

# 逐次的な解析による日本語の数量詞遊離構文の解釈

小林 昌博<sup>1</sup> 吉本 啓<sup>2</sup> 中村 裕昭<sup>3</sup> 森 芳樹<sup>4</sup>

<sup>1</sup> 鳥取大学大学教育総合センター <sup>2</sup> 東北大学大学院国際文化研究科  
<sup>3</sup> 海上保安大学校基礎教育講座 <sup>4</sup> 筑波大学大学院人文社会科学研究科

<sup>1</sup> kobayashi@uec.tottori-u.ac.jp <sup>2</sup> kei@insc.tohoku.ac.jp  
<sup>3</sup> nakamura@jcg.ac.jp <sup>4</sup> mori@lingua.tsukuba.ac.jp

## 1 はじめに

言語は線状的な性質を持っているため、我々は文を左から右へと逐次的に読み、文に解釈を与えている。日本語は英語などとは異なり、述語が文末にくるヘッド・ファイナル言語であるため、述語が読み込まれるまで文全体の構造を把握できない。例えば文頭の名詞句を読んだ段階では、述語や補文標識が読み込まれるまで、それが主節の主語であるのか、従属節の主語であるのか、あるいは連体修飾節の主語であるのか決定できない。これらの事実から、文を入力順に解析する際の理論的な難しさが指摘されてきた。本稿では、文を左から右へと逐次的に解析する Dynamic Syntax (Kempson, Meyer-Viol and Gabbay 2001) に基づく日本語の解析モデルを提案し、語順に関する現象、特に数量詞遊離構文の解釈に関して本稿の枠組みが有効であることを示す。また、本稿の枠組みの計算機上での実装を紹介する。本稿で紹介するシステムは、入力文に対して各単語を読み込んだ段階の解析結果を表示し、最終的な構造と意味を出力する。

## 2 問題の所在と先行研究

本稿で扱う数量詞の遊離現象とは、(1a)に見られるような「数量表現+の」が(1b)のように右方に移動する現象を言う。

- (1) a. 3冊の本を太郎が買った。  
b. 本を太郎が3冊買った。

Bond, Kurz and Shirai (1998)に見られるように、数量詞遊離構文の扱いは機械翻訳においても重要な問

題であるが、言語学の分野において数量詞遊離構文の問題として議論されてきた点は主に次の2点である。まずどのような名詞句が遊離を許すのか、という問題である。Shibatani (1977)はガ格とヲ格からのみ遊離が可能であると主張するが、(2)のように例外が多い。

- (2) 太郎は有名な ピアニストに2人出会った。

Miyagawa (1989)は述語の項と付加詞の統語的な違いに着目し、(2)のような例を説明している。

次に、遊離した数量詞とどの名詞句が修飾関係にあるかを決定する問題がある。Miyagawa (1989)や Saito (1985)は数量詞遊離構文における非対称性を指摘し、(3)のような文法性の違いを説明している。

- (3) a. 車を太郎は3台持っている。  
b. \* 学生が本を2人買った。

(3a)では、目的語とそれから遊離した数量詞が隣接していなくても文法的であるが、(3b)のように主語の場合は非文法的である。Miyagawa (1989)では、数量詞とそれが修飾する名詞句は相互 c-command の関係、つまり隣接関係になければならないとしている。(3a)の場合、目的語「車を」は遊離数量詞「3台」の直前の位置から移動していると仮定されるので、遊離数量詞と目的語が移動する際の痕跡 (trace) が隣接関係にある。したがって(3a)と(3b)の差異が説明できるとしている。遊離数量詞と修飾関係にある名詞句の決定の問題に関して、高見 (1998)は、Miyagawa (1989)の説明に矛盾する例(4)を挙げ、文の情報構造に基づく説明を行っている。

- (4) 今朝、学生さんがその新刊雑誌を5人買って行きましたよ。(高見 1998)

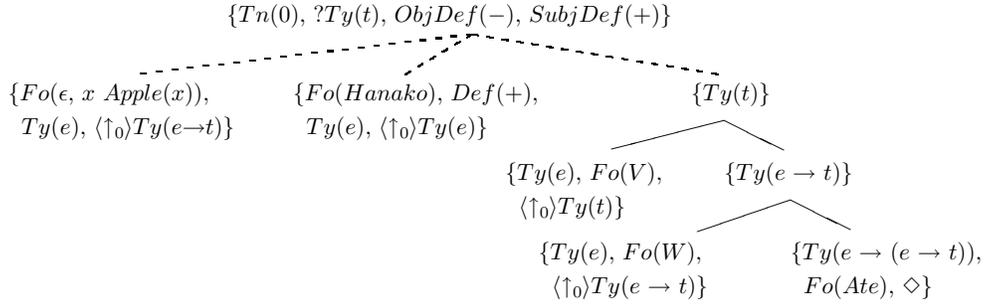


図 1: (6) の解析木

*floating quantifiers*

```

IF    {?Ty(t), ObjDef(-)},
THEN put(ObjFQ(+)); make(⟨↓*⟩);
        put(Ty(tv → tv), Fo(FQ));
        gofirst↑(Tn(0))
ELSE IF    {?Ty(t), SubjDef(-)},
THEN put(SubjFQ(+)), make(⟨↓*⟩);
        put(Ty(vp → vp), Fo(FQ));
        gofirst↑(Tn(0))
ELSE ABORT
  
```

図 2: 遊離数量詞の辞書構造

(4) では、遊離数量詞と修飾関係にある名詞句が隣接していない。高見 (1998) によると遊離数量詞は新情報を伝える不定名詞句と修飾関係を持つ。(4) は主語と数量詞が隣接していないが、目的語が定名詞句なので、主語名詞句が一義的に数量詞と修飾関係を持つと決定できる。(3b) の非文法性は、主語と目的語がともに不定名詞句であるため、どちらと修飾関係を持てばよいか一義的に決定できないため非文となると説明できる。

本稿の枠組みは、逐次的な言語解析過程のモデルであるだけでなく、上記の 2 つの問題を適切に説明することができる。次節では、本稿で用いる Dynamic Syntax (Kempson et al. 2001) の定式化について解説する。

### 3 Dynamic Syntax の定式化

Dynamic Syntax (以下 DS) と他の統語理論の大きな違いは、不完全指定 (underspecification) の概念を統語論に導入している点である。この操作により、

日本語のようなヘッド・ファイナル言語も入力順に処理できる。DS では文頭から単語を 1 語ずつ読み込み、動的に文の構造が組み立てられていく。通常の樹形図 (解析木) に相当するものはノードの集合として表現され、ノードの上下関係は tree modality により記述される。tree modality を含むノード間の関係は、Logic of Finite Tree (Blackburn and Meyer-Viol 1994) の定式化にしたがっている。ノードの集合は、解析が進むにつれ情報が付け加えられて成長していく。解析を始める前の初期状態である *axiom* は、(5) に見られる 1 つのノードである (Kempson et al. 2001: pp.76)。

(5)  $\{Tn(a), ?Ty(t), \diamond\}$

(5) の  $Tn(a)$  は、樹形図全体における当該のノードの位置を表し、項である  $a$  は変数である。 $?Ty(t)$  はこのノードのタイプが  $t$ 、つまり文であること意味する。「 $?$ 」は requirement と呼ばれ、現在は未確定であるが最終的にはこのノードのタイプが  $t$  になることを要求している。 $\diamond$  はポインタである。DS では、各単語の語彙情報が **IF**  $\alpha$ , **THEN**  $\beta$ , **ELSE**  $\gamma$  のような規則の形式をとる。これは「もしポインタを含むノードに  $\alpha$  があるならば、 $\beta$  を実行し、それ以外は  $\gamma$  を実行せよ」という意味で、これらの規則は単語を読み込む際にポインタのあるノードに適用される。具体的な解析例として (6) を考えてみる。

(6) リンゴを花子が食べた。

(6) の解析終了直前の解析木は図 1 のようになる。図 1 で示されるように、点線で結ばれているノードは、解析木において最終的な位置が固定されていない状態を表している。1 番上のノードはルートノードと呼ばれる。動詞「食べた」が読み込まれた時点で実線で表される verb-frame (Kempson et al. 2001:

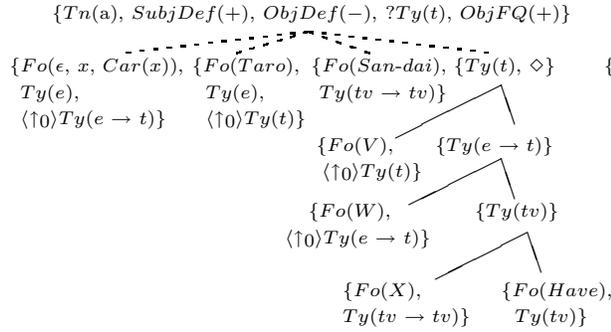


図 3: (3a) の解析木

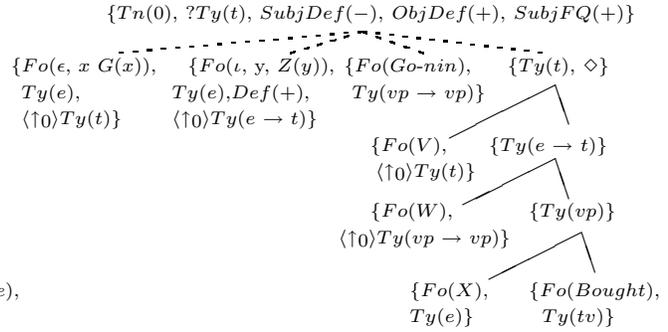


図 4: (4) の解析木

pp.72) と呼ばれる構造が導入される．点線でつながれたノードは実線で結ばれているノードのどれかにコピーされる．これはタイプのマッチングにより行われる．点線で結ばれたノードのコピーが終了し上記の requirement がすべて満たされれば，解析は成功となる．ただし，最終的な解析木は論理式に相当する情報のみを表し，実際の語順の情報を反映していない．例えば，(6) では目的語 – 主語の語順であるが，最終的な解析木ではこの情報は反映されない．

## 4 数量詞遊離構文の動的な解析

本節では，前節で解説した DS に基づく逐次的な解析手法が 2 節で述べた問題に対して有効であることを示す．具体的な方針は，(高見 1998) で主張されている情報構造を DS に取り組み，その情報と DS 独自の未指定の概念を用いて遊離数量詞と修飾関係にある名詞句を特定する，ということである．本稿では，遊離数量詞の語彙情報を図 2 のように規定した．紙面の都合上，図 2 の詳細な解説は省略するが，図 2 の辞書情報により数量詞と修飾関係にあるのが主語なのか目的語なのかに関する情報がルートノードに集まる．次に読み込まれる動詞がその情報を受け，適切な verb-frame を導入する．

図 3 は (3a) の解析終了直前の解析木である．高見 (1998) が言うところの文の情報構造は主語の定，不定を表す *SubjDef* が+か-か，目的語の定，不定を表す *ObjDef* が+か-かの組み合わせで決定される．名詞句の定，不定に関する情報がルートノードに集められ，直後に読み込まれた数量詞がその情報を受け，主語修飾か目的語修飾の数量詞であるのかをルートノードに書き込む．さらにその情報を受け，直後に読み込まれる動詞が適切な verb-frame を導入する．

図 3 において，遊離数量詞のタイプは  $Ty(tv \rightarrow tv)$  なので，verb-frame 中の  $Fo(X)$  があるノードと合致する．verb-frame 部分が最終的な解析木となるため，表層の語順とは異なり，数量詞は修飾関係を持つ目的語の直下の位置を占める．一方，図 4 は (4) の解析終了直前の解析木である．図 4 も図 3 同様，表層の語順とは異なり，タイプのマッチングにより遊離数量詞が  $Fo(W)$  がある位置にコピーされる．結果として修飾関係にある名詞句の直下の位置にコピーされることになる．図 3 と 4 に共通することは，文の情報構造が動的にルートノードに集められて，動詞がその情報を使って適切な verb-frame を導入する点である．さらに表層の語順に関わらず，遊離数量詞が修飾関係にある名詞句の直下の位置にくる．これによって，(3a) と (4) で名詞句と数量詞の修飾関係の解釈が同じ規則で解釈できる．これは Miyagawa (1989) の c-command の定義を別な形で実行している効果がある．このように，Miyagawa (1989) と高見 (1998) の説明が本稿の枠組みに矛盾なく統合されることがわかる．

## 5 DS に基づくパーサ

現在，前節までの逐次的な解析モデルを計算機上へ実装することで理論の精密さを確認し，その結果を理論にフィードバックするという作業を行っている．本節では，DS を用いた逐次的な解析システムについて紹介する．

本稿の解析システムは perl で実装されており，Tk を用いた文法開発支援のための GUI ツールを備えている．パーサは入力文に対して，各単語を読み込んだ際の解析結果と最終的な解析木及び意味に相当する論理式を出力する．パーサの設計に関しては，解

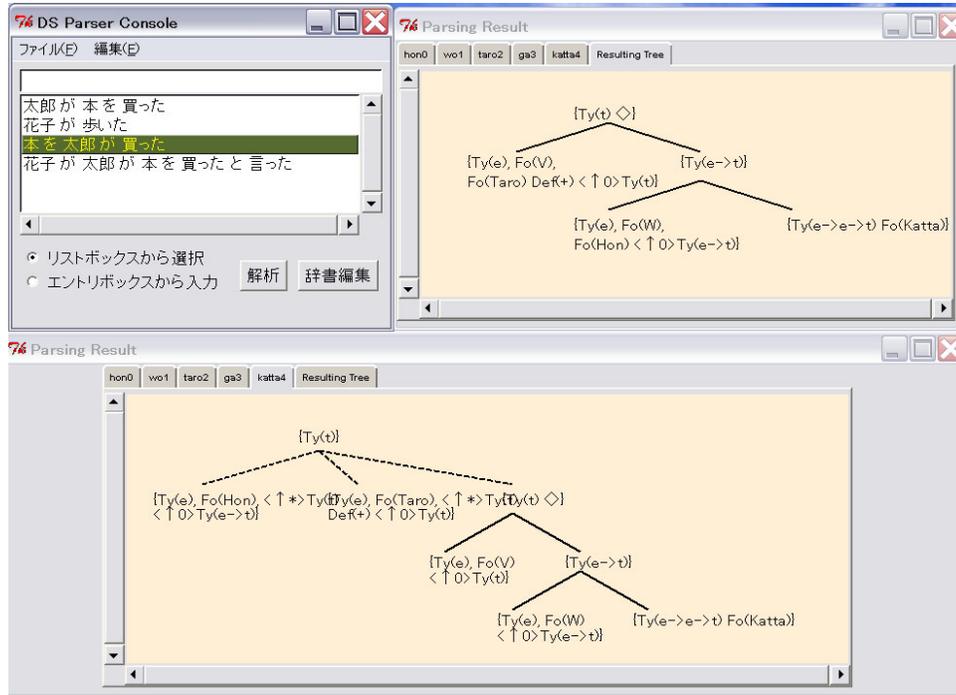


図 5: 解析結果の画面

析速度の効率化よりも理論を忠実に実装することを目指している。辞書項目は、あらかじめ形態素解析を行った結果を人手で修正し、語彙情報として使用しているが、辞書作成支援のための GUI ツールの開発も行っている。解析結果の出力画面を図 5 に示す。図 5 では、入力文に対して、単語ごとの解析結果を見ることができる。現在パーサの精密化及び扱うデータの拡張を目指している。

## 6 おわりに

本稿では、Dynamic Syntax を用いた逐次的な日本語解析モデルを提案した。具体的な現象として数量詞遊離構文を例に挙げ、情報構造を組み込んだ定式化により Miyagawa (1989) や高見 (1998) で扱われる例が同じ枠組みで解釈できることを示した。また、本稿のモデルの実装を紹介した。今後の課題としては、理論の精密化及び他言語への拡張を考えている。

## 参考文献

Blackburn, P. and W. Meyer-Viol. 1994. Linguis-

tics, Logic, and Finite trees. in *Bulletin of Interest Group of Pure and Applied Logics*. pp.2–39.

Bond, F., D. Kurz and S. Shirai. 1998. Anchoring Floating Quantifiers in Japanese-to-English Machine Translition. in *17th International Conference on Computational Linguistics: COLING-98*. pp.152–159.

Kempson, R., W. Meyer-Viol and D. Gabbay. 2001. *Dynamic Syntax: The Flow of Language Understanding*. Oxford: Blackwell Publishers.

Miyagawa, S. 1989. *Syntax and Semantics 22: Structure and Case Marking in Japanese*. San Diego: Academic Press.

Saito, M. 1985. *Some Asymmetries in Japanese and Their Theoretical Implications*. Ph.D. Dissertation. MIT.

Shibatani, M. 1977. Grammatical Relations and Surface Case. *Language*. 53: 4. pp.789–809.

高見健一. 1998. 日本語の数量詞遊離について：機能的分析 (下). 「言語」27: 3. pp.99–107.