあり方と独立して話者の心理内に存在するものとして説明し，これを意味と考えたのに対して，時枝学説では言語を対象と認識と表現の三者の関係で捉え，対象のあり方が話者の認識に反映している点を反映論で説明し，認識と表現の関係を言語規範で説明している点で本質的に異なっている。

従来の翻訳技術の研究では格文法など生成変形文法の流れをくむ考え方に基くものが多い。例えばピボット方式では，人間に共通する深層構造を仮定し，もしくは言語に共通する対象のあり方に着目して，これを言語に独立した中間言語で表現する事を仮定している[+]。また概念構造変換方式[(20)][(21)]では現実にはこのよ

うな中間言語の設計は困難であるとする立場から一歩進んで深層構造の言語依存を認め，言語依存の深層構造を中間言語として設定しているがこのような中間言語で意味を捉え切れない欠点がある。[++]いずれの場合も対象のあり方と話者の認識のあり方の関係への視点が無く，反映論が欠けている点で共通している。対象と認識の関係を考えるなら対象は認識に反映するが，それは機械的な反映ではなく，認識は対象に対して相対

+ 共通する最小ｾｯﾄをﾋﾟﾎﾞｯﾄとする考え(22)もあるがそれでは社会的（言語圏で異る）な発想の枠組みの違いを反映した自然言語を表すことはできない。

++例えば中間言語としては格構造が採用される事が多い。しかし，格文法では「どの言語においても述語の役割を果す語のすべてに敷えんできるような，数少い根元的な格概念が存在する」ことを前提とし，深層構造を格概念で説明しようとしたが，ﾌｨﾙﾓｱ自身も認めている通り（文献23），深層格の目録作り，個々の格の定義法など重要な点についての問題が解決されていない。どのような格目録を作成しても，それに該当しない概念が現れ，無理に分類すると意味の欠落を生じるため，格目録が無数生じる。

的独立性をもつことが指摘できる。すなわち，対象のあり方は共通していても人によって見方，感じ方が異り，それが表現に反映する。対象の見方，感じ方は人間の集団によっても差が生じる。この差が言語のもつ表現の枠組みの違いに影響している(24)(25)と考えることができる。従って，翻訳においても対象と認識を区分して捉え，同一の対象に対しても言語によって認識の枠組みが異ることに着目して方式を考える事が必要+ と考えられる。

そこで次節では話者の認識のあり方を見る立場から表

+ 本論文は従来の翻訳方式のもつ技術的役割まで否定するものではない。後に述べるように翻訳が原言語から目的言語への近似であることを考えるなら，格文法も近似法であり，ﾋﾟﾎﾞｯﾄ方式や概念構造変換方式も近似の為の技術と言える。しかし，より近似度の高い翻訳方式を追求するためには従来の方式の基盤となっている言語の捉え方に反省を加え，各方式のもつ限界を明らかにする必要があると考える。

現の枠組みについて考える。

２．２　話者の認識とその表現

(1)主体と客体の認識

話者が認識する対象世界は図２に示されるように，主体である話者自身とその他の対象である客体から構成される。話者は自分以外の客体と自分自身の双方のあり方を認識して表現に結びつける。時枝文法[14]によれば，日本語表現は以下の２種類の表現から構成される。

ｉ）主体的表現

話者の主観的感情や意志を直接表現するもので，日本語では助詞，助動詞などが用いられる。

ii）客体的表現

対象を概念化して捉えた表現で，日本語では名詞，動詞，形容詞などで表わされる。主体も概念化すれば客体的に表現される。

このような関係はインド・ヨーロッパ系言語においてもポールロワイヤル[26]によって指摘されている。

(2)認識と表現の結びつき

話者の認識する客体は大きく分けて実体，属性，関係の３つの概念で捉えることができる。これらはいずれも構造を持っており，その構造が認識を通して表現に反映される。実体は全体，部分などの階層的な構造をもち，属性は実体との結びつきの構造をもつ。関係も実体間の関係，属性間の関係，関係同士の関係などでそれぞれの構造をもち，実体，属性，関係が複合さ

れて総合的な構造が形成される。この構造を話者が認識するとき，話者の見方，考え方によって種々の捉え方が生まれる。各言語はこれらの話者の見方，考え方を表現する枠組みをもっている。実体を個別的，具体的に捉える時は固有名詞を用いたり，最も抽象的に捉える時は抽象名詞（「もの」，「こと」：形式名詞とも言う）を用いたりする。また，「ＡがＢからＣへ移動する。」と言う動的属性を話者の視点に合せて「Ｃに行く」，「Ｃに来る」などと表現する枠組みがある。

　このように，言語表現には対象のもつ構造が反映されているが，それは話者の認識を通して反映されるものであり，対象が直接表現に結びついているのではない。言語の違いは話者の認識の枠組みの違いであるか

ら，翻訳では原言語の枠組みの中で捉えられた主体と客体のあり方を目的言語の枠組みの中で捉えなおして表現する事が必要と考えられる。

２．３　意味の捉え方

(1)意味の解釈について

従来，言語学においては言語の意味については，カッツ[10]，グライス[18]，サアル[27]，ソシュール[1]ワイスマン[28]など数多くの説が出されているが，定説と言えるものはない。翻訳の研究においても意味の定義をしないまま意味解析や意味理解の技術を論じている場合が多い。ここでは言語過程説の立場から意味を定義するが，時枝説の一部を言語認識論の立場から改訂した三浦説[15]に従い以下の通り定義する。す

なわち，言語表現には話者の認識を通して捉えられた主体と客体のあり方が結びつけられている点に着目して，表現の意味はその表現に結びつけられた対象と認識の関係であると考える。従って，表現があって意味（関係）が生じる。辞書に記載された語の意味（語義）は厳密には意味ではなく，言語規範（広義の文法）

と解される。それが使われたとき対象と認識への結びつきが生じ，始めて意味（関係）をもつことになる。

このように考えると，意味解析は表現に話者のどんな認識が結びつけられているかを調べる事，およびそれを通じて対象がどんな存在であり，話者がそれをどう見ているかを解析することであると言うことができる。

(2)統語構造のもつ意味

話者は語の約束，句や節の約束などを用いて自己の認識を立体化して表現するが，この立体化は意味に支えられた構造化に関する文法規則に基いて行われる。すなわち，対象のあり方が話者の認識に反映し，それが構文に反映する。これは構文が対象と認識に結びついていることであり，統語構造が意味の一部であることを意味する。生成変形文法のように構造と意味が表層構造と深層構造のように対置されるものではなく，意味は表現（表層）と認識，対象の結びつきであり，表層の構造は意味の一部と言う事ができる。従って，厳密に言えば，意味を変えない変形はあり得ない事になり，通常言語処理で行われる変形は近似的な表現へ

の変形である。従来，「翻訳とはイディオムを訳すことなり」とする考え方[31]もあるが，これは統語構造と意味の一体性を述べたものと言える。統語構造のもつ意味を考えないで部分の意味から全体の意味を合成しようとする要素合成方式（原子論的方法）では，構造のもつ意味の欠落を防ぐことは困難と考えられる。+

(3)構造の意味による多義の解決について

読者が話者の認識を追体験する時の手がかりとしては２通りの知識が用いられる。１つは言語規範（文法）に関する知識であり，もう一つは話者が対象とする世界に関する知識（言語外知識：常識など）である。翻

+ 長尾等は語族の異る言語間翻訳での構造変換の必要性を指摘し，Pre-Transfer-Loopを設けている（文献

32,33）。これも構造のもつ意味の重要性に着目したものと言えるが，変換基本部が深層構造を介した構造となっている点に§2.1で指摘した問題があるため，本論文では深層構造と異る新しい意味理解の観点から構造のもつ意味を取り上げ問題の解決を図る。

訳の対象とする文で扱われる世界の知識を網羅的に取り扱うのは困難であるので，ここではまず言語知識について考える。言語知識としての言語規範を考えると図３に示されるように一語の語義についても種々の約束があり，話者がそのうちのどの約束を用いたのかを判定する必要がある。そこで構造のもつ意味を考えるならば，語と語の結びつきの中に，それぞれの語がどの約束（語義など）で使用されたかを知る情報が含まれている。句や節についても同様，それらを含むもう一段上の構造の中に多義を解決する手がかりがある。

従って，語や句や節の多義を解決するにはそれらを含

む上位の構造の中で捉えることが必要と言える。すなわち，意味と構造を一体化して扱うことには多義を絞る上でも効果が期待できる。

３．多段翻訳方式の提案

文法書と辞書で代表される言語知識の範囲で，一文単位に独立して翻訳の可能な文を一文完結型の文と呼び，この範囲$^{+}$ の実用文を対象とする日英翻訳方式として，図４に示す考え方に従い，多段翻訳方式を提案する。本方式は以下に示す２つの部分的な方式から構成される。

３．１　主体的表現／客体的表現分離融合方式

日本語は膠着言語の特徴として，主体的表現に助詞，助動詞などの単語が用いられるのに対して，英語は屈

+ 翻訳においては一般常識や文脈解析の必要性が言われているが，現実には日本語の実用文中，一文完結型とみられる文が約９割を占めるため本論文では一文完結型の文を対象に考える。

析言語の特徴として語の屈析（語形変化）を伴って主体の直接的表出が行われることが多い。従って，日本語の主体的表現の語と英語のそれとは直接的に対応しない事が多く，遂語的翻訳は困難である。そこで，日本文に表わされた話者の認識を解析するに当って，話者の主観的感情や意志を分類し，与えられた日本文の主体的表現の部分が，どのような感情や意志を表わしているかを判断する。この過程でもとの日本文は平文に変換される。平文は主体的表現を抽出したあとの客体的表現[+] である。この客体的表現は次節で述べる多

+ 主体的表現情報は目的言語の主体的表現に訳出できる範囲の分解能と精度で分類すれば良い。このため，その他の情報は客体的表現の構造を示す情報として

保存し，多段変換の枠組みを経由して目的言語に変換するものとする。具体的には変換対象となる客体的表現には，文要素の指示詞とも言える格助詞相当の語が含まれることになる。

段変換方式によって英語の客体的表現（英語素文）となる。そのあと，既に抽出されている話者の感情や意志が英語素文に対して組み込まれる。この組み込みでは，助動詞や前置詞の挿入の他，種々の語の屈析（変形）が行われる。このようにして，本方式では日本文中から分離された主体的表現情報が英文生成の段階で英文に融合される。

３．２　客体的表現の構造の抽象化と多段変換方式

主体的表現情報を抽出したあとの日本文（客体的表現）には客体のあり方が話者の見方を通して表わされている。客体に対する話者の認識も種々の構造をもち

それが客体的表現の構造に反映している。表現の変形が意味を変えること，構造が意味をもち意味と構造を一体化した扱いが必要なことを考えれば，日本語のすべての表現に対応した英語表現をもちそれらを一対一に対応させれば論理的には翻訳の近似度は向上するが，そのためには無限の数の表現の収録が必要となり工学的には困難である。そこでこの矛盾を工学的調和させるためここでは構造と意味の結びつきの強弱に着目して構造を以下の３段階に抽象化し，各レベルに応じた構造の変換方式を考える。

(1)特殊的認識構造（慣用表現変換方式）

ことわざ，慣用句，熟語など複数の語から構成される表現で一語一語の意味からだけでは表現の意味が説

明できないもの，言語特有の固定的いいまわしを言う。

通常日本語の慣用表現と呼ばれるもののほか，日本語の二語以上の表現が英語の一語に対応するもの，訳出時に英語の慣用表現に対応させるのが望ましい表現を含む。このような特殊的認識構造は要素合成方式では翻訳が特に困難であり，日本語と英語の構造を対応させた慣用表現変換規則によって文構造のもつ意味をもそっくりすくい取った変換を行う。特殊的認識構造では特定の複数の単語が用いられるため，このパターン対辞書は該当する単語の組に対してエントリーが設定され，日本語表現の中にその組み合せが現れたときは適用条件の許す限り優先的に適用される。

(2)個別的認識構造（意味的結合価パターン変換方式）

特殊的認識構造よりも若干汎用的な認識構造を言う。

具体的には特殊的認識構造が二語以上の語の字面が固定されるものであるのに対して，個別的認識構造は二語以上の組み合された表現のうち一語の字面（見出し語）が固定され，他の語はその語の単語意味属性に制約をもつものを言う。用言の字面を固定した場合はその用言と結合する文節のもつ助詞の字面と名詞の意味属性が規定される。このような個別的認識構造を捉える枠組としては，格文法の方法や結合価文法の方法が考えられるが，用言と結合する文節が深層格として抽象化されず，結合の手として助詞が明示される点が格文法[25]と異る結合価文法[34]の特徴であり，文構造を個別的に捉えるのに適している。

ここでは結合価文法に意味的制約を加えた意味的結合価文法を用いる。意味的結合価パターンでは，結合価文法の方法に比べて各文節に対する精密で排他的な意味属性体系に支えられた個別的制約条件をもつことによって，深層的格カテゴリーでは分類できない文構造の持つ意味をも英語に対応させることを可能としている。[+] 変換用辞書ではこのような日本語の個別的認識構造に対応する英語の認識構造が対にして登録されるが，英語側の記述では日本語の見出し語に対する英語訳語のほか文要素の語順，それにつく前置詞などが指定されるため，日本文の構造のもつ意味を英語に写し取ることができる。パターンの作成に当たっては見出

+ ふり返ってみれば生成変形文法では同形式異内容を説明するため深層構造を仮定し，統語構造と意味を

対立的に捉えたが，言語認識論からみれば，統語構造はより抽象化されているとは言え，対象世界の実体と属性の関係のことであり，内容の一部であるから，改めて深層と言う意味を仮定する必要性は少い。翻訳では話者の認識構造の抽象化の程度に合せて目的言語の表現が対応付けられれば良いことになる。

し語の語義毎にパターンを作成する事が必要であり，

またそのうちどのパターンを適用するかを一意に決定

するためには単語意味属性体系を十分精密に決定する

事が必要であるが，見出し語は相互に独立しており，

見出し語毎に複数のパターンが対応する構造であるた

め，変換規則相互の無矛盾性のチェックは原則として

同一見出し語内のパターン相互間に絞られるため，変

換規則を容易に成長させることができる。

(3)汎用的認識構造（汎用パターン変換方式）

上述の２つの方式では特定の語もしくはその組み合

せに着目して表現の構造をパターン化して捉えている

のに対して，ここでは語の字面は特定せずある文法的もしくは意味的カテゴリーの語のグループ毎にパターンを対応させることを考える。例えば動詞を瞬間動詞，継続動詞などに分けたり，「だ文」（「ＡはＢだ」の形式の文）をＡ＝Ｂの場合とＡ≠Ｂの場合に分けるなどにより，それぞれのパターンに応じた変換規則を設ける。このような方法は語の字面を固定しない意味で前述の２方式に比べてより抽象的で汎用的な方法と言うことができる。

以上の３段階からなる多段変換方式においては，特殊性の高い表現パターン程近似度の良い訳文が生成されるから，慣用表現変換，意味的結合価パターン変換，汎用パターン変換の順に優先して変換が行われる。パ

ターン辞書の不備によって該当する慣用表現パターンや意味的結合価パターンの無いときは汎用パターンが用いられるため，翻訳品質は低下するが，パターン辞書が完備するにつれて翻訳の品質の向上が期待される。

３．３　多段翻訳方式の構成

多段翻訳方式は図５に示すように，主体的表現／客体的表現分離融合方式と客体的表現に対する多段変換方式の２つの部分的方式を組み合わせた翻訳方式である。

本方式では図６に示す人手翻訳のプロセスと類似の手順で翻訳が行なわれる。すなわち，人手翻訳では翻訳者は与えられた表現に結びつけられた話者の認識を日本語規範に照し合せて追体験し，話者の目から見た

客体のあり方とそれに対する話者の直接的感情や意志を知る。これに対して本翻訳方式では客体のあり方は日本語素文（客体的表現），話者の感情や意志は主体的表現情報として分離される。次に，人手翻訳では客体のあり方の英語の枠組みの中での捉えなおしが行なわれ，同時に主体のあり方の融合が行なわれるが，本方式では客体的表現のもつ意味が３段階の変換からなるパターン変換によって英語の枠組みの中に写し取られ，最後にそこで得られた英語素文と主体的表現情報が融合されて，目的の英文が生成される。

３．４　多段翻訳本式の特徴

多段翻訳方式の狙いが，話者の認識に焦点を合わせ，主体と客体のあり方を解析し，文構造のもつ意味を失

わないように英語に変換することであることは既に述べた。ここではその他の特徴を示す。

(1)日本文等価的変換の必要性

特殊的認識，個別的認識構造，汎用的認識構造の順に，意味が正確に捉えられることを考えれば，翻訳の品質を上げるには慣用表現辞書と意味的結合価パターン辞書を拡充し，中でも慣用表現の適合率を上げれば良いことになる。しかし使用頻度の少いパターンをむやみに増やすことは不要な多義を発生させる点と処理の負荷を増す点からも必ずしも欲ましい事ではない。そこで英語の表現の選択において訳し分けの必要な範囲で日本語のパターンを縮退させパターン数を絞ることが大切と考えられる。このようにした場合，特に意

味的結合価パターンの適合率を上げ，汎用パターンへの流れを減少させるためには日本語内での以下の表現の縮退や書きかえが有効と考えられる。

第１は漢字表記，かな表記や送りがななどのゆらぎを取るため，システムの標準表記に合せて入力文を変換するものである。第２は主体的表現情報を抽出したあとの客体的表現をパターン辞書の登録情報を比較し易くするもので，特に複数の助詞等を組み合せた連語的表現など英文の表現を変えない範囲で，代表的な助詞的表現に縮退させるものである。第３は英語を意識した日本語内の表現書き換え$^{+}$ である。

(2)解析と変換の融合

本方式では日本文解析は慣用表現変換規則，意味的

結合価パターン変換規則の適用可能な構造を発見する

+「しか～ない」→「だけある」,「バスに乗って行く」→
「バスで行く」などの書換えが対象となる。

こと，もしくは抽出することを目標に進められる。従って，解析処理の中でこれらの辞書を参照し，適合するパターンはすべて取り出され，それを用いた解釈が実行される。この解釈においては，適合したパターンのもつ文要素は分解されず，一まとまりの構造体として扱かわれる。日本語パターンは英語パターンと対になっており，適用される日本語パターンの決定は即ち生成する英語パターンの決定を意味する。従って本方式では，日本文の解析処理と日英変換処理が融合したプログラム構造となる。

以上から，本方式は解析，変換の融合した融合方式

[35]もしくは人工知能型翻訳方式[36]と類似した方式であると言うことができる。

(3)多義解消効果

解析の多義は求める分解能の割に使用する情報の少いことから生じる。文法的情報だけでは翻訳処理で生じる多義の解消は困難であり、従来意味的情報として単語の意味属性を２単語間の関係解析として用いる方法が考えられている。これに対して本方式では単語の意味と文構造のもつ意味とが一体化された種々のパターンを持っており，これが解析の段階で使用されることにさらにより多くの多義が解消できるものと期待できる。文要素間の結びつきがパターンの中に示され，複数の文要素が一体化して扱われるため，係り受け関

係の解析では複数の要素間の関係が同時に決定され，

訳語の選択ではパターンから直接訳語が与えられたり，

排他的な意味的属性の制約から訳語が決まるなど不要

な多義の増大が防止される。重文，複文のような用言

間の文要素の取り合いにおいても，パターンの構成条

件から文要素の係り先を絞り込める効果が期待される。

(4)変換規則の相互独立性とチューンアップの容易性

慣用表現変換規則および意味的結合価パターン辞書

はいずれも万単位の規則を持つ事になるが，適用範囲

が特定されている。少くとも異る見出し語のパターン

は相互独立と考えて良いため，パターン内の相互依存

性は小さい。+ 即ち，一部の規則の変更が他の規則と

+ 慣用表現規則の充実に伴い意味的結合価ﾊﾟﾀｰﾝの一部
が不要となる場合や単語意味属性体系の一部変更に

伴い，異なる見出し語間のﾊﾟﾀｰﾝにまたがってﾁｪｯｸが必要となる場合もあるが，いずれも影響範囲は容易に(機械的に)特定できる。前者では不要となったﾊﾟﾀｰﾝを残したままでも誤動作の心配はない。

の矛盾を発生させる可能性は同一の見出し語をもつパターンに限定されるため，変換辞書のチューンアップが容易となる。従って，現存する文章の翻訳実験によって不良パターンの改良，不足パターンの追加が比較的容易に行える。

４．日英翻訳実験システムでの適用例

４．１　処理の概要

入力された日本文は図７に示すように，形態素解析のあと文要素間の係り受け関係の解析が行われる。そのうち用言の間の関係の認定の結果に基いて単位文が抽出され，さらにその中から単文が抽出される。単位

文は文を述語中心の木構造に展開したとき，最上位のレベルでの述語が単一の文であり，レベルの異る埋め込み文などを含むこともある。これに対して単文は単位文を単一の用言しか持たないレベルに分解したものを言う。これらの分解に当っては後の文の組み立て合成に備えて，用言間の接続関係情報の保存が行われる。単文の抽出後はこれが解析の単位となる。単文は述部に現われる主体的表現の解析によって様相，時制などの主体的表現情報などが抽出されたあと，平文に変換される。平文は単文から主体的表現情報を抜き取ったあとの文で，日英パターン対変換の単位となる。以上の処理は係り受け結果に多義の残るときは原則としてすべての解析候補に対して実施され，平文レベルでの

パターン対による解釈の結果に基いて最終的に一つの解釈が決定される。解釈決定時には日英変換で使用される文型パターンは決まっているため，あとはそれを適用し，英語の客体的表現である素文を生成し，様相，時制，接続などの情報を付加していけばよい。

４．２　主体的表現の情報抽出

日本語と英語の捉え方の違いによって，日本語の主体的表現が必ずしも英語の主体的表現に訳出されるとは言えない。日本語の主体的表現で表わされる情報を英語から分類すれば，法，時制，相，態，などになるが，英語表現からみれば前２者は主体的表現であり，後２者は客体的表現である。そこで，ここでは日本語の主体的表現の内容を英語側の分類に合わせて解析す

る。まず，英語では時制[+]については相と組み合せた表現が多いため，時制＆相状態に関する解析として扱い，法は態と合せて様相解析[++]として扱う。助詞のうち副助詞や終助詞などの持つ情報も様相に分類して抽出する。次に接続を考えるとこれを表す語には主体的表現と客体的表現としての用法があるが現段階では機械的判定が困難である。また事象間の時間的関係の認識の枠組みが日／英言語間で直接対応しないなどのため，ここでは接続関係を主体的表現の枠組でまとめて扱う。

(1)時制と相状態属性の決定

まず動詞の分類に従って着目する文の相（開始相，継続相など）を決める。この相と文中の副詞句，助動

+ 日本語は必ずしも時制の枠組を持つとは言えないが，主体的表現を手がかりに英語に必要な時制を推定す

る。

++客体的表現も様相を表わすものがあるが，それらは後に示す平文において変換に必要なパターンをもつためここでは主体的表現で表わされる様相を扱う。

詞相当語の組み合せによって相状態（継続直後状態，完結直後状態など１７種）を定める。その結果相状態と時制+（現在，過去，未来）をマトリックス状に組み合せた５１種のカテゴリーのうち着目する文のカテゴリーが決まる。このカテゴリーは日英の対応表によって英語側のカテゴリーのいずれかに対応させられ，英文生成で使用される。英文生成では節の内部での時制相状態生成後，節と節の関係（時の前後関係や節の時制関係など約１０種）に着目して，時制の一致など，組み換えが行なわれる。

(2)様相属性の決定

+ ここでは話者と対象間の時間的関係の直接表現を時

制として扱う。対象同士の時間的関係など英語特有の現象は英文生成の段階で扱う。

形態素解析による助動詞の文法的属性解析結果と副詞等組み合せにより受益，被役，許可，希望など約８０種に分類された様相属性のいずれの解釈をとるかを決め，これを英語の様相に対応させる。

(3)接続属性の決定

接続詞，助詞，副詞，形式名詞などの組み合せから，順接，同時，原因，理由，比況など約４０種に分類された接続属性のいずれをもつかを決める。英文生成ではこの情報に基いて接続詞の決定などが行われる。

４．３　平文のパターン照合

平文は図８に示すように用言を見出し語とする２種のパターン辞書（慣用表現変換辞書，意味的結合価パ

ターン変換辞書）と順に照合され，適切なパターンの無い時は汎用パターン変換規則の適用を受ける。変換の中心となる意味的結合価パターン辞書には１万件を超えるパターンが登録されているが，パターン間の適用基準の排他性を保つため，パターン内の文要素と文中の単語を対応づける単語意味属性として図９に示すような約２８００種の属性名からなる単語意味属性体系を作成して使用している。この単語意味属性は用言のもつ文パターンのほか，図10に示すような名詞句の訳し分けにも用いられる。

４．４　英文の生成

平文に対する変換結果を接続属性に従って組み立て，様相，時制等の情報を付加して英文とするが，より英

語らしい表現とするため，さらに以下の英文調整を行う。

①縮約処理…単純な形容詞埋め込み，並列要素の縮約など

②外置変形…it等の形式主語／目的語による変形

③受身変形＆補間…主語無し文等

④位置変形…副詞句など

⑤冠詞付与　⑥形態素調整など

以上，この実験システムの翻訳処理の流れを例によって示すと図１１の通りである。

５．むすび

従来言語の意味を深層構造にあるとする生成変形文法の立場から要素合成方式を基本とする種々の機械翻

訳方式が提案されて来たが，これらの方式のもつ基本的問題を指摘し，言語過程説の立場から，①主体と客体に関する話者の認識を分けて扱うこと，および②統語的構造のもつ意味を掬いとることの重要性を述べ，これらの課題を実現する新らしい日英翻訳方式として多段翻訳方式を提案した。また本方式の実現を目指した日英翻訳実験システムの概要を示した。

多段翻訳方式は①主体的表現／客体的表現の分離融合方式と②客体的表現に対する３段階の多段変換方式からなる翻訳方式である。統語構造と意味を一体化して扱い，品質の良い（近似精度の高い）翻訳を実現するには理想的に言えば，翻訳対象となるすべての文の対訳集をもてばよいが，自然言語の性質上それは不可

能である。この理想と現実の対立を言語認識論の立場から，①文の構造を認識の抽象化のレベルに応じてパターンとして整理すること，②整理されたパターンの適合率を上げるため表現の中から主体と客体の情報を分離することの２点で工学的に調和させたことが本論文の主たる結果である。

本論文では一文完結型の日本文を対象とする日英翻訳の方式として多段翻訳本式を提案したが本方式は英日翻訳等，他の翻訳へも適用できる。今後は文脈依存型および言語外知識依存型の翻訳への本方式の拡張について研究を進める予定である。

〈参考文献〉

（１）E.F.K.ケルナー：ソシュールの言語論，山中桂一訳，大修館(1982)

（２）J.B.ファージュ「構造主義入門」訳・加藤，１９８２，大修館

（３）ジュリオ．Ｃ．レプスキー「構造主義の言語学」訳・菅田，１９７５，大修館

（４）岩波講座「日本語６（文法Ｉ），７（文法Ⅱ）」　１９７７，岩波書店

（５）カール・オットー・アーペル「チョムスキーの言語理論と現代哲学」１９７６，大修館

（６）梶田優「変形文法理論の軌跡」１９７６，大修館

（７）チョムスキー：デカルト派言語学，訳・川本，みすず書房

（８）Chomsky,N:Language and Mind,New York(1968)

（９）Chomsky,N:Aspects of Theory of Syntax,MIT Press, Cambridge, Mass(1965)

(10) J.J.カッツ：言語と哲学，西山訳，大修館(1971)

(11) Chomsky,:Condition on Transformations, in Anderson and Kiparsky, PP.232-286(1973)

(12) 内田裕士，：言語に依存しない概念構造を中間言語表現の基本とし，常識を使う多言語向き翻訳システム」日経エレクトロニクス，1984,12/17,

PP.221-240

(13) 林木一至，「知識ベースと，言語に独立の中間表現を用いた日英機械翻訳システム」日経エレクトロニクス，1984,12/17,PP.195-220

(14) 時枝誠記：国語学原論，岩波書店（昭１６）

(15) 三浦つとむ：認識と言語の理論Ⅰ～Ⅲ，勁草書房(1967)

(16) 森田良行「日本語の発想」，Ｓ５６，冬樹社

(17) 安西徹雄：英語の発想，講談社現代新書（昭58)

(18) 三浦つとむ編：現代言語批判，勁草書房(1981)

(19) 宮下真二：英語はどう研究されてきたか，季節社（1980)

(20) 内田，他：日英機械翻訳システムＡＴＬＡＳ／Ｕ，情処自然言語処理研究会，№29-3,1982

(21) 山野，岡島：英日機械翻訳システム ATHENE における日本文生成モデルについて，情処29回 全大，4N-11(1984)

(22) 特別座談会：自動翻訳システムの開発の方向をめぐって，The English Journal 別冊翻訳辞典，PP.34-42，ｱﾙｸk.k.(1985)

(23) チャールズＪ．フィルモア：格文法の原理，三省堂(1975)

(24) 柳父章：翻訳学問批判，日本語翻訳家育成センター（１９８３）

(25) 東田千秋：直訳と言う名の誤訳，南雲堂（昭5

6)

(26) Ｃ．ﾗﾝｽﾛｰ．Ａ．ｱﾙﾉ：ポール・ロワイヤル文法，南館・訳，大修館(1972)

(27) 宮下真二：サアルの言語論，文献(11)に収録

(28) Ｆ．ワイスマン：言語哲学の原理，訳・楠瀬，大修館（１９７７）

(29) Shank,R.C.:Conceptual Information Processing, North-Holland (1975)

(30) 長尾真：言語工学，昭晃堂（昭58）

(31) 中村保男：翻訳はどこまで可能か，ジャパンタイムズ（昭58）

(32) 長尾，辻井，中村，坂本，鳥海，佐藤：科学技術庁機械翻訳プロジェクトの概要，情報処理，Vol.26,№10,PP.1203-1213(1985)

(33) 長尾，辻井：機械翻訳における訳語選択と構造変換過程，情報処理，Vol.26,№11,PP31261-1270 (1985)

(34) 石綿敏雄：文法と意味Ｉ，第２章結合価から見た日本文法，朝倉書店（１９８３）

(35) 田中穂積：解析から合成までを融合した英日機械翻訳システム，日経エレクトロニクス，1983, 8/29,PP.275-293

(36) Wilks,Y. : An Artifitisl Intelligence Approach to Machine Translation, Computer Model of Thought and Language, edited by Schank, R.C. and Colly, K.M.,Freeman and

Company（１９７３）