

「に」の構文的曖昧性とその解消

武藤 伸明 中川 裕志

横浜国立大学 工学部

概要 表現「に」は格助詞, 並立助詞, 形容動詞の連用形語尾という3つの統語範疇を有し, 曖昧性発生の原因となる。この「に」による曖昧性の発生は, 意味的に抽象性の高い名詞について特に顕著にみられる。本稿はこの「に」に関する曖昧性解消に選択制限の導入が有効であることを述べる。抽象性の高い名詞に関するもう1つの問題として, その意味出力の構築およびゼロ述語補完に意味的推論が必要である。選択制限の記述に必要である概念辞書は同時に意味的推論を行なうための知識ベースとして利用できることを示す。また, 実装のための基盤として JPSG を用いこれに意味的制約を付加し構文解析実験を行なう。

Syntactic Ambiguity of 'Ni' and its Disambiguation

Nobuaki MUTOH Hiroshi NAKAGAWA

Faculty of Engineering, Yokohama National University

Abstract Since Japanese particle 'ni' has three syntactic categories—a case marker, an NP conjunct and a suffix of an adverbial-form of adjective verbs, it causes syntactic ambiguity. Especially, this ambiguity often occurs with NPs where the concept those NPs refer to is abstractive. To resolve this kind of ambiguity, it is effective to incorporate the selectional restriction into the understanding system. There is another problem related to high-abstractive NPs. In the process of constructing the semantic output of these NPs, a reasoning including the complementation of the deleted VP is necessary. The semantic network which is needed to describe selectional restriction is also useful to this reasoning. For an experiment, we describe this kind of restriction in JPSG based parser.

1 はじめに

言語理解において、曖昧性の解消すなわち入力表現に対する不必要な出力解の削除は重要な課題の1つである。本稿は曖昧性の解消が問題になる表現として「に」という表現を考える。「に」は格助詞、並立助詞、形容動詞の連用形語尾という3つの統語範疇を有し、曖昧性発生の原因となる。本稿で注目する意味的に抽象性の高い名詞はこの3種の「に」と関わりを持ち、やはり曖昧性の発生が問題となる。これらの表現に関するもう1つの問題として、その意味を出力として得るには意味的推論およびそれによるゼロ述語補充が必要である。本稿はこれらの表現の曖昧性解消に選択制限の導入が有効であることを述べる。またこれに必要な概念辞書が同時に意味出力を得るための推論過程にも利用できることを示す。さらに文法としてJPSG[Gun87, Gun94, JWG]を用い、これに意味的制約を加え構文解析実験を行なう。

2節では、抽象度の高い名詞が名詞本来としての用法と述語としての用法を有し、それが多義性の要因となっていることをみる。これらの語彙の辞書への記述について、3節ではその統語範疇をどう定義するか、4節ではこれらの語彙1つについて幾つの辞書項目が必要であるかについて述べる。5節ではこれらの語彙項目についての多義性解消に選択制限を用いるのが有効であることを述べる。選択制限の記述に必要な概念辞書の内容について6節で示す。抽象度の高い名詞句の処理に必要な意味的推論に概念辞書の利用が有効であることを7節および8節で示す。システムの実装については9節で述べる。

2 抽象度の高い名詞の有する多義性

本稿で着目するのは次のような語彙である。

- (1) 以外, 以上, こと (形式名詞), だけ (取り立て助詞), の (準体助詞)。

これらの語彙は以下のような特徴をもつ。統語的な特徴として、これらの語彙は名詞句および文に付着することができる。また、意味的な特徴として、これらの語彙は多義性を有する。これらの語彙のもつ多義性の例を挙げれば、例えば「こと」という名詞は体言化の用法のほかに次のような用法で使われる。

- (2) 困ったことに、出張に行かなければならない。

上の例文の「ことに」の意味は、主文に対する話者の主観的見方を表すとでもいうべきものであり、けっして

体言化の「こと」+格助詞の「に」ではない(もしそうであるとしたら、述語「行く」との関わりからこの「に」はゴール格もしくは時間格を表さなければならぬ)。また、次の文は「以上」についてのものであるが、これを見ると「部長以上」という表現は部長以上の地位をもつ人と部長以上の量という2通りの意味で使用されることがわかる。

- (3) 部長以上に手紙を書いた。

読み1: 部長以上の地位をもつ人に手紙を書いた。

読み2: 部長が書く以上の量の手紙を書いた。

- (1) の語彙はこのように多義性を有し、この多義性の解消はこれらの語彙を処理する上での1つの課題である。

ここでこれらの語彙の曖昧性の発生要因について考えてみることにする。このために語彙のそれぞれの用法とそれに後接する助詞の分布についてみよう。例えば「以上」について考える。「以上」にはまず(3)の読み1のような「以上」、つまり「部長以上」が部長以上の地位をもつ人という意味になるような「以上」がある。この「以上」は、個体(の集合)を表す名詞句(この場合では「部長」)に付いて、その個体をもつ値以上の値をもつ個体の集合を表すので、この「以上」を「個体→個体」の関数である「以上」と呼ぶことにしよう。この「個体→個体」の関数の「以上」は、(4)のように格助詞を伴って述語の適当な格に入ることができる。このとき「以上」に伴う格助詞がどのようなものでなければならないかについて特に制限はない。

- (4) a. 部長以上が会議をした。

b. 部長以上を会議に呼んだ。

もう1つ、(3)の読み2の「以上」、すなわち「部長以上」が部長以上の量という意味になる「以上」がある。この「以上」は個体を表す名詞句に付着してその個体もっている値以上の値を表すので、これを「個体→値」の関数としての「以上」と呼ぶことにする。この意味での「以上」に後接する助詞についてみると、この「部長以上」に後接する助詞は「に」または「で」(ex. (5))に限定されており、他の助詞は付かない。このように「以上」についてはそれぞれの用法について後接助詞の分布が異なる。

- (5) スコアは部長以上ではない。

また、「こと」については、体言化の「こと」には任意の格助詞が付く。しかし(2)のような話者の主観的見方を表す「こと」に後接する助詞は「に」に限られる。

このような語彙の各用法に応じて後接助詞の分布が変わるという現象は(1)の各語彙についてみることができ。特徴的であるのは、後接助詞の分布のパターンは次の2つに限定されていることである。

- (6) a. 任意の格助詞が後接する。
- b. 「に」(または「に」と「で」)が後接する。

後者のパターンについて、「に」または「で」との結び付きが非常に強いということから、これは名詞+格助詞「に」という構造ではなく実は形容動詞語幹であり、「に」はその連用形語尾であろうと推測できる。このことから、(1)の語彙については名詞((6a)のパターンに対応)と形容動詞の2つの統語範疇が存在し、それぞれが別の意味を有するために多義性が生じるということが考えられる。

なお(6b)の構造が形容動詞であり名詞句+コピュラの「だ」ではないということは次のことから例証される。例えば名詞句+コピュラの構造をもつ「山だ」のような表現は「山に」の形で副詞として機能することはない。「山に」が文中に現れるときはこの「に」は格助詞である。一方「以上に」や「以外に」は副詞として機能することができ、したがってこれらの表現は名詞+コピュラという構造ではないといえる。

ある語彙について名詞と述語の分化が生じることには例えば「の」について次のような指摘がされている[Saj93, Yan93]。「の」には体言化の用法に加えて「のに」「ので」の形で接続助詞の用法がある。「のに」「ので」は、次の文のように目的と逆接の用法で用いられるが、これらは「の」が述語に転用されたものから生じたとされている。

- (7) a. 引越すのに / ので人手がいる(目的)。
- b. 雨が降っているのに傘がない(逆接)。

問題は「の」についてなぜこのような名詞と述語への分化が生じたかであるが、これについては次のように説明される。「の」は普通の名詞句に比べてその意味の抽象度が非常に高い。すなわち「の」はその意味対象(外延)を限定する力(内包)が希薄であり、実に様々な対象を意味しうる。名詞句の機能が対象を指し示すこと、すなわち外延を絞り込むことにあるとすれば、この点において「の」は名詞性が希薄であり、このために述語への転用が生じたのである。

ここで(1)の単語群を眺めてみると、すべて抽象度の高い名詞ばかりである(「だけ」は現在取り立て助詞であるが、もともとの語源は名詞「丈」)であるとされ

る。すなわち(1)の語彙はすべてもともとが名詞であり、これがこの節の表題を抽象度の高い名詞の多義性とした理由である。(1)の語群が名詞句または文に付着して使われるというのは、その広い外延を前接語句によって限定するためであり、これらの語彙について述語への転用が生じたというのは十分考えられることであると思われる。なお、名詞句または文に付着して使われる名詞すべてについて名詞と述語への分化が起こったわけではない。次の語などは名詞としての用法だけをもつ。

- (8) とき, 場合, たび, 以下, 以内, くらい。

逆に述語に転用されるうちに名詞としての機能をほとんど喪失してしまったと考えられる語彙として「よう」「同様」「たび」などがある。「よう」は名詞「様」を語源とするが、現在ではもっぱら助動詞として使われている。「たび」は形式名詞とされているが、「たび」が何らかの格に入る用法はなく、「たび」単独であるいは「たびに」の形で連用修飾する用法しかない。「たび」については、述語への転用が起こったのち名詞としての用法が失われ、さらに述語としての「たび」についても文の述部となる用法が失われ、連用修飾としての用法だけが残ったものと考えられる。

3 統語範疇について

前節では、(1)の語群が名詞としての用法と形容動詞への転用から生じた用法をもつという議論を行なった。これらはそれぞれ別の辞書項目として辞書に記述することになるが、名詞としての用法は名詞という統語範疇で記述すればよいとして、形容動詞から転じた用法についてはどのような統語範疇として記述すべきかということが問題になる。これについては次の2つの記述方法がある。

1. 形容動詞として記述する。
2. 語幹+「に」でひとまとまりにして接続助詞として記述する。

実際にはこのいずれかの記述方法を各語彙に応じて使い分けるとするのが適当であろう。その語彙について、述語から転じた用法が実際に主文の述語として用いられるかどうかがこの使いわけの基準となる。例えば理由および逆接の「のに」「ので」あるいは「たび」などは主文の述語になる用法をもたないので、これについては接続助詞として辞書に記述するのがよいと思われる。これに対して「以上」などは連用修飾する用

法と主文の述語になる用法の両方をもつので、これについては形容動詞語幹として記述するのが適当であろう。「以上」を形容動詞語幹とすると、連用修飾の「以上」と主文の述語となる「以上」は同一の辞書項目で済ませることができるので効率的である。

4 「以上」の辞書項目

2節で(1)の語彙は本来の名詞としての用法と述語として転用された用法をもつと述べた。さらにこれらの語彙は、それが名詞句に付着する用法と文に付着する用法をもつ。したがって単純に考えてもこれらの語彙は4通りの用法をもち、これらの語彙1つに対して4つの辞書項目が必要である。では実際のところこれらの語彙1つにつきどの程度の数の語彙項目が必要であるかを本節では「以上」を例にとって考えてみる。

2節で、「以上」には名詞句であり‘個体 → 個体’という関数である「以上」と、形容動詞語幹であり‘個体 → 値’という関数である「以上」があると述べた。すると‘値 → 値’と‘値 → 個体’という関数である「以上」も存在するのではないかと推測されるが、実際にこれらの「以上」は存在する。(9)の「以上」は‘値 → 値’の関数である。この「以上」には、「を」や「で」のような距離格、速度格の格助詞が付くことができるが「に」が付くことはできない。すなわちこの「以上」の統語範疇は名詞であるといえる。

- (9) a. 300km 以上を走る (距離).
b. 300km 以上で走る (速度).
c. *300km 以上に走る.

一方‘値 → 個体’という関数の「以上」であるが、これは(10)のような文にみられる。いうまでもなくこの「以上」も名詞である。名詞に付く「以上」はこのように4種類あり、まとめると表1となる。

- (10) 18歳以上が免許を取得できる。

このほかに用言に付く「以上」がある(11)。

- (11) a. 花子は太郎が働く以上に働いた。
b. 花子は太郎が稼いだ以上を使ってしまった。

この用言につく「以上」も‘事態 → 値’の関数で形容動詞であるもの(11a)と、‘事態 → 個体’という関数で名詞であるもの(11b)の2種類がある。ここまでの「以上」はすべて2つの値の順序関係ということをその意味の中核としてもつが、これらとは全く別の種類の「以上」として次の文にみられるような「以上」がある。

- (12) 太郎が来た以上、ぐずぐずしてはいられない。

この「以上」はいわば主文の前提もしくは背景情報を表す「以上」である。このように7種類の「以上」が存在し、これらはそれぞれ別の語彙として辞書に記載する必要がある。

統語範疇	隣接する名詞句の意味	全体の意味
名詞	個体	→ 個体
	値	→ 値
	値	→ 個体
形容動詞	個体	→ 値

表 1: 名詞句に付く「以上」

5 曖昧性解消の方針

前節で見たように「以上」については7通りの辞書項目が必要である。文理解システムは、入力文に含まれている「以上」がこのうちどの「以上」であるのかを決めることができなくてはならない。この曖昧性の解消のためには2種類の情報が必要である。1つはそれぞれの「以上」についての前接する要素が名詞であるか用言であるかの情報で、これは統語的制約として容易に記述できる。もう1つ必要な情報は、「以上」に前接する要素が値であるのかそれ以外の個体を表すのか、および主節の述語が補語として値をとるのかそれ以外の個体をとるのかといった知識、すなわち選択制限 [NM92] である。

このほかに「に」という表現についてはもう1つ考慮しなければならない曖昧性がある。「に」には格助詞と形容動詞語尾という範疇のほかに並立助詞という統語範疇があり、これに関する曖昧性も考慮しなければならない [MN94]。例えば次の文の「に」には、形容動詞語尾と並立助詞の2つの解釈の可能性がある。

- (13) 太郎以外に花子が来た。

「に」が形容動詞語尾と解釈されたときには、この文は太郎と花子が来たと言われる。また、これが並立助詞と解釈された場合には、この文は太郎以外の人間と花子が来たと言われるはずである。ところが実際には前者の読みしかされない。「に」については、おおよそ格助詞、形容動詞語尾 > 並立助詞という偏好があり、このような曖昧性を解消するためにはシステムが制約型の情報ばかりでなく偏好型の情報を扱うことができる必要があるが、本稿で実験に用いたシステムは制約型の情報のみを扱うことが可能であり、この偏好の導入は本稿では論じない。なお、(1)のどの語彙に関する曖昧性でも、こ

の偏好の情報と、補語に関する意味的制約を用いれば解消することができる。

6 クラス素性 / 概念辞書

選択制限を記述するためには、各言語表現について意味素性を記述する必要がある [NM92]。本稿では、このための意味素性として、各表現についてクラスという素性を与える。クラスについては次のように定義する。

定義 1 表現 X のクラスが A であるというとき、 X の意味の対象 $\llbracket X \rrbracket$ すべてについて $\llbracket X \rrbracket \subseteq A$ が成り立つ (表現が比喩的、慣習的に用いられる場合を除く)。

このような定義をすると、例えば名詞「男」のクラスは男、人、哺乳類、生物、具象物と多数存在する。辞書項目には名詞のクラス素性としてこのようなクラス群のうち最小のものを記述する。名詞「男」の辞書項目から意味素性を抜き出すと次のようになる。

$$(14) \left[\begin{array}{l} \text{意味:} \left[\begin{array}{l} \text{クラス: 男} \\ \text{型: 対象} \\ \text{属性:} \left[\begin{array}{l} \text{参照名: 男} \end{array} \right] \end{array} \right] \end{array} \right]$$

名詞句だけでなく、用言についてもクラス素性を定義する。用言に対するクラス素性は、連用修飾を扱うために必要である。われわれは JPSG [Gun87, Gun94, JWG] 上で形容詞基本連用形 (ク形) と形容動詞二形による連用修飾を次のように定式化する。¹ 例えば「明るく光る」という表現に対しては次のような素性構造を与えることにする (ただし意味に関する素性のみを抜粋)。

$$(15) \left[\begin{array}{l} \text{意味:} \left[\begin{array}{l} \text{型: 修飾} \\ \text{修飾: X} \\ \text{修飾方法:} \left[\begin{array}{l} \text{分類: 制約} \end{array} \right] \\ \text{意味主辞: Y\#} \left[\begin{array}{l} \text{関係名: 光る} \\ \text{対象: Z} \end{array} \right] \end{array} \right] \\ \text{意味制約: X\#} \left[\begin{array}{l} \text{関係名: 明るい} \\ \text{対象: Y} \end{array} \right] \end{array} \right]$$

連用修飾文で意味的に中心となるのは、主文の意味と、連用修飾語と主文との関係の2つである。前者は意味的な主辞として、後者は意味制約素性として記述するこ

¹ 活用形名については JUMAN [MKU+93] の活用形辞書での呼び方に準じた。ただし形容動詞基本連用形についてだけは二形と呼んでいる。連用修飾にはこの他に、形容詞、形容動詞、動詞のテ形による連用修飾がある。これは基本連用形による連用修飾とその意味を異にし、本稿では対象としない。

とにする。これは主文の意味素性が連用修飾語のある意味役割に収まった形をとる。この構造は、連用修飾語の下位範疇化補語の意味素性 = 連用修飾語の slash の意味素性 = 連用修飾語の依存素性 = 主文の意味素性というパイプを通して形成される。このような素性構造の形成の仕方は JPSG での連体修飾の扱い [JWG] と同じ構造である。つまり連用修飾に対し (15) のような意味素性を仮定することによって連用修飾と連体修飾を同じ構造で扱うことができる。

このように連用修飾については連用修飾語の補語に主文が収まる形をとるので、「*(光度に対する関係を表す意味での) 明るく走る」のような非文を排除するために、連用修飾語となる用言については用言に対する選択制限も記述しなければならない。用言のクラスは事態あるいはその下位クラスであるとする。事態の下位クラスとして例えば「走る」というクラスが考えられる。これは Hashida [JWG] の解釈にならない、何かが走るというイベントの集合であると解釈する。こう定義すると述語「走る」をもつ文の記述対象はクラス「走る」の部分集合であるということになり、定義1を満たす。

2つのクラス A, B の間に上位下位関係 (isa) を定義することができる。

定義 2 クラス A, B について $A \subseteq B$ ならば $\text{isa}(A, B)$ 。

Isa を用いて、例えばある名詞句が値を表すという制約は、名詞句のクラスを X とすると $\text{isa}(X, \text{値})$ と書ける。形容動詞の「以上」の場合には下位範疇化補語、隣接補語とも値ではないが、これは図1のように記述する。

実世界における isa 関係のような知識は、概念辞書 [Nom89, UK89] としてシステムにもたせる。これには isa だけでなく、ある対象がある性質をもつという関係 (is_property_of)、ある性質がどのような値 (定義域) をとるかという関係 (is_domain_of) などが記述される。

7 「以上」の意味

直観的に「以上」という語については2つの値の間の順序関係がその意味の中核としてあると考えられる。ところが4節でみたように「以上」がその補語として値を取るとは限らず、個体を表す名詞句や事態を表す用言を補語としてとる「以上」が存在する。これらの「以上」を理解するには、その個体や事態を値に結びつけるような推論が必要である。例えば (16) に現れるような形容動詞語幹である「以上」の意味について考えてみる。

(16) 太郎は次郎以上だ。

この文は直観的に、太郎と次郎それぞれに関して値 α, β が存在し、その間に $\alpha \geq \beta$ という関係が成立するという意味であると考えられる。つまり形容動詞語幹の「以上」の意味は (17) のように記述できる。²

$$(17) \text{sem(以上)} = \lambda X \lambda Y. (R_1(X, \alpha) \wedge R_2(Y, \beta) \wedge \alpha \geq \beta).$$

α と β に関しては、その間に順序関係が成立するのであるから、同じ種類の値でなければならない。すなわち α と β のクラス素性は同一でなければならないという制約が存在する。(17) で R_1 は個体 X と α 、 R_2 は Y と β を関係付けるものである。これらの関係が具体的にどのようなものであるかは文脈を参照しなければ決定することはできない。しかしこれらの関係がどのようなものかをある程度限定することは可能である。 R として、直観的に次の2通りの可能性が考えられる。まず、 X が P という性質を有し、 α は P に関する値であるというケースがある(例えば X として太郎、 P として身長、 α として長さを表す値)。このとき R は次のように書き下せる。

$$(18) R(X, \alpha) \rightarrow \text{is_property_of}(P, X) \wedge \text{is_domain_of}(\alpha, P).$$

あるいは X がある事象の意味役割の1つであり、かつその事象が意味役割の1つとして値をとる場合がある。この場合には R は次のように書ける。

$$(19) R(X, \alpha) \rightarrow \text{is_role_of}(X, \sigma) \wedge \text{is_role_of}(\alpha, \sigma) \wedge \text{isa}(\alpha, \text{値}).$$

α, β のクラス素性が同一制約や (18)(19) のような意味的制約は、意味的な非文の排除に利用できる。

8 ゼロ述語補完

このような推論は「以上」だけでなく、(1) のうち名詞句に付着するような語彙すべてについて必要である。なぜならば、このような語彙が名詞句に付着して使用される場合には、その解析の際にゼロ述語補完が必要となるためである。例えば次のような文についてゼロ述語補完が必要である。ただし ε はゼロ述語を指す。

(20) a. 太郎 ε 以外に花子が来た。

²4で7種類の以上が存在すると述べた。それぞれの「以上」について(17)のようにその意味を記述することができるが、このうち背景情報を表す「以上」を除くすべての「以上」について、その意味に $\alpha \geq \beta$ という項が共通して存在する。先に順序関係が「以上」の意味の中核であると述べたのはこのことである。

b. 太郎 ε 以上に次郎が強い。

ここでわれわれは例えば (20b) の文は次の深層構造から派生したと主張するものではない。

(21) 太郎が来た以外に花子が来た。

しかし (20b) のような入力文を受けたときに、(21) に相当する意味素性を出力することは、文理解システムとして必要であると考えられる。

ここで、どうやってゼロ述語 ε を復元するか、すなわち ε に相当する素性構造を形成する方法が問題となる。1つの方法として、統語的な構造を手掛かりにして実際に主文の意味的要素をコピーしてゼロ述語補完を行なう方法がある。しかしこの方法には問題がある。なぜならば「以上」がその意味として従属文述語と主文述語の間の何らかの統語的同一性(たとえば関係名を共有する)を要請していると仮定することになるからである。しかし「以上」がこのような要請をもたないことは、次のような文の存在によって明らかである。

(22) 太郎が上手い以上に次郎が強かった。

したがってゼロ述語補完は統語的操作ではなく純粋に意味的推論によって行なわれると考えられる。

本稿では概念辞書を選択制限などの制約の供給源として用いるだけでなく、これをゼロ述語補完に積極的に利用することを考える。「以上」に隣接する名詞句のクラスと、値のクラス—これは「以上」の意味すなわち2つの値のクラスが一致しなければならないことから与えられる—がわかっているので、概念辞書を用いて具体的にその2つの対象の間の関係を推論することができる。このとき概念辞書に、2つのクラスの対象とその間の関係が記述してあれば、その間に成立する関係を具体的に求めることができる。具体的には $\text{rel}/3$ という関係を用いてこの推論を行なう。 Rel の第1、第2引数は2つの意味的对象である。第3引数としてその間に成立する関係、例えば (18)(19) における R を返す。このようにしてゼロ述語の候補を求める提示することができるが、これだけでは候補の絞り込みに不十分であり、centering[Kam86]などの手法と組み合わせて用いることが必要である。

9 処理システム

実験に用いたシステムは、以下の要素からなる。

- 形態素解析システム JUMAN[MKU+93].
- 構文解析システム SAX[MDU93].

見出し語:	以上だ	
主辞:	品詞: 用言 品詞下位分類: 形容詞 活用型: ナ形容詞	
意味:	型: 事態 関係名: 以上 値 1: Value1 値 2: Value2	
意味制約:	型: 事態 関係名: Prop 対象: SubSem 値: Value1#[クラス: ValueClass] Rel	
下位範疇化:	意味: SubSem#[クラス: SubSemClass] 意味役割: 対象 格: が 品詞: 後置詞 文法関係: 目的語	
隣接:	主辞: 品詞: 名詞 意味: AdjSem#[クラス: AdjSemClass]	
prolog_call(nonvar(AdjSemClass),not_isa(AdjSemClass, 値)), prolog_call(nonvar(SubSemClass),not_isa(SubSemClass, 値)), prolog_call(nonvar(SubSemClass),is_property_of(Prop,SubSemClass)), prolog_call(nonvar(Prop),is_domain_of(ValueClass,Prop)), prolog_call(nonvar2(ValueClass,AdjSemClass),rel(Value2, Value Class, AdjSem, AdjSemClass, Rel))		

図 1: 形容動詞「以上だ」の語彙項目

- 制約解消システム trans[Mor94].
- IPAL 辞書 [IPA87, IPA90].
- 概念辞書.

処理の流れは以下の通りである。まず入力文について JUMAN を用いて形態素解析を行ない、入力文を各単語に切り分ける。各単語についての情報を IPAL から取り出す。取り出した各単語についての構文的情報に基づいて SAX による構文的解析がなされ、文の構文的構造が得られる。SAX に与える文法として JPSG[Gun87, Gun94, JWG] を用いる。

言語表現は素性構造 + 制約という形の情報構造に対応付けられる。例えば形容動詞語幹である「以上」の辞書項目は図 1 のようになる。素性構造以外に制約部分を設けたのは、例えば JPSG の下位範疇化原理のような制約や isa のような制約が素性構造内部では表現できないためである。構文解析時において、このよう

な構造のうち素性構造の部分に関する処理は SAX によって行なう。制約部分に関する処理は 2 段階に分かれており、trans は制約の逐次解消を、prolog 処理系は制約の遅延実行を行なう。基本的な制約解消を行なう trans に加えて制約の遅延処理系を設けたのは、概念辞書に関する制約を処理するためである。例えば図 1 の「以上」の isa や (17) のような制約解消のためには概念辞書の参照が必要である。逆にこのような制約は概念辞書を参照することによって解消可能であるがゆえに、この制約を trans で処理しようとする、trans はこの制約について逐次解消、すなわち $isa(X,Y)$ を満たすようなすべてのインスタンスを求める。しかしこの制約解消は処理時間を増大させるばかりでなく、本質的に不要である。この制約解消は、例えば (23) について太郎が次郎以上であるのはどのような値についてかを推論することに相当する。

(23) 太郎が次郎以上だ。

しかしこの文を人間が読むとき、このような推論を実際に行なっているわけではない。すなわち R に関する制約は解消されているわけではなく、かつ解消されないままでも十分に有意な情報を伝えている。したがってこの制約は直ちに解消されるべきものではなく、文脈によって具体的な値が与えられた時点まで解消が遅延されるべき制約である。図1の制約の中で、イタリックで示した部分が実質的な意味をもつ。 `prolog_call(nonvar(X),...)` は、 X に値が割り当てられるまで概念辞書の参照を遅延実行させるためのものである。

このようなシステムを用いて「以上に」を含む各文についての解析を行なった。概念辞書を用いることによって出力の候補を1ないし2に限定できる。出力解が2通り出てくる文は、(3)のように人間が読んでも曖昧な文である。

10 おわりに

本稿は抽象度の高い名詞について特にみられる「に」に関する曖昧性の解消を選択制限を用いて行なった。選択制限の記述に必要な概念辞書を用いて、これらの語彙の処理に必要である推論による意味出力およびゼロ述語補完を行なうことができる。「に」に関する処理にはこれ以外に選好型の知識が必要であろうと考えられ、選好型の知識の処理をシステムに導入することは今後の課題である。

参考文献

- [Gun87] Takao Gunji. *Japanese Phrase Structure Grammar*. D. Reidel Pub. Co., Dordrecht, 1987.
- [Gun94] 郡司隆男. 自然言語. 情報数学セミナー. 日本評論社, 1994.
- [IPA87] 情報処理振興事業協会技術センター. 計算機用日本語基本動詞辞書 IPAL(Basic Verbs), 1987.
- [IPA90] 情報処理振興事業協会技術センター. 計算機用日本語基本動詞辞書 IPAL(Basic Adjectives), 1990.
- [JWG] JPSG Working Group. JPSG: A constraint-based approach to Japanese grammar. ICOT Technical Memo, Institution for New Generation Computer Technology, in preparation.
- [Kam86] Megumi Kameyama. A property-sharing constraint in centering. In *24th Annual Meeting*, pp. 200-206. ACL, 1986.
- [Kog93] 小暮潔. 素性構造 (1). 人工知能学会誌, Vol. 8, No. 2, pp. 184-191, 1993.
- [MDU93] 松本裕治, 伝康晴, 宇都呂武仁. 構文解析システム SAX 使用説明書 version 2.0. 京都大学工学部 長尾研究室, 奈良先端技術大学院大学 松本研究室, 1993.
- [MKU+93] 松本裕治, 黒橋禎夫, 宇都呂武仁, 妙木裕, 長尾真. 日本語形態素解析システム JUMAN 使用説明書 version 1.0. 京都大学工学部 長尾研究室, 奈良先端技術大学院大学 松本研究室, 1993.
- [MN94] 武藤伸明, 中川裕志. 文接続助詞「に」に関する考察. 情報処理学会第48回(平成6年前期)全国大会, pp. 3-169, 170, 1994.
- [Mor94] 森辰則. Prolog プログラム変換に基づく否定情報の扱える制約システム. 自然言語処理研究会報告 94-NL-101-2. 情報処理学会, 1994.
- [NM92] 長尾確, 丸山宏. 自然言語処理における曖昧さとその解消. 情報処理, Vol. 33, No. 7, pp. 746-756, 1992.
- [Nom89] 野村浩郷. 自然言語理解の構造 — 理解の表現. 情報処理, Vol. 30, No. 10, pp. 1161-1168, 1989.
- [Saj93] 佐治圭三. 「の」の本質 — 「こと」「もの」との対比から. 日本語学, Vol. 12, No. 10, pp. 4-14, 1993.
- [Smo88] Gert Smolka. A feature logic with sub-sort. LILOG Report 33, IBM Deutschland, 1988.
- [UK89] 浮田輝彦, 木下聡. 自然言語理解のための知識表現と推論 — 知識表現. 情報処理, Vol. 30, No. 10, pp. 1224-1231, 1989.
- [Yan93] 柳田征司. 「の」の展開 — 古代語から近代語へ. 日本語学, Vol. 12, No. 10, pp. 15-22, 1993.