

TR-705

事態成立の確信度を生かした
発話意図の抽出

佐野 洋

October, 1991

© 1991, 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

事態成立の確信度を生かした発話意図の抽出

EFFECTS ON THE DEGREE OF REGARDNESS FOR JUDGEMENTS IN DESCRIPTION

佐野 洋

SANO, Hiroshi

(財) 新世代コンピュータ技術開発機構

INSTITUTE FOR NEW GENERATION COMPUTER TECHNOLOGY (ICOT)

e-mail : sano@icot.or.jp

はじめに

述部の無人称性と提題項目をマークする形態「ハ」を獲得している日本語は、統語上、極めて緩やかな構造制約を持っている。外心的な要素を述部が支配しているように見受けられ、述語に対する内心構造が明確に表層に現れることが寧ろ稀である。「提題-叙述」構造が「主語-述語」構造に対し優位となり、文脈や伝達の流れに著しく依存する文の連鎖構造ができあがる。

述部は、それが句であれ主節であっても、述語を核として事態の叙述の任務を負っている。冒頭に、述部の無人称性を述べたが、これは叙述における統語制約である。日本語にも人称制限はあって、それは専ら叙述した内容の伝達の局面にあって人称対立が起るとされる。この方法は幾分か巧妙で、形態上の音形対立ではなくて統語上無形を強調している。例えば、「ハ」にしろ「ガ」でマークするにしろ二人称以外の語によって主語を明示すると命令や禁止、疑問を含む勧誘などの文型を構成できなくなる。日本語の述部の構成は複雑であるといわれている。それは緩やかな構造制約と持っているからではなくて、内心構造を決める叙述の構造よりも、叙述した事態の伝達にあって、より優位の制約を持つからであろう。

本報告は、発話の中心となる「話し手」が他の人称と対立する部分に着目し、統語構造を精確とすることで統語分析から話し手の意図を抽出する試みである。この試みは、計算機によるテキストからの話し手の意図の自動抽出を最終的に目指している。さらに意図抽出の処理を統語分析の守備範囲に納めることも重要な点である。この目標達成のために、第一段階として、本稿では文の記述において意図表現に係わる形態素を整理し、こうした形態素の分布特徴から得られた統語構造を説明する。結論を言うと、文にあって事態の構造と話し手の判断の層を形態と構文の形式から区分し、話し手の判断の層に現れる意図表示のために形態素を手懸りに意図表現を人称との呼応から定めた。意図表現は話し手の判断と判断を下した確信の度合いによって示されることを利用して次の統語上の特徴を得た。(1) 発話意図は話し手の判断と判断の根拠を示す確信度に相関する。(2) 話し手と叙述の事態の主体が一致すると確信度

は開放の値を採り、話し手の意図は判断にのみ依存する。
(3) 聞き手が叙述の主体と同一の場合も(2)と同じである。

そして形態素の分布をもとに統語構造を定めた。この結果は機械処理のための文法規則に反映された。統語解析によって発話意図の抽出が可能となった。さらに文単位の統語分析の結果と実際の文の意味とが即応しない点について報告する。

次節では意図表現を担う形態素を分析する指針について、我々の採用した考え方を説明する。2節は、調査した形態素の分布について概略し、その分布特徴から得た統語構造を説明する。この構造は、従来から開発してきた機械処理のための文法枠組に組み入れられ、文法規則の一部となった。3節で、この拡張した文法規則について若干述べるのと同時に、この統語解析から得られる意図表現の特徴を挙げる。テキストベースの文単位の統語分析に基づく発話の意図抽出の不具の点は4節で述べる。

1 意図表現

1.1 統語構造と意図表現

意図表現を担う文法範疇としては、「モダリティ」がある。このことばは、文法諸家によって対応する文法現象の範囲が異なることの多い用語である。言語的にはモダリティは文を成立させる要としての機能を持つともされる。機械処理の観点から見ると、このモダリティは文末表現として一括されて個々の語あるいは特定形態の慣用的な意味用法によって区分されてきた。統語上の構造制約として機能を注視するにはさほど至っていないらしく、工学への応用では、単なる文の叙述する内容記述の抽出や、内容記述に応じた質問応答のレベルから、利用者の発話意図と知ること、より知的な応答能力をシステムに与えるために利用されている。自然言語をインタフェースとする幾つかのアプリケーションでは、発話意図を抽出するためにモダリティを注視しており、使用者の操作負担を軽減したりするガイダンスシステムへの適用が見られる。こうしたシステムでは、専らユーザの発話が、事象の叙述による事実の伝達であるのか、あるいは質問を発しているか、あるいは要

求表現であるのかなど、発話意図からその目的を同定するために利用されている。一般には、発話意図の抽出プロセスは意味処理レベルにあることが多い。文末表現や接続表現といった個々の形態素の特徴抽出によって、連鎖に見られる論理関係と文のタイプに相関する修飾関係から文章構造を同定する試みがある[1, 2]。こうしたアプローチは文のモダリティの間接的な分析であると考えられる。

しかしながら、モダリティの分析は単に発話意図といった特定の言語現象に限らず、さまざまな言語現象に対し単一の説明を与える働きを持っている。

1 形態的な接続関係

- (1) 「食べるらしい」
- (2) 「食べるのです」
- (3) 「食べるらしいのです」
- (4) * 「食べるのですらしい」

2 構文上、従属句の依存関係

- (5) 「美しく舞う」
- (6) 「朝食を食べて出勤した」¹
- (7) * 「美しくなく舞う」
- (8) * 「朝食を食べなくて出勤した」

3 構文上、人称制限や意味の違い

- (9) 「(あなたハ) (オマ) いかがですか」
(聞き手への勧誘)
- (10) 「(オハ) いかがでしたか」
(聞き手への質問)
- (11) * 「(私 / 彼 / 彼女ハ) (オマ) いかがです」
(二人称以外)
- (12) * 「(オハ) いかがでした」
(二人称以外)

4 文の接続関係における助詞の運用

- (13) 「高原に美しい秋ガやって来ました。静かさに包まれた軽井沢でいままでとちょっと違った秋休みハいかがですか」²
- (14) ? 「高原に美しい秋ガやって来ました。静かさに包まれた軽井沢でいままでとちょっと違った秋休みヲいかがですか」
- (15) ? 「高原に美しい秋ハやって来ました。静かさに包まれた軽井沢でいままでとちょっと違った秋休みハいかがですか」

こうした現象に対しモダリティの分析をもって上記の現象の認可あるいは不認可の予測を試みることは本稿の内容から逸脱する。専ら3に焦点を当てることになる。

1.2 モダリティ

機械処理のための文法規則の開発の目的から、我々の分析の指針は仁田[11]によるモダリティの規定に沿う。モダリティとは何かについて、それを暫定であるとしながらも次のように示している³。

¹ c.f. 「朝食を食べないまま出勤した」

² 軽井沢プリンスホテルの車内広告から引用した。

³ 仁田[11]からの引用である

- モダリティ 「現実とのかかわりにおける、発話時の話し手の立場からした、言表事態に対する把握の仕方、およびそれらについての話し手の発話・伝達の態度のあり方の表し分けに関わる文法表現」

文法表現であることを明示する分析は、統語分析からの拡張による手段で発話の意図抽出を行う目的に適っている。仁田の分析においては、さらに、文に記述の内容を示す「言表事態」とそれを包含する「言表態度」という重層構造を認める[11]。

文 :: 言表事態 言表態度

言表態度の表出を文成立の必須条件として扱っている。それぞれの要素は次に定められていて、「言表事態」の部分は、主語-述語(命題)となる部分と文法範疇としてテンス・アスペクトや認め方などから構成される。他方、「言表態度」は、ムードや話し手に属する認め方さらに態度が構成の要素である。

詳細な区分の仕分けには違いがあっても、文に階層を認めることは大方の一致する意見である。こうした構成的な分析は文節内分析によって形態素を調べ、構文はそれらの係り受け関係に依る方法に比べ、統語分析の精度が上がるほか、特に叙述の仁田による規定は文法表現を常に根拠に置く理由から機械処理への適合が容易である。

とりわけ伝達に係わるモダリティは、いわゆる言語能力よりも言語運用に関連する文法範疇である。文の連接を陳述の連鎖から分析した永野[12]の分析は文脈にあってモダリティの機能する範囲の分析に参考となった。以下では、仁田の分析に従ってモダリティの分析指針を手短かに説明する。

1.3 モダリティの分類

仁田[11]に従うと、モダリティには真性モダリティと擬似モダリティという下位区分がある。形態や構文に明示的に表れて曖昧性がなく発話時に話し手の立場からの態度が明確になるものが前者である。それに対し、後者は文をその置かれた文脈から取り出すと話し手の態度が不明瞭となるモダリティである。

モダリティは方向性によっても下位の区分があるとされる。この方向性の定性的な分析が本稿の報告の主眼であり、方向性という観念を機械処理のために統語分析にどのように適合させるのが問題となる。

事態を指向するモダリティは、言表事態の話し手の把握の仕方での認識のモダリティとも呼ばれる。他方、伝達のモダリティとは言表事態に対する話し手の伝達の態度のあり方であるとされる。発話時にあって話し手の立場が必須の要素となり、働き掛けのモダリティとも呼ばれる。すなわちモダリティはその赴く方向によって二区分されている。以上、仁田による分析を手短かに挙げた。

一般的には発話・伝達のモダリティが欠けると文として機能しないことが一般に言われている。伝達のモダリティは文成立の必須要素であるとされる。事態への指向のモダリティが顕著な文は、伝達のモダリティがないのではなくて空であるか、もしくは零記号による言語形式で表現されているものとする。

1.4 モダリティを示す形態素の分布

前節に挙げたモダリティの規定とその基本の分類をもとにして、専らその方向性を軸に発話の態度に係わる形態素の分布を調査した。調査の対象とした形態素は森田 [9] が挙げた助動詞相当表現から選んだ 113 語である。尚、助動詞相当語には、アスペクトを示すもの、行為の受授を示すものがあるが、いわゆる言表事態に関する表現であるから除外している⁴。また、[9] には、終助詞の働きをする形態素として 46 語の収録がある。これらについても今回は調査の対象から外した。

図 1 に形態素の分布を調べるための言表態度に相当する構造図を挙げる。

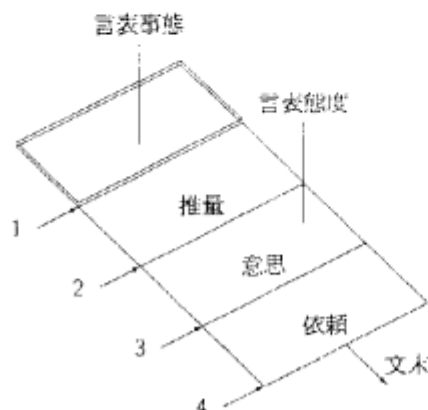


図 1: モダリティマップ

この図は、文の基本の構造を平面的に示している。これに対応する文の構造を図 2 に示す。



図 2: 文の構造

モダリティはその分析の要因として方向性が問われるが、方向性のために、この平面上に文内容を示す事態から文末へと向かう軸を設けた。事態を指向することから、次第に聞き手に働き掛けるという軸を設定した。この軸の設定のために図 1 に示す平面図形を基本とした背景は、用言の異形態として形態素がその形態でモダリティという文法範疇を明示して対立させる区分に依存している。図 1 の数値で示す位置にいわゆる用言の活用形態としてモダリティに関する文法範疇を対置させる。これは、助動詞でもなく助動詞相当語でもなく用言の異形態として示されていることが特徴である。基本の表現を表 1 に挙げる。用言の活用形態から観察すると、文の類別には表に挙げた 4 区分が

⁴但し、態変化や やりもらい といった現象もモダリティの類型とされる [3]

ある [4]。平叙表現、推量表現、意志表現と依頼表現である。実はこれは、統語分析にあって基本の文型になるとともに、前節で挙げた擬似モダリティと真性モダリティの区分の指標となっている。

表 1: 基本形態と文の類別

類別	対立形態		モダリティ上の特徴
1 平叙表現	スル	シタ	中立的な判断
2 推量表現	スルダロウ	シタダロウ	推量を伴う判断
3 意思表現	シヨウ	シマイ	意思表明
4 依頼表現	シロ	スルナ	行為の働き掛け

統語分析における文の基本区分をモダリティの面からその特徴をみると、3, 4 はテンスを分化を許さない。したがって行為の働き掛けの機能が専らである。4 は真性モダリティであり 3 は真性モダリティの度合いが大きい。行為の働き掛けには話し手が行為の主体に含まれるものと含まれないものがある。4 は行為の主体に話し手は存在し得ず、話し手は聞き手に対立する。3 は肯定的な行為の主体に話し手以外は存在し得ないし、否定的な行為の主体には基本的に話し手が設定できない。1, 2 の形式は基本的に擬似モダリティである。それは、働き掛けや意思といったモダリティに関する形態が空となっているからである。

2 意図表示の構造

2.1 形態素のマッピング

形態素の分布を見るため、分析の軸は事態の叙述に最も近接した無人称の性格を持つ形態素を始め、話し手の判断が現れる局面、そして人称が固定されて専ら伝達の表現に係わるものに大区分した後、そのグルーピングを行った。

次の図 3 は、上記の 113 語の形態素を初期のモダリティマップに展開したものである。

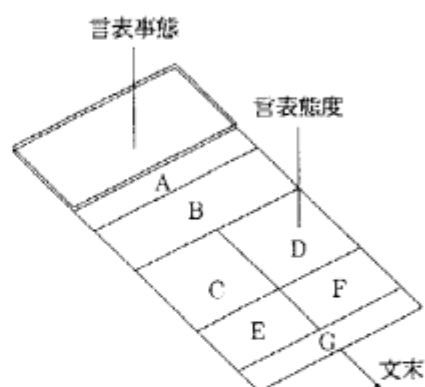


図 3: モダリティマッピング

110 以上におよぶ形態素のマッピングによって、初期のモダリティマップは修正が施された。推量は、いわゆる認識的なモダリティで、事態の成立不成立の諸層を表現仕

分けるとされる[10]。判断の確かさ不確かさも表現仕分け。さらに働き掛けにあっても確かさ不確かさの表現が可能である。事態の叙述から働き掛けへ向かう軸とは性質を異にする概念である。この扱いのために森山[10]の分析にいう確信度という考え方を採用した。

以下、図3に配置された形態素を説明する。

A は事態にもっと近く、事態が可能か不可能かを示す。形態的には「できる」「できない」に代表され「こともならない」といった表現がある。

B 位置には話し手が事態の主体を兼ねる意思判断がある。事態の可能/不可能に係わらずに、「話し手」がそれを強く主張する表現である。話し手の意図は「確信的な事態成立(あるいは不成立)の表明」である。この位置には、「ようとする」「んとする」「まいとする」「ないで(は)おかない」といった形態素がある。

C にあって話し手は主体ではなく、一人称以外の主体を事態に招き、この事態の子測に基づく事態成立(あるいは不成立)の表明である。事態の主体は話し手ではないので、事態への確信度は低下する。C 位置に所属する形態素には「でもいい」「たっていい」「かまわない」「たってかまわない」「てさしつかえない」などがある。

C 位置に平行して設けた D 部分には超意思的な判断がある。平行して設けることの適切性の根拠に欠けるものの、事態への確信度に幅がない場合である。「ことになる」「よくなる」「に至る」「仕方がない」「仕様がな」「わけにはいかない」といった形態素が見られるほか、「に過ぎない」「まです」といった極に達している表現がある。

A から D 部分までは、事態に対して肯定/否定の分化が許されている。それに対して、E から G にあってはモダリティを示す形態素自身が認め方として肯定的であるか否定的であるかの意味を内在する。

F は事態の主体を一人称以外とし、その確信度を肯定的な極へ位置付ける表現である。こうした表現は義務とか勧告、主張と呼ばれているものである。「なければならない」「なくてはならない」「ねばならない」「はずだ」「には及ばない」などの形態素が位置している。F は E にまったく平行するもので、事態の主体を一人称以外とし、その確信度を否定的な極へ位置付ける表現である。「てはいけない」「てはならない」「ことはいけない」「わけがない」「はずがない」といった表現を用いて必然を示している。

G は事態の主体に二人称をとり全く依頼の表現となる。もっともマップの外郭に位置している。命令表現と平行していて、非強制的な依頼である。「てくれないか」「てもらえないか」「てくできますか」「てください」などの形態素があって、禁止を示すために判断部分に肯定と否定の区分を設けることができる。

2.2 話し手の意図と認識的モダリティの構造上の位置関係

図3に挙げたモダリティに関する形態素の分布は、その形態がもつ意味に依って分類した。伝達を目的とした表現に人称を区分する指示体系があることが指摘されている[3]。述部における事態に対する人称の指定という観点か

ら、図3にある形態素群の分布を再構成したものが図4である。

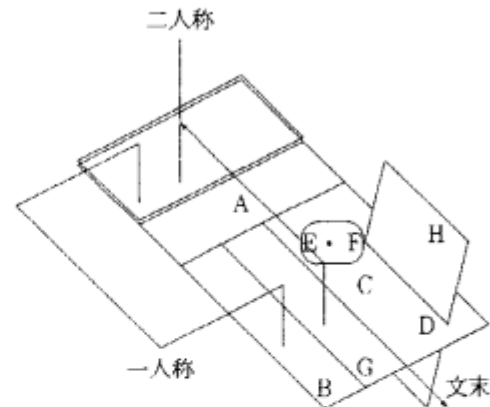


図4: 意図表示の構造

H グループにあるモダリティ表現は、モダリティの内、認識的なもの、すなわち判断への確信度を示すものである。こうした形態素は「に決まっている」「はずだ」「ことになる」「相違ない」「違いない」といった強い確信度を示すものから、「ことだろう」「わけではない」「かもしれない」という弱い確信度の表現がある。事態の叙述にあって、動作主体が話し手と一致する B グループと、動作主体が二人称に限られ話し手からの働き掛けの意味になる G グループではこうした認識的な表現はない。C, D と E, F グループのモダリティに関する表現は事態の動作主体に一人称以外をとる。さらに二人称では G グループの表現を除く場合に限り、判断の表現が可能となる。こうした判断の表現は、C, D と E, F との間に接続の制約が存在するが、双方とも H グループの接続を許す。E, F は「なければならない」と「てはいけない」のように認め方を対立させる違いが形態表現に及んでいるものであるから、意図表現としては必然とか当然といった意図の下で、それが肯定的か否かの区分を A 部分に反映する形で取り込めば良い。判断の様相は、認識の度合い(指向する判断への確信度)によって変化を受けることになる。従って、意図表現は次に示す要因によって近似的に定まる。

- 事態の主体と話し手が一致する際には確信度は開放値をとり意図を示す様相は B グループの形態素によって定まる
- 事態の主体が聞き手である際には確信度は開放値をとり意図を示す様相は G グループの形態素によって定まる
- 判断と確信度によって様相を規定し、C, D, E, F グループの形態素によって定まる
E, F は否定辞によって肯定と否定を対立させるのではなく形態素にその極性が含まれている。このグループには次のような下位分類がある。

- 話し手の意思が及ぶ能動的必然と話し手の意思がまったく及ばない受動的必然を両極とする区分である。

尚、H グループの構成群には「と思う」「と考えられる」「と見られる」も含まれる。先に挙げた、「に決まっている」「はずだ」「相違ない」などと比べると、その機能する守備範囲が違っているらしい。直感的には、前者のグループが C,D に対して付与される傾向が強いのだろうし、後者のグループは E,F に対して後接することが多いにちがいないと思われる³。

3 文法規則への適用

3.1 基本とする文法規則

我々は、機械処理に適合した文法枠組 (SFTB)[6] の提案を行い、その枠組に沿って文法規則を記述してきた。文法規則は DCG の記法に従って書かれており ICOT の成果として公表されている。

その特徴は、句構造ベースの文法規則であって、構成的な情報の汲み上げによって文構造を組む。その際の情報操作は単一化の枠組を基本としている。情報構造は素性とその値の対によって特徴づけられる。文に認められる統語構造を図5に示す。

図5においてソートという素性は、タイプを示している。“sentence”に相応する統語構造は、“叙述”というタイプであることを示している。このタイプは、“topic”、“predicate”という属性を持っている。“predicate”は文の内容記述に対応するもので、事態には2つの種類がある。その事態の中に動作主体を含む動的事態とそれを含まない静的事態である。それぞれのソート名は専ら語彙によって決まる。個々のソートが示すタイプに依存して、“complement”で示される属性の数が決まっている。

基本の文法枠組は、文に認められる階層構造を基にしている[8]。Topic は「ハ」でマークされる補語あって提題部分となる。Predicate は事態の構造であり、上述したように動的事態と静的事態の区分がある。Agent は「ガ」格でマークされる動作主体である。Complement はいわゆる格補語と呼ばれているものを指す。この文法では、いわゆる形態素解析と構文解析の区分がなく、語構成規則から通常の統語規則、ならびに発話意図の抽出までをカバーしている。現在、この文法規則は事態の記述に関する基本的な言語現象をほぼ網羅している。従属句や副詞句は任意であるのでこの図からは省略した。こうした句は、統語構造を特徴づける属性構造に適うように依存構造を構成する。

語彙に記載の情報は、用言としての性質が agent actions であるか否かと、事態の叙述にあって、その語が規定的にどのような補語を獲得するのかという情報である。

sentence :

態度 = <i>P</i>
様相 = <i>M</i>
認め方 = <i>J</i>
ソート = 叙述
topic = <i>Topic</i>
predicate = <i>Predicate</i>

Predicate ::= 動的事態 | 静的事態

動的事態 :

ムード = <i>M</i>
文法 = <i>S</i>
判断 = <i>J</i>
認め方 = <i>P</i>
アスペクト = <i>A</i>
テンスレス = 動作
ソート = “ソート名”
agent = <i>Agent</i>
complement = <i>Complement</i>
...

静的事態 :

ムード = <i>M</i>
文法 = <i>S</i>
判断 = 肯定的
認め方 = 肯定
アスペクト = 開放
テンスレス = 状態
ソート = “ソート名”
complement = <i>Complement</i>
...

図 5: 統語構造

SFTB では、文に認められる文内容の叙述と発話の意図表現を構造を区分し、ソート名を“叙述”として統語分析のレベルで認識する。「認め方」は文内容の叙述部分と意図表現の部分にそれぞれ存在していて、現在の版の文法規則では発話の意図表現のために認め方に加えて、様相と遂行の値が設定されている。この発話意図に関する素性と値の集合を以下に挙げる。

表 2: 発話意図に関する素性と値

態度	中立, 表明
様相	平叙, 推量, 意図, 依頼
認め方	肯定, 否定

3.2 統語構造の拡張

モダリティ関係して意図表示のための素性は、認め方、様相と態度である。認め方は、文が叙述の事態の成立 / 不成立を問うもので、態度は働き掛けの区分を示している。話し手の意図表示は、様相という素性によって示されるが、これまでの分析にあるように、統語分析に人称制限

³この文にあるように、著者の判断は予測の確信度としては強いが、表明の確信度は低くなっている。この検証については今後の調査に委ねられている。

と、話し手の判断とその確信度という指標によってよりきめ細かい話し手の意図抽出が行えることがわかった。この結果を基本の文法枠組に適用することで文法規則を拡張した。図4の分析図に沿って統語構造を拡張した。図6は拡張した統語構造を示す。

sentence(主張) :	$\left[\begin{array}{l} \text{態度} = P \\ \text{認識} = \text{開放} \\ \text{判断} = I_B \\ \text{participant} = \text{話し手} \\ \text{認め方} = J \\ \text{ソート} = \text{叙述} \\ \text{topic} = \text{Topic} \\ \text{predicate} = \text{Predicate} \end{array} \right]$
sentence(依頼) :	$\left[\begin{array}{l} \text{態度} = P \\ \text{認識} = \text{開放} \\ \text{判断} = I_G \\ \text{participant} = \text{聞き手} \\ \text{認め方} = J \\ \text{ソート} = \text{叙述} \\ \text{topic} = \text{Topic} \\ \text{predicate} = \text{Predicate} \end{array} \right]$
sentence(判断) :	$\left[\begin{array}{l} \text{態度} = P \\ \text{認識} = E_H \\ \text{判断} = I_{CUDU(ENF)} \\ \text{participant} = \text{Agent} \\ \text{認め方} = J \\ \text{ソート} = \text{叙述} \\ \text{topic} = \text{Topic} \\ \text{predicate} = \text{Predicate} \end{array} \right]$

図6: 拡張した統語構造

文の類型として3つの区分を設けた。事態の主体が話し手であるか、あるいは聞き手、それ以外の区分である。判断の文とされるものは、事態の Agent が participant に単一化されることによって決まる。統語解析による分析であるから、Agent は特定されることなく未束縛の場合もある。潜在的に話し手であったり、あるいは聞き手が文脈の情報から規定されていてさえも、文としては判断文として認識される。

基本の文法には、語彙情報として用言では、それが agent actions であるかどうかと補語についての情報だけであった。体言に関しては一切の情報は付与されていない。意図抽出の精確化による統語構造の拡張によって次の語彙情報を必要とする。

- 話し手と聞き手とそれ以外の別を立てた人称情報を体言に設定する
- 静的事態にあっては常に Agent は未束縛であるが、特に情意形容詞と呼ばれるものは、潜在的な主体を設定する必要がある。従って、

形容性と情態性の用言には、語の示す性質が判断属性か性質属性かの区分を設ける

認識素性は、図4の II グループの形態素が担うもので、判断を下した確信度である。この確信度は事態の主体者が話し手あるいは聞き手に統語上、同定されると開放の値をとって判断の根拠を問わない表現となる。語彙情報の追加によって、潜在的な主体を静的事態に設定した。次の現象に対応する。

- (16) *「歯が痛いだろう」
 (17) 「歯が白いだろう」
 (18) *「私は歯が白いだろう」

動的事態をとる判断の文では、C,D,E,F グループの形態素によって判断素性の値が決まり、H グループの形態素が判断の確信度を決める。次の表は拡張した統語構造を用いた文法規則で用いられる発話意図に関する素性と値の集合である。

表3: 発話意図に関する素性と値(改訂)

態度	中立, 表明
認識	開放, 推量, E_H
判断	平叙, 意思, 依頼, $I_B, I_G, I_{CUDU(ENF)}$
認め方	肯定, 否定

A ~ H までの形態素は最終頁にその一覧を挙げた。素性値は主たるものを挙げた。個々の形態素の内在する意味は十分に特定化されていない。F グループの形態素は認め方の値を内在している。G グループも含めて否定的な表現が多く、いわゆる二重否定を用いた確信度の強度の差異については考慮していない。

4 おわりに

本稿は事態の叙述に対する判断とその事態成立の根拠となる確信度を用いた発話意図の抽出を行った。それは完全に統語分析ベースの方法である。モダリティに関する情報は形態素で明示される形式を基本として、それらの分布を調べることで、事態成立の判断に関与するものと、その判断に至る根拠を現すものとに区分した。さらに、人称制限を持ち込み叙述の事態の主体者と相関させた結果、形態素という形式により示される発話意図に関するモダリティの形式上の細分類を行った。この分類は統語構造に反映され、対応する素性構造を文法規則に組み入れることで、発話意図を統語分析レベルでより精確に抽出することが可能となった。こうした情報はほとんど主幹となる用言の性質区分に依存しない。もっとも単純な形式でのモダリティ表現を分析した発話意図の抽出であると考えられる。

この分析では、語彙記述の情報として、体言には人称情報を、そして用言には、その語彙性質が agent actions か否か、さらに情態性と形容性の語彙は、その語彙性質が判断属性か性質属性かの区分が必要である。文法規則は、統語構造とモダリティに関係する形態素の構造上の分布に従って、人称制限のもとで意図表現を分析する。発話意図

は、統語素性である判断と認識、態度の値の相関から判明する。

一文単位の統語解析の範囲で文法規則は文の発話意図を抽出する。実規則は DCG 記述の文法であり、論理プログラミングの枠組にあって、どのような計算機においても動作する。辞書には上記の情報を必要とする。利用者は統語レベルでの意図抽出結果をもとに、個々の形態素の特定の意味、あるいは主語を担う語の意味記述から発話の意図をより詳細化することが可能である。

4.1 文の意味との適合度

一文単位の統語構造とモダリティをしめす形態素を基準とした形式に依る分析であることからの制限は大きい。特に計算機による統語分析のための文法規則に組み入れることから語彙依存の細かな性質を取り入れていない。聞き手、話し手の知識に状態に依存するような表現区分はその抽出ができない。

いわゆる意味処理とされるものを除いても以下に示すように統語分析の枠組と文法規則には不具の点がある。

(1) 有題文と無題文の判断は形式の上で区分が可能である。主題と主語の違いから判断される違いを本稿の枠組は取り扱っていない。従って、この区分の影響が及ぶモダリティ表現に依存する発話意図の抽出はできない。

(2) 表現類型にかかわるモダリティには疑問あるいは疑いという下位区分があるが、考慮の対象外にある。文法枠組は焦点位置の特定の手段を持たずその作用域の特定能力にも欠ける。

(3) 態度に係わるものは文末の形態に関係することが多い。本稿の分析では、いわゆる終助詞とその相当語句は調査の対象外であったために、素性値だけの設定が統語構造に行われた。

(4) 副詞と文末形式の相関関係が考慮外にある。

(5) 態表現や やりもらい もモダリティの一種であり、話し手の視点に依存する。(2)と同様に表現能力に欠けている。

4.2 課題

分析の不具の点とした(3)は、本報告で提示した枠組の上で、文末に位置する形態素 46 語の分布調査によって判明するだろう。文法枠組としては(1)の表示のための能力を持つから、人称を考慮した有題文と無題文の関係を調べることで不具の解消を図ることが可能である。(1),(5)については目下、統語表示上の記述方法を知らず、この点から改善が必要とされる。(4)は言語現象の調査から始まる。

謝辞

形態素の分類に関しては長澤陽子氏に多くの指摘を受けました。本稿の作成にあたっては米田順美氏に協力を得ました。両氏に感謝します。

参考文献

- [1] 木下聡、小野賢司、浮田峰彦、天野真家、「日本語テキスト理解における文脈構造抽出法」、談話理解モデルとその応用シンポジウム論文集、情報処理学会、1989
- [2] 福本淳一、「筆者の主張に基づく日本語文章の構造化」、自然言語処理研究会資料(NL 78-15)、情報処理学会、1991
- [3] 金水敏、『代名詞と人称』、「講座 日本語と日本語教育 第4巻 日本語の文法文体(上)」、北原保雄編、明治書院、1989
- [4] 佐野 洋、福本文代、「発話の意図と疑問の焦点」、言語理解とコミュニケーション研究会資料(NRC 90-17)、電子通信学会、1990
- [5] 佐野 洋、「統語文型と文の階層」、自然言語処理研究会資料(NL 82-1)、情報処理学会、1990
- [6] 佐野 洋、「局所単一化文法 LUG による日本語文法の記述」、第9回 第五世代コンピュータに関するシンポジウム・発表予稿集、新世代コンピュータ技術開発機構、1991
- [7] 益岡隆志、「モダリティの文法」、くろしお出版、1991
- [8] 南不二男、「現代日本語の構造」、大修館書店、1974
- [9] 森田良行、松本正恵、「日本語表現文型」、株式会社アルク、1989
- [10] 森山卓郎、「認識のモードとその周辺」、「日本語のモダリティ」、くろしお出版、1989
- [11] 仁田義雄、『現代日本語のモダリティの体系と構造』、「日本語のモダリティ」、くろしお出版、1989
- [12] 永野賢、「文章論概説」、朝日書店、1986
- [13] 市川孝、「国語教育のための 文章論概説」、教育出版、1978
- [14] トマス・ワソー、原田康也訳、「文法理論」、JPSG ワーキンググループ資料、1991

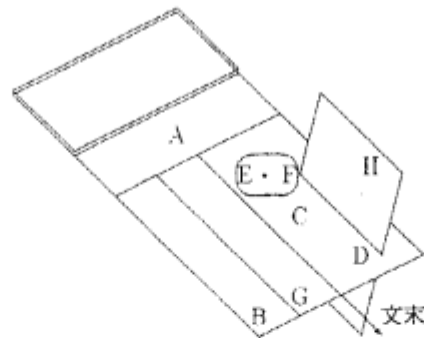


図 7: 意図表示の構造 (再載)

表 4: 調査形態素一覧

グループ	形態素	判断素性の主たる値
A	ことができる, ことができない, こともならない	(認め方)
B	て(は) いられない, わけに(は) いかない (よ) うとする, んとする, まいとする, とこにする, ようにする, つもりだ, てみせる てならない, ていけない, ずに(は) おかない, ないで(は) おかない ずに(は) いられない, ないで(は) いられない てたまらない, (たい) ものだ	意向 決定 自発 自発 意向 消極的意向
C	ばいい, といい, がいい, ほうがいい てもいい, たっていい, て(も) かまわない, たってかまわない て(も) さしつかえない, とよい	消極的願望 許可 許可
D	ことになる, ようになる, に至る, て仕方(が) ない, て仕様が, ざるを得ない (より) ほか(は) ない, より(ほか)に 仕方がない, にほかならない, まだ, までのことだ に過ぎない, に足りない, ても仕方がない, ても仕様が (とい)ったらない, だけのことはある	自然成立 限定付き成立 程度付き成立 程度付き成立
E	なければならない, なくてはならない, なければいけない, わけだ, のだ, はずだ, ことだ べくもない, どころではない, とは限らない, に及ばない, までもない	必然 勧告
F	てはいけない, てはならない, ことはいけない わけがない, はずがない, ことはないて(は) ならない, てはいけない	必然 (認め方が否定)
G	たらどう, ではどう, (よ) うではないか, てほしい て(も) けっこうだ てくれ(ないか), てもらえないか, てくださいませんか, ていただけますか, ください お〜ください, お〜願います, てほしい, てもらいたい, ていただきたい, てちょうだい ねばならぬ, なくてはいけない, なければだめ(だ), べきだ ざるを得ない, よりほか(は) ない, ものだ ことはならない, てはだめ(だ), たらだめ(だ), べからず, ものではない もので(は) ない, わけではない, わけにはいかない, べきで(は) ない	消極的願望 許可 要求 依頼 必然 必然 禁止 禁止
グループ	形態素	認識素性の主たる値
H	かもしれない, かもわからない に違いない, に相違ない, に決まっている, はずだ ところだ, のだろう, ことだろう ことになる, にきまっている	推定 確信 推定 確定