

TM-0994

文の様相区分の文型射影への試み

佐野 洋、福本 文代

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

文の様相区分の文型射影への試み

佐野 洋 福本文代

(財) 新世代コンピュータ技術開発機構

概要

文の法的分類は、形態的な切れ続きの側面に基礎を置く文型に必ずしも対応していないことはしばしば指摘される。例えば、シロ・スルナの対立は形態上の構成からは、動詞の命令形と動詞の終止形に否定の終助詞である「ナ」の接承というまったく異なる形態論的手続きによって実現される。しかし、この2つの小さな文型は命令という法的区分に属し、しかも肯定・否定、言い換えれば依頼と禁止という対応をもっている。例えば、疑問文と平叙文とは不確定項目の有無という違いにおいて平行性を持っていて、法的には同族にある文型である。また、語彙区分を見ると、言い切る音が*i*であることと*i*であることの特徴から「ある」が動詞に配置され「ない」が形容詞に区分されている。この2つの語彙は意味の面を考慮すると同族に配されてよいばかりか、区分ける根拠がない。

本稿では、文の様相区分を中心として、その形態上の構成手続きをもとに文型射影することで、従来、形態素解析と構文解析としてその守備範囲に違いのあった文の構成手続きを纏めようとする試みである。その中では、従来の品詞分類にかわり、統語分析には主要範疇を導入し、語彙の基本要素には語幹よりもその構成規模の小さい語根あるいは語基と呼ばれる単位を導入した。文の基本構成の単位を形態的な切れ続きではなく、様相区分に基づく意味分類に置いた。この意味分類は文の成立の要件を文内容の記述態度の顕在化においたもので、極めて認識的である。この認識的な文の構図の中に、意志否定や基本アスペクト、事態の否定、さらに述語の派生ならびに態変化や文体変化などを配置した。文を示す構図は階層的構造を持ち、その射影として上記の文法的側面を説明する形態論上の手続きを述べて、語彙論と統語論が統一的に結ばれることを示す。

他方、計算言語的な側面からの議論のため、当該枠組を論理文法の枠組において実証し、検証も併せて行っている。実用であることの基準が文法の性質とどう関係するかについての議論がないために評価を定量的に述べるができない。しかしながら、基本的な言語現象を中心に、現在、約700程度の規則を持つ文法をDCG記述の枠組を用いて実現した。

1 はじめに

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、意味論がその構成要素として挙げられる。音韻論では、発声という物理的な生理現象を取り巻く要因から音の并別にわたる領域を扱う。音の組合せは一つの言語を特徴づける要素である。一般に区分される音は音素とされ音韻論における基本分析単位となっている。音素列の言語記号への射影であると解釈できる。意味的な機能を含まない音韻論は他と区分されることが多い。

ひとたび言語記号が与えられると、形態論や統語論、あるいは意味論がその分析対象の領域となる。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するといわれている。与えられた形と意味の間にその成因を必然とする関係はなく、恣意的にかたちと意味の対応が存在している。いわゆる形態素はこれら概念や観念を示す標識に使われ、形態論の手続き－語を構成するための手続き－より事物を表す。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。これは語彙論の範疇に属する。

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。意味まとまりを示すための形態素に加え、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。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ている。ここに統語論上のカテゴリーが見い出される。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の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といえる。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。意味論は語の意味の研究であるとされるが、ことばの持つ意味作用についても含まれ

て論ぜられることが多い。そうして意味論は言語研究という対象領域においてさえも心理学をはじめ、論理学や認知の科学とも密接に関係を持つ。音韻論とは別の動機をもって独立させる領域だろう。

言語記号の上で形態論を共通の手段としてもつ統語論と語彙論は近未来の統合化の範囲にあるといえる。形態素解析から統語解析さらに語用論的分析という一連の流れは、統語論と語彙論の対比において統合化が可能である。本稿では自然言語における統語化の一つの試みとして、文の語用論的な意味の類型を文型に射影する。本稿での文型とは、いわゆる語彙構成をも含み、単なる単語の認定の手続きにとどまらない。すなわち言語記号の連鎖を対象とする統語分析の試みである。

2 語彙論と統語論における統合化

一般に、言語学においては弁別が重要な要素である。第一次の分節は内容表現のため、後者は二重分節機能として関係表現と言表記述のために専ら特徴づけられる。文法上の現象は多く対立することで自らを規定しているが、効率の問題から、対立は片方においてのみ顕在化することが多い。その故に、分析の困難さは規定値の無標に対する対立の相手の存在を明確に位置づけることにある。

2.1 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ため、例えば、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。自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。

例えば、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。解釈の多様性は難形からの逸脱において説明されるべきである。

2.2 仮説

様相表現を中心とした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ことにある。そうして、主体的発話行為のもくろみを文の外形に求め、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、確かに文の中核を構成していて、それを叙述する客体的な表現がある。しかし、事態の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する。

3 文型 - 述部複合層への文型射影の試み -

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。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]。

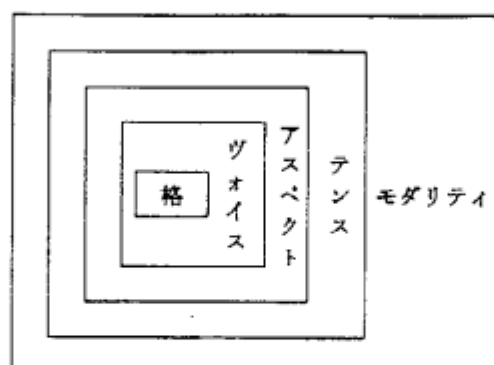


図 1: 文の層状構造「拡大語彙論的統語論」より抜粋

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、その述部のム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。平叙文の持つ多義性は、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。そうして、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。

3.1 認識文型

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 [5]。本稿では、発話行為論の成果を受け入れて、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。

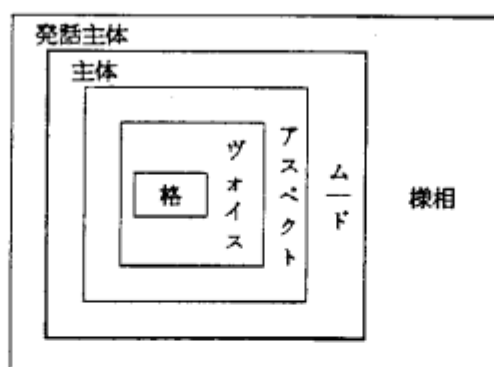


図 2: 述部の階層構造

本稿の提案する述部構造は、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。さらにモダリティを発話主体の認識階層へ移行させた。この模相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。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。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標の形態として認められる。

3.2 形態文型

前節では、述部の認識上の複合階層を仮定した。これは、統語範疇を形態論的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。表1は述語の形態により示される外形特徴と述部の階層構造で示される意味で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。

表 1: 文の基本類型

1 平叙文	{	*	未完了	肯定 / 否定	平叙
			完了	肯定 / 否定	
2 推量文	{	*	未完了	肯定 / 否定	推量
			完了	肯定 / 否定	
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	

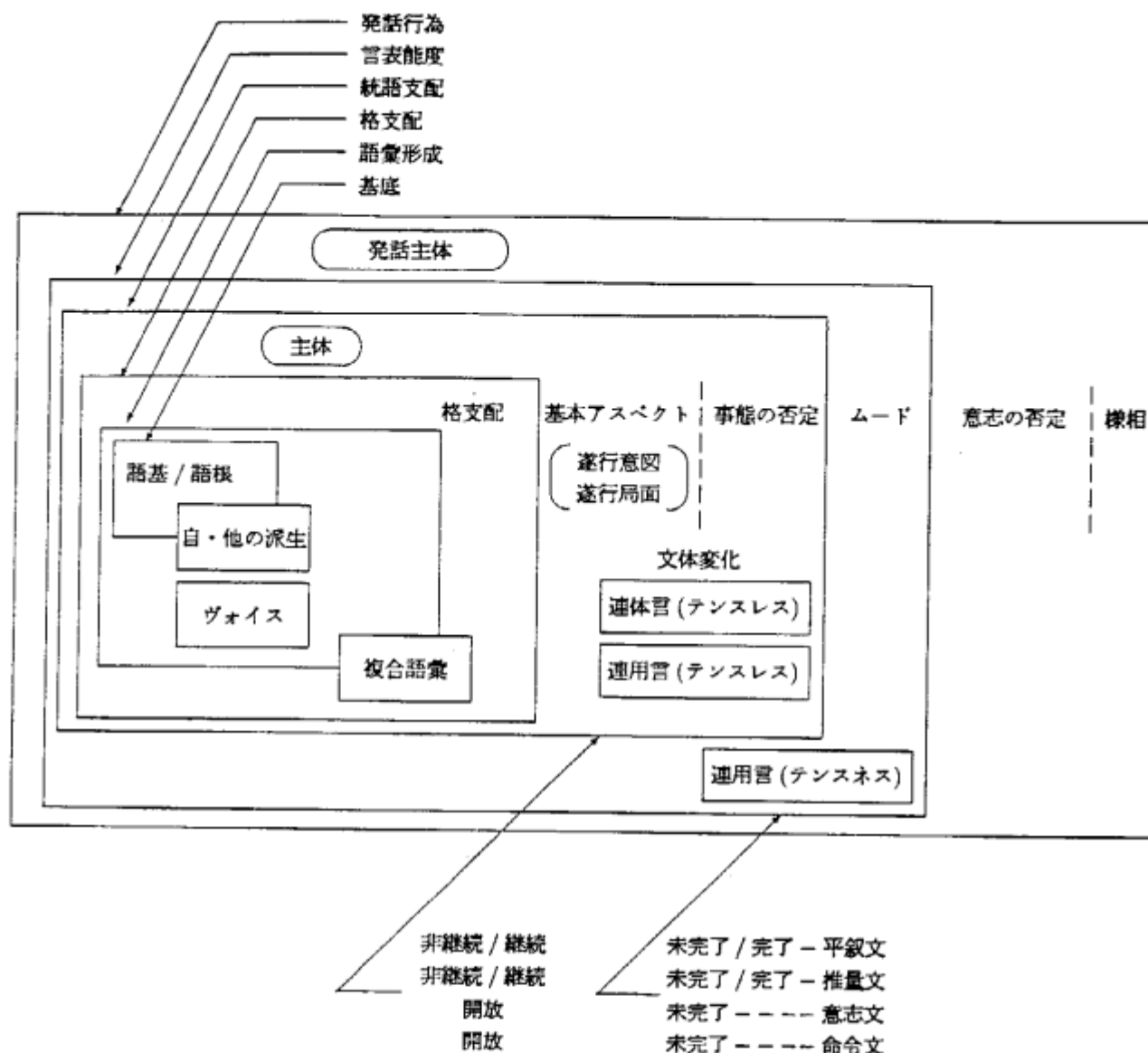
1の平叙文は、基本的に物事を認識し、確定的な意味で事態を表現する文である。事態が自らの裁断の範囲内で顕在し、在在することを主張する。日本語では、テンス、アスペクトという文法範疇が派生的な形態変化として動詞文に認められている。動詞自身は事態の包括的な確定の力が弱いのである。このかたちは本来的に事態の顕在の主張が基本機能である。2は推断の表現であり不確かな事態の顕在の主張である。事態を表出しているが、確定の態度は保留されていると言える。

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、推量文とは逆に事態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、文で表出される事態に対する確定や認定の作用はなく相手に対する事態の行動要求である。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。尚、表中における,*は文の記述内容、ムードよりも内部の文構造を略した表現である。

3.3 統合化した構造表示

次の図は表1の基本類型の形態に述部の階層構造(図2)をあわせ纏めたものである。対立の所在を構造的に明確にした。1と2についてはテンスの分化が許されておりアスペクチュアル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主体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。これは述語の認識文型と形態文型を合せた述部の統合表示である。



この統語表示枠組の利点を手短かに述べる。

述部のもっとも核となる部分が語基に相当する。派生現象もこの表示の中で扱われるほか、態変化もこの枠組の中におさまリ、語基の派生現象として扱われる。基本アスペクトは「スル・シテイル」の対立として考察されるが、形態文型がテンスの分化を基本として扱っているので「シタ・シテイタ」も自然に枠組の中に取り入れることができる。

また、基本アスペクトの他に2次アスペクトとして、遂行局面と遂行意図を取り入れた。遂行局面は、時間に関するアスペクトー「ハジメル」「ツツク」ーなどにかかわる。遂行意図は、これまで、特異なアスペクトとして語彙に依存する現象として個別に議論されてきた「カケル」「イク」「クル」などを扱う。この両者は基本アスペクトの階層を越えて機能することのないことが言語現象の調査において判明している。この基本アスペクトは事態の否定の層において閉じることになる。日本語の否定形は通常「ナイ」で示されるが、この「ナイ」は属性の語であってアスペクトの分化を許さない。あるいは固定とする機能を持つ。

事態の否定の階層を囲むかたちでムード階層が存在する。この表示におけるムードとは、極めて狭い意味を持つ。それはアスペクトに近いもので、完了かあるいは未完了かの値をもつものである。基本的にテンスの分化を許すかどうかの指標である。例えば形態文型の意志形や命令形においてはこの階層は固定の値をとり、テンスの分化が許されないことを示す。ここで、事態に含まれる主体の作用域が基本的に分かれる。これらを認定する形で認識主体が存

在する。この認識主体は基本的に言語表現には現れない。

認識主体にも否定の標識が存在する。「ホシイ」や「ハズダ」などの表現が事態の肯定と否定とは別に事態を認めるか否かの区分があることから伺える標識である。この標識のなかにあって、様相の標識は認識主体の事態全体に対する態度を表示する。様相標識については、未調査の部分が多くあるがそれでも形態区分を発展させる形で扱う。

4 素性群

素性はある範疇の形態や統語上の性質のみならず意味上の特徴をも表現するものである。文型の性質は、語や構文のもつさまざまな特徴によって表現され記述される。本稿では、統合化した統語表示を特徴を表す属性の類別とそれが持つ値について種類の列挙にとどめる。

4.1 主要範疇

主要範疇は体言、用言、連体詞、連用詞の範疇区分と述部階層表示を表す値をもって示される。直感的に体言はいわゆるモノの表象であり、用言はコトの表象手段である。また、体言と関係を構成する範疇は連体詞であり、用言は連用詞と関係を構成する。

4.2 語彙素性

4.2.1 語彙

言語記号として使用された形態を直接記述する。

4.2.2 範疇区分

範疇区分は語彙に属する範疇名である。先に挙げた主要範疇が統語論上、認められるものであるのに対して、範疇区分は形態論上の性質区分のための範疇名である。語彙に所属し、その語を区分仕付ける指標である。日本語では、語基が語の構成要素の中で容易に見分けられる。語基とは意味を有する語彙上の構成単位である。文の分析に際しての分析単位を語彙の最少単位に限定する必然性はない。統語論上の単位は主要範疇のいずれかに所属すればよく、文法の終端記号－すなわち語彙項目は－範疇区分に属していればよい。

一般の概念や観念をしめす語彙項目を記載する辞書記述に必要な範疇区分は体言語基と用言語基である。概ね、用言語基は動詞では語幹に相当し、形容詞や形容動詞では、さらに小さな形態素に対応する。表2は、文法利用に際し使われる代表的な語基区分と一般流通辞書における品詞との対応を示す。

表 2: 語基区分

語基区分	品詞分類上の対応
用言語基	動詞語幹
体言語基	名詞
副用言基	副詞の不変部分
形容語基	形容詞語幹
情態語基	形容動詞語幹

4.2.3 系統

系統とは用言に属する語が屈折し何らかの機能を有する際にみられる固有の形態変化の区分を示す。用言の異形態や派生形態を特徴づける値である。

例えば、用言という範疇において動詞も形容詞も名詞述語も区分しない理由に、文法項目は必ずしも形態顕在化の方法によって示されないことが挙げられる。肯定と否定の対立は、文法項目としては2値の対立で示されるが、形態上では前者が無標であって、後者は有標である。つまり、肯定を示す積極的な表示手段をもたない。さらに、否定を示すために「ナイ」という否定辞へ派生するわけであるが、この否定辞は形容詞と同じ形態変化のシステムを持っている。その意味で「試す」は動詞であり「試さない」は形容詞である。「試ス」という述語について肯定と否定は統

語彙上の対立を示すが、形態変化のシステムはウ系とイ系にわかれる。形態を基準とし区分した体系であれば、従来の品詞分類は有用である。本稿の類型化は文の成り立ちを基本とする体系であるため、語基区分と系統をもって用言や体言の性質を表現する。範疇区分において語彙を代表する標識が割り当てられているので、従来の品詞分類はその必要性を失う。

次の表は用言の形態システムとしての区分である。用言はその形態変化のシステムとしてウ系統とイ系統とダ系統に区分される。表3に示すように下位区分をもち派生時の音韻変化によって区分する。流通辞書にある品詞分類を用いて比較する。尚、比較の対象は岩波辞書の付録から引用した。

表 3: 用言系統

範疇	系統区分	下位区分	いわゆる活用
用言	ウ系統	ウ系弱, ウ系スル, ウ系ズル, ウ系クル, ウ系マス, ウ系ナサル ウ系カ, ウ系ガ, ウ系サ, ウ系タ, ウ系ナ, ウ系バ, ウ系マ, ウ系ラ, ウ系ワ	上一段 / 下一段動詞 サ変動詞 カ変動詞 (特殊変化動詞) 四段動詞 (カ行, ガ行, サ行, タ行, ナ行) 四段動詞 (バ行, マ行, ラ行, ワ行)
	イ系統	イ系ア, イ系イ, イ系ウ, イ系オ	形容詞
	ダ系統	ダ系ダ, ダ系デス, ダ系ナ	名詞 + ダ (デス), 形容動詞

4.3 構文素性

4.3.1 構文関係

連用句や連体句の機能の特徴づける統語情報である。専ら統語規則において付与され、使用される。

4.3.2 アスペクト

ここでのアスペクトとは広く用言に見られる形態アスペクトである。この値は平叙文と推量文では「継続」あるいは「非継続」をもって示す。「継続」は有標の値に対応し「非継続」は無標の値に付けられる規定値である。意志文と命令文では時間から開放されているため基本的にアスペクトはなく「開放」の値をとる。

4.3.3 文体

用言は丁寧さによって対立する形態をもっている。この形態の対立は先述のアスペクトや後述の認め方と同様に2つの対立形態が顕在化されるのではなく、他方の顕在化により、その対立関係が認められるものである。したがって狭義の文体と捉えて差し支えない。これは発話者の態度にかかわり文章構成の一部となっている。

4.3.4 認め方

認め方は一般に否定の派生辞が接続し示される。「肯定」と「否定」の2値を取る。

4.3.5 ムード

ムードの値は用言の主節に対する定形形態に基づく形態ムードである。これは、事態のアスペクチュアルな面を捉えたもので完了、あるいは未完了の値を持つ。この値はテンスの分化が許されている文型に与えられる選択枝である。意志形や命令形のムードは「未完了」に固定されている。

4.4 認識素性

4.4.1 認め方

発話者に属する認め方の類型である。意志否定と呼ばれるものがこれによって示される。例えば「私は食べない」は意志の否定であって事態の否定ではない。「肯定」と「否定」の2値を持つ。

4.4.2 様相

文の様相を表す指標である。文の法的分類に近く話者の意志表示にかかわる値をもつ。基本的に文の述部の形態と終助詞の接承による組合せに対応している。

基本的にはこれが文の様相区分となる。当該類型には「叙述」「推量」「意志」「依頼」「疑い」「規定」「情態」などが挙げられる。

5 実現

計算言語学の視点から、この統語化した文法枠組を論理文法の枠組で実現した。汎用性への考慮から DCG 記述を採用し、特定の計算機システムや特定の枠組に依存しない実現を試みた。解析のために BUP トランスレータを用い、Prolog の実行環境であれば、ほとんどどのような計算機においても動作する。現在、約 700 の規則を持つこの文法は、従来の形態素解析と構文解析を区分なく行う。著者は現在 PSI-II を用い、実験を行っており動作を確認している。また SUN 上での動作確認も報告されている。また、部分的ではあるが並列マシン MULTI-PSI 上での動作実験報告もある。

このような動作上の透明性に加え、当該文法は研究成果として開放されている。

6 おわりに

本稿で提案する統語化の試みは、日本語における弁別と対立の手続きの所在と機能から文法の法則を明らかにしてゆくことにあった。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意味上の機能を中心に文の法則を捉えた。形態論の手続きを手がかりに分析を行い統語文型へ射影した。

文法の実現の一つの手立てとして、論理文法の枠組を用いて規則記述を行った。紙面の制約上、詳しくは触れなかったが、DCG 記述を採用し、約 700 程度の規則を持つ文法を構築している。尚、この文法は開発途中であるとともに評価も併せて行っている。いわゆるベタ入力を受けつけ、いわゆる形態素解析と構文解析、さらに一部の意味的な処理を一つの文法でおこなう。

本稿は効率の面において議論に欠ける。しかし、分析効率とは統合化を考慮した場合、統語分析だけの速度や解析候補の数では計測することができないように思われる。例えば当該文法では文節という単位が認められない以上、一文の文節数において比較の考慮外となる。また、単語認定の単位が異なるためにその形態素の数においても若干となる。全体の言語遂行の測度の設定が今後の自然言語処理研究に望まれる。

本稿では文の様相表現を文型に射影することで統合化の枠組を提案した。これは従来のいわゆる形態素解析と統語解析の区分をなくし、意味類型を文の形に射影したものである。

しかしながら、このレベルでさえも、大きな危険も侵していることを指摘したい。それは、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している。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る。ことば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。ことばの列の意味は、ある解釈の体系のもとに決定されるが、それ以上に伝達者の非記号的な行動と、受動者の感性や知識にも依存する。従って、全体の中での占めることばの位置がそのおかれる状況によって変化する。

今後の統合化の方向は認識主体の中での文の占める位置を見出して文型に射影することであろう。

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, 東京, 1989.
- [2] 『国文学解釈と鑑賞』, 第 51 巻 1 号, 至文堂, 1986.
- [3] 佐野洋.「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. (1990), 「局所化した単一文法とその表現」, 第 41 回情報処理学会全国大会. 1990.
- [5] 仁田義雄, 益岡隆志編.『日本語のモダリティ』, くろしお出版, 東京, 1989.
- [6] ビーター・セルズ, 郡司隆男, 田窪行則, 石川彰 訳.『現代の文法理論』, 産業図書, 東京, 1988.
- [7] 福本文代, 佐野洋.「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.

- [8] 益岡隆志, 田窪行則. 『基礎日本語文法』, くろしお出版, 東京, 1989.
- [9] 松本裕治, 清野正樹, 田中穂積. 「BUP トランスレータ」, 電総研彙報, Vol.47, No.8, 1983.
- [10] Fernando C.N. Pereira, David H.D. Warren. *Definite clause Grammars for Language Analysis - A Survey of the Formalism and a Comparison with Augmented Transition Networks*, Artificial Intelligence 13(3): pp231-278, 1980.