

1E-8

新潟大学大学院自然科学研究科

計算機による自然言語処理において、言語の形式のみの把握だけでは十分とはいえない。より高度な自然言語の意味理解の実現には、自然言語の本質についての考察が必要である。そのため、人間の言語活動である「対象 $\leftrightarrow$ 認識 $\leftrightarrow$ 表現」という時枝の言語過程説を発展的に継承した三浦の説 [1]、三浦説に基づき英語文法を構築した宮下の考え [2] に基づき、言語の表現構造と認識された対象世界を把握する認識構造を抽出する英語文パーザを試作した。一般化 LR 法（富田法）[3] を Prolog 上に実現した拡張型の SGLR パーザ [4] である SGLR-plus[5] 上に、英語文法を補強項付きの DCG 形式で記述することにより、英語表現から認識構造を抽出する英語文パーザのプロトタイプを試作した。

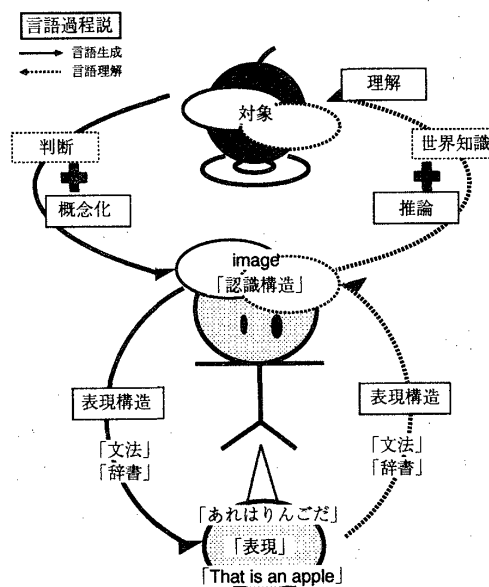


図 1：言語の過程的構造

人間の言語活動には言語生成と言語理解の2つがある。言語生成の場合、表現しようとする対象の客観的な姿を概念として表す。かつ、その対象に対し主観的で様々な判断を重ね合わせる。これにより人間は頭の中に認識構造を形成する。そして、その認識構造から発話、記述するなりして表現しようとするが、そこには必ず社会的な約束である言語規範（文法、辞書、…）で拘束される。このようにして人間は言語を表現する。また逆に、ある言語表現を言語規範に基づいて聞いたり読んだりして、認識構造を頭の中に形成する。そして、その認識構造から主体の持つ世界知識と重ね合わせ対象の姿を推論することによって、言語表現は理解される。この様子を図1に示す。

人間の主な言語表現は文字や発話である。それらは「文 = Sentence」として存在し、表現と呼ばれる。それが各言語において、文法上の構造について考えたものが表現構造である。その構造は「対象を客観的に捉えたものを概念として表す客体的表現」と「それに対する人間の主観的な判断、感情を表す主体的表現」にわかれる。この分割作業は人間が言語規範によって聞いたり読んだりする過程で行なわれ、それを本能的に噛み砕いて「客体認識」と「主体判断」に変換する。そして認識構造を形成する。その様子を図2に示す。

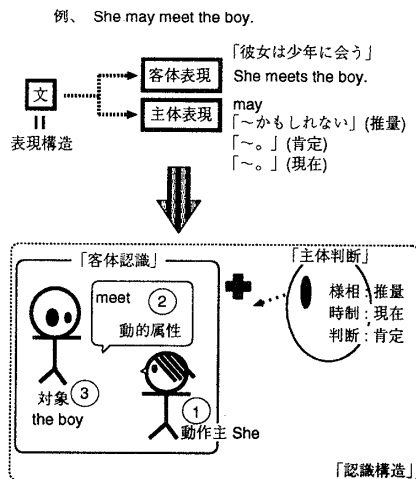


図2：表現構造から認識構造へのながれ

4 解析例

「The boy "that she met" may run」("～"は関係代名詞節)を例として、表現構造(図3)と認識構造(図4)を抽出した。表現構造については、関係代名詞節内を先行詞と同じ構造が来るべき場を捜す痕跡機能によって、< \$gap >のように表示される。認識構造については、対象の概念として" The boy "、動的属性として" run "が客体認識として存在する。そこへ主体判断である「内容の肯定」、「時制は現在」、「様相は推量」が重ねられる。次に関係代名詞節に対する先行詞" The boy "が持つ対象認識世界についてだが、" she "と" meet "と" the boy "が客体認識として存在する。そして先行詞である" the boy "があるべき" 対象_Object "の場所には、痕跡機能によって先行詞である" 対象_Agent "と同じ構造を出力し、全体の構造をわかりやすいものへ近付けた。また < 把握 >は、話者が対象をどのような観点で概念化したかを表したものである。

5 おわりに

本稿では、計算機による自然言語処理において、言語表現の持つ本質的な意味理解の実現を目指し、人間の言語活動である「対象 ⇄ 認識 ⇄ 表現」という過程的構造に着目した時枝誠記による言語過程説を発展的に継承した三浦つとむの説に基づいて、自然言語の本質について検討した。その検討に基づき SGLR-plus パーザを用い、認識構造を抽出する英語文パーザのプロトタイプを試作した。今後、多言語間での機械翻訳に有用のものへと検討する必要がある。また文法上でのいくつ

かの曖昧性解消のため意味解析との融合も必要となってくる。

```

|- 平叙表現 1
  |- 名詞句 1
    |- 直接目的語 0 . . . . .
    |   |- 名詞句 0
    |     |- 冠詞 -- < the >
    |     |- 普通名詞 -- < boy >
    |- 関係代名詞 -- < that >
    |- 平叙表現 1 <対応している!>
      |- 人称代名詞 -- < she >
      |- 肯定判断辞_emp
      |- 動詞句 0
        |- 動詞 -- < met >
        |- 直接目的語 0 -- < $gap >
      |- 助動詞 -- < may >
      |- 肯定判断辞_emp
      |- 動詞句 0
        |- 動詞 -- < run >
  
```

図3：「The boy that she met may run」の表現構造

```

|- 認識世界
  |- 客体認識
    |- 対象_Agent
    |   |- 把握 -- < the >
    |   |- 概念 -- < boy >
    |   |- 認識世界属性
    |     |- 客体認識
    |       |- 対象_Agent
    |         |- 概念 -- < she >
    |         |- 動的属性 -- < meet >
    |       |- 対象_Object
    |         |- 把握 -- < the >
    |         |- 概念 -- < boy >
    |- 主体判断
    |   |- 時制 -- < 過去 >
    |   |- 判断 -- < 肯定 >
    |- 動的属性 -- < run >
  |- 主体判断
    |- 時制 -- < 現在 >
    |- 様相 -- < may >
    |- 判断 -- < 肯定 >
  
```

図4：「The boy that she met may run」の認識構造

参考文献

- [1] 三浦つとむ：日本語とはどういう言語か、講談社学術文庫(1976)
- [2] 宮下真二：英語とはどういう言語か、季節社(1985)
- [3] 田中穂積：“自然言語解析の基礎”，産業図書，1993。
- [4] 沼崎，田中：SGLR：逐次型一般化 LR パーザの Prolog による実現，情報処理学会論文誌，Vol.32，No.3，pp.396-403 (1991)
- [5] 五百川，宮崎：痕跡処理のための逐次型一般化 LR パーザ SGLR の拡張，言語処理学会第4回年次大会発表論文，pp.314-317 (1998)
- [6] 中島太郎：英語の仕組み，日本実業出版社(1994)