

TM-0995

エキスパートシステムにおける自然言語
インタフェースについて
(日本語文法-SFTB-の枠組と
その開発支援環境-LINGUIST)

佐野 洋

December, 1990

© 1990, ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03)3456-3191 ~5
Telex ICOT J32964

Institute for New Generation Computer Technology

エキスパートシステムにおける自然言語インタフェースについて

日本語文法-SFTB-の枠組とその開発支援環境-LINGUIST-

佐野 洋

(財) 新世代コンピュータ技術開発機構

E-mail: sano@icot.or.jp

平成2年12月1日

はじめに

自然言語処理技術はワードプロセッサの普及をはじめとして機械翻訳の商用化に見られるようにしだいに我々の日常に浸透しつつある。このように自然言語の扱いは機械処理の可能性が見出されて以来、その言語自身の処理技術の発展だけでなく適用技術の発展にもない応用の可能性も議論されてきた。さらに実システムの開発も試みられてきた。当研究会においても以下の表の示すように自然言語処理を取り巻く技術について少なからぬ研究活動が散見される。

表 1: 活動状況

1985	6/21	田中穂積	「ブログにおける自然言語処理とその応用」
1985	7/26	村木一至	「統合自動翻訳システム PIVOT について」
1985	12/20	田中穂積	「知的インタフェース-自然言語処理-」
1987	11/3	横井俊夫	「電子化辞書について」
1988	2/1	内田 謙士	「電子化辞書における概念辞書について」
1988	2/1		「法的概念辞書の格(関係詞)の検討」
1988	12/20	橋田浩一	「談話理解実験システム DUALS について」
1989	1/23	牧野武則	「自然言語理解のための概念の表現について」
		ディスカッション	法的概念辞書の検討
1990	2/20	野村浩郷	「自然言語理解について」
1990	6/9	田中裕一	「談話理解と自然言語インタフェース」

このほかエキスパートシステムへの機械処理の応用という立場をみると自然言語処理だけでなく、論理プログラミングとかかわりあいも多く議論されている。本稿は、エキスパートシステムのインタフェースとして機能すると期待される自然言語処理技術について、工学的なアプローチよりもことばの基本的性質やことばを基本的に律する文法の構築という立場から管見の範囲ながら幾つかの話題について報告する。

自然言語の機械処理技術あるいは、辞書の概念構成方法、あるいは論理表現とのかかわりなど、工学的アプローチの幾多がすでに報告されている。そこで、本稿は、自然言語処理における話題の重複を避ける目的から、ことばの基本的な性質について概観した後、日本語という言語の性質を手短かに述べる。その後、機械処理を可能とする文法枠組について述べ、その開発環境についても若干の説明を加える。従って工学的アプローチの色彩が幾分か薄れ言語現象からの説明がなされることを予め承知いただきたい。

そうして、言語を処理することは何をすることなのか、どのように利用することが効果的であるかについて何らかの得るところがあることを希望する。さらに、現在の自然言語処理技術のかかえる問題点を指摘し、エキスパートシステムへの応用の可能性の有無あるいは難易を議論したい。

また、私の立場からは、これらの議論を通じて我々自然言語の処理を目指すものにとって大きな関心であるその適用領域における諸々の問題点について意見を伺えることができれば幸いである。

尚、説明の至らぬ点については随時問い合わせ頂こうお願いする。

1 ことばの法則

1.1 機能

言語記号から成ることばは、分節の機能によって概念や観念を示しその持つ意味を表象しうる。記された形と意味の間にその成因を必然とする関係はなく、すこぶる恣意的に対応が存在する。形態素はこれら概念や観念を示す標識に使われ、形態論的手続き - 語を構成するための手続き - より事物を表す。例えば「法律」という言語記号と概念としての「法律」の必然的な関係は基本的にない。表音文字を採用する言語 - 英語やハングル語 - においてこれは顕著な性質である。表記文字を採る言語では個別の文字の持つ概念にしばしば影響を受けることもあるが、基本的には必然とされる関係はない。

またことばは生産的でもある。接辞は生産的な付加的形態素として機能して一群の語を構成するほか、さまざまな語を臨時に作りだす。例えば「法」は「司法」「立法」のようにすでにその生産的性質が意識に昇らない語をはじめ「道路交通法」など身近にある。実は「山」も形態素の分析によれば「や」と「ま」に区分される。「やま」「はま」「すま」などを比較することでその語彙の作られた歴史をかいま見ることができる。これは語彙論という言語の範疇に属する。

ことばの機能には表現された各要素に存在する関係を示す働きがある。語彙に示されるように意味まとまりを示すための形態素があることに加え、関係を示す形態素がある。意味の観念や概念の表示に加え、形態論の手続きは関係の観念を示す。例えば「罰則を適用する」という文の「罰則を」が「適用する」に対して対格の位置を占めることは「を」という形態素の機能による。基本的に「司法」「立法」「道路交通法」という語彙中における「法」という形態素とは違う機能である。

ことばの二重分節といわれる機能はこれを指す。語彙論に対置し、統語論の領域が形態論の手続きの中に展開される。ここに統語論上のカテゴリーが見い出される。いわゆる秋葉の文法はこの現象の分析を指すことが多く記号列の中の語彙の関係構成の類型をその研究の対象としている。格、態やアスペクトといった要素間の関係とムードや遂行の類型であるところに表現形式にかかわる。形態論は概念にせよ関係にせよ表示のための手段となり、語彙論のための形態論があり、統語論のための形態論がある。

その意味で不幸にも日本では形態素解析と構文解析という対立区分があり多くの誤解を招いている。つまり形態素解析の後に構文解析がなされるという言語処理のフローはいわばミスリーディングに近いといえよう。

分節に与ることばは恣意的に定められた規律のもとに運用される。この規律はことばの法則であり、すなわち文法なのである。この意味で、文法は語彙が示す内容と語彙間の関係による表現を律する法則である。弁別することで自らの意味内容を示し、対立によって自己の表現の類型化を行う。語彙論においても統語論においても形態素は多価値である。コミュニケーションの手段たることばには規則性が存在するが、運用の効率の面から形態論的手段において不均衡で例外的な手続きを残す。文法とは、ことばの法則であり恣意的に定められていわば必然的に運用される。

1.2 文法の体系

一つの言語の体系はいくつもの要素から組織されている。一般には、音韻論、形態論、統語論、意味論をその構成要素として挙げるが、前節にあるように統語論と対比するのは語彙論である。音韻論も対比させられる要素であるけれども言語記号の面からは、独立させることができる。意味的機能を含まない音韻論は他と区分されることが多い。

言語学に厳密な科学の方法を持ち込むことはできない。だからといって、例えば「言語学に関する要素が互いに支え合い体系をなしている」という説明ではあまりに抽象的である。確かに、文法という法則には、それが見出された後にも異義をさしはさむ余地がある。いわゆる自然科学における定理と大きく性質を分かるところである。

言語学における法則は言語事実から見極める。精緻な観察と抽象化を経て編まれた法則は、再び事実を照らして検証されなければならない。ことばは、わたしたちに客観的な基礎のうえに立って事実を観察してゆく努力を求めているのかのごとくふるまっている。

説明的記述に基礎を置く文法-SFTB-、すなわち我々が開発している日本語文法枠組は、日本語をその分析対象としたことばの法則である。日本語における弁別と対立の手続きの所在と機能を見つめ、その法則を明らかにしてゆく。その分析は形態を中心に外形特徴に言及し、文を構成する要素の機能を鑑みて文の法則を捉える。従って、意味的な解釈過程を暗黙に仮定すると、その意味は単純であるような文が、実は複雑な文として位置付けられることがある。形態論的手続きを手がかりに分析と評価をするのである。同時に、これは、大きな危険も侵している。というのも、本来、ことばは音声において伝達される事実である。さらに発話の状況や記述の文脈を捨象することへの警戒心

をもたなければならない。ことばの運用という全体的な場の中で、場を構成する文法や発話の状況といったメンバー間には絶対的な序列関係があるわけではなく、場に応じて相対的な関係で結ばれている。言語記号を使った言語行動に加え、音調やしぐさ、表情といった非記号の信号がある。言語記号の列の意味は、ある解釈の体系のもとに決定されることは前述したが、実は、それ以上に伝達者の非記号的な行動と、受動者の感受性の強さや見識の深さに左右されるのである。従って、全体の中での占める位置がその役割によって変化するといえる。このような点を踏まえ敢えてことばの見えるかたちに注目していることを意識として持つ必要がある。

1.3 法律文への概観

法律文や特許文は自然言語処理の歴史において比較的初期の頃から、その適用の可能性が議論され試みられた文類型に属するのではない。すなわち言語記号の運用において周囲の状況の影響をできるだけ排した、あるいは排することを旨とした文となっていると思われるからである。詳細な分析の経験がないから軽々な記述を避けるべきであるが、法律文を機械処理することや法律文の示す表現を抽象化することに科学の道具が使われることは当然であるようだ。そこに現在の自然言語処理技術の応用の一側面が見出されるようだ。

この問題については意見をうかがいたい。

次の章では我々の文法-SFTB-が仮定する基本的な文の構造について説明し、日本語の文が統語的にどのような記法構造をもっているかを概観する。

2 文のかたち

2.1 仮説

ある形態列を眺めて文だと考える研究者がいる。研究者の石頭が文を創りだしているとすれば、余計なのは研究者の石頭である。恣意的であることは個人に属するものの、運用においては必然である以上、何かしら「文」という文法上の単位の認定が求められる。では、一体何を具備するものを文と仮定するのか。描写が中心となる日本語では、述語が主要語となり、しかも文末に位置するという統語特徴が知られている。では、動詞や助動詞などの述語の終止の形をもって文という単位を認めるのだろうか。終止の形にそれほどの負担を負わせてもよいものかいさきか心もとない。例えば、用言の終止形は近代以降に造られた形であるし、“。”は過去において period に対応するよう設定された経緯をもつ。

SFTB では「終止の形と“。”によって終止して纏まる単位を、文法上の纏まりとする」という立場を基本的に避ける。流通する文法教科書では、その規範として文の単位の認定の要として強要しているが、造られた機能や形には自ずと限界がある。SFTB では文という単位を形態の弁別からの要請だけで認可することはないのである。

文は多くの情報を伝達する。基本的にはさきほど述べたように語彙に属する意味によって基本的な概念を示す。そうして各々の概念はある関係において結ばれる。これは統語構造と呼ばれるもので表される。概念と関係の構成において文の内容は成立する。しかしながらこれまでの自然言語の研究から得られた重要な事実、文はその内容がそれを出発する話者の意志のもとにおいてのみ成立するのであるということである。

いかなる文もこれを例外としない。

例えば「佐野」ということばを発する際に、戸口から現れた人を指して言う場合には「佐野だ」という意味かあるいは「佐野か」という意味の表現は文脈に依存する。しかし、いずれにしても話者の意志が現れる。このように通常、文とはその意味内容が問われることが多いけれども、実は話者の意志が常に反映されているのである。その意志の類型は形態に示されたり、表情で実現されたりあるいは態度に現れることがある。

フォーマルな疑問文は、通常「か」の文末への付与によって実現されるけれども、日常会話ではイントネーションによる実現が専らである。さらに「か」の付与は疑問や質問の意味だけでなく反語や勧誘の意味をもつ。多語義が語彙の分析の難易度を示す例でしばしば引き合いに出されるように、語彙論においても統語論においても形態素は多価である。

これは言語効率の問題である。意味の伝達において、つまり、話し者-対-聞き手の対応において最大効率での意味伝達が達成されればよいのである。文がすべての意味を包含するという誤解はすでに過去のものとなっている。そして今、我々は話者の意志において文が成り立ち、しかも文の内容が成立することを認識している。

このような視点からの考察は、次のような指摘を生じさせる。論説調の文であるとか法律文のように話者の意志をできるだけ排した、きわめて客観的に記述されている文は、比較的分析しやすく、機械処理が容易であるという、直感への疑問である。話者の意志を表出させることばの運用がもっとも負荷の軽減された使用法であるとするれば、客観的な叙述は統語構造上、無理を生ぜしめている可能性がある。とすれば論理的にものごとを記述し人の読みにおいて難解さを露呈する文は機械にたいしてもその処理が難しいかもしれない。

すくなくとも直感による回答は保留せざるを得ない。
次の章はSFTBにおける文の構成の基本を述べる。

2.2 文の外形と意味との関係

文という単位は何を具備するものなのか。いわゆる希求文や感嘆文は、文の意味のすべてが言葉で表現されるわけではない。希求や感嘆のモードにおいて発せられた文の外形は、それ自身では文として自立できないことが知られている。他方、文の外形によって意味を比較的忠実にあらわすのは、いわゆる平叙文である。事態を描きとることが第一義の目的であるから、事態の認定についての外形の変種を持ちえる。この外形の変種が許されていることが実は重要なことで、平叙文では、テンスの分化が許されている。そして、確認や未確認、あるいは完了や未完了のモードを持つ。白らの心理状態になるだけ中立であろうとする事態の認定の行為である。

しかしながら、この分析はあくまで一般的であって、外形は平叙文の形態を持ちながら相手に対する行動要求を述べたり、時には情報提供を促すことがある。しかしながら上述の表現は、平叙文の意味を示す文の基本形態とは考えない。文の意味を表現するために用意された基本の外形、すなわち語の形態変化や構造についての第一義の分析を求めるためである。それは、個々の発話の状況や個人に属する話者の心理状態に、できる限り依存しないことを心掛けることが要請される。

SFTBでは、まず基本的解釈を設定する。言わば文の外形と意味との関係の類形的な分析を試みるものである。プロトタイプともいうべき型を見出し、プロトタイプの最大投射を求めることがまずもって要請される。様々な文型は雛形からの逸脱において説明されるのである。

まず、文の捉え方を明らかにする。文認定の基本前提は、話者の意志(主体的発話行為)のない所では文は自立性を持たず、文として成立しないということである。そうして、主体的発話行為の意志を文の外形に求め、形態変化による文の外形を文の意味に結び付けて述部の分析とする。当然のことながら、文の中心部に位置する記述は主体的なものではない。文は事柄の記述や関係を伝えるものであるし、簡かに文の中核を構成して、それを叙述する客体的な表現がある。客体表現は有形無形の主体的発話の意図や意志により支えられており、その支えを失うと文として機能しないことを仮説として用いる。

2.3 基本文型

構文の違いや語の形態変化といった統語特徴から基本的に文の区分けを行う。すなわち文に固有の外形と、その外形を持つ文が表す意味の違いで類別を試みるのである。これまでの日本語研究の成果は次のような述部の構造を明らかにしている [1]。

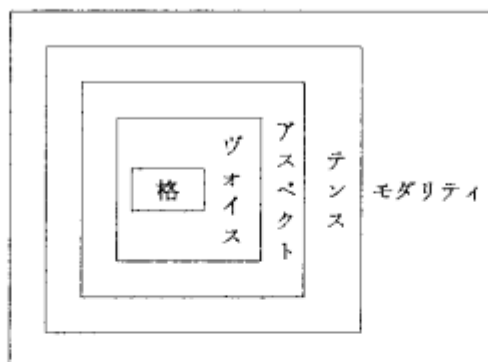


図 1: 文の層状構造「拡大語彙論的統語論」より抜粋

モダリティとは記述の内容には基本的に依存しない。文には発話者が義務的に仮定されるにもかかわらず、言語的にはそれを必然的に示さない。しかし、推量文は、その述部のモード相に発話者に属する形態が顕在化される。命令を示す文も同類である。平叙の持つ多義性は、この発話者の態度が形態に明示的に示されないことに依存する。敢えて、平叙文にも発話者の態度を認め、無形によって示していると解釈するのである。これは言語にとっては一般的なことである。例えば、否定文は否定の形態マーカを顕在させるが、肯定文は通常文と変わるところがない。使役文は非使役と対立するが、非使役文と肯定文の区分はない。また、日本語には直接テンスを示す表現を欠くために、テンス層を第一次の文法範疇として認めるににくい。

基本的に文には基本的に事態に存在する主体と発話主体の二者を認めることができる。発話行為としての文を考えた際に、それが形態的にはほとんど顕在化されないにもかかわらず発話主体の導入は重要である。文の末尾に集中する形態素の多くは、従来、陳述の模様を規定するものとして議論されてきた。近年、発話の力と関連して分析のすすむこれらの形態素は話者の態度が直接反映される。SFTB では発話行為論の成果を受けて、統語表示の中に発話主体の存在を常に想定する。そのために、発話主体に属する認識事態に関する階層を設定し、形態論的手続きにおいて規定される少なからぬ範疇を設定した。

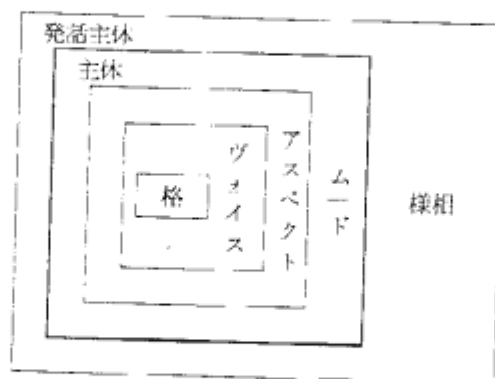


図 2: 述部の階層構造

SFTB において仮定される述部構造は、テンス層を副次的な文法範疇として降格し、ムード層を狭義のムードとして、その対象範囲を事態に属する範囲に限定した。そうして、モダリティを発話者の認識階層へ移行させた。このモダリティの階層には発話を遂行するための諸々の性質が記述される。この点が従来の統語表示と大きく性質を分かるところである。統語表示は拡大解釈され、言語の遂行分析をもそのスコープに含めるのである。形態論の手続きは語彙の認識から、統語的な範疇の表示を行うとともに文の遂行にかかわる部分においても機能している。この事実こそは、統語表示に発話主体にかかわる層の認知と存在を動機づけるものなのである。事態の記述の階層と、事態の認識の階層を区分したもので、複数(基本的には2者)の主体を常に設定している。尚、発話者自身は、規定値として言語化されることはない。一人称としてことばにあらわれると有様の形態として認められる。

2.4 形態文型

前節では、述部の複合階層を仮定した。これは、統語範疇を形態論の手続きに規定される制限に従って階層化したものである。形態論の手続きは統語範疇のすべてを効率や慣用のため顕在化しない。このことから、基本となる形態文型を設定することは困難を覚めるのだが、SFTB においては次のように形態文型を定める。述部の階層構造にできるかぎり忠実に沿った文形態を基本の文型とする。

表 2: 基本形態文型

	文の類型	形態特徴
1	平叙文	ル形-タ形
2	推量文	ル形-タ形
3	意志文	肯定-否定
4	命令文	肯定-否定

個々の文は、構文や形態変化で特徴づけられる外形を伴って、それに応じた意味を表す。効率面から、ことばの形は必ずしもことばの意味を直接に反映しないことは先に述べたとおりである。表 2 は述語の外形特徴と述部の階層構造で示される意味をもって相関させた文の基本類型である。文により表出される叙述内容が、文の外形をもって忠実に述べられることを基本とする。

1の平叙文は、基本的に物事を認識し、確定的な意味で事柄を表現する文である。事柄が自らの裁断の範囲内で顕在し、在ることを主張する。日本語では、テンス、アスペクトという文法範疇が派生的な形態変化として動詞文に認められている。動詞自身は事柄の包括的な確定の力が弱いのである。このかたちは本来的に事柄の顕在の主張が基本

機能である。2は推測の表現であり不確かな事柄の顕在の主張である。事柄を表出しているが、確定の態度は保留されていると言える。文の外形としてル形・タ形の対応を形態上持っていることは示唆的である。つまり、表現している事柄が不確かであるにも関わらず事柄を認識し、確定する傾向が強いのである。

3の意志文は、文の外形の点から肯定・否定の対立関係を持っている。要求表現である命令文に近くて、推量文とは逆に事柄を認定しようとする機能は弱い。時間に依存する形態を第一義に示さないことから、より心的行為の傾向が強いと解釈される。1と2においてのみテンスの分化が許されているのである。4は、文で表出される事柄に対する確定や認定の作用はなく相手に対する事柄の行動要求である。心的な行為が含まれないという点においては平叙文とパラレルな関係がある。3と4は肯定と否定において対立していて、記述の内容に時間の設定は認可されていない。

日本語では、記述の内容と発話者の態度を区分するのは、第一次のムードの値である。これは完了事態であるが、未完了事態であるかの区分によっておこなわれる。テンスの分化の許されない事態は規定値として未完了のムードを設定する。

次の図は基本文型の形態に述部の階層構造を加味し纏めたものである。対立の所在を明確にしている分類で、1と2についてはテンスの分化が許されていて時間的な面での対立を見ることができる。3ならびに4のムードはテンスの分化を許さず発話者の態度において認め方の対立を示している。上記は主節における基本形態を示したものである。従属節についても統語論上の対立関係がある文法範疇において認められる。尚、*は記述の事態について略した表現である。

表 3: 文の基本類型

1 平叙文	*	未完了 肯定 平叙
		完了 肯定
2 推量文	*	未完了 肯定 推量
		完了 肯定
3 意志文	*	未完了 肯定 意志
		未完了 否定
4 命令文	*	未完了 肯定 命令
		未完了 否定

3 表示

3.1 述部階層の概念表示

SFTBでは、文がその文内容としての客体表現とその意志表である主体表現による階層構成をなしていて、述部階層は形態も含めて統語上5つのレベルに区分できる立場を採っている。述部階層表示の値は、各々の階層において形態上の振るまいや統語上の構造が制限されるという文分析の仮定に基づいて導入されたものである。依存関係を間接的に表現する手段である。

一般に文法といえば形態の分析や統語上の解析を問わずそれらのすべてを網羅する枠組を与えるのが普通である。SFTBは形態上の語の構成から文の成立までの過程をまとめ説明しようとする文法である。詳細は説明しないが用言の活用区分や派生語の扱いについても同枠組が適用される。

簡略に各々の階層の機能を紹介する。構文構成よりも語の構成に関する階層がある。LEVEL0,1は最少の意味を有する単位とされる基本の語構成に対応する現象が見られる層である。LEVEL1,2,3,4にかけては複数の関係要素間の関係を記述する。事象間の従属節との関係や他の副詞句との関係がある。LEVEL4,5については、主題-叙述、あるいはその在り方が問われている層であって、取り立てのことばでマークされた語の扱いや接続表現に対応した関係が認められる。LEVEL5に至って文と認められる。

前述の仮定を設けることと各層に現れる要素を列挙してその扱いを制限することで、文法の立場から見た分析の基準を比較的容易に設定できる。

各層における分析対象となる要素を挙げる。

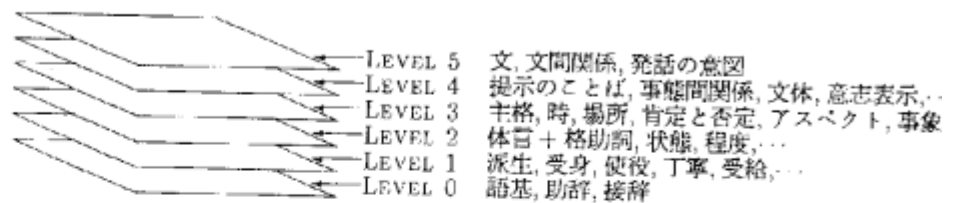


図 3: 述部階層表示

法律文については詳しく言及できないため、いわゆる論説文といわれる比較的客観的な文をその分析範囲として階層のもつ性質を述べたい。

LEVEL5 は、話者の意志が反映される層であり、この層の分析を詳しくすると意志表現を細分化けすることができる。問い掛けや依頼、あるいは意志伝達などの感情表出がみられる。客観的な文ではこの層の役割が比較的小さいように思われる。

LEVEL4 を代表するものは提題と接続関係である。提題は話題と結びつき、どのような適用分野においても比較的多用される。その提題化される対象は、分野ごとに異なるものの、あらかじめ予想されるならば、それをもとに多くの情報を得ることができる。いわゆる文脈構造に多く影響を示す層でもある。接続関係は論旨をつなぐ際にや結論や帰結、原因や結果を明示したいときに運用される層である。形態では「すなわち」「から」「ので」など、いわゆる clue となることばが現れる。接続関係は単に論理関係に限定されない。例えば、料理についての叙述の文であれば、もっぱら接続関係は時間のシーケンスを説明することにある。論説文では因果関係がその注目の対象になるだろう。そうして同じ形態が多くの意味で使用されるところに分析の困難さと自然言語処理技術の特定分野への依存性がある。

LEVEL3 は中心となる事態を示すことばに關係する他の従属する事態の記述をおこなう。レベル4では関係の構成される事態の間は比較的、対等である。この階層では、中心語の概念にたの語が支配されるといってもいい。そうした関係が記述される。おなじ原因・結果の関係でも関与する主体がおなじであることが多い。

「彼が汽車でついたので、ぼくが車で出発した」はレベル4であるけれども「汽車がついたので、降りた」というのはレベル3になる。こうした、事態間の関係構成の従属度の強さは多分に構文に反映される。また、記述する分野に現れる動詞の性質にもかかわるだろう。さらにこの層では時間と場所の修飾がなされる。「いつ」「どこで」「だれが」という記述はこの層に属する。

LEVEL2 では、いわゆる格関係を記述する。もっとも述語表現にマッピングしやすいレベルである。この層を集中的に分析することでいわゆる述語表現の抽出が可能となるだろうし、事実、これまでの文の論理表現の記述はこの層の分析においておこなわれてきた。

しかしながら、文の構成の面から見るとこの階層は説明してきたように文の下層である。こうした下層の表現はしばしば上層の影響を受けやすく、そのため文内容の抽出がその目的であっても適切な内容抽出が難しい。

LEVEL1 は語彙の構成にかかわる層である。ここでは概念構成に關係する現象がみられる。臨時に作られる語彙の多くはこの層での分析において判明する。しかし、現在の知識表現の記述は概念を構成的に作り上げることができないようである。予め概念を固定のものとすれば、この層の分析は不要である。

3.2 適用の可能性

1つの分野にチューニングされた言語処理の手続きは容易にはたの分野へは移植しにくい。かといって、SFTB のように中立性を重視すると適用の分野がないということになる。

歩み寄りの方向は、文法を提供しようとするものは中立性の指摘とある分野への適用例の提示をしなければなら

い。応用する側は、応用分野における語彙で示される概念の特定化、並びに関係記述の限定である。また、意志表現が多用されるか否かの指摘も含まなければならない。

4 実現

4.1 文法の一実現例

前章では文法の階層区分に基づく枠組に着いて述べ、さらに各々の階層における言語特質を論説文を例として説明した。この枠組を実現する一手立てとして、論理文法を応用した。記述の枠組として DCG と呼ばれる論理文法を用いている。すでに多くの機会において説明されていることと思われるので詳しくは述べてない。

その使用の意図は論理型言語と呼ばれる Prolog を用いることで、簡便な使用環境を提供しようとするものである。DCG 記述を使用することで、Prolog の動作する計算機環境であればパソコンと呼ばれる機種から、ワークステーション、あるいは大型機にいたるまでその計算機環境をもっても文法の利用が可能である。特定パーザを仮定しないことは、一面文法の脆弱さを露呈する危険性があるものが、それ以上に多くの人の要求に応じられる答えられる可能性をもつ。

自然言語処理技術を応用する際に、すくなくとも使用計算機の制限からは基本的に開放される。確かに特定分野へのチューニングに対する労力は依然としてこのころが、文法の考えかたの透明性からもわかるように相互の歩み寄りの可能性は低くはないはずである。

4.2 文法的能力と評価

文法的能力の測度や評価の基準についてはまとまった議論がまだないと言っても過言でないだろう。基本的に文法が公開されていないことも原因の1つである。

SFTB を基本とした文法は DCG によって記述されており、これは ICOT の成果の1つである。基本的には要求に応じてリリースされる。そのことは文法的能力向上には極めて重要である。実使用の経験によって文法は改善され、能力が向上することが期待される。さらに、同じ文法の使用による比較評価の土俵が作られることになる。

例えばエキスパートシステムの能力の測度においても多くの問題があろうが、すくなくとも自然言語インタフェース部分の能力差による不当な性能比較の危機から開放されるだろう。(反面、インタフェースの不具合を起因とするシステムの能力が制限されることもあるかもしれない。)

奇妙なことだがこれまで文法的能力の紹介において分析不能である文例が体系的に示されたことがない。分析可能な例については、詳細な説明とまことしやかな解析結果が付随している。SFTB では、分析可能な例文と出来ない例文の区分が明確になっている。そうして使用者は文法的能力を正しく認識できる。

SFTB ではさらに能力の段階化を実現した。例えば、ある能力レベルの文法ではこれこれの言語現象が分析が可能である、ということができし、さらに能力レベルを上げると、さらに分析能力が拡大する、という説明が可能である。文法利用者は適切なレベルの文法を選択することが可能で、能力過剰に起因する曖昧性の増加の問題点の解消の1つの手立てでもある。

5 LINGUIST

文法開発者にとって文法を評価し、動作を観察し検証する環境は、文法を構築する上で大変重要である。我々はこれまで述べてきた文法枠組と実現の手立てに従い、文法開発支援システムを構築している。このシステムは論理型自然言語実験システム-LINGUIST-と呼ばれる。これは直接に、文法使用者の関与するものではないが文法開発のためのサポートシステムの1つとしてここに紹介する。

5.1 動作例

LINGUIST

(c) copyright 1990

Institute for New Generation Computer Technology

ALL RIGHTS RESERVED

ONODERA, Hiroshi + OHARA, Masami + SASAKI, Hiroshi
Executive Consultant TANAKA, Yuichi
Executive Producer SAN0, Hiroshi
Designed by SAN0, Hiroshi

Produced by ICOT 6th Research Laboratory
A LKB TOOL



systems report

JPN ENG

E X I T

Notify all stations. Prepare for use.

Engine

Invoke

Debugger

ANALYSIS				GENERATION			
INPUT	WINDOW	SYSTEM	FILE	INPUT	WINDOW	SYSTEM	FILE
OUTPUT	PRINT	SYSTEM	FILE	OUTPUT	PRINT	SYSTEM	FILE
TOPLEVEL	ON	BU	PROCESS	TOPLEVEL	ON	TDP	PROCESS
REDO	OFF	SUCCESS	!!	REDO	OFF	SUCCESS	!!
TOKEN POOL				STRUCTURE POOL			
【太郎, が, 問題, を, 解, いた】 ◎ 【太郎, が, ゆっくり, 問題, を, 解, いた】				【系統 (ウ系カ), 範疇 (用言語基), 語彙 (解)】 ◎ 【系統 (ウ系カ), 範疇 (用言語基), 語彙 (解)】			
GO SET DEL RESET				GO INSPECT SET DEL RESET			
【系統 (ウ系カ), 範疇 (用言語基), 語彙 (解)】 ◎ 【系統 (ウ系カ), 範疇 (用言語基), 語彙 (解)】				◎ 【太郎, が, ゆっくり, 問題, を, 解, いた】			
INSPECT ADD MOVE CLEAR				ADD MOVE CLEAR			

systems report		JPN	ENG	EXIT	
				Engine	Invoke Debugger

inspector (top menu)	
** exit **	
** quit **	
構文構造	系統 (ウ系カ)
	範疇 (用言語基)
	語彙 (解)
	表出 (A, 叙述)
言語構造	ムード (A, 完了)
	認め方 (A, 肯定)
	アスペクト (A, 非継続)
	{*, *, *}
	inmanner (A, C)
	{*, *, *}
	{*, *, *}
	指標 (A)
	ソート (A, 解く)

LINGUIST Ver. 1.2 [Sentence Processor]			
Bottom up		Top down	Freeze
4130	連用詞_2-->使言_2	格助辞	
3370	用言_2-->連用詞_2	用言_2	
3700	用言_1-->用言_0		
3350	用言_2-->用言_1		
3290	用言_3-->用言_2		
用言_3 連用詞_3 体言_2 格助辞 体言_0 太郎 用言_3 用言_2 連用詞_2 体言_2 格助辞 が 用言_2 連用詞_2 体言_2 格助辞 ゆっくり 用言_2 連用詞_2 体言_2 格助辞 問題 用言_2 用言_1 を 用言_2 用言_1 解 いた			
		step jump goal leap skip redo notrace fail abort inspector	
E-Exit 用言_2 (用言_3)>step E-Exit 連用詞_2 (用言_3)>step S-Exit 用言_3>step EXPAND 用言_3 (sentence)>			
systems report		JPN ENG	E X I T
		Engine	Invoke
		Debugger	

LINGUIST Ver. 1.2 [Sentence Processor]			
Bottom up		Top down	Dic Library
FREEZE			
%%%%%%%%%% % 用言_3 % %%%%%%%%%%		全範囲指定 範囲指定 指定解除	
3290::用言_3 (SYN, X, REL, F, PRO) --> 用言_2 (SYN, X, REL, F, PRO).		追加・修正 削除	
%%%%%%%%%% % 主格 % %%%%%%%%%%		BUP TDP	
3300::用言_3 (VSYN, VX, [(構文関係 (subj, CASE) AdvSYN], AdvREL, AdvF) VREL], [(CASE, subj, AdvX) VAL], RL), (AdvP, PT)) --> 連用詞_3 (構文関係 (subj, CASE) AdvSYN], AdvX, AdvREL, AdvF, (AdvP, VP)), 用言_3 (VSYN, VX, VREL, (VAL, VRL), (VP, PT)), {assignment (VRL, (CASE, subj, AdvX), RL)}.		FREEZE	
%%%%%%%%%% % 副詞句 % %%%%%%%%%%			
3310::用言_3 (VSYN, VX, [attach (VX, AdvX), (構文関係 (副用言) AdvSYN], AdvREL, AdvF) VREL], VF, (AdvP, PT)) --> 連用詞_3 (構文関係 (副用言) AdvSYN], AdvX, AdvREL, AdvF, (AdvP, VP)), 用言_3 (VSYN, VX, VREL, VF, (VP, PT)).			
systems report		JPN	ENG
E X I T		Engine	Invoke
		Debugger	

参考文献

- [1] 久野すすむ、柴谷方良編,「日本語学の新展開」,くろしお出版,1989.
- [2] 福本文代,佐野洋.(1990),「制限依存文法とその表現」,第41回情報処理学会全国大会,1990.
- [3] 佐野洋.(1989),「述部の階層分析と文脈情報」,『談話理解モデルとその応用シンポジウム』,情報処理学会,1989.
- [4] 佐野洋,福本文代,田中裕一.(1990),「説明的記述に基礎を置く文法-SFTB」,ICOITテクニカルメモ, TM-0885 1990.
- [5] 佐野洋,福本文代.(1990),「発話意図と疑問の焦点」,言語理解とコミュニケーション研究会,電子情報通信学会,1990.
- [6] 佐野洋,福本文代.(1990),「局所化した単一化文法とその表現」,第41回情報処理学会全国大会,1990.