

## 自然言語処理における意味解析と意味理解

池原 悟

NTT 情報通信網研究所

最近、自然言語処理においては、意味処理、意味解析、意味理解などに関する研究が盛んであるが、意味そのものについて、考察や議論を提起した研究はまれである。自然言語処理は社会的産物である自然言語が研究対象であるため、従来の自然科学と異なる困難さがあり、統一的な見解が得られない状況にあるが、意味処理の研究を促進するには、常識的な感覚に頼るだけでなく、言語表現の意味について言語処理の観点からあらためて考察を加え、定義を明確にして研究することが望まれる。

本稿では、言語過程説の立場から、従来の言語哲学の分野での議論を振り返り、言語表現の意味とその処理について考察する。具体的には、言語表現には、対象の姿とそれに対する話者の認識が対応づけられていることに着目して、「対象」と「認識」「表現」の3者の関係を意味と考える。そして「意味処理」を、表現に使用された言語上の約束を特定するための「意味解析」と、言語表現と対象世界の対応付けを行う「意味理解」の二つの過程に分けることを提案する。また、この内の「意味解析」の例として、日英機械翻訳システムALT-J/Eの翻訳方式と意味辞書の関係について紹介する。

## Meaning Analysis and Meaning Understanding in Natural Language Processing

Satoru Ikehara

NTT Network Information Systems Laboratories

From the view point of Constructive Process Theory, traditional meaning theories of philology are reviewed and new idea for meaning processing is propounded. Thinking the fact that shapes of objects and speakers' recognitions are combined to Language expressions, relations between "Object", "Speaker's recognition" and "Expression" are considered as meanings. And meaning processing is segregated into "meaning analysis" that is the process which designates the rules or conventions used in linguistic expression and "meaning understanding" that is the process of identifying the recognition of the speaker and the status of the object that is tied in with the expression.

The meaning processing in machine translation system ALT-J/E is shown as for an example.

## 1. はじめに

1960年代の自然言語処理においては、チョムスキーの生成文法<sup>(1,2)</sup>など、統語論を中心とする文法がクローズアップされ、機械翻訳などへの応用が期待されたが、計算機処理にとって現実の言語は想像以上に難解な対象であることが分かる<sup>(3)</sup>につれて、研究が低調となっていた。しかし、知識をプログラム化した例<sup>(4,5)</sup>など、人工知能グループからの刺激もあって、1970年代後半には研究は再び活性化してきた。従来の統語論に加えて、ここ10年、意味論に関する研究が盛んになり、意味解析、意味理解、言語理解の研究や意味処理のための辞書の研究<sup>(6,7,8)</sup>など、様々な角度から意味に関する多くの研究が行われるようになった。しかし、意味論と言えど、意味そのものについて考察を加えたり、議論を提起した例は見当たらない。

言語の研究は自然科学の研究と異なり、人間の精神的活動の産物である言語が研究の対象であるため、その見方によって様々な説が生まれる。その中でも「意味の意味」については哲学的にも定説のない状況にあるが、意味処理を目指すには、対象とする「意味」の明確化は避けて通れない課題と考えられる。

本稿では、従来の意味論を振り返り、それらの問題点を考察すると共に、注目すべき言語理論として言語過程説(時枝文法)<sup>(9)</sup>を取り上げ、三浦つとむ<sup>(10,11)</sup>によって提案された「関係意味論」を紹介する。次に、この考えに基づき、「意味処理」を、表現に使用された言語上の約束を特定するための「意味解析」と、言語表現と対象世界の対応づけを行う「意味理解」の二つの過程に分けて扱うことを提案する。また、このうちの「意味解析」の例として、日英機械翻訳システムALT-J/E<sup>(12,13)</sup>の翻訳方式と意味辞書との関係、及び意味解析によって実現された新しい翻訳機能の一部を紹介する。

## 2. 言語表現の意味

### 2.1 言語の過程的構造について

自然言語処理では、チョムスキーに始まる生成文法の考えが理論的背景であると考えられて来た。この理論では、表層構造のほかに深層構造の存在を仮定して、これを表層構造の意味とし、両者の関係で言語をモデル化した点に特徴がある。その後、GB理論やHPSGなど<sup>(14,15,16)</sup>多くの改良された理論が提案されている。これらの理論では、必ずしも深層構造だけで意味を説明しているわけではないが、深層構造の存在を仮定する点は変わらない。

これに対して、本居直長の文法を継承発展させた言語過程説(時枝文法)によれば、言語は「対象」とそれを見て話す人の「認識」及び「表現」の関係で捉えられる。これは、言語も自然と同様、原因と結果が多重に織りなす過程の複合体とみなすもので、言語においては「対象」が原因となって話者の「認識」と言う結果が生まれ、その結果がまた原因となって「表現」と言う結果が生まれると考えるものである。

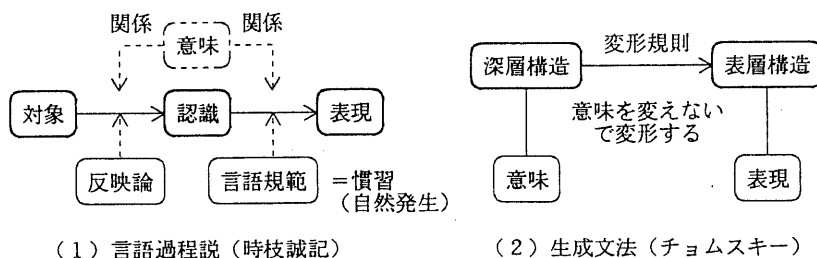


図1. 言語モデルの比較

両者の違いをモデル的に示すと図1となる。生成文法では、深層構造なる意味から、意味を変えない変形によって表層構造が生成される。これに対して、言語過程説では、「対象」が話者の「認識」に反映される過程は「反映論」によって説明され、「認識」が「表現」に対応づけられる(変換ではない)過程は言語規範(広い意味での「文法論」)によって説明される。後者の過程で使用される文法は、自然発生的に成長した社会的な約束であり、慣習の一つである。「認識」と「表現」が同一言語社会に共通する慣習に支えられていることによって、表現から話者の認識が追体験できると説明する点に特徴がある。

## 2.2 従来の意味論の分類

言語表現の扱う「対象」には、実在する対象と実在しない観念的な対象があるが、いずれも話者が取り上げたものである。話者はこれを自己の感性の中で、自分の見方、捉え方で「認識」し、それを言語の約束（言語規範）に従って超感性的な言語表現に結びつける。逆に、「聞き手」は話者と共通した言語規範を手がかりに表現から話者の認識のあり方を見、その認識を通して対象のあり方を知る。このような言語の過程的構造からみれば、言語を構成する実体は「対象」、「話者」、「表現」および「聞き手」の4種である。

したがって、「意味」をある実体概念で捉えようとする限り「言語表現の意味」は「対象」、「（話者の）認識」、「表現」および「（聞き手の）理解」の4つの実体概念のいずれか、およびそれらの折衷案以外にはありえないことになる\*1。そこで、多少大胆ではあるが、この考えに従って、従来の意味論を図2のように分類する。

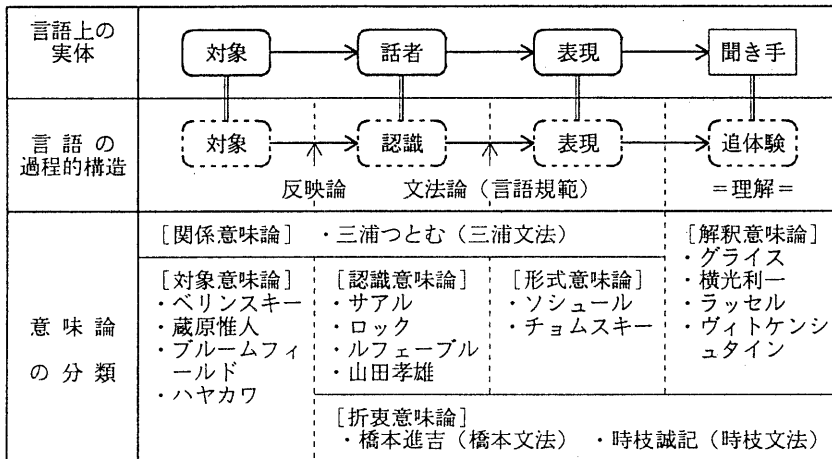


図2 意味論の分類

### (1) 対象意味論

表現の対象が表現の意味であるとする意味論にはベリンスキーや蔵原惟人の説がある。ブルームフィールド（構造言語学）<sup>(17)</sup>も実際の出来事が意味だと考えた点でこの説に分類されるが、彼は実際の出来事は人間の持つ科学的知識がもっと豊富にならないと議論できないとして、それ以上の議論を避けた。

対象が現実に存在する時はこの説も一見もっともらしい点があるが、フィクションの世界では解釈が困難となる。そこで、実在する対象については「外在の意味」、観念的对象については「内在の意味」と区別する説(S. I. ハヤカワ)が生まれた。

### (2) 認識意味論

認識意味論は話者の言いたいこと、話者の頭の中にある理論や思想が意味であるとするものであるから、上記の内在の意味には認識意味論への接点が見える。ルフェーブルを始め土屋文明、中野重治が「作品の質や創造の良否を決めるのは認識の内容だ」と主張するなど芸術作品の場合はこの説を取る人が多い。

山田孝雄（山田文法）<sup>(18)</sup>は「言語の内容は思想であり、外相は音声だ」として認識意味論を取っている。サールの説も話者の側に意味を位置づけているが、話者の認識と言うよりも「話者の意図」を取り上げ、これを意味としている。

### (3) 表現（形式）意味論

ソシュール<sup>(19, 20)</sup>は言語を生得的な言語の単位とするラング（所記）とその運用であるパロール（能記）から説明し、ラングの中で思想と表現（音）が意味の単位を作っているとして、意味を表現の側から説明

\*1 言語では、話者の認識と表現の間、及び表現と聞き手の理解の間が、言語世界で共通する言語規範に媒介されているのに対して、芸術作品の場合はこれらの関係が直接、感性によって媒介される。この点を除けば言語も芸術も意味論については同様に考えることができる。

している。ソシュール学派によれば、言語から直接受け取るのは思想でなく、音の側面であり、思想は音と結合した透明なものであるから、現実の音声の裏側に結びついた言語の辞書的な意義（本来は言語規範）が意味と言うことになる。

意味論に関する構造言語学の限界を克服することを狙ったチョムスキーは統語的深層構造を仮定しこれを意味と考へた。深層構造が先験的に与えられるものとする点では認識意味論に見えるが、これは言語表現の統語的側面を取り出してそれを意味と考へた点では、やはり形式意味論と見なされる。深層構造意味論が破綻し、二元論に変わった後は表層構造にも意味があるとしており、より表現意味論に接近したと言える。最近の心の研究<sup>(21)</sup>は神秘主義的にすら見える。

#### (4) 解釈意味論

言語表現に形式が存在することは疑い得ないが、文字にせよ音声にせよ、インクのシミがあるか空気の粗密波かの違いはあっても、それ以外の何も存在しない。そこに内容が存在するように思っているだけではないかと考えれば、「意味」とは聞き手の頭の中に得られた認識のことだと言う説になる。

ラッセル（論理分析学派）<sup>(22)</sup>やヴィトケンシュタイン（日常言語学派）<sup>(23)</sup>は話者の認識を離れ、聞き手の側で意味を解釈しようとした。日常言語学派のグライスは言語規範と「場の理論」を区別せず、「聞き手の行動」を意味と説明したが、聞き手は一人とは限らないし、誰もが同じ行動を取ってくれるわけではない。そこで、サール<sup>(24)</sup>は、話者の意図を意味と考へるべきだと反論した。

#### (5) 折衷型意味論

以上の意味論の持つ弱点を補う立場から、折衷的な意味論がいくつか提案されている。学校文法でお馴染みの橋本文法（橋本進吉）<sup>(25)</sup>は、意味内容説に聞き手の理解をも加えて、話者の言っていることと聞き手の理解していることとの共通性に着目する立場から、両者の頭の中にあるものを意味と考へている。しかし、話者と聞き手が通じ合うのは両者が同じ言語規範を手がかりとしているためと考へるべきで、ソシュールと同じ問題を引き継いでいると言える。

ソシュールが本来の言語規範を概念として捉え、この概念と音声との結合単位をラングで説明したのに対して、時枝誠記は表現対象を概念とした上で、意味は概念などのような実体ではないとして、実体意味論を排した。この点では従来あらの飛躍が伺えるが、意味とは話者と聞き手の双方にある「把握の仕方」だと説明することにより、従来の機能論のレベルに後退してしまった。

### 2.3 言語過程説から見た意味論

解釈意味論では与えられた表現の形式から意味が生まれると言う観念的な説明になる。逆に、形式は内容から生まれると見れば、対象意味論か認識意味論となるが、対象も認識もいつまでもそのまま存在するとは限らないし、「誤って書かれた文でも意味は正しいのか」と言う問題がある。

言語過程説を踏襲した三浦つとむは、時枝の意味論を修正し、関係意味論を提案した。すなわち、「音声や文字にはその背景に存在した対象から認識への複雑な過程的構造が関係づけられている。このような音声や文字の種類に結びつき固定された客観的な関係を言語の意味という。」と説明している。対象や話者の認識は意味そのものではなく、意味を形成する実体だとする説明である。これによれば、意味は話者や聞き手の側にあるのではなく、表現そのものに客観的に存在するのであって、表現（音声や文字）の消滅と共に、そこに言語規範によって固定されていた対象と認識の関係、すなわち意味も消滅すると言うことになる。

言語上の実体ではなく、それらの関係で意味を定義する考へ方は、最近の状況意味論<sup>(26)</sup>と共通する点がある。しかし、状況意味論が「言語規範に媒介された表現の意味」と「表現の置かれた（発話された）場の意味」とを区別せず、むしろ「場の表現」の側から意味を説明しているのに対して、三浦文法は両者を分けている。言語の個性と特殊性に着目する立場からは、言語にも記号にも共通する「場の意味論」は、分けて扱うのが望ましいと考へる。

### 3. 意味処理の分類

表現には、話者の認識と対象の姿が言語規範によって関係づけられているから、意味理解とはその関係

を手繰って、表現に結び付けられた認識と対象の姿を追体験することとされる。そこで「意味処理」を2つの過程に分けて考えることにする。その第1は、言語表現に使われた約束を特定する過程である。言語の約束は複雑で、同一の語や構造に対していくつもの解釈が生まれるが、言語処理では、まず「話者はどの約束を使ったのか」を調べなければならない。使われた約束を特定するこの処理をここでは「意味解析」と呼ぶことにする。第2の過程は、その約束を通して表現に結び付けられた話者の認識と対象の姿を知る過程である。ここではこの処理を「意味理解」と呼ぶことにする。

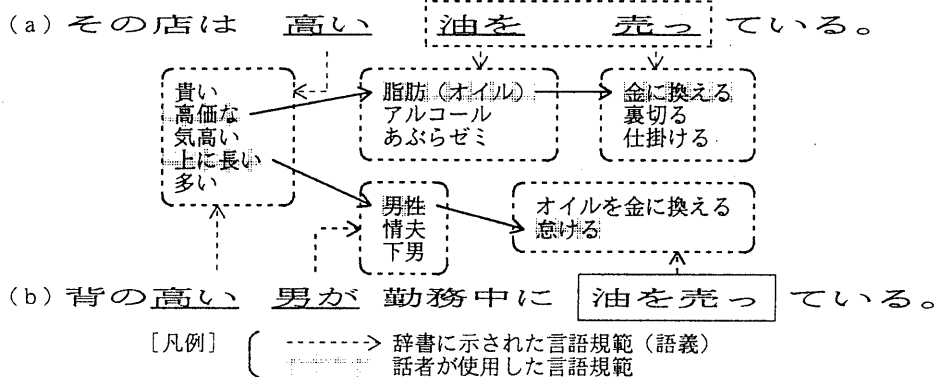


図3. 統語構造に支援された意味解析

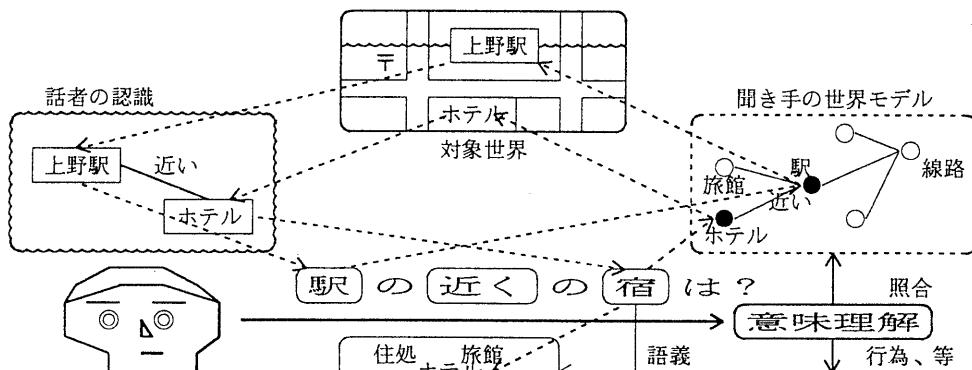


図4. 世界知識に支援された意味理解

### 3.1 意味解析

図3に「意味解析」の例を示す。文(a)に使用された単語「高い」は、「貴い」、「高価な」等の語義を持つが、この文では「高価な」の疑義で使われている。「油」にもいろいろの約束があるが、ここでは「オイル」を表すという約束が使われている。このように、言語上、沢山ある約束の中から、話者が文中で実際に使った約束を特定する処理を意味解析と言う。

語義は文を単語に分けて、その一つづつを眺めたただけで決まるとは言えない。文(a)の「高い」の語義が「油」との関係で決まるのに対して、文(b)の「高い」の語義は「背の」との関係で決まる。また、文(b)では「油を売る」が約束の一つの単位で語義は「怠ける／さぼる」であることは前後関係で判断される。約束の単位は言語によっても異なる。英語では単にtallと言う一語で「背の高い」と言う語義を持つ。

この例に見られるように、言語の約束は単語単位に考えていただけでは決まらない。表現の意味は表現と認識と対象の関係であるとする立場から見ると、対象の構造は認識に反映し、さらに表現に対応づけられることから考えて、表現の構造も意味の一部と言える。従って、意味解析においても、構造と意味の関係を考えることが必要であり、要素合成法から脱却する方法が望まれる。

### 3.2 意味理解

前に述べたように、意味理解とは表現に結び付けられた話者の認識と対象の姿を知ることである。話者の認識と対象の姿を知るためには、聞き手は、話者が描いている世界とある程度共通する世界を持つことが必要である。自分の持っている世界モデルの中に、話者の表現と対応する世界を再構成するためには、表現に現れた要素と自分の世界モデル内の要素との間に、対応づけを行う必要がある(図4)。

すなわちこのような意味で、意味理解とは、言語表現と聞き手の世界モデル内の各要素との対応付けであると考えることができる。この対応付けによって、聞き手の世界には新しい部分的な世界が構築される。

### 4. 機械翻訳と意味解析

意味解析に必要な基本的知識は言語知識、すなわち言語上の約束であるのに対して、意味理解では世界知識が必要となる(図5)。

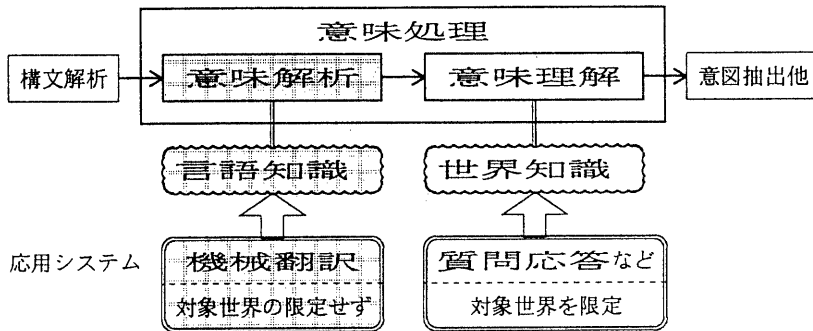


図5. 意味処理と機械翻訳

言語知識に比べて、世界知識は膨大であるから、意味理解の研究は大変な困難さがつきまとうことになる。特に、機械翻訳の対象とする内容は多岐に亘るので、原文の意味理解に必要な世界知識をあらかじめ用意することは困難である。しかし、機械翻訳では、計算機がテキストの内容を全て理解していなくても、翻訳の結果が、最終的には人間に理解できれば良いと考えることができる。

そこで、日本語で使われた言語上の約束を、英語の約束に置き換えて人の判断を待つ。これが我々の機械翻訳研究の基本的な立場である。すなわち、意味解析で出来る範囲の翻訳をまず実現する。次の段階で、意味解析では翻訳できなかった表現や思想を抽出して、限定的な世界モデルを対象に意味理解を試みると言うことである。

これに対して、電話の番号案内や、データベースの検索などの場合は、対象世界が比較的小さく絞れるので、世界モデルを構築し、意味理解の研究を行っている。

#### 4.1 多段翻訳方式の考え方

人手による翻訳を考える。翻訳しようとする表現には、話者の目を通して認識された対象の姿と、対象に対する話者の意識が結び付けられている。日本語や英語と言った言語の違いによって、話者の認識を表現に結びつけるための約束が異なるだけでなく、「認識」にも相違が生じる。

そこで、人間の翻訳家による翻訳を考えると、まず、翻訳家は表現を通じて話者の経験を追体験する。そして、追体験した内容を、目的言語の枠組みの中で、捉えなおして表現する。このとき、翻訳される内容は、話者の目を通して見た対象の姿と、それに対する話者の意識の二つである。このプロセスを模擬したのが多段翻訳方式である。

多段翻訳方式は、図6に示すように、大きくみて、話者の意識、すなわち主体的表現<sup>\*1</sup>を翻訳するパス

\*1主体的表現：時枝文法の言葉で、話者の概念化されない感情、意志、判断などの認識を直接表現する言葉を使う。話者の感情や意志も概念化されたときは客体的表現が用いられる。日本語では通常、助詞、助動詞などが使用される。

と、対象の姿、すなわち、客体的表現\*1を翻訳するパスの2つのパスから構成される。客体的表現のパスは、対象の抽象化の程度によって、さらに3つのパスに分かれる。

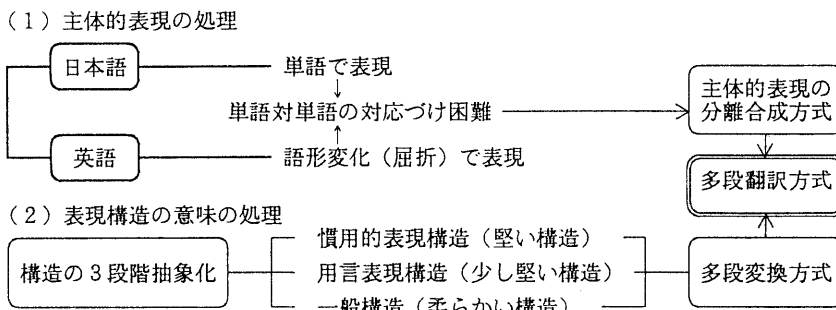


図6. 多段翻訳方式の概念

工学的な立場から見れば、これらのパスは以下の意味を持つ。既に述べたように、言語表現では表現の構造が大きな意味を持っており、言語表現と話者の認識の関係は単語単位に見ただけでは捉えきれない。従って、意味をなるべく失わないように、日本語の表現を英語の表現に対応づけるには、日本語の全ての表現に対応する英語表現のリストを持てば良いが、それは工学的に不可能である。そこで、日本語の表現を意味を失わないように段階的に抽象化して、英語表現に対応づけることになるが、客体的表現は構造が比較的抽象化しやすいのに対して、日本語と英語では、主体的表現は枠組みが大きく異なるため、両者を一体化したまま日英言語間で対応づけるのは困難である。そこで、与えられた日本語の表現から主体的表現を抜き取り、得られた客体的表現を変換した後、英語側で日本語に相当する主体的表現を付加することにしたものである。

#### 4.2 意味解析用言語知識の構成

以上から、多段翻訳方式を実現するために必要な言語知識、すなわち言語上の約束に関する知識は、主体的表現に関するものと、客体的表現に関するものの二つに分けられる。主体的表現に関する知識は、客体的表現に関する知識に比べて、比較的規模が小さいため、プログラムの的に解決したが、客体的表現に関する知識は規模も膨大であるため、辞書の形式に整理することとした<sup>(27)</sup>。

客体的表現で表される世界は、実体概念\*2と属性概念\*3から構成される。ALTでは、実体概念に関する約束を「単語辞書」、属性概念に関する約束を「構文意味辞書」として、客体的表現に関する言語知識

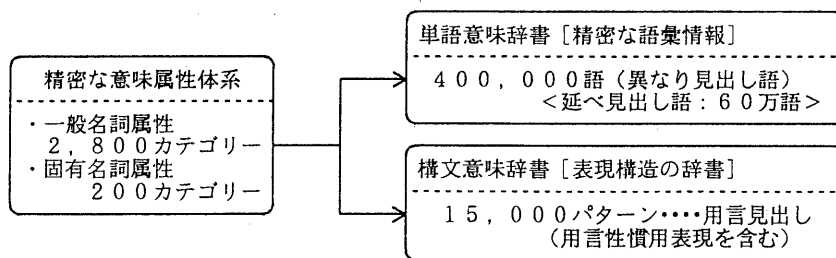


図7. 意味解析のための言語知識

\*1客体的表現：時枝文法の言葉で、話者によって概念化された対象を表現する言葉を使う。日本語では、通常、名詞、動詞、形容詞などの自立語が用いられる。  
 [備考] 客体的表現と主体的表現は、本居宣長の単語分類「玉」と「緒」を継承した概念で、時枝文法の単語分類では「詞」と「辞」に対応する。  
 \*2実体概念：対象化された「もの」と「こと」の概念で、名詞で表現される概念。  
 \*3属性概念：実体に付属する概念で、実体の動き、性質、状態などの概念。通常、静的な属性概念は形容詞で、動的な属性概念は動詞で表される。但し、これらの概念も「美しい→美しさ」、「登る→登り」のように実体概念がすることもできる。

を、図7の通り2つの辞書に整備する。「単語辞書」は名詞を中心とする単語の意味的な用法を収録するものであるのに対して、「構文意味辞書」は用言の用法を文型としてまとめるものである。すなわち、単語単位の知識は「単語辞書」に、抽象化された表現構造の知識は「構文意味辞書」に整理する。構文意味辞書は Valenz パターンの形式で、見出し語は動詞と形容詞である。

次に、これらの言語知識を記述する方法であるが、計算機に語の持つ概念を人間と同じように獲得させる方法は不明である。しかし、日英機械翻訳ではまず、日本語の語義に対応する英語側の表現が求めれば良い。一般に、日本語の単語や表現に対応する英語の表現は複数あり、1対1の関係にないが、日本語での用法を見れば対応する英語表現を決定できる場合が多い。そこで、言語知識を記述する手段として、単語の意味の用法を意味属性体系としてまとめた。例えば「学校」と言う単語は、「場所」として使われたり、「組織」や「建造物」として使われたりする。このような使われ方を単語意味属性として体系化した。

#### 4.3 言語知識分解能の評価

各単語の持つ語義の数は通常、漢語系の単語に比べて和語系の単語の方が多い。特に「行く」「上がる」「取る」など、抽象化の程度の大きい概念を示す語ほど、その数は多い。従って、意味属性をどれだけ細かくすれば良いかは、これらの単語の語義の訳し分けの能力を調べることによって推定できる。ALTではまず、約500の分類での訳し分けを試みたが、訳し分けできない動詞が続出するため、一般名詞意味属性として2,800通り、固有名詞意味属性として200通りに分類した。

実験の結果によれば、「する」「なる」など特に語義の多い機能動詞を除けば、この分類精度で用言の訳し分けはほぼ可能であることが分かった。機能動詞では、別の訳し分けのための別の手段を設けることにした。

ここで、私たちの作製した単語意味属性体系を、従来の類似のシステムの場合と比較するため、以下の3つの場合を考える。

第1の場合は、商用化されている代表的なシステムの例で、意味属性の数はおよそ30~50の場合である。第2の場合は、意味属性数が約500の場合で、およそEDRの計画<sup>(2)</sup>の分解精度に匹敵する。そして、第3の場合は、ALT-J/Eの例である。これらの3つの場合について、図8の関係を仮定して、意味属性数を意味属性の深さに換算し、構文意味辞書の文型記述に使用された意味属性の深さから、動詞の意味による訳し分けのルールが、それぞれの場合でどれだけ書けるかを比較した。

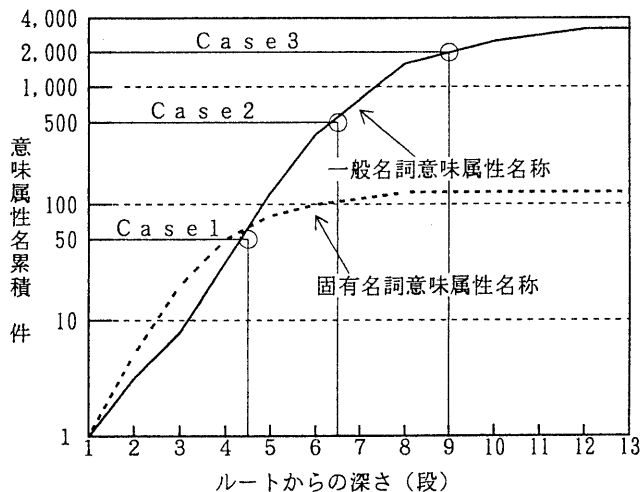


図8. 意味属性名称の数の累積分布

表1. 意味属性精度と言語知識記述能力の関係

比較項目	場合	場合1	場合2	場合3	
意味属性のノード総数		30~50	約500	2,800	
ノードの実行的な深さ		約4段	約6段	約9段	
単語辞書	一般名詞	3%	22%	100%*	
	用言性名詞	21%	25%	100%*	
構文意味辞書	一般文型	和語動詞	31%	59%	100%*
		漢語動詞	57%	86%	100%*
		形容詞	38%	60%	100%*
		形容動詞	67%	84%	100%*
	慣用文型	80%	95%	100%*	
専門文型	10%	34%	100%*		

\*場合3(ALT-J/E)の値を100%と仮定し、相対値を評価



比較の結果を表1に示す。比較においては、ALT-J/Eの場合の記述能力を100%と仮定している。この結果から、文型ルールの記述能力は、意味属性の分解精度に大きく依存することが分かる。その依存度は、予想通り、漢語系の用言に比べて、和語系の用言に大きいことが分かる。和語系の用言は、漢語系の用言に比べて種類は少ないが、実際の文での使用頻度は高いため、その訳し分け能力はより重要である。

#### 4.4 意味解析による新機能の実現例

以上のような意味解析技術を実現した日英機械翻訳システムALT-J/Eで可能となった翻訳機能として、動詞と名詞の訳し分けの例を図9に示す。

##### (1) 動詞「掛ける」の訳し方

彼女は花に水を掛けた。	She <u>poured water</u> on a flower.
母は神様に願を掛けた。	A mother <u>made a vow</u> to God.
私は彼らに迷惑を掛けた。	I <u>caused them trouble</u> .
彼は二階に梯子を掛けた。	He <u>placed a ladder</u> up to the second floor.
建設省はここに橋を掛けた。	The Ministry of Construction <u>built a bridge</u> here.
彼は椅子に腰を掛けています。	He is <u>sitting down</u> on a chair.
彼らは砂をふるいに掛けた。	They <u>sifted</u> sand.
彼女は毎日廊下に雑布を掛けている。	She <u>mops up</u> the corridor every day.
彼女はプレゼントにリボンを掛けた。	She <u>tied ribbon</u> around a gift.
彼女は食卓にテーブルクロスを掛けた。	She <u>spread a tablecloth</u> on a dining table.
あの喫茶店はモダンジャズを掛けている。	That coffee shop is <u>playing modern jazz</u> .

##### (2) 名詞「群れ」の訳し方

狼の群れが羊の群れを追った。	A <u>pack of a wolf</u> chased a <u>flock of a sheep</u> .
鯨の群れが魚の群れを追った。	A <u>school of a whale</u> chased a <u>shoal of a fish</u> .
牛の群れが蜂の群れに襲われた。	A <u>bevy of cattle</u> was attacked by a <u>swarm of a bee</u> .
人の群れが暴徒の群れに変わった。	A <u>group of people</u> changed to a <u>mod of a mob</u> .
少女の群れが少年の群れと遊んでいる。	A <u>bevy of a girl</u> is playing with a <u>crowd of a boy</u> .

図9. 動詞と名詞の意味による訳し分けの例

意味属性体系とそれによって記述された単語意味辞書、構文意味辞書はこれらの単語の語義による訳し分けだけでなく、意味解析を応用した多くの新しい翻訳機能の実現に有効である。これらの樹段の実現によって、翻訳のルールが従来の文法的属性と単語の字面だけでなく、意味属性でも記述できるようになったため、ルールの柔軟性と記述能力が大幅に向上した。

これらの仕組みを使用して、実現した新しい翻訳機能の例を以下に示す。

##### (1) 日本語自動書き換え型翻訳機能<sup>(29)</sup>

英語に直接対応する表現がなくて直訳できない表現や直訳できても英語としてみれば不適切な表現を翻訳し易い日本語に自動的に書き換えて翻訳する機能。書き換えの規則の適用条件が精密に書けるようになり、副作用の心配が減少したため、効果的な書き換えが可能となった。

##### (2) 複合語翻訳機能<sup>(30)</sup>

一般名詞を結合したもの、接頭、接尾辞を持つもの、サ変名詞を含むもの、固有名詞を含むものなど多彩な複合語を一般語の組み合わせで翻訳する機能。

##### (3) 文脈処理による文要素補完型翻訳機能<sup>(31)</sup>

用言を中心とする表現構造が意味の単位として捉え、用言間の関係としての文脈を追跡することにより、省略された格要素を補完する機能。

## 5. あとがき

従来の言語表現にの意に関する議論を振り返り、言語過程説の立場から、「意味処理」を「意味解析」と「意味理解」に分けることを提案した。また「意味解析」を実現した例として日英機械翻訳システムALT-J/Eを取り上げ、意味解析に必要な知識（言語知識）を単語意味属性を媒介とする単語意味辞書と構文意味辞書の2つの辞書にまとめる方法を示した。

単語意味属性の分解精度と用言文型の記述能力の関係では、従来の約50種または約500種の意味属

性分類では、特に和語系の動詞について、記述不能の文型が続出するのに対して、約3,000種程度に分解精度を上げた場合は、機能動詞などの一部の用言を除いて、訳し分けに必要な文型がほぼ記述できることが分かった。

単語意味属性を介して記述されたこのような単語辞書と構文意味辞書の知識は、用言と体言の意味の訳し分けに効果があるだけでなく、複合語解析や文脈型処理、翻訳し易い日本文への自動書き換えなどのルール記述においても有効であることが分かった。

現在のALITでは、客体化された概念と語義との間の関係知識と、属性概念と実体概念間の関係知識の2種類を、言語知識(言語規範)として、辞書化してきた。対象の概念化のレベルを第1回の概念化、対象間の関係の概念化を第2回の概念化と見ると、現状のALITは1.5回概念化のレベルに相当する。名詞句の解釈などの曖昧性解消を考える場合は、第2回の概念化のレベルまで言語知識を体系化する必要があると考えられる。

[謝辞]共に議論を進めてきた白井、小倉、横尾の各主任研究員ほか、翻訳研究グループの皆様に感謝する。

#### ◇ 参考文献 ◇

- (1)Apel他:「チョムスキーと現代哲学」井口編訳,大修館,1976
- (2)C. フックス, P. ル・ゴフィック:「現代言語学の諸問題」田島他監修,三修社,1983
- (3)Language and Machines: Computer in Translation and Linguistics, National Academy of Sciences, National Research Council, 1966 (ALPAC report)
- (4)T. Winograd, A Procedural Model of Language Understanding, " In Computer Models of Thought and Language, Edited by R. Shank and K. Colby. San Francisco: W. H. Freeman, 1973
- (5)W. A. Woods, Procedural Semantics for Question-Answering Machine, Proc. of FJCC, N. J.: AFIPS Press, 1968
- (6)遠藤, 高岡: 図形処理との協調に基づくドリルテキストの解析, 第42回情処全大, 2Q2, 1991
- (7)荒木, 桃内: 名詞述語文における意味概念の学習, 情処学会研究報告, Vol. 90, No. 77, 1990
- (8)奥村, 田中: 自然言語解析における意味的曖昧性を増進的に解消する計算モデル, 人工知能学会誌, Vol. 4, No. 6, PP. 65-72, 1989
- (9)時枝誠記: 「国語学原論」, 岩波書店, 1941
- (10)三浦つとむ: 「認識と言語の理論, Vol. 1~3」, 勁草書房, 1967
- (11)三浦つとむ: 「言語学と記号学」勁草書房, 1977
- (12)池原, 宮崎, 白井, 林: "言語における話者の認識と多段翻訳方式", 情処論, Vol. 28, No. 12, pp. 1269-1279, 1987
- (13)S. Ikehara, "Multi-Level Machine Translation System", Future Computer Systems, Vol. 1, No. 3, pp. 261-274, 1989
- (14)N. Chomsky, Knowledge of Language, Ms. MIT Press., 1984
- (15)C. Pollard, Lecture on HGPS, ms. Stanford Univ.
- (16)郡司隆男, Generalized phrase structure grammar and Japanese reflexivization, Linguistics and Philosophy, Vol. 6, No. 1, 1983
- (17)C. ヴォース: 「構造主義の言語学」菅田訳, 大修館, 1975
- (18)山田孝雄: 日本文法論, 宝文館, 1908
- (19)ツシュール: 「言語学序説」山内訳, 勁草書房, 1971
- (20)E. F. K. カルナー: 「ツシュールの言語論」山中訳, 大修館, 1982
- (21)N. Chomsky, 言語と知識—マゲアナ講義録—, 産業図書, 1989
- (22)B. Russell, 意味と真偽性, 毛利可信訳, 分科評論出版, 1972
- (23)奥雅博訳: ヴェイトケンシュタイン全集第2巻「哲学的考察」, 大修館, 1976
- (24)J. R. サール: 「言語行為」坂本・土屋, 勁草書房, 1986
- (25)橋本進吉: 国語学言論, 岩波書店, 1946
- (26)J. M. Gawron and S. Peters, Anaphora and Quantification in Situation Semantics, CSLI Lecture Note No. 19, 1990
- (27)池原, 宮崎, 横尾: 日英機械翻訳のための意味解析辞書, 信学情処合同研究会, '91. 07. 20
- (28)技術報告: 「概念辞書」第3版, 日本電子化辞書研究所, 1991
- (29)白井論, 日本文自動書き換えによる構文多義の解消, 情処全大4S-6, 1990
- (30)植田, 小松, 横尾, 宮崎: 部分複合語による複合名詞構造解析, 情処全大, '91. 10. 20
- (31)H. Nakaiwa and S. Ikehara: Zero Pronoun Resolution in a Japanese to English Machine Translation System using Verbal Semantic Attributes, Applied Computational Linguistics, '92. 04. 1-4