Internal Use Only (非公開)

TR-I-0334

日本語解析文法の開発の軌跡

Development History of the Unification-Based Spoken-Style Japanese Grammar for ASURA

永田昌明　田代敏久　衛藤純司*　坂口明子*
Masaaki NAGATA, Toshihisa TASHIRO,
Junji ETOH, Akiko SAKAGUCHI

1993.3

概要

本報告では、音声翻訳システムのために、約4年間に渡って継続的に行われた、日本語解析文法の開発の経緯を述べる。まず、音声翻訳システムが対象としたコーパス、および、このコーパスに対応する日本語文法の版(Version)について解説する。次に、90年4月から92年9月までの間、半年ごとに作成された文法開発の進捗報告に基づいて、各版とその直前の版との違いについて解説する。

ATR 自動翻訳電話研究所
ATR Interpreting Telephony Research Laboratories

©ATR 自動翻訳電話研究所
©ATR Interpreting Telephony Research Laboratories

---

*日本アイアール株式会社

目次

1 はじめに 1
1.1 本報告の構成 1
1.2 音声翻訳実験システムが処理対象とするコーパス 1
1.3 日本語解析文法の開発の経緯 3
1.4 音声翻訳実験システムと日本語解析文法の関係 4

2 粗粒度文法から中粒度文法への書き換え 7
2.1 概要 7
2.2 述語の生成および述語要素の承接 7
2.3 助詞の承接 12
2.4 副詞類の修飾先の指定 15
2.5 複合名詞と数量表現 17
2.6 その他 19
2.7 解析実験結果 20
2.8 処理速度の向上のために考えられること 23
2.9 補遺 24
2.10 さらなる改良の試み 26

3 会話 6 〜会話 10 の文法 31
3.1 概要 31
3.2 新たに処理対象となる言語現象と、その処理方法 32
3.2.1 連用形接続 (連用中止) 32
3.2.2 不定語 32
3.2.3 真偽疑問文の補足節 33
3.2.4 取立助詞の意味 34
3.2.5 可能表現のガ / ヲ交替 34
3.2.6 複合形容名詞 36
3.2.7 音形のない (ポーズによる) 並列表現 36
3.2.8 「– になる」 37
3.2.9 名詞を下位範疇化する引用の助詞 38
3.2.10 多文節名詞連続複合語 (助詞 / 動詞の省略) 38
3.3 その他の変更 39
3.3.1 提題助詞「は」と取り立て助詞「も」の扱い 39
3.3.2 複文の曖昧性の処理 39

3.3.3 関係節の解析: 内の関係と外の関係 40
3.3.4 サ変名詞、尊敬転成名詞を修飾する副詞 40
3.4 解析実験結果 42
3.4.1 会話 6 〜会話 10 の解析実験結果 42
3.4.2 旧文法と新文法の比較 45

4 A ランクの文法 49
4.1 はじめに 49
4.2 A ランクの実験文のために追加した規則 49
4.2.1 文接続の規則 49
4.2.2 副詞句の規則 49
4.2.2.1 副詞的に使われる相対名詞 49
4.2.2.2 形容詞の連用中止法 50
4.2.3 助動詞の規則 50
4.2.3.1 「〜ではありません」「〜ではございません」 50
4.2.3.2 形容詞語幹 + 助動詞 50
4.2.4 助詞類の規則 51
4.2.4.1 総主文 51
4.2.4.2 動詞に後続する助詞 51
4.2.4.3 形容名詞に直接かかる必須格補語 52
4.2.4.4 数量名詞 + 取立助詞 52
4.2.5 連体詞の規則 52
4.3 残された課題 53
4.4 資料 1: A ランク解析実験文の解析所要時間 54
4.5 資料 2: A ランク解析実験文でまだ解析ができていないもの 62
4.5.1 機械のメモリ不足で現在では解析不可能な実験文 62
4.5.2 交差現象のために解析不可能の文 62
4.5.3 まだ解析規則を記述していないために解析できない文 63
4.5.3.1 総主構文 63
4.5.3.2 受身文 (助詞のコントロール) 63
4.5.3.3 変則的な使役文 (助詞のコントロール) 63
4.5.3.4 並列表現 63
4.5.3.5 形容詞の連体形接続 63
4.5.3.6 従属文の述語がない 63
4.5.3.7 意味記述の誤り 63
4.6 資料 3: 解析結果において曖昧性が生じた文の分析 64
4.7 資料 4: 旧文法と新文法による会話 1 〜会話 10 の結果比較 70

5 B・C ランクの文法 75
5.1 はじめに 75
5.2 活用語と活用語尾の扱い 75
5.3 時制の助動詞「た」「う」の扱い 76
5.4 副詞的名詞 76

5.5 形式副詞 77
5.6 取立助詞 77
5.7 疑問代名詞・疑問副詞 78
5.8 接頭辞・接尾辞の規則 78
5.8.1 接頭辞の規則 78
5.8.1.1 構文規則 78
5.8.1.2 意味規則 79
5.8.2 接尾辞の規則 81
5.8.2.1 構文規則 81
5.8.2.2 意味規則 81
5.9 数詞の規則 82
5.9.1 桁読み型の数詞 82
5.9.2 棒読み型の数詞 82

6 E セット (機能試験文 600 文) の文法 83
6.1 文法規則の変更 83
6.1.1 スタートルールの変更 83
6.1.2 句点を扱う規則の削除 83
6.1.3 述語接続規則の変更 83
6.1.4 形式副詞から構成される副詞句のカテゴリ名の変更 84
6.1.5 助動詞「だ」「です」のカテゴリ名の変更 84
6.1.6 変化の構文 85
6.2 SEM 部 (意味表現) の変更 86
6.2.1 名詞句 86
6.2.1.1 数詞表現 88
6.2.1.2 数量表現 89
6.2.1.3 時間の表現 90
6.2.1.4 取立助詞が後続する名詞句 90
6.2.1.5 人名、地名など固有名詞の名詞句 91
6.2.1.6 接頭辞、接尾辞類が後続する名詞句 92
6.2.1.7 形式名詞型の意味表現 92
6.2.1.8 形式名詞型の意味表現をとる接尾辞 92
6.2.1.9 副詞修飾の副詞 93
6.3 PRAG 部の変更 93

7 M セット (モデル会話文) の追加作業 95
7.1 音声認識誤り検出用の意味情報の付加 95
7.1.1 述語のとる必須格補語の意味素性の制限 95
7.1.2 格助詞のとる補語の意味素性の制限 95
7.1.3 任意格要素の修飾先の制限 96
7.1.4 形式副詞の修飾先の制限 96
7.1.5 感動詞を修飾する副詞の制限 96

参考文献 97

第 1 章

はじめに

1.1 本報告の構成

本報告では、最終的に ASURA と呼ばれることになった、音声翻訳実験システムの作成のために、約 4 年間に渡って継続的に行なわれた日本語対話文解析用文法 (以下では「日本語解析文法」と呼ぶ) の開発の経緯を述べる。

本報告の目的は、次の世代の研究者のために、過去の文法開発の歴史を記述しておくことである。実際の文法開発の過程は、試行錯誤の連続である。しかし、出来上がった文法から、そこに至るまでの経緯を読み取ることは、非常に困難である。従って、本報告が、日本語解析文法がどうして現在のような形になっているかを理解する上での手助けになれば幸いである。

また、日本語に関する大規模な文法を開発する際の問題点について述べた資料は、これまで、ほとんど存在していない。これは、現在の日本では、大規模な自然言語処理システムに関する研究が、主に企業の研究所で行なわれているために、このような技術資料が公開される機会が少ないことに原因があると思われる。その意味でも、本報告は、文法開発の「現場」を記録した貴重な資料であるといえるだろう。

次に本報告の構成を述べる。まず、第 1 章では、音声翻訳実験システムが処理対象としたコーパス、および、このコーパスに対応する日本語文法のバージョンについて、音声翻訳実験システムの開発の歴史を辿りながら説明する。

第 2 章以降では、文法開発に際して、90 年 4 月から 92 年 9 月までの間、半期ごとに作成された進捗報告を掲載する。これらの進捗報告は、すべて直前の版の文法に対する差分の形で記述されている。そして、最初の進捗報告が比較対象にしているのは、90 年 3 月現在の日本語解析文法[永田・久米, 90] である。

なお、93 年 3 月現在の日本語解析文法に関しては、これを包括的に記述したテクニカルレポートを別途作成する[永田 ほか, 93]。従って、日本語解析文法の内容を系統的に知りたい読者は、このテクニカルレポートを参照して頂きたい。

1.2 音声翻訳実験システムが処理対象とするコーパス

ATR の音声翻訳実験システムが処理対象としている文は、大きく次の二つに分けられる。

モデル会話 (M-set) それぞれが特定のテーマ、例えば、参加申し込み、会議の概要、観光ツアー、ホテルの予約、会場への行き方、キャンセル、などを持った会話形式のデータ。全部で 12 会話 (262 文) ある。

(機能) 試験文 (E-set) それぞれが特定の言語表現、例えば、受動表現、使役表現、依頼表現、などの典型例となっている文。これらは、国際会議予約の領域に現れる、話し言葉の文となるように作られているが、会話形式にはなっていない。全部で 600 文 (あるいは追加分を入れて 638 文) ある。

「モデル会話」および「機能試験文」は、それぞれ、通称、"M-set" および "E-set" と呼ばれている[1]。また、M-set を対象とした音声翻訳システムを「700 語バージョン」、E-set を対象とした音声翻訳システムを「1500 語バージョン」と呼ぶこともある[2]。

ただし、音声翻訳実験システムのターゲットが、M-set および E-set として固まったのは、92 年 6 月頃のことである。歴史的には、次の 4 つの試験文の集合が、音声翻訳実験システムのターゲットとして用いられた。

- モデル会話 A,B
- モデル会話 1 ～ 10
- 語彙の重要度に基づく試験文 (ランク A,B,C)
- 表現の重要度に基づく試験文 (E-set)

モデル会話 A,B とモデル会話 1 ～ 10 は、文としては多少重複している部分がある。しかし、会話としては別の内容なので、別々に扱っている。また、モデル会話は、番号が大きくなるほど、複雑な言語表現を含むように設計されている。

E-set と呼ばれる試験文が出来るまでは、通称「ランク A,B,C」と呼ばれる試験文が使われていた。これは、ATR 言語データベース[江原ほか, 90]の品詞別の出現頻度統計から、「基本語彙セット」と語彙の重要度を決め、各語彙の典型的な用例となるような試験文を作成したものである。

その後、音声翻訳システムが対象とする言語表現を選定する非公式の委員会により、「対象言語表現」を決定する努力が続けられた。この委員会は、91 年 11 月頃から、日本語学や英語学に造詣の深い大学の先生方を交えて、「翻訳評価サブワーキンググループ」として活動を始めた。そして、「目的指向型」の電話会話によく使われる表現を重要度に応じて 5 段階にランク付けしたものが、92 年 3 月頃に「目的指向型電話会話表現形」としてまとめられた。

この「目的指向型電話会話表現形」の上位 3 ランクを含むように、「ランク A,B,C」の試験文を大幅に修正・変更して出来上がったのが、E-set の試験文である[3]。E-set の試験文は、この後も若干変更されている。また、翻訳評価のために更に追加された試験文もある。最終的な「対象言語表現」および「試験文」の一覧については、まもなくテクニカルレポートとしてまとめられる予定である。

---

[1] "M" および "E" が何を表すのかについては諸説あり、定かではない。

[2] システム内部の辞書の語数は、音声認識、言語解析 (日本語)、言語生成 (英語)、などの各モジュールで異なるので、700 とか 1500 という数値は単なる目安でしかない。

[3] この当時の M-set および E-set の一覧が、[永田・衛藤, 92c]の付録に掲載されている。このテクニカルレポートは、当時の問題となっていた、試験文の日本語表記に関する混乱を収拾する目的で作成された。

1.3 日本語解析文法の開発の経緯

表1.1に、日本語解析文法の開発の進捗状況を示す。表の中の「対象」という項目は、新たに処理対象となった試験文を示している。文法は、常に、以前の処理対象を包含するように拡張されている。

表 1.1: 日本語解析文法の開発の進捗状況

期間	対象	版	コメント
～89年2月	モデル会話 AB	1.x	理論の検証が目的
89年3月～90年3月	モデル会話1～5	2.x	SL-TRANS の最初の文法
90年4月～90年9月	モデル会話1～5	3.x	粗粒度から中粒度へ
90年10月～91年3月	モデル会話6～10	3.x	
91年4月～91年9月	ランク A	3.x	
91年10月～92年3月	ランク BC	4.x	接辞・取り立て助詞などの扱い
92年4月～92年9月	E-set	4.x	意味表現の大幅な改良

最初に作られた文法[吉本, 88]は、モデル会話 AB を対象としていた。これを Version 1.x と呼ぶことにする。この文法の開発がいつ頃始まったのかは、良く分らない。

報告者の一人が文法の開発に携わるようになったのは、89年3月からである。89年3月～90年3月の間に、モデル会話1～5を対象として、文法の大幅な拡張が行なわれた[永田・久米, 90]。これを Version 2.x と呼ぶことにする。

90年4月～90年9月にかけて、日本語解析部の処理効率の向上のために、Version 2.x の句構造規則を詳細化する作業が行なわれた。これにより、それ以前は 20 個程度だった句構造規則が 200 個程度になった。我々は、前者を粗粒度文法、後者を中粒度文法と呼んでいる[永田, 91][4]。この詳細化以降の文法を Version 3.x と呼ぶことにする。この作業に関する報告書を、第2章に示す。

90年10月～91年3月にかけて、モデル会話6～10を対象とする文法が作成された。この作業に関する報告書を、第3章に示す。また、この頃から、ランク A,B,C の試験文を対象とした文法開発のための基礎データを得る目的で、「日本語基本辞書」[永田・衛藤, 92a][永田・衛藤, 92b]の作成が開始された。

91年4月～91年9月にかけて、ランク A を対象とする文法が作成された。この作業に関する報告書を、第4章に示す。この版までの文法は、基本的に、90年4月～90年9月にかけて作られた文法を拡張して作られているので、Version 3.x の系統に入る。

91年10月～92年3月にかけて、ランク B,C を対象とする文法が作成された。この作業に関する報告書を、第5章に示す。この文法を作成する際に、一部の助動詞や接頭辞・接尾辞の扱いを大幅に変更し、複合名詞・数量名詞・取り立て助詞などを本格的に扱うようになった。従って、これ以降の文法は、Version 4.x と呼ぶことにする。

92年3月～92年9月にかけて、E-set を対象とする文法が作成された。なお、E-set の機能試験文が決定したのは、この期間の前半なので、この期間の作業は、最初、ランク ABC

[4] この文献の中では、詳細化された句構造規則を「細粒度」としているが、後に、「中粒度」と呼ぶように改めた。

の文法の改良という名目で始められた。92年6月〜8月にかけて、複合名詞・数量名詞・取り立て助詞などの扱いを改良するために、日本語の意味表現の大幅な変更を行なった。しかし、この変更は、意味計算の記述のみ影響し、句構造規則はあまり変更されなかったので、依然として、この文法は、Version 4.x の系統である。この作業の報告書を、第6章に示す。

文法のバージョン管理が始まったのは、実は、92年6月頃からである。この時、当時のM-set 対応版を Version 3.1 、E-set 対応版を Version 4.0 とすることが決まった。

当初、Version 3.x は接辞や数量名詞の取り扱いが不十分なので obsolete とし、日本語文法は E-set 対応の Version 4.x で一本化する予定であった。しかし、自動翻訳電話プロジェクト最終年度にあたり、92年7月(COLING-92)、92年11月(研究発表会)、93年1月(自動翻訳電話国際共同実験)という3つの大きなデモンストレーションを成功させるために、十分に動作の安定した日本語解析文法が要求されていたので、M-set 対応の Version 3.1 は、引続きサポートされることになった。この作業の報告書を、第7章に示す。

その後、M-set 対応の Version 3.1 は、音声認識誤りを排除するための意味記述が強化されて、Version 3.2 となった。また、E-set 対応の Version 4.0 は、テンプレートの体系が整理されて、Version 4.1 となった

1.4 音声翻訳実験システムと日本語解析文法の関係

最後に、ATR の音声翻訳実験システムと日本語解析文法のバージョンとの関係について述べておく。日本語解析文法は、次の三つの(音声)翻訳実験システムの中で使用された。

- 対話翻訳実験システム NADINE (88年11月〜)
- 音声言語日英翻訳実験システム SL-TRANS (89年5月〜)
- ATR 音声翻訳システム ASURA (92年11月〜)

NADINE[Kogure et. al, 90] は、端末間の対話のように、キーボード入力による対話文の翻訳システムである。このシステムには、モデル会話 AB を対象とする Version 1.x の日本語文法が用いられた。

SL-TRANS[小倉ほか, 89] は、89年秋に、ATR が初めて公開した音声翻訳実験システムである。このシステムには、最初、モデル会話1〜5を対象とする Version 2.x の日本語文法が用いられた[5]。その後、90年後半から、中粒度文法である Version 3.x が使われるようになった。

91年5月に、天皇陛下が ATR を御訪問になった。それ以前のシステムは、予め録音された音声を用いていたが、この時に初めて、マイクによる音声入力が可能になり、真に「音声翻訳システム」と呼べるものが出来上がった。これ以降のシステムを、SL-TRANS 2 と呼ぶこともある。

ASURA[竹沢ほか, 93] は、ATR の自動翻訳電話プロジェクトの研究成果を結集して作られた、世界初の自動翻訳電話システムである。このシステムには幾つかのバージョンがあ

[5]ただし、89年5月、精華町の現研究所への移転直後のオープンハウスで、暫定的に公開したデモシステムは Version 1.x を使用していた。

るが、日本語文法に関しては、93 年 3 月現在、M-set に対しては Version 3.2 が、E-set に対しては Version 4.1 が用いられている。

なお、93 年 1 月 28 日の自動翻訳電話国際共同実験で用いられた日本語文法は、基本的には、Version 3.x の系統であるが、この実験が対象とするコーパスのために特別に用意された。従って、この文法は "C-STAR Version" と呼ばれている。

第 2 章

粗粒度文法から中粒度文法への書き換え

2.1 概要

日本語対話文の解析用文法を、中規模 (164 個) の句構造規則に展開し、これによって不必要となった素性を削除するとともに、語彙規則の中の選言をできるだけ減らした。

今回は、次のテーマについて、主として単語間の連接に関する制約の構造化を行なった。

- 述語の生成および述語要素の承接
- 助詞の承接
- 副詞類の修飾先の指定
- 複合名詞と数量詞

以下に具体的に述べる。

2.2 述語の生成および述語要素の承接

述語は、動詞または名詞述語にいくつかの助動詞、補助動詞が連接して生成される。これは、次の 2 つの規則で扱われていた。

$$V \rightarrow (N \quad AUXV) \tag{2.1}$$
$$V \rightarrow (V \quad AUXV) \tag{2.2}$$

「名詞 + だ、です、etc.」「サ変名詞 + する、いたす、etc.」「形容名詞 + だ、です」の区別は、規則 (2.1) の注釈に選言で、また、助動詞と補助動詞の承接は、規則 (2.2) の注釈に modl 素性で、記述されていた。いずれも、かなり大きく複雑な注釈である。

これを句構造規則で表すために、述語要素の承接関係に図 2.1のような階層があることを利用した。

すなわち、ボイスの助動詞は他のすべての助動詞、補助動詞に先行し、かつ、使役の「(さ)せる」が受身の「(ら)れる」に先行する。

動作様相の補助動詞「(て)いる、(て)おる、etc.」と受益、待遇の補助動詞「(て)もらう、(て)やる、etc.」は、ボイスの助動詞に後続し、ムードの助動詞に先行する。そして、

>	ボイス	>	アスペクト	>	ムード1	>	否定	>	テンス	>	ムード2	>	テンス
	(さ)せる		(て)いる		たい		ない		た		らしい		た
	(ら)れる		(て)もらう		たがる		ん		だ		です		う
					ます				だろう				

図 2.1: 日本語の述語要素の階層

これら 2 つの補助動詞は互いに位置を交換することができる。「- ていてもらう」「- てもらっている」というように。そこで、これらをまとめて、アスペクト要素とする。

話者の主体的表現を表すムードの助動詞を、否定やテンスの助動詞との承接関係を考慮して、ムード 1 のグループとムード 2 のグループに分ける。ムード 1 には、願望の助動詞「たい」や義務の助動詞「べき、なくてはいけない、etc.」がある。また、丁寧の助動詞「ます」もこの中に入る。ムード 2 には、心的様相の助動詞「らしい、そうだ、ようだ、etc.」と丁寧の助動詞「です」が入る。また、推量の助動詞「だろう、かもしれない」もこの中に入る。

否定の助動詞「ない、ん」は、ボイス、アスペクト、ムード 1 の助動詞補助動詞に後続し、テンス、ムード 2 の助動詞に先行する。

過去を表す助動詞「た、だ」は、やや特異なふるまいをする。すなわち、ボイス、アスペクト、ムード 1、否定の助動詞、補助動詞に後続し、ムード 2 の助動詞に先行する一方で、ムード 2 の助動詞に後続することもできる。(このことから、「句構造文法にもとづく日本語文の解析」[吉本, 88]に述べられたような、述語要素の相互承接関係の例外が生じる)。

意思や推量を表す助動詞「う」は、他のすべての助動詞、補助動詞に後続する。

以上のことを考慮して、述語のタイプを表 2.1のように分類することにする。

表 2.1: 述語の分類

V-kernel	本動詞
V-voice	ボイスの述語
V-aspect	アスペクトの述語
V-mood1	ムード 1 の述語
V-negt	否定の述語
V-mood2	ムード 2 の述語
V-tense	テンスの述語

また、助動詞、補助動詞を表 2.2のように分類する。

すると、述語の生成規則は次のようになる。

表 2.2: 助動詞・補助動詞の分類

AUXV-caus	使役	せる、させる
AUXV-deac	受身、可能、自発、尊敬	れる、られる
AUXV-aspc	動作様相	ている、てある、ておく
AUXV-dont	受益、待遇	てもらう、ていただく、てくれる
AUXV-optt	願望、義務、丁寧	たい、てほしい、べき、ます
AUXV-negt	否定	ない、ん
AUXV-tense	テンス	た、だ、う
AUXV-evid	心的様相	だろう、らしい、そうだ、ようだ
AUXV-polt-aux	「のです」「ませんです」の「です」	
AUXV-copl	判定	だ、です、である
AUXV-npred	名詞述語形成	する、いたす、なさる、くださる

```
V-voice -> (V-kernel  AUXV-voice)
V-aspect -> (ADV  AUXV-aspc)
V-aspect -> (ADV  AUXV-dont)
V-mood1 -> (V-kernel  AUXV-optt)
V-mood1 -> (V-voice  AUXV-optt)
V-mood1 -> (V-aspect  AUXV-optt)
V-negt -> (V-kernel  AUXV-negt)
V-negt -> (V-voice  AUXV-negt)
V-negt -> (V-aspect  AUXV-negt)
V-negt -> (V-mood1  AUXV-negt)
V-tense -> (V-kernel  AUXV-tense)
V-tense -> (V-voice  AUXV-tense)
V-tense -> (V-aspect  AUXV-tense)
V-tense -> (V-mood1  AUXV-tense)
V-tense -> (V-negt  AUXV-tense)
V-mood2 -> (V-kernel  AUXV-evid)
V-mood2 -> (V-voice  AUXV-evid)
V-mood2 -> (V-aspect  AUXV-evid)
V-mood2 -> (V-mood1  AUXV-evid)
V-mood2 -> (V-negt  AUXV-evid)
V-mood2 -> (V-tense  AUXV-evid)
V-tense -> (V-mood2  AUXV-tense)
```

ただし、AUXV-voice の生成規則として次の規則が必要である。

```
AUXV-voice -> (AUXV-caus)
AUXV-voice -> (AUXV-deac)
AUXV-voice -> (AUXV-caus  AUXV-deac)
```

さらに、本動詞、ボイス、アスペクト、ムード1、否定、過去、ムード2、推量の述語をVに上げる規則として次の規則が必要である。

```
V => (V-kernel)
V => (V-voice)
V => (V-aspect)
V => (V-mood1)
V => (V-negt)
V => (V-tense)
V => (V-mood2)
```

終助詞は文の終りにつく。そこで、次の規則。

```
V-sfp -> (V  AUXV-sfp)
```

終助詞の承接関係に関する規則として、次の規則。

```
AUXV-sfp -> (AUXV-sfp-1)
AUXV-sfp -> (AUXV-sfp-2)
AUXV-sfp -> (AUXV-sfp-3)
AUXV-sfp -> (AUXV-sfp-1  AUXV-sfp-2)
AUXV-sfp -> (AUXV-sfp-1  AUXV-sfp-3)
AUXV-sfp -> (AUXV-sfp-2  AUXV-sfp-3)
AUXV-sfp -> (AUXV-sfp-1  AUXV-sfp-2  AUXV-sfp-3)
```

本動詞のうち「名詞 + だ、です」「サ変名詞 / 転成名詞 + する、いたす、なさる、etc.」「形容名詞 + だ、です」は、前述のように、(2.2) の規則とその注釈によって記述していた。これを句構造規則で表すために名詞を表 2.3のように分類する。

「名詞 + だ、です」および「形容名詞 + だ、です」は次の規則によって生成する。

```
V-kernel -> (N-com  AUXV-copl)
V-kernel -> (N-prop  AUXV-copl)
V-kernel -> (N-num  AUXV-copl)
V-kernel -> (N-quant  AUXV-copl)
V-kernel -> (N-temp  AUXV-copl)
V-kernel -> (N-anaph  AUXV-copl)
V-kernel -> (N-wh  AUXV-copl)
V-kernel -> (N-adj  AUXV-copl)
```

N.B. V-kernel -> (N AUXV-copl) とすると、「(尊敬) サ変名詞 / 転成名詞 + です」に曖

表 2.3: 名詞の分類

N-com	普通名詞
N-prop	固有名詞
N-num	数詞
N-quant	数量詞
N-temp	時詞
N-anaph	代名詞
N-wh	疑問詞
N-adj	形容名詞
N-sahen	サ変名詞
N-verb	転成名詞
FN	形式名詞

昧性が生じる。

「(尊敬) サ変名詞 / 転成名詞 + する、いたす、なさる、です」は次の規則によって生成する。

```
V-kernel -> (N-sahen  AUXV-npred)
V-kernel -> (N-sahen-honor  AUXV-npred)
V-kernel -> (N-sahen-honor  AUXV-polt-aux)
V-kernel -> (N-verb  AUXV-npred)
V-kernel -> (N-verb-honor  AUXV-npred)
V-kernel -> (N-verb-honor  AUXV-polt-aux)
```

2.3 助詞の承接

助詞は、名詞に格助詞が連接して述語補語を、副詞に提題助詞や取り立て助詞が連接して連用修飾語をつくる。また、格助詞が提題助詞や取り立て助詞に交替したり、格助詞に提題助詞や取り立て助詞、連体助詞が連接することもある。これらの現象は、次の規則で扱われていた。

$$P \rightarrow (N \quad POSTP) \tag{2.3}$$
$$ADV \rightarrow (ADV \quad POSTP) \tag{2.4}$$
$$P \rightarrow (P \quad POSTP) \tag{2.5}$$
$$V \rightarrow (P \quad V) \tag{2.6}$$
$$V \rightarrow (ADV \quad V) \tag{2.7}$$

助詞の連接可能性は、これらの規則の注釈に表層形を記述して指定していた。
これを句構造文法で表すために、助詞を表 2.4のように分類する。

表 2.4: 助詞の分類

POSTP-oblg	必須格助詞	が、を、に
POSTP-optn	任意格助詞	で、から、より、へ、と、 etc.
POSTP-topic	提題助詞	は
POSTP-contr	取り立て助詞	も、こそ、すら、さえ、 etc.
POSTP-quote	引用助詞	と、か
FATT	形式連体詞 (連体助詞)	の、と (並列)、や (並列)
FADV	形式副詞 (接続助詞)	ば、たら、なら、から、 etc.

また、後置詞句のタイプを修飾の仕方によって表 2.5のように分類する。

表 2.5: 後置詞句のタイプの分類

P	動詞の必須格になるもの
ADV	動詞の任意格になるもの
ATT	名詞の修飾句になるもの

すると、後置詞句は次の規則によって生成される。

```
P -> (N  POSTP-oblg)
ADV -> (N  POSTP-optn)
ATT -> (N  FATT)

P -> (N  POSTP-topic) ; 提題助詞、取り立て助詞は格助詞の代
P -> (N  POSTP-contr) ; わりに用いられる。

P -> (V  POSTP-quote) ; 文に引用助詞の「と」がつくと、「思
                      ;  う」などの必須格となる。

ADV -> (V  FADV) ; 文に形式副詞が接続して、副詞句をつくる。
```

必須格後置詞句は、本動詞の補語、ボイスの述語の補語、ムード1の述語の補語になる。

```
V-kernel -> (P  V-kernel)
V-voice -> (P  V-voice)   ; 登録用紙が送られる
V-mood1 -> (P  V-mood1)   ; 私は参加したい
```

任意格後置詞句は、本動詞の補語になる。

```
V-kernel -> (ADV  V-kernel)
```

連体修飾後置詞句は名詞を修飾する。被修飾名詞の分類に従って、次の一連の規則を設ける。

```
N-com -> (ATT  N-com)
N-com-honor -> (ATT  N-com-honor)
N-prop -> (ATT  N-prop)
N-num -> (ATT  N-num)
N-quant -> (ATT  N-quant)
N-temp -> (ATT  N-temp)
```

助詞相互の連接は、これまで (2.5) の規則とその注釈で表層形を指定することで処理していた。これを、最初に助詞同士を結合して、いわば助詞句を形成し、しかるのちに名詞その他に接続させることにした。この方が、処理時間が短くなる。

```
POSTP-oblg -> (POSTP-oblg POSTP-topic)  ; には
POSTP-oblg -> (POSTP-oblg POSTP-contr)  ; をも、にも
POSTP-oblg -> (POSTP-contr POSTP-oblg)  ; だけが(を、に)

POSTP-optn -> (POSTP-optn POSTP-topic)  ; では
POSTP-optn -> (POSTP-optn POSTP-contr)  ; でも
POSTP-optn -> (POSTP-contr POSTP-optn)  ; ばかりへ

POSTP-quote -> (POSTP-quote POSTP-oblg) ; (--する)かが
POSTP-quote -> (POSTP-quote POSTP-topic); (--する)とは
POSTP-quote -> (POSTP-quote POSTP-contr); (--する)とも

POSTP-contr -> (POSTP-contr POSTP-contr); さえも
```

2.4 副詞類の修飾先の指定

副詞類が文末の述語を修飾するが、それは次の規則によって解析していた。

$$V \rightarrow (ADV \quad V) \tag{2.8}$$

そして、副詞が文末のどの述語要素を修飾するかについては、modl 素性を副詞の coh 素性に記述していた。ただし、否定の副詞を除いて、すべて、!caus- すなわち、本動詞を修飾するように記述されていた。

これを、句構造文法で表すために、副詞を表 2.6のように分類する。

表 2.6: 副詞の分類

述語修飾副詞	ADV-mann	様態副詞	ゆっくり、至急
	ADV-degr	程度副詞	ちょっと、全然
	ADV-tloc	時間副詞	もうすぐ、いま
	ADV-freq	頻度副詞	しばしば、たまに
文修飾副詞	ADV-state	陳述副詞	たぶん、ぜひ、別に
	ADV-uttr	発言副詞	実は、失礼ですが
	ADV-conj	接続副詞	では、ところで

様態副詞と程度副詞は本動詞を修飾する。したがって、次の規則を設ける。

```
V-kernel -> (ADV-mann  V-kernel)
V-kernel -> (ADV-degr  V-kernel)
```

時間副詞は、テンスやアスペクトに作用する。このうち、テンスは動詞の基本形と過去形の対立によって、アスペクトは動詞の基本形と過去形、「ている」「てしまう」など動詞のテ形に後接する補助動詞によって表される。したがって、次の規則を設ける。

```
V-kernel -> (ADV-tloc  V-kernel)
V-voice -> (ADV-tloc  V-voice)
V-aspect -> (ADV-tloc  V-aspect)
V-tense -> (ADV-tloc  V-tense)
```

N.B. 時間副詞の係り受けについては、したがって、最大 4 通りの曖昧性が生じる。

頻度副詞は、アスペクトの層の外側にあって、これらの現れを規定する。したがって、次の規則を設ける。

```
V-kernel -> (ADV-freq  V-kernel)
V-voice -> (ADV-freq  V-voice)
V-aspect -> (ADV-freq  V-aspect)
V-tense -> (ADV-freq  V-tense)
```

陳述副詞、発言副詞、接続副詞などの文修飾副詞は、後続の文全体を修飾する。したがって、次の規則を設ける。

```
V -> (ADV-state  V)
V -> (ADV-uttr  V)
V -> (ADV-conj  V)
```

副詞が終助詞を含む述語に係らないようにするために、V-sfp を設け、次のスタートルールをあらたに設けた。

```
START start=>v (V-sfp)
```

2.5 複合名詞と数量表現

複合名詞や数量表現は、氏名や住所、年月日を除いて、基本的に解析していなかった。氏名、住所、年月日については、次の規則と ncat 素性で扱っていた。

$$N \rightarrow (N \quad N) \tag{2.9}$$

これを、句構造規則で表すために、氏名、住所、年月日を表す名詞を、表 2.7のように細分類した。

表 2.7: 氏名・住所・年月日を表す名詞の分類

N-family	姓
N-name	名
N-city	市
N-ward	区
N-town	町
N-chome	丁目
N-banchi	番地
N-temp	年、月、日

生成規則は次のとおりである。

```
N-prop -> (N-family  N-name)   ; 日本人名
N-prop -> (N-name  N-family)   ; 外国人名

N-ward -> (N-city  N-ward)     ; 大阪市北区
N-town -> (N-ward  N-town)     ; 大阪市北区茶屋町
N-chome -> (N-town  N-chome)   ; 大阪市北区茶屋町二丁目
N-banchi -> (N-chome  N-num)   ; 大阪市北区茶屋町二丁目二三

N-temp -> (N-temp  N-temp)     ; 八月二日
```

数量詞を句構造文法で解析するために、助数詞を表 2.8のように分類した。

表 2.8: 助数詞の分類

PREFIX-quant	前置助数詞	約、およそ、のべ
SUFFIX-quant	後置助数詞	人、回、円、パーセント
SUFFIX-num	助数詞承接	頃、以上、未満、ほど

生成規則は次のとおりである。

```
N-quant -> (N-num  SUFFIX-quant)  ; 千円
N-quant -> (PREFIX-quant  N-quant); 約千円
N-quant -> (N-quant  SUFFIX-num)  ; 千円以上
```

数字の棒読みと桁読みの区別は、大きな数字になると曖昧性が非常に増えるので、今回は扱わなかった。

また、固有名詞を受けて尊敬を表す接尾辞、普通名詞 / サ変名詞 / 転成名詞について丁寧さを表す接頭辞を図 2.9のように設けた。

表 2.9: 接頭辞・接尾辞の分類

SUFFIX-prop	固有名詞承接	様、さん、君
PREFIX-honor	敬意添加型	ご、お

生成規則は次のとおり。

```
N-prop -> (N-prop  SUFFIX-honor)  ; 鈴木様

N-com-honor -> (PREFIX-honor  N-com)     ; お名前
N-sahen-honor -> (PREFIX-honor  N-sahen) ; ご参加
N-verb-honor -> (PREFIX-honor  N-verb)   ; お申し込み
N-adj-honor -> (PREFIX-honor  N-adj)     ; ご必要
```

2.6 その他

時詞のあるものや数量詞は、格助詞「に」を伴わずに述語を修飾できる。これについては、数量詞についてのみ、次の規則で解析することにした。

```
V-kernel -> (N-quant  V-kernel)
```

これに対して、時詞は、基本的に N-temp の品詞を与え、「に」を伴わずに副詞的に用いられることが多いものに ADV-tloc の品詞を与えることにして、次のような規則は設けなかった。

```
V-kernel -> (N-temp  V-kernel)
```

次のような文で、曖昧性が生じるからである。

```
 ---------------------------
| ----                      |
||    |                     |
三月 二十日 までに 要約を 提出する
```

2.7 解析実験結果

モデル会話の 1 から 5 までを解析実験して、旧文法と新文法の処理時間を比較してみた。ただし、旧文法での実験は SUN-4 で、新文法での実験は SPARC-STATION で行なったので、それぞれの処理速度のちがいを考慮しなければならない。SPARC-STATION の方が SUN-4 より、1.3 倍ほど速いようである。新文法の右側に補正した数字を示す。

新文法 1 は aspect の規則を含む版 (aspect) であり、新文法 2 は aspect の規則を含まない版 (non-aspect) である。

dialogue-1

日本語	旧文法	新文法 1	新文法 2
もしもし	1	0.7/0.9	0.7/0.9
そちらは会議事務局ですか	40	25/33	17/22
はい	1	0.6/0.8	0.6/0.8
そうです	9	6/8	4/5
どのようなご用件でしょうか	32	36/47	27/35
会議に申し込みたいのですが	69	37/48	22/29
どのような手続きをすればよろしいのでしょうか	122	84/109	73/95
登録用紙で手続きをして下さい	23	37/48	19/25
登録用紙は既にお持ちでしょうか	43	29/38	24/31
いいえ	1	1/1.3	0.8/1
まだです	8	5/7	4/5
分かりました	10	12/16	5/7
それでは登録用紙をお送り致します	44	103/134	73/95
ご住所とお名前をお願いします	32	73/95	40/52
住所は大阪市北区茶屋町二十三です	77	35/46	24/31
名前は鈴木真弓です	40	19/25	12/16
分かりました	10	12/16	5/7
登録用紙を至急送らせて頂きます	85	178/231	122/159
よろしくお願いします	17	18/23	12/16
それでは失礼します	14	46/60	27/35

dialogue-2

日本語	旧文法	新文法 1	新文法 2
はい	1	0.6/0.8	0.6/0.8
こちらは会議事務局です	26	17/22	13/17
会議の参加料について教えて頂きたいのですが	239	638/829	583/758
いま会議に申し込めば参加料はいくらですか	343	180/234	68/88
はい	1	0.6/0.8	0.6/0.8
参加料は現在お一人三万五千円です	45	36/47	28/36
来月お申込みになりますと四万円です	121	310/403	84/109
参加料には予稿集代と歓迎会費が含まれています	243	307/399	166/216
わたしは情報処理学会の会員なのですが	124	40/52	40/52
参加料の割引はないのですか	237	44/57	33/43
今回は割引を行っておりません	40	123/160	64/83
そうですか	13	8/10	5/7
参加料はどのようにお支払いしたらよいのですか	272	242/315	132/172
参加料は銀行振り込みです		15/20	12/16
案内書に記載されている口座番号に振り込んで下さい			
また期限は今年いっぱいです	44	25/33	18/23
分かりました	11	12/16	5/7

どうもありがとうございました	5	3/4	3/4
どういたしまして	3	2/3	2/3
分からない点がございましたらいつでもお聞き下さい	386	655/892	168/218
失礼致します	8	7/9	3/4

dialogue-3

日本語	旧文法	新文法 1	新文法 2
はい	1	0.6/0.8	0.6/0.8
こちらは会議事務局です	25	17/22	12/16
会議に論文を発表したいと思っているのですが	380	680/884	900/1170
会議の内容について教えて下さい	41	171/222	78/10
今回の会議は通訳電話に関連する広範な研究分野を含んでいます言語学や心理学を専攻する方にも参加して頂く予定です			
分かりました	10	12/16	5/7
ところで会議での公式言語は何ですか	134	43/56	30/39
英語と日本語です	24	14/18	11/14
わたしは日本語が全然分からないのですが		85/111	53/69
発表が日本語で行われる場合英語への同時通訳はあるのですか		831/1080	659/857
はい	1	0.6/0.8	0.7/0.9
英語への同時通訳を用意しております	61	83/108	52/68
分かりました	11	12/16	5/7
どうもありがとうございました	5	3/4	3/4
さようなら	2	3/4	2/3

dialogue-4

日本語	旧文法	新文法 1	新文法 2
こちらは会議事務局です	28	17/22	12/16
会議について詳しいことを教えて下さい	84	268/348	111/144
会議の案内書はお持ちですか	78	30/40	24/31
いいえ	1	1/1.3	0.8/1
持っていません	12	40/52	26/34
そうですか	14	8/10	5/7
会議は八月二十二日から二十五日まで京都国際会議場で開催されます		650/845	757/984
参加料は四万円です	30	18/23	14/18
発表を希望されるのでしたら三月二十日までに要約を提出して下さい			
会議の案内書をお送り致しますのでそれをご覧下さい	163	99/129	65/85
失礼ですがお名前とご住所をお願いします	82	77/100	43/56
アダムスミスです	30	7/9	5/7
住所は大阪市東区玉造二丁目二十七の七です	272	83/108	62/81
分かりました	13	12/16	5/7
電話番号もお聞きしたいのですが	369	88/114	58/75
はい	1	0.6/0.8	0.6/0.8
三七二の八〇一八です	29	15/20	10/13
三七二の八〇一八でございますね	28	21/27	12/16
はい	1	0.6/0.8	0.6/0.8
そうです	10	6/8	4/5
それではよろしくお願いします	30	25/33	16/21
失礼します	7	8/10	5/7

dialogue-5

日本語	旧文法	新文法 1	新文法 2
はい	1	0.6/0.8	0.6/0.8
こちらは会議事務局でございます	20	26/34	14/18
ちょっとお願いがあるのですが	78	33/43	20/26
私は会議に申込みをした者です	292	1038/1349	168/218
参加を取り消したいのですが	309	61/79	37/48
お名前をお伺いできますでしょうか	101	73/95	46/60
はい	1	0.6/0.8	0.6/0.8
ベル研のジムワイベルです	47	19/25	13/17
既に登録料の八万五千円を振り込まれておられますね	275	495/644	263/342
はい	1	0.7/0.9	0.6/0.8
そうです	10	6/8	3/4
登録料を払い戻して頂けますか	71	127/165	67/87
お気の毒ですができません	11	10/13	7/9
案内書にも書いていますが	128	358/465	182/237
九月二十七日以後の取り消しに対する払い戻しはできません		1681/2185	2084/2709
後日プログラムと予稿集をお送り致します	68	80/104	50/65
では誰かが私の代わりに参加することはできますか	758	306/398	220/286
それは別に問題ありません	11	23/30	12/16
代理人が参加する場合はあらかじめこちらまでお知らせ下さい	72	239/311	71/92
分かりました	10	12/16	5/7
代理人が決まりましたらお知らせ致します	145	180/234	97/126
では失礼します	7	11/14	5/7

以上の結果を見ると、期待に反して、あまり処理速度が上がっていないことがわかる。その理由として、次のようなことが考えられる。

1. 文末の述語が短いもの、すなわち、述語要素の少ないものは、句構造文法の総数が増えたことによる計算量の増大を打ち消すほどには、単一化の回数が減少しない。

2. 数量表現や尊敬名詞は、これまで全体を語彙登録していたが、今回は接辞による解析を行なったので、その分だけ処理速度が遅くなった。

2.8 処理速度の向上のために考えられること

処理速度を向上させるための方法として、次のようなものが考えられる。

1. 述語補語の語順や省略の自由を、現在は Subcat-Slash Scrambling を用いて扱っているが、これを句構造文法で扱う。

2. 述語要素には固定的に用いられるものがある。たとえば、アスペクトの「(て) いる、(て) おる、 etc.」、受益、待遇の「(て) いただく、 (て) 下さる、 etc.」、形式名詞につく「(– の) です」など。これらを、一語として登録する。

3. 助動詞の「です」は会話においてもっとも頻繁に用いられるものだが、意味計算のための選言がまだ残っている。そこで、これをさらに細分類することが考えられる。

2.9 補遺

以下は、新文法 / 旧文法の値を小さい順に並べたもの。

0.2 参加料の割引はないのですか
0.3 アダムスミスです
0.3 電話番号もお聞きしたいのですが
0.3 参加を取り消したいのですが
0.4 わたしは情報処理学会の会員なのですが
0.4 ところで会議での公式言語は何ですか
0.4 住所は大阪市東区玉造二丁目二十七の七です
0.5 会議の案内書はお持ちですか
0.5 ベル研のジムワイベルです
0.5 では誰かが私の代わりに参加することはできますか
0.6 住所は大阪市北区茶屋町二十三です
0.6 名前は鈴木真弓です
0.6 ちょっとお願いがあるのですが
0.7 会議に申し込みたいのですが
0.7 いま会議に申し込めば参加料はいくらですか
0.7 そうですか
0.7 三七二の八〇一八です
0.8 そちらは会議事務局ですか
0.8 はい
0.8 はい
0.8 こちらは会議事務局です
0.8 はい
0.8 そうですか
0.8 また期限は今年いっぱいです
0.8 どうもありがとうございました
0.8 はい
0.8 英語と日本語です
0.8 はい
0.8 どうもありがとうございました
0.8 こちらは会議事務局です
0.8 参加料は四万円です
0.8 会議の案内書をお送り致しますのでそれをご覧下さい
0.8 はい
0.8 はい
0.8 そうです
0.8 はい
0.8 はい
0.8 そうです
0.9 もしもし
0.9 そうです
0.9 どのような手続きをすればよろしいのでしょうか
0.9 登録用紙は既にお持ちでしょうか
0.9 まだです
0.9 こちらは会議事務局です
0.9 お名前をお伺いできますでしょうか
0.9 はい
1.0 参加料は現在お一人三万五千円です
1.0 どういたしまして
1.0 三七二の八〇一八でございますね
1.0 それではよろしくお願いします
1.1 失礼致します
1.2 参加料はどのようにお支払いしたらよいのですか
1.2 失礼ですがお名前とご住所をお願いします
1.2 分かりました
1.2 お気の毒ですができません
1.3 いいえ
1.3 いいえ
1.4 よろしくお願いします
1.4 失礼します
1.5 どのようなご用件でしょうか

1.5 分かりました
1.5 分かりました
1.5 後日プログラムと予稿集をお送り致します
1.6 分かりました
1.6 分かりました
1.6 参加料には予稿集代と歓迎会費が含まれています
1.6 分かりました
1.6 分かりました
1.6 代理人が決まりましたらお知らせ致します
1.7 こちらは会議事務局でございます
1.8 英語への同時通訳を用意しております
2.0 さようなら
2.0 では失礼します
2.1 登録用紙で手続きをして下さい
2.3 分からない点がございましたらいつでもお聞き下さい
2.3 会議に論文を発表したいと思っているのですが
2.3 既に登録料の八万五千円を振り込まれておられますね
2.3 登録料を払い戻して頂けますか
2.7 登録用紙を至急送らせて頂きます
2.7 それは別に問題ありません
3.0 それでは登録用紙をお送り致します
3.0 ご住所とお名前をお願いします
3.3 来月お申込みになりますと四万円です
3.5 会議の参加料について教えて頂きたいのですが
3.6 案内書にも書いていますが
4.0 今回は割引を行っておりません
4.0 持っていません
4.1 会議について詳しいことを教えて下さい
4.3 それでは失礼します
4.3 代理人が参加する場合はあらかじめこちらまでお知らせ下さい
5.0 私は会議に申込みをした者です
5.4 会議の内容について教えて下さい
- 参加料は銀行振り込みです
- 案内書に記載されている口座番号に振り込んで下さい
- 今回の会議は通訳電話に関連する広範な研究分野を含んでいます
- 言語学や心理学を専攻する方にも参加して頂く予定です
- わたしは日本語が全然分からないのですが
- 発表が日本語で行われる場合英語への同時通訳はあるのですか
- 会議は八月二十二日から二十五日まで京都国際会議場で開催されます
- 発表を希望されるのでしたら三月二十日までに要約を提出して下さい
- 九月二十七日以後の取り消しに対する払い戻しはできません

2.10 さらなる改良の試み

報告書で、処理速度を向上させるための方法として述べたことのうち 2) と 3) を実装した。さらに、新文法のアスペクト計算を、V-kernel, V-voice, V-aspect, etc. が V になるときに行なうように修正した。このようにしても、旧文法と同等のアスペクト解析が行なわれるはずである。

以上の修正を行なったのち、改めて解析実験を行なった。

dialogue-1

日本語	旧文法	新文法
もしもし	1	0.7/0.9
そちらは会議事務局ですか	40	18/23
はい	1	0.6/0.9
そうです	9	2/3
どのようなご用件でしょうか	32	31/40
会議に申し込みたいのですが	69	30/39
どのような手続きをすればよろしいのでしょうか	122	81/105
登録用紙で手続きをして下さい	23	31/40
登録用紙は既にお持ちでしょうか	43	29/38
いいえ	1	0.8/1
まだです	8	2/3
分かりました	10	9/12
それでは登録用紙をお送り致します	44	94/122
ご住所とお名前をお願いします	32	55/72
住所は大阪市北区茶屋町二十三です	77	29/38
名前は鈴木真弓です	40	15/20
分かりました	10	10/13
登録用紙を至急送らせて頂きます	85	98/127
よろしくお願いします	17	15/20
それでは失礼します	14	39/51

dialogue-2

日本語	旧文法	新文法
はい	1	0.6/0.8
こちらは会議事務局です	26	14/18
会議の参加料について教えて頂きたいのですが	239	152/198
いま会議に申し込めば参加料はいくらですか	343	157/204
はい	1	0.6/0.8
参加料は現在お一人三万五千円です	45	30/39
来月お申込みになりますと四万円です	121	107/139
参加料には予稿集代と歓迎会費が含まれています	243	
わたしは情報処理学会の会員なのですが	124	40/52
参加料の割引はないのですか	237	40/52
今回は割引を行っておりません	40	44/57
そうですか	13	3/4
参加料はどのようにお支払いしたらよいのですか	272	159/207
参加料は銀行振り込みです	14/18	
案内書に記載されている口座番号に振り込んで下さい		
また期限は今年いっぱいです	44	22/29
分かりました	11	10/13
どうもありがとうございました	5	3/4

どういたしまして	3	2/3
分からない点がございましたらいつでもお聞き下さい	386	202/263
失礼致します	8	5/7

dialogue-3

日本語	旧文法	新文法
はい	1	0.6/0.8
こちらは会議事務局です	25	15/20
会議に論文を発表したいと思っているのですが	380	305/397
会議の内容について教えて下さい	41	64/83
今回の会議は通訳電話に関連する広範な研究分野を含んでいます		
言語学や心理学を専攻する方にも参加して頂く予定です		
分かりました	10	10/13
ところで会議での公式言語は何ですか	134	33/43
英語と日本語です	24	13/17
わたしは日本語が全然分からないのですが		75/98
発表が日本語で行われる場合英語への同時通訳はあるのですか		544/707
はい	1	0.7/0.9
英語への同時通訳を用意しております	61	56/73
分かりまじた	11	10/13
どうもありがとうございました	5	3/4
さようなら	2	2/3

dialogue-4

日本語	旧文法	新文法
こちらは会議事務局です	28	15/20
会議について詳しいことを教えて下さい	84	91/118
会議の案内書はお持ちですか	78	27/35
いいえ	1	0.9/1.2
持っていません	12	15/20
そうですか	14	3/4
会議は八月二十二日から二十五日まで京都国際会議場で開催されます		
参加料は四万円です	30	15/20
発表を希望されるのでしたら三月二十日までに要約を提出して下さい		
会議の案内書をお送り致しますのでそれをご覧下さい	163	89/116
失礼ですがお名前とご住所をお願いします	82	58/75
アダムスミスです	30	6/9
住所は大阪市東区玉造二丁目二十七の七です	272	67/87
分かりました	13	10/13
電話番号もお聞きしたいのですが	369	77/100
はい	1	0.6/0.0.8
三七二の八〇一八です	29	13/17
三七二の八〇一八でございますね	28	14/18
はい	1	0.6/0.8
そうです	10	2/3
それではよろしくお願いします	30	22/29
失礼します	7	8/10

dialogue-5

日本語	旧文法	新文法
はい	1	0.6/0.8
こちらは会議事務局でございます	20	17/22

ちょっとお願いがあるのですが	78	29/38
私は会議に申込みをした者です	292	121/153
参加を取り消したいのですが	309	53/69
お名前をお伺いできますでしょうか	101	67/87
はい	1	0.6/0.8
ベル研のジムワイベルです	47	15/20
既に登録料の八万五千円を振り込まれておられますね	275	210/273
はい	1	0.6/0.8
そうです	10	2/3
登録料を払い戻して頂けますか	71	68/88
お気の毒ですができません	11	8/10
案内書にも書いていますが	128	95/124
九月二十七日以後の取り消しに対する払い戻しはできません		
後日プログラムと予稿集をお送り致します	68	68/88
では誰かが私の代わりに参加することはできますか	758	196/255
それは別に問題ありません	11	13/17
代理人が参加する場合はあらかじめこちらまでお知らせ下さい	72	147/191
分かりました	10	10/13
代理人が決まりましたらお知らせ致します	145	144/187

以下は、新文法 / 旧文法の値を小さい順に並べたもの。

0.2 参加料の割引はないのですか
0.2 参加を取り消したいのですが
0.3 そうです
0.3 そうですか
0.3 ところで会議での公式言語は何ですか
0.3 そうですか
0.3 アダムスミスです
0.3 住所は大阪市東区玉造二丁目二十七の七です
0.3 電話番号もお聞きしたいのですが
0.3 そうです
0.3 そうです
0.3 では誰かが私の代わりに参加することはできますか
0.4 まだです
0.4 わたしは情報処理学会の会員なのですが
0.4 会議の案内書はお持ちですか
0.4 ベル研のジムワイベルです
0.5 住所は大阪市北区茶屋町二十三です
0.5 名前は鈴木真弓です
0.5 ちょっとお願いがあるのですが
0.5 私は会議に申込みをした者です
0.6 そちらは会議事務局ですか
0.6 会議に申し込みたいのですが
0.6 いま会議に申し込めば参加料はいくらですか
0.6 三七二の八〇一八です
0.6 三七二の八〇一八でございますね
0.6 後日プログラムと予稿集をお送り致します
0.7 こちらは会議事務局です
0.7 また期限は今年いっぱいです
0.7 分からない点がございましたらいつでもお聞き下さい
0.7 英語と日本語です
0.7 こちらは会議事務局です
0.7 参加料は四万円です
0.7 会議の案内書をお送り致しますのでそれをご覧下さい
0.8 はい
0.8 はい
0.8 どうもありがとうございました
0.8 はい
0.8 会議の参加料について教えて頂きたいのですが
0.8 こちらは会議事務局です
0.8 どうもありがとうございました
0.8 はい
0.8 はい
0.8 はい
0.8 はい
0.8 はい
0.8 参加料はどのようにお支払いしたらよいのですか
0.9 もしもし
0.9 はい
0.9 どのような手続きをすればよろしいのでしょうか
0.9 登録用紙は既にお持ちでしょうか
0.9 参加料は現在お一人三万五千円です
0.9 失礼致します
0.9 はい
0.9 失礼ですがお名前とご住所をお願いします
0.9 お名前をお伺いできますでしょうか
0.9 案内書にも書いていますが
1.0 いいえ
1.0 どういたしまして
1.0 会議に論文を発表したいと思っているのですが
1.0 分かりました
1.0 それではよろしくお願いします
1.0 お気の毒ですができません
1.1 来月お申込みになりますと四万円です

1.1 こちらは会議事務局でございます
1.1 既に登録料の八万五千円を振り込まれておられますね
1.2 分かりました
1.2 よろしくお願いします
1.2 分かりました
1.2 分かりました
1.2 いいえ
1.2 登録料を払い戻して頂けますか
1.2 英語への同時通訳を用意しております
1.3 どのようなご用件でしょうか
1.3 分かりました
1.3 分かりました
1.3 分かりました
1.3 代理人が決まりましたらお知らせ致します
1.4 今回は割引を行っておりません
1.4 失礼します
1.4 では失礼します
1.4 会議について詳しいことを教えて下さい
1.5 さようなら
1.5 それは別に問題ありません
1.5 登録用紙を至急送らせて頂きます
1.7 持っていません
1.7 登録用紙で手続きをして下さい
2.0 会議の内容について教えて下さい
2.2 ご住所とお名前をお願いします
2.4 それでは登録用紙をお送り致します
2.7 代理人が参加する場合はあらかじめこちらまでお知らせ下さい
3.6 それでは失礼します
- 参加料には予稿集代と歓迎会費が含まれています
- 参加料は銀行振り込みです
- わたしは日本語が全然分からないのですが
- 発表が日本語で行われる場合英語への同時通訳はあるのですか
- 案内書に記載されている口座番号に振り込んで下さい
- 今回の会議は通訳電話に関連する広範な研究分野を含んでいます
- 言語学や心理学を専攻する方にも参加して頂く予定です
- 会議は八月二十二日から二十五日まで京都国際会議場で開催されます
- 発表を希望されるのでしたら三月二十日までに要約を提出して下さい
- 九月二十七日以後の取り消しに対する払い戻しはできません

改善方法のうち 3) の「です」の意味計算の選言をなくしたことが、かなり有効であった。これに反して、2) の「ている、て下さる」などを一語として登録したのは、あまり有効ではなかった。

アスペクト計算を V-kernel, V-voice, V-aspect, etc. が V になるときに限ったことによって、処理速度が改善された文がある。

新文法 / 旧文法の平均値は、0.8 である。やはり、あまり向上していない。処理速度が低下した文を見ると、次のような特徴がある。

1. 「ている、て下さい」などを含む文。

2. 「それでは」など、ADV の係り先に曖昧性のあるもの。

3. 複文で、主語の係り先が主文の述語と副文の述語の 2 通りあるもの。

4. 尊敬名詞、数量名詞など、旧文法では一語だったものを含む文。

5. 「- ません」は、旧文法では一語だったが、新文法では分解した。

6. 「分りません」など、慣用句とそうでないものとの曖昧性があるもの。

また、解析実験結果を見ればわかるとおり、素性構造に曖昧性のあるものがいくつかある。これらは、バグの可能性がある。

第 3 章

会話 6 〜会話 10 の文法

3.1 概要

日本語対話文 (サンプル会話文 6 - 10) を解析できるように、文法および辞書を拡張した。新たに加えた CFG ルールは 26 ルール、語彙は 55 語である。

今回、新たに処理対象となった言語現象の主なものは次のとおりである。

- 連用形接続 (連用中止)
- 不定語
- 真偽疑問文の補足節
- 取り立て助詞の意味
- 可能表現のガ / ヲ交替
- 複合形容名詞
- 音形のない (ポーズによる) 並列表現
- 「- になる」
- 名詞を下位範疇化する引用の助詞
- 多文節名詞連続複合語 (助詞 / 動詞の省略)

以上のほかに、次の諸点で変更を行った。

- 提題助詞「は」、取立助詞「も」の扱い
- 複文の曖昧性の処理
- 関係節の外の関係と内の関係
- サ変名詞、尊敬転成名詞を修飾する副詞

3.2 新たに処理対象となる言語現象と、その処理方法

3.2.1 連用形接続 (連用中止)

d8-7 こちらで審査を行ない五月二十日までに結果をお送り致します

「連用形接続」を単純に実装しようとすると、次のような規則を加えなければならない。

```
V -> (V  V)
```

ところが、この規則は計算効率を著しく低下させるうえ、次の文のように、「で」が現れるところでは必ず、助動詞「だ」の連用形「で」と格助詞「で」の間の曖昧性が生じる。

d6-20 参加料は当日集合場所でお支払い下さい

d8-21 契約はどのような書式で書けばいいんですか

そこで、連用中止の場合は、次のようにその直後に読点「、」を挿入して、そうでない場合と区別することにする。

d8-7 こちらで審査を行ない、五月二十日までに結果をお送り致します

そして、「、」を PUNC として、次の規則を設ける。

```
V -> (V  PUNC)
```

ただし、読点の挿入は音声認識でポーズを検出することに相当するが、技術的にこれが可能かどうか疑問ではあるが。

なお、連用形接続の意味表現は次のとおりである。

```
[sem [[reln  連用形接続]
      [arg-1 [[動詞節1]]]
      [arg-2 [[動詞節2]]]]]
```

N.B. 連用中止法をまともに文法化するとすれば、動詞を意味分類して、連用中止の動詞がそれ以降に現れるどの動詞と結びつくかを記述することになるだろう。たとえば、Mu では、表3.1のような意味分類を行っている。Mu は科学技術論文の英訳が目的であるから、それにふさわしい分類がなされているのであろう。NADINE では、これとはまた別の分類を行わなければならない。

3.2.2 不定語

d6-13 講演者の何人かは参加することになっています

d7-17 他に何か(が)ございますか

d10-5 そちらでどこか(を)紹介していただけますか

表 3.1: Mu における動詞の意味分類

記述動詞	例: 述べる、報告する、記述する、実証する
準記述動詞	例: 明らかにする
意志動詞	例: 決定する、使用する
意志: 無意思動詞	例: 検出する、変える
無意思動詞	例: 比例する、低下する、放出する

不定語を構成する「か」は、構文的には取立助詞と同じで、格助詞の「が」「を」と交替することがある。そこで、次の規則で処理する。

```
P -> (N  POSTP-contr)
```

ただし、疑問詞は WH- 素性の値を持っていて、単独では文末の終助詞「か」と呼応して INFORMREF のタイプの疑問文となるが、これに不定語の「か」がつくと、INFORMIF のタイプの疑問文となる。そこで、この規則の WH- 素性の計算式を次のようにする。

```
(<!m syn wh> == !empty-dlist)
```

意味表現は次のとおりである。

```
[sem [[reln  か-indefinite]
      [arg-1 [[前接要素]]]]]
```

3.2.3 真偽疑問文の補足節

d9-5 会議場へどうやって行ったらいいか教えて欲しいんですが

d10-28 お部屋が取れるかどうか調べます

疑問文が補足節になる場合、「か」「か / どうか」の品詞をどう考えるかという問題がある。終助詞、取立助詞、引用助詞、準体助詞などの可能性があるが、現在の文法と照らしあわせてみると、それぞれに矛盾を来す。

終助詞は、現在の文法では助動詞の一種とみなされ、文末述語の最後について文をつくることになる。したがって、「か、かどうか」を終助詞とすると、次のような規則が必要になる。

```
V -> (V  V)
```

この規則は、連用中止のところで述べたとおり、きわめて副作用の大きい規則である。取立助詞は、現在の文法では名詞句のみを下位範疇として取る。したがって、「か、どうか」を取立助詞とすると、次の規則が必要になる。

```
ADV -> (V  POSTP-contr)
```

ある種の取立助詞は、動詞句を下位範疇として取ることがあるので(例: わざわざ言っただけの価値はあった)、いずれはこの規則も設けなければならないだろう。

準体助詞は、現在の文法では形式名詞として動詞句を下位範疇にとり、名詞句をつくることになっている。しかし、名詞句が格助詞・取立助詞なしで、単独で動詞句にかかる規則はない。したがって、「か、かどうか」を準体助詞とすると、次の規則が必要になる。

```
V -> (N  V-kernel)
```

会話ではしばしば助詞が省略されることがあるので(例: わたし冷蔵庫のアイスクリーム食べちゃったわ)、いずれはこの規則も設けなければならなくなるかもしれない。

引用助詞は動詞句を下位範疇として取り、思考動詞や表現動詞の補語になる。「か、どうか」を引用助詞とすると、少なくとも現在のところ新しい規則を設ける必要はない。そこで、とりあえず引用助詞として、次の規則で処理することにする。

```
P -> (V  POSTP-quote)
```

ただし、次のテンプレートを作って、「教える」「調べる」に記述する。

```
(GA/SUBJ-KA/OBJ-VT (%RELN  %SUBJ-ROLE  %OBJ-ROLE))
(GA/SUBJ-KADOUKA/OBJ-VT (%RELN %SUBJ-ROLE %OBJ-ROLE))
```

3.2.4 取立助詞の意味

d6-7 八月五日の午後に清水寺、金閣寺、龍安寺などを見学します

d9-10 京都駅からタクシーで会議場までいくらぐらいかかりますか

取り立て助詞それ自体の意味は、SEM-素性にではなく、PRAG-素性に次のように記述する。

```
(<!m prag nounif-terms> == (:dlist [[mode  など]
                                     [object  リステイング表現]]))
(<!m prag euphemism-terms> == (:dlist [[mode ぐらい]
                                       [scope ?sem]]))
```

3.2.5 可能表現のガ / ヲ交替

d9-8 そこから国際会議場へ行くバスが利用できます

可能表現では、格助詞の交替(「が」→「に」「を」→「が」)が起こる。

例) 会員がバスを利用する → 会員にバスが利用できる

従来の方法では、次のように、「できる」の辞書に格助詞をコントロールするルールを記述するのが普通であった。

```
(:or
      ((<!m syn head grfs subj> ==
        ?(subj !(simple-postp-agr-が-with-semf ?subj-sem ?subj-semf)))
       (<!m syn head grfs obj> ==
        ?(obj !(simple-postp-agr-を-with-semf ?obj-sem ?obj-semf)))
       !(aux-sc-sl-3 ?comp ?subj ?obj)
       (<!m synhead grfs subj semf> == [[anim +][hum +]])
       (<!m sem> == [[reln できる-possible]
                     [expr ?subj-sem]
                     [obje ?pred-sem]]))
      ((<!m syn head grfs obj> ==
        ?(obj !(simple-postp-agr-が-with-semf ?obj-sem ?obj-semf)))
       (<!m syn head grfs subj> ==
        ?(subj !(simple-postp-agr-に-with-semf ?subj-sem ?subj-semf)))
       !(aux-sc-sl-3 ?comp ?obj ?subj)
       (<!m syn head grfs subj semf> == [[anim +][hum +]])
       (<!m sem> == [[reln できる-possible]
                     [expr ?subj-sem]
                     [obje ?pred-sem]])))
```

今回、新たに (v-poss +/-) という素性を設け、「が」「に」の語彙記述を次のように変更することにした。

```
(DEFLEX-NAMED が-POSTP が POSTP-oblg
  !(lex-phon-orth "ga" "が")
  (:or
    ((<!m syn head coh syn head v-poss> == -)
     !(simple-postp-with-semf が))
    ((<!m syn head coh syn head v-poss> == +)
     (:or
       (!(simple-postp-with-semf が))      ; バスが利用できる
       (!(simple-postp-with-semf を)))))  ; バスを利用できる
  (<!m prag aspe> == !empty-dlist))

(DEFLEX-NAMED に-POSTP-oblg に POSTP-oblg
  !(lex-phon-orth "ni" "に")
  (:or
    ((<!m syn head coh syn head v-poss> == -)
    !(simple-postp-with-semf に))
    ((<!m syn head coh syn head v-poss> == +)
     (:or
       (!(simple-postp-with-semf が))     ; 会員が利用できる
       (!(simple-postp-with-semf に))))) ; 会員に利用できる
  (<!m prag aspe> == !empty-dlist))
```

この方が、処理速度がかなり速くなる。

3.2.6 複合形容名詞

d6-6 まだ参加可能です

複合名詞は、これまで、名前と住所と年月日に限って生成規則を設けていた。今回、サ変名詞と形容名詞から形容名詞を作る規則を新たに設ける。

```
N-adj -> (N-sahen  N-adj)
```

3.2.7 音形のない (ポーズによる) 並列表現

d6-7 八月五日の午後に清水寺、金閣寺、龍安寺などを見学します

この読点は、構文的には並列助詞の「と」「や」と同じふるまいをする。そこで、次の辞書を設ける。

```
(deflex-named 、 -fatt 、 fatt
  !(lex-phon-orth "period" "、 ")
  (<!m> == [[syn [[head [[pos p]
                         [coh  [[syn [[head [[pos n]]]
                                      [lex +]]]
                                [sem ?sem-2]
                                [semf ?semf]]]]]
                  [subcat (:list [[syn [[head [[pos n]]]
                                        [subcat !empty-list]]]
                                  [sem ?sem-1]
                                  [semf ?semf]])]
                  [slash !empty-dlist]
                  [wh    !empty-dlist]]]
            [sem [[reln pause-coordination]
                  [arg-1 ?sem-1]
                  [arg-2 ?sem-2]]]
            [semf ?semf]]) )
```

N.B. [lex +] という素性は、三つ以上の語が並列する場合の曖昧性を除くためのもので、それ以上解析されない単独の語を意味する素性である。次のように用いられる。

```
N-com -> (ATT  N-com)
   (<!m syn lex> == -)
```

たとえば、次のような文。

d6-19 集合場所は会議場の受付の前になっております。

これは、次のように解析される。

((会議場の受付) の 前)

3.2.8 「− になる」

d6-13 講演者の何人かは参加する予定になっています

d6-19 集合場所は会議場の受付の前になっております

これまで、動詞「なる」は、SUBCAT-素性に「〜が」「〜に」を指定して、それぞれ、OBJE、RESL という SEM-素性を与えていた。これによると、d6-13、d6-19 の文の意味表現は次のようになる。

```
[[RELN  なる-1]
 [OBJE (:LIST 講演者の何人か)]
 [RESL (:LIST 参加する予定)]]

[[RELN  なる-1]
 [OBJE (:LIST 集合場所)]
 [RESL (:LIST 会議場の受付の前)]]
```

しかし、これは d6-13、d6-19 の文の意味表現としては正しくない。「講演者の何人か」が「参加する予定」に変化するわけではないからである。正しくは次のようであるべきであろう。

```
[[RELN  だ-STATEMENT]
 [OBJE (:LIST 講演者の何人かが参加する予定)]]

[[RELN  だ-IDENTICAL]
 [OBJE (:LIST 集合場所)]
 [IDEN (:LIST 会議場の受付の前)]]
```

そこで、今回、「− になる」全体を一語とみなし、「− だ / です」と同じ扱いをすることにする。辞書登録は次のとおり。

```
(deflex-named になる-identical/statement にな auxvstem-copl-com
      :
  (:or
    ((<!m syn subcat first syn subord> == -)
     (<!m !synhead grfs> == [[comp ?(comp !(simple-noun-agr ?comp-sem))]
                             [subj ?(subj !(simple-postp-agr-が ?subj-sem))]])
     !(aux-sc-sl-2 ?comp ?subj)
     (<!m sem> == [[reln だ-identical]
                   [obje ?subj-sem]
```

```
                [iden ?comp-sem]]))
  ((<!m syn subcat first syn subord> == +)
   (<!m syn subcat> == (:list [[syn [[head [[pos n]]]
                                     [subcat !empty-list]]]
                              [sem ?noun-sem]]))
   (<!m !synsl > == !empty-dlist )
   (<!m sem> == [[reln だ-statement]
                [obje ?noun-sem]]))
        :
```

ただし、こうすると、「変化」を意味する「なる」が正しく解析できなくなる。だから、本来的には "or" を用いて両方の意味が解析できるように記述すべきであるが、今のところ、例文に「〜だ」の意味しか出現していないので、こちらだけを記述しておく。

3.2.9 名詞を下位範疇化する引用の助詞

d6-17 ケンブラウンと申します

この「ケンブラウンと」を、「申す」の必須格補語とみなす。したがって、「と」を必須格助詞として登録する。

なお、この場合の「- と申す」は「- だ」と同じ意味である。そこで、「申す」の意味を次にように記述する。

```
[sem [[reln  だ-identical]
      [obje ?subj-sem]
      [iden ?obj-sem]]]
```

3.2.10 多文節名詞連続複合語 (助詞 / 動詞の省略)

d7-13 送り先は大阪市東区城見二の一の六十一渡辺明です

住所と名前のそれぞれの生成規則は存在するが、例文のように2つが連続する場合の規則はない。そこで、次の規則を設ける。

```
N-prop -> (N-address  N-name)
    :
   (<!m sem> == [[reln  pause-coordination]
                 [arg1 <!a-dtr sem>]
                 [arg2 <!h-dtr sem>]])
```

3.3 その他の変更

3.3.1 提題助詞「は」と取り立て助詞「も」の扱い

これまで、「は」「も」のついた主題化文は、[orig-form (: が を に)] の素性を用いて、Slash-Consuming によって処理していた。ところが、これに要する単一化は負荷が大きく、処理効率の悪化を招いている。そこで、これを次のように変更する。

「は」「も」を、名詞に後接する場合と、助詞、副詞に後接する場合に分け、それぞれに form- 素性を次のように記述する。

```
(:or
   ((<!m syn subcat> == (:list [[syn [[head [[pos n]] ]
                                        [subcat !empty-list]
                                   [wh     !empty-dlist]]]
                                [sem ?topic]
                                [semf ?semf]]))
    (<!m syn head form> == (:or が を に)))
   ((<!m syn subcat> == (:list [[syn [[head [[pos (:or p adv)]]]]]
                                [sem ?topic]
                                [semf ?semf]]))
    (<!m syn head form> == は))
   )
```

このようにすると、名詞に連接する場合は、通常の必須格補語と同じように Subcat-Consuming で扱うことができる。これによって、処理速度がおよそ 20 〜 30 パーセントほど向上する。

3.3.2 複文の曖昧性の処理

これまで、複文において、従属節内部の補語が主節の述語に係るという曖昧性を排除できなかった。

今回、[subord +/-] という素性を新たに設けて、この問題の解決を計った。すなわち、次の一連の規則において、親ノードに [subord +] を付加する。

```
N -> (V  FN)            ; 連体修飾節 + 形式名詞
N-com -> (V  N-com)     ; 連体修飾節 + 普通名詞
ADV -> (V  FADV)
```

そして、後置詞句が述語に係る規則に次の制限を加える。

```
V-kernel -> (P  V-kernel)
      :
  (:or
    ((<!c-dtr syn subord> == +)
     (<!m syn complex> == +))
    ((<!c-dtr syn subord> == -)
```

```
    (<!m syn complex> == -)))
  (<!h-dtr syn complex> == -)
```

3.3.3 関係節の解析: 内の関係と外の関係

d6-13 講演者の何人かは参加する予定になっています。

d7-3 会議で扱う話題に関して質問したいんですが。

d8-3 ちょっとお聞きしたいことがあるんですが。

d6-13 は外の関係、d7-3、d8-3 は内の関係である。外の関係の特徴は、被修飾名詞が内容節をとる名詞であることである。そこで、名詞を、内容節をとりうるか否かで分類し、[content +/-] を与える。そして次のような素性式によって解析する。

```
N-com -> (V N-com)
     :
 (:or
   ((<!h-dtr syn head content> == -)     ; 内の関係
    !standard-subcat-fp-ah
    !v-n-slash-fp-ah
    !v-n-sem-fp-ah
    !v-n-semf-fp-ah)
   ((<!h-dtr syn head content> == +)    ; 外の関係
        !standard-subcat-fp-ch
    !standard-slash-fp-ch
    (<!m sem parm> == <!h-dtr sem>)
    (<!m sem restr reln> == 外の関係の連体修飾)
    (<!m sem restr arg-1> == <!h-dtr sem>)
    (<!m sem restr arg-2> == <!a-dtr sem>)
    !standard-semf-fp-ch)
    )
```

3.3.4 サ変名詞、尊敬転成名詞を修飾する副詞

副詞句は V-kernel 以下の述語を修飾する。ところが、サ変名詞、尊敬転成名詞は「する、いたす、できる、etc.」の助動詞、補助動詞が連接してはじめて述語になる。

そこで、次のような不都合が生じることになる。

d9-9 北大路駅ではタクシーも利用できます

```
[[SEM [[RELN  できる -POSSIBLE]
 [ASPT  UNRL]
 [EXPR ?X06[]]
 [OBJE [[RELN  利用する -1]
        [AGEN ?X06]
```

```
        [OBJE [[PARM ?X05[]]
               [RESTR [[RELN  タクシー-1]
                       [ENTITY ?X05]]]]]]]
[SLOC [[PARM ?X07[]]
        [RESTR [[RELN  NAMED]
                [ENTITY ?X07]
                [IDEN  北大路駅-1]]]]]]]
```

すなわち、「北大路駅では」は意味的には「利用」を修飾するにもかかわらず、実際にはそうは解析されないのである。

そこで、次の規則を設けることにする。

```
N-sahen -> (ADV N-sahen)
```

当然のことながら、d9-9 のような文には 2 つの曖昧性が生じることになる。

同様にして、種々の副詞 (句) の修飾規則に次のものを加える。

```
N-verb-honor -> (ADV N-verb-honor)
N-sahen -> (ADV-mann N-sahen)
N-verb-honor -> (ADV-mann N-verb-honor)
N-sahen -> (ADV-degr N-sahen)
N-verb-honor -> (ADV-degr N-verb-honor)
```

3.4 解析実験結果

3.4.1 会話 6 〜会話 10 の解析実験結果

dialogue-6

	日本語	文字長	単語長	CPU 時間
1	はい	2	1	1
2	こちらは会議事務局ですが	12	5	17
3	会議の間に市内観光があるそうですが	17	9	87
4	まだ参加できますか	9	5	28
5	はい	2	1	1
6	まだ参加可能です	8	4	10
7	八月五日の午後に清水寺、金閣寺、龍安寺などを見学します	27	15	239
8	参加なさいますか	8	4	13
9	参加料はいくらですか	10	5	15
10	八千円です	5	2	8
11	参加料には夕食代も含まれています	16	10	140
12	講演者も参加されるのですか	13	8	83
13	講演者の何人かは参加する予定になっています	21	13	194
14	そうですか	5	3	6
15	それでは参加したいと思います	14	9	92
16	ではお名前と人数をお願い致します	16	10	52
17	ケンブラウンと申します	11	5	34
18	家内と参加します	8	5	26
19	集合場所は会議場の受付の前になっております	21	12	84
20	参加料は当日集合場所でお支払い下さい	18	8	40
21	分かりました	6	3	10
22	ありがとうございました	11	4	2
23	ではお待ちしております	11	7	14

dialogue-7

	日本語	文字長	単語長	CPU 時間
1	はい	2	1	1
2	こちらは会議事務局です	10	4	16
3	会議で扱う話題に関して質問したいんですが	20	11	286
4	はい	2	1	1
5	何でしょうか			
6	機械翻訳という話題が案内書に載っていますが	21	12	79
7	具体的にこれはどういう内容のものなんですか	21	13	49
8	申し訳ありませんがこちらでは専門的な質問にお答えできません	29	18	
9	第二版の案内書に会議で発表される論文の題目が載っております	29	19	
10	そちらを見て頂けないでしょうか	16	9	98
11	いいですよ	5	3	1
12	それでは早急にその案内書を送って下さい	21	11	58
13	送り先は大阪市東区城見二の一の六十一渡辺明です	23	12	96
14	大阪市東区城見二の一の六十一渡辺明様ですね	22	12	61
15	はい	2	1	1
16	早速送らせて頂きます	10	6	63
17	他に何かございますか	10	7	17
18	いいえ	3	1	1
19	ありません	5	3	3
20	ありがとうございました	11	4	2
21	失礼します	5	3	9

dialogue-8

	日本語	文字長	単語長	CPU 時間
1	はい	2	1	1
2	会議事務局です	7	2	9
3	ちょっとお聞きしたいことがあるんですが	19	11	125
4	私は今度の会議に発表したいと思っているんですが	23	16	452
5	どのような手続きをすればよろしいでしょうか	21	10	89
6	先ず二百字の要約を三月二十日までにこちらまでお送り下さい	28	15	96
7	こちらで審査を行ない、五月二十日までに結果をお送り致します	29	16	190
8	投稿が受理された場合原稿用紙を同封いたします	22	12	313
9	六月三十日までに原稿の送付をお願いします	20	12	72
10	分かりました	6	3	10
11	契約はどのような書式で書けばいいんですか	20	12	156
12	所定の申込み用紙がありますのでそれに記入して下さい	25	13	102
13	それでは申込み用紙を送りますので送り先をお願いします	26	14	195
14	分かりました	6	3	10
15	人工知能研究所のジョージオハラです	17	5	18
16	住所は東京都豊島区東池袋三丁目二番五号です	21	12	46
17	人工知能研究所のジョージオハラ様ですね	19	7	22
18	ご住所は東京都豊島区東池袋三丁目二番五号でよろしいですね	28	16	102
19	はい	2	1	1
20	そうです	4	2	4
21	それでは申込み用紙の送付をよろしくお願いします	23	12	66
22	はい	2	1	1
23	分かりました	6	3	10
24	では早速お送り致します	11	6	57
25	失礼致します	6	3	7

dialogue-9

	日本語	文字長	単語長	CPU 時間
1	そちらは会議事務局ですか	12	5	21
2	はい	2	1	1
3	会議事務局です	7	2	9
4	何のご用件でしょうか	10	7	42
5	会議場へどうやって行ったらいいか教えて欲しいんですが	26	15	
6	いま京都駅にいるんです	11	6	21
7	地下鉄で北大路駅まで行って下さい	16	7	42
8	そこから国際会議場へ行くバスが利用できます	21	10	
9	北大路駅ではタクシーも利用できます	17	8	74
10	京都駅からタクシーで会議場まで行くにはいくらぐらいかかりますか	31	14	248
11	京都駅からですとおよそ六千円かかります	19	9	34
12	では北大路駅からですといくらぐらいかかりますか	23	10	60
13	北大路駅からですとおよそ九百円です	17	8	40
14	分かりました	6	3	10
15	どうもありがとうございました	14	5	6
16	いいえ	3	1	1
17	どういたしまして	8	4	3

dialogue-10

	日本語	文字長	単語長	CPU 時間
1	もしもし	4	1	1
2	はい	2	1	1
3	会議事務局でございます	11	4	13
4	会議の宿泊施設についてお尋ねしたいのですが	21	11	129
5	そちらでどこか紹介して頂けますか	16	10	212

6	はい	2	1	1
7	私どもでご紹介できるホテルは京都ホテルと京都プリンスホテルです	31	11	
8	一人部屋の値段は一晩七千円から一万円です	20	12	81
9	二人部屋の値段は九千五百円から六万円です	20	11	52
10	そうですか	5	3	6
11	どちらのホテルが会議場に近いですか	17	9	75
12	京都プリンスホテルが会議場には近いんですが	21	9	77
13	それでは京都プリンスホテルを予約したいのですが	23	11	104
14	ホテルの手配もして頂けるのですか	16	10	154
15	はい	2	1	1
16	京都ホテルと京都プリンスホテルは予約できます	22	7	49
17	そうですか	5	3	6
18	では京都プリンスホテルの七千円の一人部屋をお願いします	27	13	68
19	はい	2	1	1
20	京都プリンスホテルの七千円の一人部屋ですね	21	9	34
21	はい	2	1	1
22	そうです	4	2	4
23	いつからお泊まりになりますか	14	8	19
24	八月四日の夜からです	10	6	16
25	八日の朝までお願いします	12	8	51
26	分かりました	6	3	10
27	少々お待ちください	9	5	12
28	お部屋が取れるかどうか調べます	16	8	81
29	お部屋をお取りできます	11	7	28
30	ではお名前とご住所をお願いします	16	11	77
31	中村一雄です	6	3	7
32	住所は東京都港区新橋一丁目一番三号です	19	12	46
33	電話番号もお願いします	11	6	24
34	電話番号は三三一の二五二一です	15	6	24
35	分かりました	6	3	10
36	京都プリンスホテルに八月四日から八日まで一人部屋をお取りしました	32	15	120
37	どうもありがとうございました	14	5	6
38	失礼します	5	3	9

3.4.2 旧文法と新文法の比較

新たに文法・辞書を加えたことによって、処理速度がどのように変化したかを調べるために、会話１～会話５の文を解析・比較してみた。左端の数値を見ると、大体 1～1.3 倍の範囲におさまっているものが大部分である。したがって、文法・辞書が増えた分を考慮すると、文法の処理性能はそれほど落ちていないと考えられる。(旧文法の左側の数値は、今回の変更を行う前の文法による解析結果、旧文法の右側の数値は、前回の変更を行う前の文法による解析結果である。)

dialogue-1

新 / 旧	日本語	旧文法	新文法
0.7	もしもし	0.7/0.9	0.5
1	そちらは会議事務局ですか	18/23	19
0.7	はい	0.6/0.9	0.44
1.3	そうです	2/3	2.6
1.2	どのようなご用件でしょうか	31/40	37
1.3	会議に申し込みたいのですが	30/39	38
1	どのような手続きをすればよろしいのでしょうか	81/105	77
0.7	登録用紙で手続きをして下さい	31/40	21
1	登録用紙は既にお持ちでしょうか	29/38	27
1	いいえ	0.8/1	0.8
1.4	まだです	2/3	2.7
1	分かりました	9/12	10
0.6	それでは登録用紙をお送り致します	94/122	51
1.3	ご住所とお名前をお願いします	55/72	72
1.1	住所は大阪市北区茶屋町二十三です	29/38	32
1	名前は鈴木真弓です	15/20	15
1	分かりました	10/13	10
1	登録用紙を至急送らせて頂きます	98/127	94
1.9	よろしくお願いします	15/20	26
0.4	それでは失礼します	39/51	14

dialogue-2

新 / 旧	日本語	旧文法	新文法
0.8	はい	0.6/0.8	0.45
1.1	こちらは会議事務局です	14/18	16
1.5	会議の参加料について教えて頂きたいのですが	152/198	227
0.8	いま会議に申し込めば参加料はいくらですか	157/204	120
1.3	はい	0.6/0.8	0.8
1.1	参加料は現在お一人三万五千円です	30/39	34
0.8	来月お申込みになりますと四万円です	107/139	90
-	参加料には予稿集代と歓迎会費が含まれています		230
0.8	わたしは情報処理学会の会員なのですが	40/52	31
1	参加料の割引はないのですか	40/52	38
0.9	今回は割引を行っておりません	44/57	39
2	そうですか	3/4	6
1	参加料はどのようにお支払いしたらよいのですか	159/207	163
1	参加料は銀行振り込みです	14/18	15
-	案内書に記載されている口座番号に振り込んで下さい		
1	また期限は今年いっぱいです	22/29	24
1	分かりました	10/13	10
2	どうもありがとうございました	3/4	6

1.5	どういたしまして	2/3	3
1.1	分からない点がございましたらいつでもお聞き下さい	202/263	230
1.2	失礼致します	5/7	6

dialogue-3

新 / 旧	日本語	旧文法	新文法
1.3	はい	0.6/0.8	0.8
1	こちらは会議事務局です	15/20	16
1.1	会議に論文を発表したいと思っているのですが	305/397	328
1	会議の内容について教えて下さい	64/83	61
-	今回の会議は通訳電話に関連する広範な研究分野を含んでいます		
-	言語学や心理学を専攻する方にも参加して頂く予定です		
1	分かりました	10/13	10
1.2	ところで会議での公式言語は何ですか	33/43	38
1.3	英語と日本語です	13/17	17
1	わたしは日本語が全然分からないのですが	75/98	73
0.9	発表が日本語で行われる場合英語への同時通訳はあるのですか	544/707	468
1.4	はい	0.7/0.9	1
0.9	英語への同時通訳を用意しております	56/73	48
1	分かりました	10/13	10
2	どうもありがとうございました	3/4	6
1	さようなら	2/3	2

dialogue-4

新 / 旧	日本語	旧文法	新文法
1	こちらは会議事務局です	15/20	16
1.2	会議について詳しいことを教えて下さい	91/118	109
1	会議の案内書はお持ちですか	27/35	30
1	いいえ	0.9/1.2	1
0.9	持っていません	15/20	13
2	そうですか	3/4	6
-	会議は八月二十二日から二十五日まで京都国際会議場で開催されます		656
1	参加料は四万円です	15/20	17
-	発表を希望されるのでしたら三月二十日までに要約を提出して下さい		2040
1	会議の案内書をお送り致しますのでそれをご覧下さい	89/116	96
1.3	失礼ですがお名前とご住所をお願いします	58/75	77
1	アダムスミスです	6/9	7
1	住所は大阪市東区玉造二丁目二十七の七です	67/87	67
1	分かりました	10/13	10
0.9	電話番号もお聞きしたいのですが	77/100	71
1.3	はい	0.6/0.0.8	0.8
1.2	三七二の八〇一八です	13/17	15
1.5	三七二の八〇一八でございますね	14/18	21
1.3	はい	0.6/0.8	0.8
2	そうです	2/3	4
1.2	それではよろしくお願いします	22/29	27
1	失礼します	8/10	9

dialogue-5

新 / 旧	日本語	旧文法	新文法
1.3	はい	0.6/0.8	0.8
1.3	こちらは会議事務局でございます	17/22	22

1	ちょっとお願いがあるのですが	29/38	30
2	私は会議に申込みをした者です	121/153	254
1	参加を取り消したいのですが	53/69	58
1	お名前をお伺いできますでしょうか	67/87	67
1.3	はい	0.6/0.8	0.8
1.2	ベル研のジムワイベルです	15/20	18
0.9	既に登録料の八万五千円を振り込まれておられますね	210/273	193
1.5	はい	0.6/0.8	0.9
2	そうです	2/3	4
0.9	登録料を払い戻して頂けますか	68/88	59
1	お気の毒ですができません	8/10	9
1	案内書にも書いていますが	95/124	92
-	九月二十七日以後の取り消しに対する払い戻しはできません		422
1	後日プログラムと予稿集をお送り致します	68/88	73
1.4	では誰かが私の代わりに参加することはできますか	196/255	284
1.2	それは別に問題ありません	13/17	15
1.5	代理人が参加する場合はあらかじめこちらまでお知らせ下さい	147/191	210
1	分かりました	10/13	10
0.7	代理人が決まりましたらお知らせ致します	144/187	103
1.6	では失礼します	8/10	13

第 4 章

A ランクの文法

4.1 はじめに

A ランクの実験文は、会話 1 〜会話 10 の文法をベースにして語彙を拡大するという 方針で作成した。しかし、語彙を登録するのに伴って、文法そのものも拡大・修正しな ければならない場合もあった。以下では、それらについて説明する。

4.2 A ランクの実験文のために追加した規則

4.2.1 文接続の規則

A-458 学校へ行って、そして図書館へ寄った

「学校へ行って」は本来的には ADV となる。ところが、ADV は V-kernel にしか係らないように制限されている。一方、「そして」は ADV-conj であり、V に係るようになっている。そこで、この文は交差によって解析できない。そこで、「学校へ行って、」を次の規則で解析することにして、

```
V -> (ADV  PUNC)
```

文全体は、連用形接続を扱う次の規則で解析するようにした。

```
V -> (V  V)
```

4.2.2 副詞句の規則

4.2.2.1 副詞的に使われる相対名詞

A-360 お忙しい中申し訳ありません

「お忙しい中」は、副詞句であり、本来的には ADV となるべきである。しかし、この文の「申し訳ありません」は IDIOM として V と登録してあるので、これにかけることはできない。そこで、「中」を N-adv とし、「お忙しい中」全体に繰り上げることにして、N-adv を直接 V にかけることにした。

```
N-adv -> (V-kernel N-adv)                    ; お忙しい+中
V -> (N-adv V)                               ; お忙しい中申し訳ありません
```

現在の文法では、副詞句を V-kerenl, V-voice, etc. にかかるものと、V にかかるものに分けているが、これが交差や曖昧性の問題を生む原因となっている。文末の述語要素連続の問題とあわせて、再検討しなければならない。

4.2.2.2 形容詞の連用中止法

A-356 10月になればかなり涼しくなります

この場合の形容詞は副詞的用法だと考えて次の規則を新たに設けた。

```
ADV => (V-kernel)
```

ただし、娘ノードに [CTYPE I] の制限を加えた。

4.2.3 助動詞の規則

4.2.3.1 「〜ではありません」「〜ではございません」

A-387 複雑な手続きは必要ではありません

A-389 そういう事情なのではございません

まともに形態素解析すれば、助動詞「だ」、係助詞「は」、補助動詞「ある」「ござる」、助動詞「ます」「ん」となる。しかし、そのまま解析すると処理効率がおそろしく悪い。そこで、「はありません」「はございません」を AUXV-negt として、次の規則で処理することにした。

```
V-negt -> (V-kernel AUXV-negt)
V-negt -> (V-mood2 AUXV-negt)
```

従来の文法では、AUXV-negt は V-mood2 には連接しないことになっていたので、後の規則を新たに設けた。

4.2.3.2 形容詞語幹 + 助動詞

A-206 あの古そうな建物は何ですか

形容詞語幹の「古」に助動詞「そうだ」の連体形が連接している。これを解析するために、次の規則を設けた。

```
V-mood2 -> (VSTEM AUXV-evid)
```

ただし、娘ノード1に [CTYPE I] の制限を加える。

4.2.4 助詞類の規則

4.2.4.1 総主文

A-130 京都は朝が早い

A-28 用紙って特別な書式があるのですか

従来の文法では、必須格補語が主題となったもののみを次の規則で扱っていた。

```
P -> (N POSTP-topic)
V-kernel -> (P V-kernel)
```

総主文の主題は、必須格補語＋動詞が作る命題の外にあって、文全体にかかると考えるべきなので、上の規則では扱えない。そこで、総主文の主題となる「は」のためには、POSTP-contr を与えて別の語彙とし、TOPIC というカテゴリーを導入して、次の規則を設けることにした。

```
TOPIC -> (N POSTP-contr)
V -> (TOPIC V)
```

A-130 の文の意味表現は次のようになる。

```
[[RELN 早い-1]
 [TOPIC [[PARM ?X1]
         [RESTR [[RELN 京都]
                 [ENTITY ?X1]]]]]
 [OBJE [[PARM ?X2]
        [RESTR [[RELN 朝-1]
                [ENTITY ?X2]]]]]]
```

ただし、必須格補語が主題となる文では、従来の解釈に加えて、総主文と解釈される曖昧性が生じることになる。

4.2.4.2 動詞に後続する助詞

A-453 発表できるほど研究が進んでいません

A-349 開催にいたるまで大変でした

動詞に取立助詞「ほど」や格助詞「まで」が後続して、副詞句を作っている。これを解析するために、次の規則を設けた。

```
ADV -> (V POSTP-contr)
ADV -> (V POSTP-optn)
```

ただし、動詞に後続しうる取立助詞・格助詞が限られているので、娘ノード 2 に FORM の制限を加える。

4.2.4.3 形容名詞に直接かかる必須格補語

A-54 私は彼と同じくらいの年です

「彼と」は「同じ」の必須格だが、現在の文法では、必須格補語は、形容名詞が助動詞「だ・です」と結合して V-kernel となった後にかかるようになっている。そのために、上の文では、直接「同じ」にかかることはできなくて、「同じくらいの年です」にかかってしまう。そこで、「彼と」を形容名詞「同じ」に直接にかける次の規則を設けた。

```
N-adj -> (P N-adj)
```

「同じくらい」は、たとえば「今年の会議は去年と同じくらい盛況でした」のように、副詞句として用いられる。そこで、次の規則を設けた。

```
ADV -> (N-adj POSTP-contr)
```

さらに、「同じくらいの年」を解析するために、副詞に形式連体詞が後続して連体詞を作る次の規則を設けた。

```
ATT -> (ADV FATT)
```

4.2.4.4 数量名詞 + 取立助詞

A-404 ツアーの参加者は 200 人もいます

数量名詞は、助詞を伴わずに動詞を直接修飾することができた。

```
V-kernel -> (N-quant V-kernel)
```

上の文のように、数量名詞に取立助詞「も」が後続しても、動詞に対する関係は変わらない。そこで、次のように N-quant をそのまま繰り上げる。

```
N-quant -> (N-quant POSTP-contr)
```

4.2.5 連体詞の規則

A-211 かなりの量だ

形式連体詞「の」が副詞につく例は、今回初めて出現した。そこで、次の規則を設けた。今回は 程度副詞だけだが、用例次第で他の副詞のための規則も設けることにする。

```
ATT -> (ADV-degr FATT)
```

4.3 残された課題

Aランクの実験文に含まれながら、今回は解決できなかった問題として、次のものがある。

並列文

A-230 都ホテルと京都ホテルとどちらがよいですか

A-455 彼は優れた学者でありまた組織家でもある

従属文の述語が省略されている文

A-340 会議のテーマが何か知りません

A-345 次の会議がいつかご存じですか

交差現象を含む文

A-152 ご用件があればどうぞ事務局に申しつけてください

A-174 会議にぜひ出席してください

A-175 参加者はおそらく二千人くらいでしょう

A-317 あの星はまるでダイヤモンドのようだ

A-319 上野駅はまるで迷路みたいだ

また、任意格やアスペクトにまつわる曖昧性も、今回は完全には処理できていない。これらについては、B・Cランクの語彙を登録する際に解決したい。

4.4 資料1: A ランク解析実験文の解析所要時間

A ランク解析実験文の解析にかかった時間を以下に示した。時間の欄が – となっているものは、解析不可能だった文である。

		(単位: 秒)
A-1	バスが来た	35.15
A-2	事務局に問合わせる	70.21
A-3	登録用紙を送ります	52.58
A-4	大勢の研究者がアメリカから参加します	168.21
A-5	会議は十時から始まります	47.24
A-6	では登録方法から説明しましょう	130.50
A-7	このセッションは四時までです	91.34
A-8	プロシーディングを十五日までに印刷します	183.77
A-9	京都駅から北大路までバスが出ています	381.43
A-10	参加資格は六十歳までです	53.94
A-11	日本へ行くのはこれが初めてです	–
A-12	プリンスホテルの方がロイヤルホテルより会場に近いです	335.89
A-13	シンポジウムは午前十時より開かれます	1638.19
A-14	北大路駅よりバスがあります	34.00
A-15	参加者は受付の前に集まってください	–
A-16	会議は十時に始まります	46.38
A-17	新幹線は三時に大阪に到着します	242.67
A-18	会期中に観光ツアーがあります	69.61
A-19	先生と相談します	102.57
A-20	同僚と参加したいと思います	355.20
A-21	原稿を京都駅でなくしたらしい	307.53
A-22	郵送で申し込みました	73.85
A-23	このセッションは五時で終わります	167.11
A-24	長尾先生は組織委員長として出席されます	–
A-25	会議の登録について知りたいのですが	322.37
A-26	私は名古屋の交通の便について心配しているのです	1248.63
A-27	加藤先生は機械翻訳について講演します	125.48
A-28	用紙って特別な書式があるのですか	–
A-29	国際会議において技術協力について話合われた	–
A-30	委員会には十一人のメンバーがいる	128.16
A-31	会場にはOHPの設備がある	81.62
A-32	委員会には女性のメンバーがいない	148.32
A-33	会場にはOHPの設備がない	116.41
A-34	登録用紙を送ってあげる	192.60
A-35	登録用紙を送ってくれるように頼みました	822.31
A-36	登録用紙を送ってもらいたいのですが	509.10
A-37	申し訳ありませんが登録用紙を送ってください	131.72
A-38	それは有難いお話です	120.18
A-39	よろしくお願いいたします	116.81
A-40	ホテルの予約が取れなくて困っています	878.12
A-41	立派な会場に驚きました	226.47
A-42	それはおもしろい企画ですね	104.61
A-43	ありがとうございます	3.16
A-44	さようなら	17.08
A-45	京都へ一度行ってみたい	528.98
A-46	登録用紙が欲しい	26.51
A-47	どういたしまして	85.19
A-48	プリンスホテルの方がロイヤルホテルより安いです	180.76
A-49	ロイヤルホテルよりプリンスホテルの方が安いです	173.28
A-50	タクシーよりも電車の方が速いです	506.75
A-51	タクシーよりは電車の方が速いです	314.81
A-52	スライドを上映する方がむしろいいと思います	–
A-53	会議についてもっと詳しいことが知りたい	–
A-54	私は彼と同じくらいの年です	285.27
A-55	医者に行った方がよい	–
A-56	これ以上申し上げることはありません	337.20
A-57	一番速い交通機関は新幹線です	264.20

A-58	ロイヤルホテルに比べてプリンスホテルの方が安いです	–
A-59	ロイヤルホテルと比べてプリンスホテルの方が安いです	2256.38
A-60	ロイヤルホテルに比べればプリンスホテルは安いです	1043.53
A-61	ロイヤルホテルと比べればプリンスホテルの方が安いです	833.40
A-62	今年から参加料は十万円になります	94.50
A-63	雨から雪になった	94.32
A-64	彼は学者になった	113.82
A-65	シンポジウムは非常な混乱となった	568.73
A-66	まだすべての原稿が集まっていません	181.78
A-67	電話番号を言いましょうか	39.76
A-68	この論文はデータベースについて言っています	551.79
A-69	まだ登録を受付けています	122.90
A-70	登録用紙を送ってください	121.23
A-71	午前十時から開会式を行ないます	75.22
A-72	会議の申込みの方法を教えてください	313.05
A-73	参加人員は五百人以上になると思います	–
A-74	私はそう思いません	42.89
A-75	これでシンポジウムを終わります	47.11
A-76	大きなスクリーンが壁に掛かっています	485.39
A-77	ご住所を書いてください	103.94
A-78	論文のテーマを考える	35.81
A-79	会議に関する問合わせは事務局にお願いします	1227.37
A-80	以前田中先生の講演を聞きました	131.52
A-81	コンピュータに関する国際会議があると聞いているんですが	1440.01
A-82	午前九時までに受付に来てください	1080.42
A-83	会場までバスが来ています	160.15
A-84	論文の内容が人に知れてしまった	–
A-85	論文の投稿は学会の規約に従う	125.65
A-86	気温が三十度を越した	66.01
A-87	部屋が空いているかどうかホテルに問合わせます	483.62
A-88	講演者の声が後ろまで届きませんでした	250.99
A-89	登録用紙がまだ届いていないのですが	239.75
A-90	会場の近くのホテルに泊っています	861.27
A-91	すみませんがその荷物を取ってください	224.42
A-92	分からないことがありましたら事務局にお尋ねください	1363.76
A-93	発表にはスライドを使いたいのですが	242.73
A-94	いま北大路駅に着きました	104.38
A-95	論文集が出来ました	55.19
A-96	京都駅を出るとすぐタクシー乗り場です	215.16
A-97	この会議のことが新聞に出ました	186.40
A-98	ホテルの予約をお願いします	422.72
A-99	今月一杯に登録料をお支払ください	85.37
A-100	観光ツアーがあることを知りませんでした	672.15
A-101	私はノーマン教授を知っています	216.40
A-102	山田先生にインタビューをしたい	359.15
A-103	会議場行きのバスに乗る	72.99
A-104	オープニングセッションは朝九時に始まります	85.88
A-105	彼は四カ国語を話します	373.33
A-106	会場に入る	33.63
A-107	事務局のスタッフと演題について話した	897.18
A-108	参加料を払う	14.48
A-109	窓を開いてください	40.19
A-110	登録料は昼食代を含む	30.46
A-111	登録料は銀行口座に振込んでください	119.76
A-112	いま北大路駅でバスを待っています	184.51
A-113	会議で発表する論文をまとめています	–
A-114	展示場を回りました	36.80
A-115	京都のお寺を見たいのですが	412.83
A-116	会議の内容についてはプログラムを見てください	162.01
A-117	医者に見てもらった方がよい	897.48
A-118	会議への参加を申込みたいのですが	116.13
A-119	マイクを手に持っている	123.13
A-120	私の代わりに助手をやります	58.49
A-121	京都の方に行きたい	394.86

A-122	先生には事務局から連絡が行っていると思います	3242.22
A-123	シャンク先生が参加するかどうか分りません	273.40
A-124	受付で名札を渡します	78.06
A-125	京都見物は秋がよろしいと思います	–
A-126	先生の講演がとてもよかった	65.70
A-127	京都には寺が多い	51.37
A-128	暇がない	23.45
A-129	彼は京都の地理に詳しい	97.26
A-130	京都は朝が早い	–
A-131	会場から大阪城が近い	43.57
A-132	登録料が高い	16.45
A-133	京都タワーは高い	24.08
A-134	登録料が安い	13.98
A-135	日程が短い	13.80
A-136	新幹線は速い	17.40
A-137	会議の紹介が細かい	32.46
A-138	日程が長い	14.47
A-139	あるいは遅れるかもしれません	–
A-140	一応理解しました	283.95
A-141	そう思います	20.49
A-142	海外からの参加者もかなりいます	168.69
A-143	締切は今日です	26.74
A-144	案内書にはこう書いてあるのですが	318.01
A-145	それでは早速飛行機の予約をしましょう	479.65
A-146	実際にご自分でご覧ください	75.50
A-147	会議に出席できるかどうか至急お知らせください	440.57
A-148	京都の市内観光では寺をたくさん回ります	117.98
A-149	八万円はちょっと高い	37.48
A-150	この会議は特に言語処理の研究のために開かれます	2472.07
A-151	どう思いますか	26.73
A-152	ご用件があればどうぞ事務局に申付けてください	–
A-153	どうも病気になったらしい	479.38
A-154	どうもありがとうございます	9.09
A-155	英語は少々話します	388.87
A-156	バスはすぐ来ます	50.01
A-157	このセッションは少し難しい	43.13
A-158	まあいいでしょう	31.00
A-159	まず登録用紙を送ってください	160.03
A-160	登録料をまだお支払いいただいておりません	160.25
A-161	まったく知りません	34.63
A-162	今年の登録期限はもう過ぎました	161.14
A-163	もし登録が済んでいれば	80.81
A-164	電話番号をもう一度お願いします	245.86
A-165	もちろん出席します	62.04
A-166	やはり高いと思います	142.15
A-167	どうぞよろしくお願いします	75.14
A-168	では案内書を再度送ります	61.48
A-169	後ほど電話をいたします	76.38
A-170	予稿集は会場で直接渡します	149.03
A-171	なるほどそうですか	425.52
A-172	すぐ手配します	64.00
A-173	予稿集はすぐになくなります	91.25
A-174	会議にぜひ出席してください	–
A-175	参加者はおそらく二千人くらいでしょう	–
A-176	私はまったく知りませんでした	236.82
A-177	会議の内容について詳しく教えてください	732.47
A-178	事務局長とよく相談します	133.68
A-179	どうぞよろしくお願いいたします	104.85
A-180	多くの参加者がいた	106.76
A-181	参加者が二百名集まった	99.57
A-182	案内書を二百通送った	122.34
A-183	お気の毒ですが締切が過ぎています	89.06
A-184	運悪く雨が降った	31.55
A-185	その用紙を送ってください	124.01

A-186	どの用紙を送るのですか	147.26
A-187	この用紙を送るのですか	147.70
A-188	実際の会議はいつですか	92.56
A-189	当の事務局では予定を変更しました	1132.56
A-190	当初の計画を変更しました	492.23
A-191	通常の会議ですと参加者は二百人くらいです	159.89
A-192	そんな事情は知りませんでした	183.59
A-193	こんな会議は初めてだ	30.99
A-194	いろんな分野の研究者が集まります	86.28
A-195	これは小さな会議です	53.61
A-196	これは大きな会議です	53.64
A-197	それはよい方法ですね	100.07
A-198	詳しいことは分りません	119.32
A-199	早い人は半年前から会議に申込んでいます	–
A-200	会議場に近いホテルを見つけました	948.12
A-201	京都タワーという高い塔があります	94.90
A-202	一般の参加者より安い登録費で参加できます	–
A-203	新幹線は短いトンネルをいくつも通過する	88.98
A-204	もっと速いバスがあります	70.90
A-205	会議の細かい計画を立てた	137.35
A-206	あの古そうな建物は何ですか	179.04
A-207	彼のような学者はまれだ	72.42
A-208	会議の案内書を送ります	55.59
A-209	できれば京都駅の近くが好ましいのですが	159.88
A-210	先生のご提案に関しては事務局の方で検討させていただきます	–
A-211	かなりの量だ	11.19
A-212	会議に関する資料を送る	44.13
A-213	会議についての資料を送る	42.87
A-214	五十件もの論文を書いている	175.93
A-215	集合場所は受付の前です	58.71
A-216	集合場所は受付の後ろです	65.65
A-217	登録用紙は机の上にあります	86.69
A-218	大きなビルとビルの間に小さなレストランがある	142.26
A-219	ご住所とお名前をお願いします	995.89
A-220	それが登録用紙ですか	70.55
A-221	それは結構なお話ですね	–
A-222	こちらは国際会議事務局です	59.34
A-223	そちらを見ていただけないでしょうか	620.63
A-224	そちらで探していただけないでしょうか	728.00
A-225	テーマは何ですか	42.14
A-226	テーマはなんですか	43.76
A-227	私が議長です	33.44
A-228	これをお渡しいたします	72.93
A-229	いつがご都合よろしいでしょうか	205.44
A-230	都ホテルと京都ホテルとどちらがよいですか	–
A-231	どちらさまでしょうか	109.15
A-232	どなたさまをお呼びすればよろしいでしょう	202.65
A-233	そこにいてください	47.06
A-234	そこまでは分っています	157.17
A-235	どれが正式の用紙ですか	471.19
A-236	みなさん優秀な通訳です	–
A-237	誰を最初のスピーカーにしますか	555.98
A-238	会議場はどこですか	56.29
A-239	あれは新製品です	41.82
A-240	ここから京都まで三時間掛かります	110.96
A-241	飛行機と新幹線のいずれが速いですか	305.51
A-242	他の研究論文から教えられることが多い	–
A-243	他の研究論文に教えられることが多い	–
A-244	他の研究論文によって教えられることが多い	–
A-245	多くの講演者が質問された	917.52
A-246	多くの聴衆から質問された	–
A-247	情報処理学会は十一月に大阪で開催されます	517.57
A-248	私の名前を彼に使わせました	–
A-249	代理に助手を行かせましょう	–

A-250	学生に勉強させる	93.40
A-251	事故が交通を渋滞させる	–
A-252	OHP を使うことができますか	188.25
A-253	おかげさまで本当に楽しい旅行ができました	866.09
A-254	残念ですが吉田先生は急病で出席できません	192.64
A-255	その論文は信頼できますか	123.34
A-256	武田先生に連絡を取ることは可能です	2397.40
A-257	会場までバスで行けますか	139.29
A-258	登録用紙を送っていただけるのでしょうか	1085.49
A-259	受付に申込用紙があります	30.04
A-260	発表になれます	158.66
A-261	論文を書きます	38.00
A-262	すでに論文を書いてある	117.75
A-263	すでに論文を書いてあった	211.98
A-264	いま論文を書いている	144.45
A-265	何度も書直している	85.68
A-266	すでに多くの論文を書いている	211.00
A-267	すでに多くの論文を書いていた	466.39
A-268	彼は日本を二度訪れている	160.82
A-269	登録料は八万円に設定されています	1194.19
A-270	閉会式のあとにフェアウェルパーティーを予定しています	380.02
A-271	会議の内容は了解しています	254.74
A-272	論文を書いておく	89.99
A-273	論文を書いておいた	140.60
A-274	私はノーマン教授を知っております	188.17
A-275	若干の先生方がその企画に反対しておられます	994.93
A-276	登録料の件でお電話しております	582.61
A-277	きのうまでずっとホテルは満室でした	99.49
A-278	すでに多くの論文を書いている	385.51
A-279	雨が降った	35.27
A-280	登録用紙を送りました	82.97
A-281	その二つの出来事は同時に起きた	207.56
A-282	すべての発表が順調に進んだ	76.18
A-283	基調講演はジョーンズ博士だ	13.84
A-284	これは学会誌に掲載された論文です	–
A-285	十時以降は会場に入らないでください	–
A-286	もう会場に入ってもいいです	873.06
A-287	銀行振り込みでも構いません	249.40
A-288	登録料は小切手でも結構です	166.83
A-289	個人でホテルの予約をしてください	660.36
A-290	登録用紙を送ってくれませんか	184.02
A-291	登録用紙を送ってもらえますか	359.33
A-292	登録用紙を送っていただけないでしょうか	1056.66
A-293	登録用紙を送ってほしい	136.64
A-294	登録用紙を送ってほしいのですが	306.39
A-295	試験の時くらい勉強するものだ	234.24
A-296	子供はいたずらするものだ	437.49
A-297	学生時代はよく勉強したものだ	187.42
A-298	よく勉強しなければならない	463.64
A-299	いつまでに登録しなくてはならないのでしょうか	–
A-300	いつまでに登録しなくてはいけないのでしょうか	2247.57
A-301	いつまでに登録しないといけないのでしょうか	2595.91
A-302	早く登録した方がいいです	–
A-303	それでは登録用紙を送ります	76.71
A-304	参加しようと思います	217.45
A-305	それでは次の議題に移りましょう	79.56
A-306	参加するつもりです	33.91
A-307	それでは登録用紙を送りましょうか	111.53
A-308	あす相談しましょうか	107.11
A-309	京都に一度行ってみたいのです	1358.43
A-310	展示用のパーティションを確保したいのですが	387.90
A-311	会議のために何とか日程をアレンジしたいと思います	–
A-312	新製品を会議で紹介したいと思っています	3065.81
A-313	直接送ってほしいのですが	219.31

A-314	用紙を送ってもらいたいのですが	292.99
A-315	有意義な会議になるだろう	875.64
A-316	五百人ほどの参加者が出席するでしょう	835.72
A-317	あの星はまるでダイヤモンドのようだ	–
A-318	彼は来るようです	53.56
A-319	上野駅はまるで迷路みたいだ	–
A-320	どうも風邪をひいたみたいです	83.52
A-321	今年の会議には一般人が参加するそうだ	387.58
A-322	今年の会議には一般人が参加するそうです	623.88
A-323	ではちょっと事務局に問合せてみましょう	485.14
A-324	来月の会議に登録したいのですが	743.80
A-325	来月の会議に登録したいのですけれども	742.04
A-326	いいお天気ですね	78.60
A-327	ちょっとお聞きしたいことがあるのですが	998.47
A-328	天候によってはバスの運行が中止になることがあります	–
A-329	会議に出席するのは難しいです	585.42
A-330	私が会議に参加するのは確かです	279.40
A-331	ホテルの予約が取れなくて困っているところです	1381.01
A-332	いま出張から戻ったところです	264.67
A-333	登録する前に出席者の人数をもう一度確かめた	–
A-334	夕食を取った後に京都の町に出掛けます	–
A-335	会議に出席する際にどのような手続きが必要ですか	1283.44
A-336	学会に出席するために三日ほど留守にします	–
A-337	出張から戻ったばかりです	117.10
A-338	登録証を紛失した場合は再発行はできません	2449.20
A-339	論文主旨を拝見した上で決めたいと思います	–
A-340	会議のテーマが何か知りません	–
A-341	出席するかどうか分かりません	205.13
A-342	どこか空いているホテルを知りませんか	1359.25
A-343	何か分からない点がございましたらお問合わせください	–
A-344	誰かが私の論文を悪用したようです	351.92
A-345	次の会議がいつかご存じですか	–
A-346	会場にいらした時名札を差上げます	168.39
A-347	学会に出席される際お立寄りください	–
A-348	会議が催されている間雨が降っていた	–
A-349	開催にいたるまで大変でした	–
A-350	会議に出席するためアメリカに行きます	1827.84
A-351	参加者が多かったためホテルの手配ができませんでした	–
A-352	担当者がすぐ来ますからお待ちください	103.38
A-353	ツアーの参加者が多いとバスは二台になります	1351.46
A-354	登録証を受取ったら手続きは終わりです	148.74
A-355	会議で京都にお越しになるならぜひお立寄りください	–
A-356	十月になればかなり涼しくなります	–
A-357	京都は暑いですけれども会場には冷房が入っています	1076.71
A-358	会議が終わった後でパーティーがあります	–
A-359	登録証を紛失した場合再発行はできません	639.51
A-360	お忙しい中申し訳ありません	79.78
A-361	多くの講演がありますが機械翻訳に関するものは注目されています	–
A-362	せっかく京都へ来たので遊ぼうと思います	1083.11
A-363	昨日読んだ本は面白かった	693.15
A-364	議長が公表した事実に驚いた	–
A-365	登録用紙をお持ちですか	52.04
A-366	登録用紙をお持ちでしょうか	55.06
A-367	講演者も懇親会に参加なさるというのは本当ですか	–
A-368	申込みの期限はいつですか	155.19
A-369	会場の近くに郵便局はありませんか	91.50
A-370	はい、ございます	4.67
A-371	はい、あります	8.69
A-372	いいえ、ありません	18.68
A-373	ええ、あります	8.62
A-374	いいえ、ありません	18.74
A-375	いいえ、ございません	25.08
A-376	参加料はいくらでしょうか	51.60
A-377	二万五千円です	22.34

A-378	参加されるのですか	409.87
A-379	先生は出席されると思いますか	3446.53
A-380	会場に出席しない	141.89
A-381	まだお送りしてないのですが	1626.14
A-382	そう言ってるんじゃないのです	397.56
A-383	そうじゃない	27.84
A-384	そう言ってるのではない	–
A-385	複雑な手続きは必要ではないと思います	260.71
A-386	欲しくないです	41.12
A-387	複雑な手続きは必要ではありません	60.32
A-388	そうではございません	19.44
A-389	そういう事情なのではございません	49.53
A-390	そういう事情ではございません	91.61
A-391	学会には出席しません	141.74
A-392	そうだとは思いません	58.36
A-393	そうではないと思います	167.16
A-394	彼は決して来ない	72.65
A-395	嫌がる様子は少しもない	85.28
A-396	八万円の登録料は全然高くない	126.10
A-397	登録用紙はお持ちですか	62.13
A-398	参加者はすでに二百人を越えました	130.71
A-399	登録用紙って送ってもいいのですか	–
A-400	あなたも会議に出席されますか	1011.62
A-401	会議で使われる言葉は日本語だけですか	–
A-402	参加者は二百人くらいです	100.63
A-403	参加者は二百人ぐらいです	30.57
A-404	ツアーの参加者は二百人もいます	148.99
A-405	では登録用紙をお送りいたします	101.86
A-406	登録用紙は登録費と一緒に返送します	230.27
A-407	スケジュールについてはこちらから連絡します	224.96
A-408	事務局は会場を手配します	80.22
A-409	次の議題に移りましょう	67.40
A-410	登録用紙をお送りしましょうか	116.34
A-411	そうです	15.69
A-412	会場は広いです	44.28
A-413	広い会場が必要です	42.80
A-414	原稿を準備する時間が必要です	373.60
A-415	いつ登録なさいましたでしょうか	462.24
A-416	会場は広いでしょうか	71.78
A-417	広い会場が必要でしょうか	78.40
A-418	会場は広いでしょうね	56.84
A-419	広い会場が必要でしょうね	52.32
A-420	すでに登録用紙をお送りでしょうね	88.11
A-421	いつおいでになりますか	43.91
A-422	初日の夜に親睦会がございます	36.48
A-423	すでに原稿は出来ております	160.99
A-424	あすそちらに参ります	65.25
A-425	あすそちらに伺います	65.70
A-426	失礼いたします	61.88
A-427	申込書をお送りいたします	77.15
A-428	それでは登録用紙と振込用紙を郵送いたします	165.40
A-429	名前は田中一郎と申します	103.22
A-430	こちらが申し上げることは何もありません	330.87
A-431	送っていただけるのでしょうか	251.19
A-432	送ってくださるのでしょうか	186.49
A-433	非常の場合は右手の非常口から逃げてください	129.88
A-434	口座番号を申し上げますのでメモを用意してください	379.87
A-435	もう一度そちらの電話番号を確認させてください	693.65
A-436	最終原稿の形態についてはそちらで判断してください	532.51
A-437	お送り願えますか	66.65
A-438	市内観光には参加なさらないのですか	330.52
A-439	この資料を読まれましたか	293.22
A-440	先生は何時に会場に来られますか	2023.69
A-441	先生はすぐ京都にいらっしゃいます	96.12

A-442	京都大学の教授でいらっしゃいます	141.12
A-443	案内書をご覧になりましたでしょうか	119.93
A-444	先生に日時をお伝えしました	108.94
A-445	近いうちにご連絡いたします	679.76
A-446	発表の日時の変更をご了承いただきたいのです	392.99
A-447	簡単な履歴書を同封していただけると有難いのですが	–
A-448	通訳料が高いと思われるかもしれません	–
A-449	でもこちらの通訳は優秀です	65.22
A-450	それから開場は九時ですから	81.24
A-451	出張していてそれで準備がまだ出来ていないのです	–
A-452	それでは失礼いたします	86.54
A-453	発表できるほど研究が進んでいません	587.93
A-454	じゃ次の回にお願いします	128.48
A-455	彼は優れた学者でありまた組織家でもある	–
A-456	タクシーでおいでになりますか	56.54
A-457	それともバスですか	48.05
A-458	学校へ行って、そして図書館へ寄った	–
A-459	会議中に地震があった	175.98
A-460	そしたらみんな逃出した	33.54
A-461	では会議でお目にかかるのを楽しみにしています	410.33
A-462	実はまだ研究が終わっていないのです	257.27
A-463	ただ会議場は少し不便な所にあるのです	–
A-464	先生の奥様はアカンパニーパーソンになることができます	2386.82
A-465	しかし助手の方は駄目です	66.68

4.5 資料2: A ランク解析実験文でまだ解析ができていないもの

A ランク解析実験文でまだ解析ができていない文は、以下の3種類に分類される。

1. 機械のメモリ不足で現在では解析不可能な実験文、
2. 交差現象のために解析不可能の文、
3. まだ解析規則を記述していないために解析できない文、

である。

1. については機械上の問題なので、それさえ解決されれは、大半は解析可能だろうと思われる。

4.5.1 機械のメモリ不足で現在では解析不可能な実験文

A-29 国際会議において技術協力について話合われた
A-52 スライドを上映する方がむしろいいと思います
A-53 会議についてもっと詳しいことが知りたい
A-55 医者に行った方がよい
A-58 ロイヤルホテルに比べてプリンスホテルの方が安いです
A-73 参加人員は五百人以上になると思います
A-84 論文の内容が人に知れてしまった
A-113 会議で発表する論文をまとめています
A-139 あるいは遅れるかもしれません
A-199 早い人は半年前から会議に申込んでいます
A-202 一般の参加者より安い登録費で参加できます
A-210 先生のご提案に関しては事務局の方で検討させていただきます
A-221 それは結構なお話ですね
A-236 みなさん優秀な通訳です
A-242 他の研究論文から教えられることが多い
A-243 他の研究論文に教えられることが多い
A-244 他の研究論文によって教えられることが多い
A-284 これは学会誌に掲載された論文です
A-299 いつまでに登録しなくてはならないのでしょうか
A-302 早く登録した方がいいです
A-328 天候によってはバスの運行が中止になることがあります
A-333 登録する前に出席者の人数をもう一度確かめた
A-334 夕食を取った後に京都の町に出掛けます
A-336 学会に出席するために三日ほど留守にします
A-339 論文主旨を拝見した上で決めたいと思います
A-343 何か分からない点がございましたらお問合わせください
A-347 学会に出席される際お立寄りください
A-348 会議が催されている間雨が降っていた
A-351 参加者が多かったためホテルの手配ができませんでした
A-355 会議で京都にお越しになるならぜひお立寄りください
A-361 多くの講演がありますが機械翻訳に関するものは注目されています
A-384 そう言ってるのではない
A-401 会議で使われる言葉は日本語だけですか
A-447 簡単な履歴書を同封していただけると有難いのですが
A-448 通訳料が高いと思われるかもしれません
A-451 出張していてそれで準備がまだ出来ていないのです
A-458 学校へ行って、そして図書館へ寄った
A-463 ただ会議場は少し不便な所にあるのです

4.5.2 交差現象のために解析不可能の文

A-152 ご用件があればどうぞ事務局に申付けてください
A-174 会議にぜひ出席してください
A-175 参加者はおそらく二千人くらいでしょう
A-317 あの星はまるでダイヤモンドのようだ
A-319 上野駅はまるで迷路みたいだ

4.5.3 まだ解析規則を記述していないために解析できない文

4.5.3.1 総主構文

規則化は済み。ただし、必須格補語のある文では曖昧性を生じるという問題がある。

A-11 日本へ行くのはこれが初めてです
A-15 参加者は受付の前に集まってください
A-28 用紙って特別な書式があるのですか
A-125 京都見物は秋がよろしいと思います
A-130 京都は朝が早い
A-285 十時以降は会場に入らないでください

4.5.3.2 受身文 (助詞のコントロール)

A-246 多くの聴衆から質問された

4.5.3.3 変則的な使役文 (助詞のコントロール)

A-249 代理に助手を行かせましょう
A-251 事故が交通を渋滞させる

4.5.3.4 並列表現

A-230 都ホテルと京都ホテルとどちらがよいですか
A-455 彼は優れた学者でありまた組織家でもある

4.5.3.5 形容詞の連体形接続

規則化は済み。ただし実験途中。

A-356 十月になればかなり涼しくなります

4.5.3.6 従属文の述語がない

A-340 会議のテーマが何か知りません
A-345 次の会議がいつかご存じですか

4.5.3.7 意味記述の誤り

「なさる」は AGEN を SPEAKER としているため。

A-367 講演者も懇親会に参加なさるというのは本当ですか

4.6 資料 3: 解析結果において曖昧性が生じた文の分析

今回、A ランクを解析するにあたり、文法および辞書を大幅に変更したため全 465 文中、134 文において曖昧性が生じた。原因は大別して 6 つに分類される。

1. 最も多かったのは格要素が正しく解析されず、同じ素性をもつ他の格 (大半は任意格) として解析されたもので、55 例である。しかし、「と思う」の「と」を [COND] に解析してしまう、といった曖昧性は何らかの制限を設ければ解決されるかもしれない。

2. 同じく多かったのは、述語が複数の語義をもつ場合、辞書に記述された語義がすべて解析結果として出力されるという例である。これは 27 例だった。動詞だけでなく、助動詞「れる、られる」の尊敬や受身などのすべての意味、「た」「ている」などのアスペクトの過去や過去完了などのすべての意味も出力されていた。それぞれ 3 例と 35 例である。

3. 1 文が複数の述語を含んでいる場合、格要素 (副詞を含む) の係り先が誤っているものが 15 例あった。

4. 副詞や助動詞を含む語列では、形態素解析を誤っているものが 13 例あった。

5. 名詞類に関しては、複数の語義をもつものがほとんどないので、曖昧性は生じないが、唯一の例外が形式名詞「こと」である。「こと」に記述した制限のために、「内の関係」と「外の関係」が両方出力されているものが 4 例あった。

6. また、原因はよく分からないが、タグの付与が誤っているもの、また [SEM] で正しく解析されていなかったり、結果が出ていないものが合計で 11 例あった。

各文における曖昧性の原因を以下に列挙した。全解析結果の数を括弧内に示した。結果数と曖昧性の原因の数とは必ずしも一致しない。1 文中に複数の原因がある場合もあり、複数の原因が組み合わさって解析結果を出しているからである。

A-2 (3) 事務局に問合わせる
「事務局に」が [LOCT] として解析されている
A-6 (2) では登録方法から説明しましょう
「登録方法から」が [SOUR] として解析されている
A-8 (2) プロシーディングを十五日までに印刷します
SPEAKER が [AGEN] として解析されていない
A-12 (2) プリンスホテルの方がロイヤルホテルより会場に近いです
「会場に」が [LOCT] として解析されている
A-13 (2) シンポジウムは午前十時より開かれます
助動詞「れる」、「られる」の複数の意味表現が解析結果に出ている
A-17 (2) 新幹線は三時に大阪に到着します
「大阪に」が [LOCT] として解析されている
A-19 (2) 先生と相談します
SPEAKER が [AGEN] として解析されていない
A-20 (2) 同僚と参加したいと思います
SPEAKER が [AGEN] として解析されていない
「同僚と」の係り先の誤り (「思います」にかかっている)
A-23 (2) このセッションは 5 時で終わります。
「セッションは」が [AGEN] として解析されている
A-25 (3) 会議の登録について知りたいのですが
「知る」の複数の意味表現が解析結果に出ている
A-35 (2) 登録用紙を送ってくれるように頼みました
アスペクトの複数の意味表現が解析結果に出ている
A-40 (2) ホテルの予約が取れなくて困っています
「て」が [MANN] [COND] の両方に解析されている

A-41 (2) 立派な会場に驚きました
「会場に」が [LOCT] として解析されている
A-45 (3) 京都へ一度行ってみたい
「行く」の複数の意味表現が解析結果に出ている
A-54 (3) 私は彼と同じくらいの年です
「私は」の係り先の誤り
「彼と」が [ACCM] として解析されている
A-56 (2) これ以上申し上げることはありません
「こと」が [外の関係の連体修飾] として解析されている
A-59 (2) ロイヤルホテルと比べてプリンスホテルの方が安いです
「て」が [COND] として解析されている
A-60 (3) ロイヤルホテルに比べればプリンスホテルは安いです
「ロイヤルホテルに」が [LOCT] として解析されている
A-64 (2) 彼は学者になった
「なる」の複数の意味表現が解析結果に出ている
A-65 (4) シンポジウムは非常な混乱となった
アスペクトの複数の意味表現が解析結果に出ている
「シンポジウムは」の係り先の誤り
A-75 (2) これでシンポジウムを終わります
「これで」が [METH] として解析されている
SPEAKER が [AGEN] になっていない
A-79 (2) 会議に関する問合わせは事務局にお願いします
「事務局に」が [LOCT] として解析されている
A-80 (2) 以前田中先生の講演を聞きました
アスペクトの複数の意味表現が解析結果に出ている
A-81 (3) コンピュータに関する国際会議があると聞いているんですが
「と」が [COND] として解析されている
「聞く」の複数の意味表現が解析結果に出ている
A-82 (4) 午前九時までに受付に来てください
表記の誤り (「九」が「9」と出ていた)
「来る」の複数の意味表現が解析結果に出ている
A-83 (2) 会場までバスが来ています
アスペクトの複数の意味表現が解析結果に出ている
A-85 (3) 論文の投稿は学会の規約に従う
SPEAKER が [AGEN] になっている (誤り)
「投稿は」が [GOAL] として解析されている
「規約に」が [LOCT] として解析されている
A-90 (3) 会場の近くのホテルに泊っています
SPEAKER が [AGEN] になっていない
「会場の近くのホテル」の係り方が誤り
A-91 (2) すみませんがその荷物を取ってください
「来る」の複数の意味表現が解析結果に出ている
A-92 (4) 分からないことがありましたら事務局にお尋ねください
「こと」が [外の関係の連体修飾] として解析されている
「分かる」の複数の意味表現が解析結果に出ている
A-94 (2) いま北大路駅に着きました
「北大路駅に」が [LOCT] として解析されている
A-99 (2) 今月一杯に登録料をお支払ください
タグの付与の仕方が異なる
A-106 (2) 会場に入る
「入る」の複数の意味表現が解析結果に出ている
A-107 (4) 事務局のスタッフと演題について話した
アスペクトの複数の意味表現が解析結果に出ている
「話合う」の複数の意味表現が解析結果に出ている
「スタッフと」が [ACCM] として解析されている
A-121 (4) 京都の方に行きたい
「京都の方に」が [LOCT] として解析されている
「行く」の複数の意味表現が解析結果に出ている
A-122 (6) 先生には事務局から連絡が行っていると思います
「と」が [COND] として解析されている
アスペクトの複数の意味表現が解析結果に出ている
「行って / いる」と解析されている
A-126 (2) 先生の講演がとてもよかった
アスペクトの複数の意味表現が解析結果に出ている
A-129 (4) 彼は京都の地理に詳しい
動作主が OBJE になっている
「地理に」が [LOCT] として解析されている
A-142 (2) 海外からの参加者もかなりいます
「海外からの参加者も」が [SEM] ではなく [PRAG] で解析されている
A-145 (2) それでは早速飛行機の予約をしましょう
「早速」の係り先誤り (「う」にかかっている)

A-146 (2) 実際にご自分でご覧ください
「ご自分で」が [METH] で解析されている
A-147 (2) 会議に出席できるかどうか至急お知らせください
「会議に」の係り先誤り(「お知らせ」にかかっている)
「至急」の係り先誤り(「ください」にかかっている)
A-150 (4) この会議は特に言語処理の研究のために開かれます
助動詞「れる」、「られる」の複数の意味表現が解析結果に出ている
A-153 (5) どうも病気になったらしい
「なる」の複数の意味表現が解析結果に出ている
「病気に」が [LOCT] として解析されている
A-158 (2) まあいいでしょう
「まあ」の係り先の誤り
A-161 (3) まったく知りません
「知る」の複数の意味表現が解析結果に出ている
A-164 (2) 電話番号をもう一度お願いします
「もう一度」が「もう」が [TLOC] として、「一度」が [DEGR] として解析されている
A-166 (4) やはり高いと思います
「高い」の複数の意味表現が解析結果に出ている
SPEAKER が [AGEN] になっていない
「と」が [COND] として解析されている
A-171 (7) なるほどそうですか
「高い」の複数の意味表現が解析結果に出ている
「なるほど」が動詞「なる」と助詞「ほど」として解析されている。しかも
「なる」の複数の意味表現が解析結果に出ている
A-176 (8) 私はまったく知りませんでした
SPEAKER が OBJE になっている
「知る」の複数の意味表現が解析結果に出ている
アスペクトの複数の意味表現が解析結果に出ている
A-177 (3) 会議の内容について詳しく教えてください
「教える」の複数の意味表現が解析結果に出ている
A-178 (2) 事務局長とよく相談します
「事務局長と」が [ACCM] として解析されている
A-180 (2) 多くの参加者がいた
「多くの」が [多く] の連体修飾として解析されている
A-188 (2) 実際の会議はいつですか
「です」が [だ -IDENTICAL] として解析されている
A-198 (2) 詳しいことは分りません
「詳しい」の複数の意味表現が解析結果に出ている
A-200 (3) 会議場に近いホテルを見つけました
「会議場に」が [LOCT] として解析されている
A-207 (2) 彼のような学者はまれだ
変則的な解析規則により、係り先の誤った結果が出ている(「彼のような」が
「まれ」にかかっている)
A-209 (3) できれば京都駅の近くが好ましいのですが
「できれば」が副詞として登録されているため、[INFMANN] として解析されて
いる
A-214 (2) 五十件もの論文を書いている
アスペクトの複数の意味表現が解析結果に出ている
A-234 (2) そこまでは分っています
「分かる」の複数の意味表現が解析結果に出ている
A-235 (2) どれが正式の用紙ですか
タグの誤り
A-241 (2) 飛行機と新幹線のいずれが速いですか
「飛行機と / 新幹線のいずれ」と解析されている
A-245 (4) 多くの講演者が質問された
アスペクトの複数の意味表現が解析結果に出ている
「多くの」が [多く] の連体修飾として解析されている
A-253 (2) おかげさまで本当に楽しい旅行ができました
「本当に」が [LOCT] として解析されている
A-254 (2) 残念ですが吉田先生は急病で出席できません
「吉田先生」が [LOCT] として解析されている
A-263 (2) すでに論文を書いてあった
SPEAKER が [AGEN] になっていない
アスペクトの複数の意味表現が解析結果に出ている
A-264 (2) いま論文を書いている
SPEAKER が [AGEN] になっていない
アスペクトの複数の意味表現が解析結果に出ている
A-266 (2) すでに多くの論文を書いている
SPEAKER が [AGEN] になっていない
アスペクトの複数の意味表現が解析結果に出ている
A-267 (4) すでに多くの論文を書いていた

アスペクトの複数の意味表現が解析結果に出ている
「多くの」が [多く] の連体修飾として解析されている
A-269 (2) 登録料は八万円に設定されています
助動詞「れる」、「られる」の複数の意味表現が解析結果に出ている
A-273 (3) 論文を書いておいた
アスペクトの複数の意味表現が解析結果に出ている
A-275 (2) 若干の先生方がその企画に反対しておられます
「その企画に」が [LOCT] として解析されている
A-276 (2) 登録料の件でお電話しております
「登録料の件で」が解析されていない (バグと思われる)
A-277 (2) きのうまでずっとホテルは満室でした
アスペクトの複数の意味表現が解析結果に出ている
A-278 (4) すでに多くの論文を書いている
「多くの」が [多く] の連体修飾として解析されている
アスペクトの複数の意味表現が解析結果に出ている
A-281 (4) その二つの出来事は同時に起きた
アスペクトの複数の意味表現が解析結果に出ている
A-286 (4) もう会場に入ってもいいです
「入る」の複数の意味表現が解析結果に出ている
アスペクト表現が結果として出ていない
A-287 (2) 銀行振り込みでも構いません
「銀行振り込みでも」が [INST] として解析されている
A-289 (3) 個人でホテルの予約をしてください
「する」の複数の意味表現が解析結果に出ている
A-297 (2) 学生時代はよく勉強したものだ
アスペクトの複数の意味表現が解析結果に出ている
A-304 (2) 参加しようと思います
「と」が [COND] として解析されている
A-305 (2) それでは次の議題に移りましょう
「議題に」が [LOCT] として解析されている
A-307 (2) それでは登録用紙を送りましょうか
タグの誤り
A-309 (3) 京都に一度行ってみたいのです
「京都に」が [LOCT] として解析されている
A-312 (2) 新製品を会議で紹介したいと思っています
SPEAKER が [AGEN] になっていない
「会議で」が解析されていない
A-315 (5) 有意義な会議になるだろう
「なる」の複数の意味表現が解析結果に出ている
「会議に」が [LOCT] として解析されている
A-320 (2) どうも風邪をひいたみたいです
アスペクトの複数の意味表現が解析結果に出ている
A-322 (2) 今年の会議には一般人が参加するそうです
タグの誤り
A-323 (12) ではちょっと事務局に問合せてみましょう
アスペクトの複数の意味表現が解析結果に出ている
「ちょっと」の係り先が誤り
「問合わせる」の複数の意味表現が解析結果に出ている
「事務局に」が [RECP] として解析されている
A-327 (4) ちょっとお聞きしたいことがあるのですが
「ちょっと」の係り先が誤り
「こと」が [外の関係の連体修飾] として解析されている
A-331 (2) ホテルの予約が取れなくて困っているところです
「て」が [COND] として解析されている
A-332 (4) いま出張から戻ったところです
「出張から」が [SOUR] として解析されている
A-337 (4) 出張から戻ったばかりです
「出張から」が [SOUR] として解析されている
アスペクトの複数の意味表現が解析結果に出ている
A-338 (2) 登録証を紛失した場合は再発行はできません
アスペクトの複数の意味表現が解析結果に出ている
A-344 (2) 誰かが私の論文を悪用したようです
アスペクトの複数の意味表現が解析結果に出ている
A-346 (4) 会場にいらした時名札を差上げます
「会場に」が [LOCT] として解析されている
アスペクトの複数の意味表現が解析結果に出ている
A-350 (4) 会議に出席するためアメリカに行きます
「アメリカに」が [LOCT] として解析されている
「行く」の複数の意味表現が解析結果に出ている
A-353 (6) ツアーの参加者が多いとバスは二台になります
「なる」の複数の意味表現が解析結果に出ている

「二台に」が [PURP] として解析されている
A-354 (2) 登録証を受取ったら手続きは終わりです
「終わる」の複数の意味表現が解析結果に出ている
A-359 (2) 登録証を紛失した場合再発行はできません
アスペクトの複数の意味表現が解析結果に出ている
A-362 (10) せっかく京都へ来たので遊ぼうと思います
アスペクトの複数の意味表現が解析結果に出ている
「と」が [COND] として解析されている
SPEAKER が [AGEN] になっていない
「ので」の係り先の誤り
「せっかく」の係り先の誤り
「京都へ」の係り先の誤り
A-363 (4) 昨日読んだ本は面白かった
アスペクトの複数の意味表現が解析結果に出ている
「昨日」の係り先の誤り
A-368 (2) 申込みの期限はいつですか
「だ」の複数の意味表現が解析結果に出ている
A-379 (5) 先生は出席されると思いますか
「と」が [COND] として解析されている
「先生」が [LOCT] として解析されている
「先生」の係り先の誤り
A-381 (3) まだお送りしてないのですが
「お送りして / ない」と解析されている
「て」が [COND] と [MANN] として解析されている
A-382 (2) そう言ってるんじゃないのです
「言う」の複数の意味表現が解析結果に出ている
A-385 (2) 複雑な手続きは必要ではないと思います
「と」が [COND] として解析されている
A-386 (2) 欲しくないです
アスペクトが出ていない
A-389 (2) そういう事情なのではございません
「ございません」が本動詞として解析されている
「ので」が [CAUS] として解析されている
A-390 (2) そういう事情ではございません
「ございません」が本動詞として解析されている
「事情で」が [INST] として解析されている
A-392 (2) そうだとは思いません
「と」が [COND] として解析されている
A-393 (3) そうではないと思います
「と」が [COND] として解析されている
「そう」の係り先誤り
「そう / では / ない」として解析されている
A-400 (2) あなたも会議に出席されますか
「会議に」が [RECP] として解析されている
A-407 (2) スケジュールについてはこちらから連絡します
「連絡する」の複数の意味表現が解析結果に出ている
A-409 (2) 次の議題に移りましょう
「議題に」が [LOCT] として解析されている
A-411 (2) そうです
「そうです」がイディオムとして解析されている
A-415 (2) いつ登録なさいましたでしょうか
アスペクトの複数の意味表現が解析結果に出ている
A-417 (2) 広い会場が必要でしょうか
タグの誤り
A-423 (2) すでに原稿は出来ております
アスペクトの複数の意味表現が解析結果に出ている
[できる -POSSIBLE] と [出来る] の両方で解析されている
A-424 (2) あすそちらに参ります
「そちらに」が [LOCT] として解析されている
A-429 (2) 名前は田中一郎と申します
「田中一郎と」が [ACCM] として解析されている
A-430 (2) こちらが申し上げることは何もありません
「こと」が [外の関係の連体修飾] として解析されている
A-435 (2) もう一度そちらの電話番号を確認させてください
「もう一度」が「もう」が [TLOC] ととして、「一度」が [DEGR] として解析されている
A-438 (2) 市内観光には参加なさらないのですか
「市内観光に」が [PURP] として解析されている
A-440 (6) 先生は何時に会場に来られますか
アスペクトの複数の意味表現が解析結果に出ている
アスペクトが出ていない

A-441 (2) 先生はすぐ京都にいらっしゃいます
「京都に」が [LOCT] として解析されている
A-442 (2) 京都大学の教授でいらっしゃいます
「教授で」が [LOCT] として解析されている
A-445 (2) 近いうちにご連絡いたします
「連絡する」の複数の意味表現が解析結果に出ている
A-449 (4) でもこちらの通訳は優秀です
アスペクトの出る位置が誤り
A-450 (2) それから開場は九時ですから
タグの誤り
A-457 (2) それともバスですか
「それとも」の係り先誤り
A-459 (2) 会議中に地震があった
アスペクトの複数の意味表現が解析結果に出ている
A-464 (2) 先生の奥様はアカンパニーパーソンになることができます
「なる」の複数の意味表現が解析結果に出ている
A-465 (2) しかし助手の方は駄目です
アスペクトが出ていない

4.7 資料 4: 旧文法と新文法による会話 1 〜会話 10 の結果比較

A ランクの解析実験文のための文法・辞書を追加したことによって、処理速度がどのように変化したかを調べるために、会話 1 〜会話 10 の文を再度解析し、旧文法による実験結果と比較した。左端の数値を見ると、最低が会話 7-3 の 0.8 倍、最高が会話 7-11 の 30.6 倍で、1 倍〜 3 倍が大半を占めているが長い文ではかなりの倍率になっているので、全平均では 4.1 倍である。

文法規則はそれほど追加していないので、辞書を大幅に追加したことが、処理性能の落ちた原因と考えられる。

dialogue-1

新 / 旧	日本語	旧文法	新文法
3.5	もしもし	1.08	3.82
2.7	そちらは会議事務局ですか	29.15	77.70
1.6	はい	0.82	1.31
6.1	そうです	4.37	26.80
4.7	どのようなご用件でしょうか	53.80	254.40
2.8	会議に申し込みたいのですが	56.55	157.65
6.8	どのような手続きをすればよろしいのでしょうか	118.85	813.78
19.6	登録用紙で手続きをして下さい	34.33	673.21
3.1	登録用紙は既にお持ちでしょうか	38.23	117.63
8.8	いいえ	1.71	15.06
6.2	まだです	4.28	26.58
1.7	分かりました	14.79	24.43
2.4	それでは登録用紙をお送り致します	73.46	178.95
1.7	ご住所とお名前をお願いします	103.25	170.47
3.7	住所は大阪市北区茶屋町二十三です	36.75	134.67
3.5	名前は鈴木真弓です	22.09	77.09
1.7	分かりました	14.73	24.44
5.2	登録用紙を至急送らせて頂きます	135.45	707.73
1.1	よろしくお願いします	33.10	37.28
3.3	それでは失礼します	25.18	82.83

dialogue-2

新 / 旧	日本語	旧文法	新文法
1.5	はい	0.82	1.27
2.4	こちらは会議事務局です	23.08	54.59
2.9	会議の参加料について教えて頂きたいのですが	394.05	1145.15
5.8	いま会議に申し込めば参加料はいくらですか	180.06	1052.58
1.5	はい	0.85	1.29
5.1	参加料は現在お一人三万五千円です	44.46	224.84
16.3	来月お申込みになりますと四万円です	139.49	2277.23
2.3	参加料には予稿集代と歓迎会費が含まれています	241.88	564.96
2.2	わたしは情報処理学会の会員なのですが	42.56	95.41
1.8	参加料の割引はないのですか	56.08	101.73
2.2	今回は割引を行っておりません	53.18	118.14
4.9	そうですか	6.85	33.31
3.6	参加料はどのようにお支払いしたらよいのですか	316.14	1131.94
4.8	参加料は銀行振り込みです	20.94	100.50
-	案内書に記載されている口座番号に振り込んで下さい	–	–

1.6	また期限は今年いっぱいです	32.32	51.64
1.7	分かりました	14.81	24.44
4.8	どうもありがとうございました	5.23	25.17
25.4	どういたしまして	4.02	102.14
–	分からない点がございましたらいつでもお聞き下さい	534.18	–
–	失礼致します	7.37	–

dialogue-3

新 / 旧	日本語	旧文法	新文法
1.8	はい	0.74	1.32
2.2	こちらは会議事務局です	23.12	51.65
3.2	会議に論文を発表したいと思っているのですが	649.86	2067.47
3.1	会議の内容について教えて下さい	86.96	266.86
-	今回の会議は通訳電話に関連する広範な研究分野を含んでいます	–	–
-	言語学や心理学を専攻する方にも参加して頂く予定です	–	–
1.6	分かりました	14.77	23.68
1.9	ところで会議での公式言語は何ですか	55.34	104.52
2.2	英語と日本語です	21.59	48.35
3.1	わたしは日本語が全然分からないのですが	109.86	343.99
2.4	発表が日本語で行われる場合英語への同時通訳はあるのですか	978.28	2305.23
1.7	はい	0.87	1.49
2.1	英語への同時通訳を用意しております	67.54	138.75
1.6	分かりました	14.86	23.58
4.8	どうもありがとうございました	5.27	25.17
9.9	さようなら	2.65	26.15

dialogue-4

新 / 旧	日本語	旧文法	新文法
2.2	こちらは会議事務局です	23.23	51.63
–	会議について詳しいことを教えて下さい	2987.47	–
2.3	会議の案内書はお持ちですか	43.41	100.04
9.1	いいえ	1.68	15.32
2.6	持っていません	19.17	49.55
5.0	そうですか	6.97	34.78
1.4	会議は八月二十二日から二十五日まで京都国際会議場で開催されます	872.47	1229.94
3.7	参加料は四万円です	21.77	80.25
–	発表を希望されるのでしたら三月二十日までに要約を提出して下さい	2833.78	–
1.5	会議の案内書をお送り致しますのでそれをご覧下さい	140.75	208.61
–	失礼ですがお名前とご住所をお願いします	112.28	–
4.3	アダムスミスです	10.46	44.78
4.3	住所は大阪市東区玉造二丁目二十七の七です	59.29	255.80
1.7	分かりました	14.72	24.29
3.1	電話番号もお聞きしたいのですが	123.77	386.75
1.6	はい	0.80	1.26
2.2	三七二の八〇一八です	17.36	38.03
3.5	三七二の八〇一八でございますね	18.59	65.65
1.6	はい	0.83	1.31
6.4	そうです	4.36	27.74
1.7	それではよろしくお願いします	50.03	83.01
3.1	失礼します	12.02	37.72

dialogue-5

新 / 旧	日本語	旧文法	新文法
1.5	はい	0.84	1.30
2.5	こちらは会議事務局でございます	31.00	77.00
2.0	ちょっとお願いがあるのですが	48.44	95.90
–	私は会議に申込みをした者です	383.78	–
1.5	参加を取り消したいのですが	85.45	131.89
2.8	お名前をお伺いできますでしょうか	96.03	272.79
1.6	はい	0.85	1.35
3.7	ベル研のジムワイベルです	25.57	94.25
1.9	既に登録料の八万五千円を振り込まれておられますね	262.49	511.49
1.6	はい	0.83	1.29
6.5	そうです	4.30	27.97
1.8	登録料を払い戻して頂けますか	86.78	152.54
3.7	お気の毒ですができません	11.77	43.94
3.9	案内書にも書いていますが	89.72	350.85
6.8	九月二十七日以後の取り消しに対する払い戻しはできません	127.91	876.05
1.6	後日プログラムと予稿集をお送り致します	101.27	162.27
–	では誰かが私の代わりに参加することはできますか	2201.75	–
2.6	それは別に問題ありません	17.38	45.75
14.1	代理人が参加する場合はあらかじめこちらまでお知らせ下さい	260.18	3672.94
1.6	分かりました	14.82	23.84
3.8	代理人が決まりましたらお知らせ致します	147.78	568.18
2.8	では失礼します	17.31	49.17

dialogue-6

新 / 旧	日本語	旧文法	新文法
1.6	はい	0.82	1.33
2.2	こちらは会議事務局ですが	23.68	52.17
3.4	会議の間に市内観光があるそうですが	113.75	382.53
2.4	まだ参加できますか	39.12	94.24
1.5	はい	0.83	1.29
4.1	まだ参加可能です	14.27	58.36
2.1	八月五日の午後に清水寺、金閣寺、龍安寺などを見学します	425.59	919.73
3.5	参加なさいますか	16.22	56.79
3.3	参加料はいくらですか	19.03	62.36
3.8	八千円です	9.76	37.55
2.1	参加料には夕食代も含まれています	157.78	338.91
3.6	講演者も参加されるのですか	139.01	500.86
–	講演者の何人かは参加する予定になっています	629.21	–
4.9	そうですか	7.08	34.48
4.0	それでは参加したいと思います	121.03	487.06
1.8	ではお名前と人数をお願い致します	74.36	131.04
1.7	ケンブラウンと申します	49.44	83.75
2.2	家内と参加します	35.75	80.41
17.4	集合場所は会議場の受付の前になっております	111.28	1939.23
1.5	参加料は当日集合場所でお支払い下さい	64.51	99.45
1.6	分かりました	14.79	23.54
8.3	ありがとうございました	1.85	15.32
7.2	ではお待ちしております	18.68	134.39

dialogue-7

新 / 旧	日本語	旧文法	新文法
1.6	はい	0.82	1.28
2.3	こちらは会議事務局です	22.88	51.68
0.8	会議で扱う話題に関して質問したいんですが	1437.86	1104.79

1.5	はい	0.90	1.38
3.7	何でしょうか	16.98	63.09
2.5	機械翻訳という話題が案内書に載っていますが	109.57	269.49
2.4	具体的にこれはどういう内容のものなんですか	70.04	167.84
–	申し訳ありませんがこちらでは専門的な質問にお答えできません	–	–
–	第二版の案内書に会議で発表される論文の題目が載っております	–	–
2.3	そちらを見て頂けないでしょうか	143.15	322.31
30.6	いいですよ	1.70	52.04
2.2	それでは早急にその案内書を送って下さい	83.65	184.73
9.6	送り先は大阪市東区城見二の一の六十一渡辺明です	93.56	899.55
12.5	大阪市東区城見二の一の六十一渡辺明様ですね	67.73	845.03
1.6	はい	0.87	1.39
4.1	早速送らせて頂きます	90.58	368.75
2.6	他に何かございますか	18.67	48.30
8.7	いいえ	1.74	15.06
2.4	ありません	3.23	7.85
7.8	ありがとうございました	1.99	15.43
3.2	失礼します	11.89	37.57

dialogue-8

新 / 旧	日本語	旧文法	新文法
1.6	はい	0.81	1.32
1.9	会議事務局です	14.27	27.65
17.1	ちょっとお聞きしたいことがあるんですが	174.93	2992.16
–	私は今度の会議に発表したいと思っているんですが	1023.56	–
4.8	どのような手続きをすればよろしいでしょうか	129.27	615.50
5.1	先ず二百字の要約を三月二十日までにこちらまでお送り下さい	137.88	704.34
5.0	こちらで審査を行ない、五月二十日までに結果をお送り致します	249.51	1239.70
3.2	投稿が受理された場合原稿用紙を同封いたします	618.70	1986.25
6.9	六月三十日までに原稿の送付をお願いします	104.95	724.61
1.6	分かりました	14.89	23.50
4.7	契約はどのような書式で書けばいいんですか	295.28	1373.14
2.3	所定の申込み用紙がありますのでそれに記入して下さい	197.99	450.00
2.6	それでは申込み用紙を送りますので送り先をお願いします	356.11	919.18
1.6	分かりました	14.71	23.55
3.5	人工知能研究所のジョージオハラです	26.15	91.58
4.1	住所は東京都豊島区東池袋三丁目二番五号です	65.20	267.17
3.8	人工知能研究所のジョージオハラ様ですね	31.20	119.05
4.2	ご住所は東京都豊島区東池袋三丁目二番五号でよろしいですね	143.04	604.37
1.5	はい	0.86	1.28
6.3	そうです	4.39	27.79
2.9	それでは申込み用紙の送付をよろしくお願いします	108.50	312.44
1.5	はい	0.84	1.25
1.6	分かりました	14.86	23.66
1.4	では早速お送り致します	80.31	112.05
1.7	失礼致します	7.57	13.11

dialogue-9

新 / 旧	日本語	旧文法	新文法
2.6	そちらは会議事務局ですか	28.96	75.95
1.4	はい	0.84	1.21
1.9	会議事務局です	14.25	27.66
4.9	何のご用件でしょうか	59.78	293.57
–	会議場へどうやって行ったらいいか教えて欲しいんですが	1518.15	–
2.3	いま京都駅にいるんです	30.04	68.75

3.8	地下鉄で北大路駅まで行って下さい	55.40	207.93
–	そこから国際会議場へ行くバスが利用できます	1289.22	–
1.8	北大路駅ではタクシーも利用できます	79.74	142.47
–	京都駅からタクシーで会議場まで行くにはいくらぐらいかかりますか	407.61	–
2.9	京都駅からですとおよそ六千円かかります	47.73	140.18
–	では北大路駅からですといくらぐらいかかりますか	85.15	–
4.3	北大路駅からですとおよそ九百円です	53.47	231.42
1.6	分かりました	14.76	23.73
4.8	どうもありがとうございました	5.27	25.05
9.1	いいえ	1.66	15.06
25.8	どういたしまして	3.97	102.33

dialogue-10

新 / 旧	日本語	旧文法	新文法
4.1	もしもし	0.96	3.90
1.5	はい	0.86	1.29
2.6	会議事務局でございます	17.91	46.35
3.1	会議の宿泊施設についてお尋ねしたいのですが	140.48	442.50
–	そちらでどこか紹介して頂けますか	385.33	–
1.6	はい	0.85	1.32
–	私どもでご紹介できるホテルは京都ホテルと京都プリンスホテルです	3147.51	–
5.8	一人部屋の値段は一晩七千円から一万円です	113.45	661.68
9.4	二人部屋の値段は九千五百円から六万円です	70.08	661.68
5.1	そうですか	6.83	34.66
5.1	どちらのホテルが会議場に近いですか	109.89	562.81
5.3	京都プリンスホテルが会議場には近いんですが	54.07	287.72
3.0	それでは京都プリンスホテルを予約したいのですが	162.14	485.80
4.1	ホテルの手配もして頂けるのですか	329.69	1343.35
1.6	はい	0.86	1.38
1.9	京都ホテルと京都プリンスホテルは予約できます	70.19	134.11
5.0	そうですか	6.91	34.60
3.4	では京都プリンスホテルの七千円の一人部屋をお願いします	94.69	325.84
1.5	はい	0.84	1.24
3.5	京都プリンスホテルの七千円の一人部屋ですね	49.44	173.03
1.6	はい	0.82	1.33
6.4	そうです	4.33	27.90
3.8	いつからお泊まりになりますか	23.88	89.61
5.8	八月四日の夜からです	29.05	167.45
3.3	八日の朝までお願いします	72.86	240.58
1.6	分かりました	14.85	23.61
3.0	少々お待ちください	11.50	34.27
2.2	お部屋が取れるかどうか調べます	116.10	259.37
2.3	お部屋をお取りできます	66.21	149.36
2.1	ではお名前とご住所をお願いします	114.48	237.22
4.3	中村一雄です	10.52	44.82
4.3	住所は東京都港区新橋一丁目一番三号です	65.16	279.37
3.2	電話番号もお願いします	37.55	121.49
2.7	電話番号は三三一の二五二一です	30.30	81.91
1.6	分かりました	14.60	23.62
8.2	京都プリンスホテルに八月四日から八日まで一人部屋をお取りしました	286.82	2345.07
4.8	どうもありがとうございました	5.30	25.20
3.2	失礼します	11.78	37.46

第 5 章

B・C ランクの文法

5.1 はじめに

B・C ランクの語彙を登録するのに伴って、次のような文法事項について見直しを行った。

- 活用語と活用語尾の扱い
- 時制の助動詞「た」「う」の扱い
- 副詞的名詞の扱い
- 形式副詞の扱い
- 取立助詞の扱い
- 疑問代名詞・疑問副詞の扱い
- 接頭辞・接尾辞の扱い
- 数詞の扱い

以下に、それぞれについて説明する。

5.2 活用語と活用語尾の扱い

従来、動詞・形容詞・形容動詞などの活用語は、先に活用語尾を結合して V-kernel とした後で補語・付加語をかけていた。そのため、一段動詞の未然形や連用形のように活用語尾を持たないものを V-kernel として語幹とは別に登録していた。

また、サ変動詞もサ変名詞と補助動詞を結合して V-kernel とした後で補語・付加語をかけていた。そのため、「〜できる」のような表現では、意味表現が正しく出力されないという欠陥があった。

そこで、次のように規則を変更して、以上の欠陥を是正する。

```
VSTEM -> (P VSTEM)
V-kernel -> (VSTEM VINFL)
V-kernel => (VSTEM) ; 一段動詞の未然形・連用形

N-sahen -> (P N-sahen)
V-kernel -> (N-sahen AUXV-npred)
```

5.3 時制の助動詞「た」「う」の扱い

時制の助動詞「た」「う」は、AUXV-tense として、次のような階層関係で扱っていた。

KERNEL > VOICE > ASPECT > MOOD1 > NEGT > TENSE > MOOD2 > TENSE

このため、テンス・アスペクトの計算において、活用形や時間の副詞の情報が使いにくい。

そこで、「た」「う」を活用語尾と同じようにみなして、次のような規則で扱うことにする。

```
V-KERNEL -> (VSTEM VINFL)
AUXV-x -> (AUXVSTEM-x VINFL)
VINFL -> (VINFL VINFL)
```

5.4 副詞的名詞

- 「間」には名詞的な用法と副詞的な用法がある。

～の間…する	例) 食事の間しゃべる
～の間に…する	例) 食事の間にしゃべる
～する間…する	例) 食べる間しゃべる
～する間に…する	例) 食べる間にしゃべる

- 「間」を N-adv として、次の規則を設ける。

```
;;; 単独で副詞的に用いられる場合
N-adv -> (ATT  N-adv)
N-adv -> (V  N-adv)
V-kernel -> (N-adv  V-kernel)

;;; 格助詞を伴う場合
N => (N-adv)
P -> (N  POSTP-oblg)
ADV -> (N  POSTP-optn)
```

- 意味表現は次のとおり。

```
[[RELN しゃべる-1]
      [AGEN []]
      [TLOC [[PARM ?X[]]
             [RESTR [[RELN 間-1]
                     [ENTITY ?X]
                     [IDEN [[PARM ?Y[]]
                            [RESTR [[RELN 食事-1]
                                    [ENTITY ?Y]]]]]]]]]]
[[RELN しゃべる-1]
      [AGEN []]
      [TLOC [[PARM ?X[]]
             [RESTR [[RELN 間-1]
                     [ENTITY ?X]
                     [IDEN [[RELN 食べる-1]
                            [AGEN []]
                            [OBJE []]]]]]]]]
```

ただし、「… 間」と「… 間に」とは意味が少し異なるが。

5.5 形式副詞

- 多様な補足成分をとって、副詞句をつくる。

  例) びっくりするほど美しい。
  すごいほど美人だ。
  静かなほどいい。
  ショウほど素敵な商売はない。

- 句構造規則の基本形は次のとおりである。

```
ADV -> (X  FADV)
```

N.B. 奥津は「いわゆる日本語助詞の研究」[奥津ほか, 86]の中で、副助詞と接続助詞の区別に触れて、従来副助詞とされてきたものを副詞的用法のものとそうでないもの分け、前者を接続助詞とともに形式副詞とし、後者を取立助詞としている。上記の句構造規則はこの考えに基づいている。

5.6 取立助詞

- 分布が自由であり、なくても文の成立に支障がない。つまり、前接する品詞を変えない。

例) 太郎 (N-com) だけが来た。 太郎に (ADV) までお菓子をくれた。 花子はしっかりと (ADV-kernel) だけしているわけではない。 外国人が 200 人 (N-quant) も参加する。 試験の時 (N-adv) くらい勉強しなさい。

- 句構造規則の基本形は次のとおり。

```
X -> (X  POSTP-contr)
```

5.7 疑問代名詞・疑問副詞

- 従来、疑問代名詞や疑問副詞は N-wh という特別な品詞を与えて、専用の句構造規則を設けていた。しかし、
  - 疑問代名詞 (誰、何、どこ) は、普通名詞と同じ統語的振る舞いをする

    例) 誰が来るか
    何を食べるか
    どこに行くか
  - 疑問副詞 (いつ、どう) は、副詞と同じ統語的振る舞いをする

    例) いつ来るか
    どうするか
- そこで、疑問代名詞には N-com、疑問副詞には ADV-kernel というカテゴリーを与えて、普通の代名詞・副詞と同じ扱いにする。

5.8 接頭辞・接尾辞の規則

接頭辞・接尾辞を含む複合語の構文規則を、次のように想定する。

```
X -> (PREFIX- α  Y)
X -> (Y  SUFFIX- β)
```

X は派生語の品詞、Y は承接語の品詞である。α-prefix、β-suffix はそれぞれ、接頭辞の品詞、接尾辞の品詞である。できるだけ少数の規則で一義的に構文解析ができるように、α と β を決めることが大切である。

接頭辞・接尾辞の意味表現は、原則的に、派生語の品詞と、接頭辞・接尾辞と承接語の関係 (補語構造か付加構造か) によって決まる。しかし、この原則では記述できない特殊な意味構造を持っているものも多く、その場合はそれぞれの特殊性に応じた経路方程式によって記述する。

5.8.1 接頭辞の規則

5.8.1.1 構文規則

接頭辞は一般に派生語の品詞のあり方に影響しないので、承接語の品詞が何であるかによって分類することができるはずである。α は、この分類基準によって決まる。

接頭辞には、名詞につくもの、動詞につくもの、形容詞につくものがある。そこで、名詞につく接頭辞を n-prefix、形容詞につく接頭辞を adj-prefix、動詞につく接頭辞を verb-prefix として、一般的に次のような構文規則を設ける。

```
N-x -> (n-prefix  N-x)          ; お (名前)
VSTEM -> (adj-prefix  VSTEM)    ; ま (新しい)
VSTEM -> (verb-prefix  VSTEM)   ; ぶち (壊す)
```

否定の意味を表す接頭辞には、名詞について形容名詞を派生させるものがある。これを n-prfix-negt として、次のような構文規則を設ける。

```
N-adj -> (n-prefix-negt  N-x)     ; 無 (神経)
```

N.B.1 尊敬名詞・尊敬サ変名詞・尊敬転成名詞には、現在 N-com-honor、n-sahen-honor、N-verb-honor という品詞を当てているが、尊敬・非尊敬で統語的に特別なふるまいをするわけではない。また、サ変名詞と転成名詞とで統語的に区別があるわけでもない。そこで、以後は、N-com と N-sahen にまとめることにする。

N.B.2 毎 (回)、幾 (人) などは、接尾辞につくとすべきかもしれない。しかし、そうすると、規則が複雑になるので、さしあたって接尾辞を除いて構文規則を構築することにする。その後で、コーパスを参照して、必要ならば規則を加える。

5.8.1.2 意味規則

接頭辞の意味表現は、承接語に対する統語的役割、および、その接頭辞が何に由来するかということによって決まる。すなわち、名詞を修飾する接頭辞は連体修飾の、動詞・形容詞を修飾する接頭辞は連用修飾のタイプの意味表現を与える。

また、連体詞に由来する接頭辞、形容詞に由来する接頭辞、副詞に由来する接頭辞、動詞に由来する接頭辞、etc. には、それぞれの意味表現と同じタイプの意味表現を与える。

たとえば、「各国」「新情報」「再開発」「反原発」には、それぞれ次のような意味表現を与える。

```
[[PARM ?X02 [[PARM ?X01[]]
             [RESTR [[RELN 国-1]
                     [ENTITY ?X01]]]]
 [RESTR [[RELN 各-1]
         [ARG ?X02]]]]

[[PARM ?X02 [[PARM ?X01[]]
             [RESTR [[RELN 情報-1]
                     [ENTITY ?X01]]]]
 [RESTR [[RELN 新しい-1]
         [OBJE ?X02]]]]
```

```
[[RELN 開発する-1]
 [MANN [[PARM ?X01[]
        [RESTR [[RELN 再-1]
                [ENTITY ?X01]]]]]]

[[RELN 反-1]
 [OBJE [[PARM ?X01 []]
        [RESTR [[RELN 原発-1]
                [ENTITY ?X01]]]]]]
```

B・Cランクの語彙リストにある、「第(一)」「毎(朝)」「全(世界)」「不(慣れ・確か・合格)」には、次の意味表現を与える。

```
[[PARM ?X02[[PARM ?X01[]
            [RESTR [[RELN 一-1]
                    [ENTITY ?X01]]]]]
 [RESTR [[RELN 第-1]
         [ARG1 ?X01]]]]]

[[PARM ?X02[[PARM ?X01[]]
            [RESTR [[RELN 朝-1]
                    [ENTITY ?X01]]]]]
 [RESTR [[RELN 毎-1]
         [ARG1 ?X02]]]]

[[PARM ?X02 [[PARM ?X01[]]
             [RESTR [[RELN 世界-1]
                     [ENTITY ?X01]]]]
 [RESTR [[RELN 全-1]
         [ARG ?X02]]]]
```

「幾(人)」は、数量名詞の「八万五千円」と同じ次の意味表現を与える。

```
[[RELN 人数]
 [UNIT 人]
 [QUANT [[PARM ?X01[]]
         [RESTR [[RELN 幾-1]
                 [ENTITY ?X01]]]]]]
```

「不(慣れ・確か・合格)」には、否定の助動詞「ない」と同じ次の意味表現を与える。

```
[[RELN NEGATE]
 [OBJE [[RELN 慣れる-1]
        [EXPR []]
```

```
        [OBJE []]]]]

[[RELN NEGATE]
 [OBJE [[RELN 確か-1]
        [OBJE []]]]]

[[RELN NEGATE]
 [OBJE [[RELN 合格する-1]
        [EXPR []]
        [GOAL []]]]]
```

5.8.2 接尾辞の規則

5.8.2.1 構文規則

接尾辞は、それが付加してできる派生語がどの品詞に属するかを決定する。したがって、まず、どの品詞を派生するかによって分類することができる。さらに、承接語の品詞が何であるかによっても、分類することができる。β は、この2つの分類基準の組み合わせによって決まる。

接尾辞には、名詞を派生するもの、形容詞を派生するもの、動詞を派生するもの、連体詞を派生するもの、副詞を派生するものがある。これらを n-suffix、adj-suffix、verb-suffix、att-suffix、adv-suffix とする。また、承接語には名詞の他に、形容詞、動詞などがある。そこで、-adj、-verb を付加して下位区分する。

```
N-com -> (N-com  n-suffix)              ;(子供) たち
N-com -> (VSTEM  n-suffix-adj)          ;(美し) さ
N-com -> (N-adj  n-suffix-adj)          ;(確か) さ
N-com -> (V-kernel  n-suffix-verb)      ;(送り) 先
N-com -> (N-sahen n-suffix-verb)        ;(帰省) 先
N-quant -> (N-num  n-quant-suffix)      ;(100) ドル
N-quant -> (N-quant  n-suffix-quant)    ;(100 ドル) 以上
N-temp -> (N-num  n-temp-suffix)        ;(十月、十日、十時)
N-adv -> (N-com  n-adv-suffix)          ;(会議) 中
VSTEM -> (V-kernel  adj-suffix)         ;(持ち) やすい
```

5.8.2.2 意味規則

接尾辞は派生語の意味タイプを決める。すなわち、名詞を派生する接尾辞は、名詞タイプの意味表現を与え、形容詞・動詞を派生する接尾辞は述語タイプの意味表現を与える。また、連体詞・副詞を派生する接頭辞は、派生語がさらに修飾する名詞・動詞・形容詞との関係を考慮して意味表現を与える。

承接語との意味関係は多様で、連体・連用・etc. のように規定することができないので、基本的には FIELD 素性によって表現することにする。

例) テレビ局

```
[[PARM [[PARM ?X01[]]
        [RESTR [[RELN 局-1]
                [ENTITY ?X01]]]]]
                [FIELD [[PARM ?X02[]]
                        [RESTR [[RELN テレビ-1]
                                [ENTITY ?X02]]]]
```

5.9 数詞の規則

5.9.1 桁読み型の数詞

桁読み型の数詞は、漢数字を用いたものに限定する。「万」「千」「百」「十」のような、位取りを表す語を digit-suffix として、次のような規則を設ける。

```
N-num -> (N-num  digit-suffix)
N-num -> (N-num  N-num)          ; 二万 / 三千、
                                 ; 二万三千 / 四百
                                 ; 二万三千四百 / 五十
N-num-> (N-num  N-num)           ; 二万三千四百五十 / 六
```

後の二つの規則は、「万」から「十」の位までを連接する規則と、「一」のを連接する規則である。

5.9.2 棒読み型の数詞

日本語対話文では、棒読み型の数詞は、電話番号のもののみであり、三桁のものと四桁もののがある。そこで、次のような規則を設ける。

```
N-num -> (N-num  N-num  N-num)
N-num -> (N-num  N-num  N-num  N-num)
```

第 6 章

E セット (機能試験文 600 文) の文法

6.1 文法規則の変更

Ver4.0 では、構文規則の大幅な変更は行わず、意味表現を中心に変更した。

6.1.1 スタートルールの変更

感動詞 (V) が文 (V) を修飾する、V → (V V) という構文規則は、同じカテゴリが連続するために、解析効率を低下させる。そこで、感動詞のカテゴリを V-interj に変更し、V → (V V) の規則を削除した。そして次の規則を設けた。

```
V -> (V-interj V)            ; いいえ協賛の条件はございません
V-sfp -> (V-interj V-sfp)    ; あのーちょっとお聞きしたい事があるのです
が
```

また、スタートルール (V) の下に SENT というカテゴリを新たに設け、V, V-sfp, V-interj がそこを経由するようにした。

```
SENT => (V)
SENT => (V-sfp)
SENT => (V-interj)
```

6.1.2 句点を扱う規則の削除

音声認識部で句点に相当する記号を扱わないことになったため、句点の辞書項目および句点を扱う構文規則を削除した

```
V -> (V PUNC)            ; PUNC = 句点 (、)
V -> (ADV PUNC)
V-sfp -> (V-sfp PUNC)
```

6.1.3 述語接続規則の変更

AUXV-evid が接続していた各階層の述語を V に統一した。

これにより、AUXV-evid の受ける V のアスペクト素性と、AUXV-evid のアスペクト素性は別々に表現される。

V-mood2 → (V AUXV-evid) に統合された規則は以下の通り

```
V-mood2 -> (V-kernel AUXV-evid)
V-mood2 -> (V-voice AUXV-evid)
V-mood2 -> (V-aspect AUXV-evid)
V-mood2 -> (V-mood1 AUXV-evid)
V-mood2 -> (V-negt AUXV-evid)
V-mood2 -> (V-tense AUXV-evid)
```

意味表現は次のようになる

p-43 今出張から戻ったところです

```
[sem [[reln  ところだ -1]
      [obje [[reln 戻る -1]
             [agen []]
             [tloc [[parm !X01]
                    [restr [[reln 今 -1]
                            [entity !X01]]]]]
             [dest []]
             [aspt past]          ←「戻った」の [aspt]
             [dept [[parm !X02]
                    [restr [[reln 出張 -1]
                            [entity !X02]]]]]
      [aspt stat]]]                ←「ところです」の [aspt]
```

6.1.4 形式副詞から構成される副詞句のカテゴリ名の変更

形式副詞によって構成される副詞句のカテゴリ名を、解析効率のために、ADV から ADV-fadv に変更した。

```
ADV-fadv -> (X FADV)
```

これに伴い、以下の規則を追加した

```
Vstem -> (ADV-fadv Vstem)
N-sahen -> (ADV-fadv N-sahen)
N-adj -> (ADV-fadv N-adj)
```

6.1.5 助動詞「だ」「です」のカテゴリ名の変更

従来は、名詞句を補語とする助動詞の「だ」「です」には AUXV-copl-com というカテゴリ名を、副詞句を補語とする「だ」「です」には AUXV-copl-adv というカテゴリ名を与えていた。

p-196 このセッションは四時までです

しかし、上の文では、「セッションは」「四時まで」の両方を必須格補語としてとっている。このような文を解析するために、副詞句と補語とする「だ」「です」のカテゴリ名であったAUXV-copl-advを廃止し、AUXV-copl-comに統合した。

```
V-kernel -> (ADV AUXV-copl-adv)            ; 北大路駅からです
V-kernel -> (ADV-kernel AUXV-copl-adv)     ; まだです

 ↓

Vstem -> (ADV AUXV-copl-com)            ; 北大路駅からです
Vstem -> (ADV-kernel AUXV-copl-com)     ; まだです
```

これにより、意味表現は次のようになる

```
[sem [[reln  だ-adverbial]
      [aspt stat]
      [obje [[parm !X05[[parm !X01]
                        [restr [[reln セッション-1]
                                [entity !X01]]]]
             [restr [[reln この-1]
                     [obje !X05]]]]]
      [tdes-1 [[parm !X02]
               [restr [[reln time]
                       [entity !X02]
                       [hour [[reln number]
                              [iden [[place-1 四-1]]]]]]]]]]
```

6.1.6 変化の構文

新たに変化の構文を扱った。

p-81 会議場は五百人ほどの参加者でいっぱいになった

基本的な意味表現は次のとおり

```
[[reln なる]
 [obje !X03 [[parm !X01]                ←タグを張る
            [restr [[reln 会議場-1]
                    [entity !X01]]]]]
 [resl [[reln だ-identical]
        [iden [[parm !X02]
               [restr [[reln いっぱい-1]
                       [entity !X02]]]]]
        [obje !X03]]]]                  ←タグを張る
```

「なる」が N-adj を受ける場合には、次のような意味表現になる。

p-81 これから言語研究が盛んになっていくでしょう

```
[[reln なる]
 [obje !X02 [[parm !X01]                    ←タグを張る
            [restr [[reln 言語研究-1]
                    [entity !X01]]]]]
 [resl [[reln 盛んだ-1]
        [obje !X02]]]]                      ←タグを張る
```

「に」は AUXV-copl-com として、また、「なる」は AUXV-dont として扱っている。これは、以下のような述語の階層関係

KERNEL > VOICE > ASPECT > MOOD1 > NEGT > TENSE > MOOD2 > TENSE

において、「に」「なる」を助動詞とし、それぞれ AUXV-copl-com, AUXV-dont に設定する方法が、作成する規則数の面からみて最も効率的なためである。

変化の構文を含む文を解析するために次の規則を追加した。

```
AUXV-aspc -> (AUXV-dont AUXV-aspc)   ; 盛んになっ / ていく
V-mood2 -> (V-aspect AUXV-polt-aux-masu)
                                     ; 盛んになっていく / でしょう
V-aspect -> (ADV V-aspect)           ; 天候によっては〜中止になる
V-aspect -> (ADV-kernel V-aspect)    ; これから〜盛んになっていく
```

ただし、「なる」が現れる普通の文 (雨が雪になった) と変化の構文 (会議場がいっぱいになった) では、意味的な違いだけしかなく、統語的に制限するのは不可能である。必然的に曖昧性が増え、解析効率を下げるという問題点がある。

6.2 SEM 部 (意味表現) の変更

6.2.1 名詞句

基本的な変更点

1. 従来の名詞句構造では、それぞれの名詞の [[parm [[restr]]]] 構造を保存したまま名詞句全体を構成していた。

   例: 田中一郎

```
[[parm !X03]
 [restr [[reln name]
         [first-name [[parm !X01]
                      [restr [[reln named]
                              [iden 一郎]
```

```
                                    [entity !X01]]]]]
          [entity !X03]
          [family-name [[parm !X02]
                        [restr [[reln named]
                                [iden 田中]
                                [entity !X02]]]]]]]]
```

E セットではこれを変更し、1 個の [[parm [[restr]]]] 構造にラベルを与えることによって、次のような意味表現を構成する。

```
[[parm !X01]
 [restr [[reln name]
         [entity !X01]
         [family-name 田中-1]
         [last-name 一郎-1]]]]
```

2. 従来、「の」によって接続される名詞句の場合、「の」の前項に arg-1, 後項に arg-2 というラベルを与えていたが、これを iden, obje と変更した。番地を表す名詞句 (例: 「二の三」)、または前項、後項に修飾関係がない名詞句 (例: 「銀行振込または現金」) のみ、arg1, arg2 を残した。

例: トピックの一覧

```
[[parm !X03 [[parm !X01]
             [restr [[reln 一覧-1]
                     [entity !X01]]]]]
 [restr [[reln の-連体修飾]
         [arg-1 [[parm !X02]
                 [restr [[reln トピック-1]
                         [entity !X02]]]]]
         [arg-2 !X03]]]]
      ↓
[[parm !X03 [[parm !X01]
             [restr [[reln 一覧-1]
                     [entity !X01]]]]]
 [restr [[reln の-連体修飾]
         [iden [[parm !X02]
                [restr [[reln トピック-1]
                        [entity !X02]]]]]
         [obje !X03]]]]
```

大幅な名詞句の意味表現の変更により、E セットは M セットの拡張版ではなく、全く異なる意味構造をもつ文法になっている。

以下にそれぞれの変更について述べる。

6.2.1.1 数詞表現

日本語対話文解析用文法 (1992 年 2 月 27 日付) で、数詞の規則 (桁読み型) は、N-num → (N-num N-num) と表わされている。しかし、解析時間がかかるため、CFG 規則に展開した。

数詞のカテゴリを次のように定めた。

```
N-digit   一 ～ 九        ; 1 の位の数字を表し、かつ各桁の接尾辞に
                          ; 接続する

Digit-suffix-zyuu   十   ; それぞれ、十、百、千、万の桁を表す
Digit-suffix-hyaku  百   ; それぞれ二～九 (N-digit) が接続して
Digit-suffix-sen    千   ; 各桁の数字を表す
Digit-suffix-man    万   ;

N-num-zyuu   十      ; それぞれ、十、百、千、万という数字を表す
N-num-hyaku  百      ; 各桁の頭に一が接続していると考えられる
N-num-sen    千      ;
N-num-man    万      ;
```

構文規則は以下の通り

```
N-num-zyuu -> (N-digit Digit-suffix-zyuu)       ; 二十
N-num-hyaku -> (N-digit Digit-suffix-hyaku)     ; 二百
N-num-sen -> (N-digit Digit-suffix-sen)         ; 二千
N-num-man -> (N-digit Digit-suffix-man)         ; 二万

N-num-zyuu -> (N-num-zyuu N-digit)              ; 十二
N-num-hyaku -> (N-num-hyaku N-num-zyuu)         ; 百十
N-num-sen -> (N-num-sen N-num-hyaku)            ; 千百
N-num-man -> (N-num-man N-num-sen)              ; 二万二千

N-num => (N-digit)
N-num => (N-num-zyuu)
N-num => (N-num-hyaku)
N-num => (N-num-sen)
N-num => (N-num-man)
```

数詞は、当該の桁より下の桁に空の値を代入している。

```
例: 千 (N-num-sen)
[[parm ?x]
 [restr [[reln number]
         [entity ?x]
         [iden [[place-1000 一-1]
```

```
                [place-100 []]
                [place-10 []]
                [place-1 []]]]]]]
```

これに他の桁の数詞が接続すると、次のような意味表現になる。

例: 千九百九十三

```
[[parm !X01]
 [restr [[reln number]
         [entity !X01]
         [iden [[place-1000 一-1]
                [place-100 九-1]
                [place-10 九-1]
                [place-1 三-1]]]]]]
```

棒読み型の数詞の意味表現について

電話番号のように、三桁、四桁の数字が連続した文字列の意味表現は以下の通り。

例: 一一〇八

```
[[parm !X01]
 [restr [[reln 棒読み型の数詞]
         [digits [[in [[first 一-1]
                       [rest [[first 一-1]
                              [rest [[first 〇-1]
                                     [rest [[first 八-1]
                                            [rest !X02]]]]]]]]]
                  [out !X02]]]
         [entity !X01]]]]
```

棒読み型の数詞のために N-num-tel という品詞カテゴリを設け、〇から九の数詞を登録した。

構文規則は以下の通り。

```
N-num -> (N-num-tel N-num-tel N-num-tel)  ; 3つの数字の連続
N-num -> (N-num-tel N-num-tel N-num-tel N-num-tel)
                                          ; 4つの数字の連続
```

6.2.1.2 数量表現

数量詞の意味表現は、数詞表現に、数量を表す単位辞が接続して構成される。

基本的な意味表現は次の通り。

例: 二百三人

```
[[parm !X01]
```

```
 [restr [[reln number]
         [entity !X01]
         [unit 人]                ←単位辞の場合
         [iden [[place-100 二-1]
                [place-10 []]
                [place-1 三-1]]]]]]
```

6.2.1.3 時間の表現

時間の表現は、数量表現と同じ構文規則を使用するが、意味表現は異なっている。これは、数量詞が連続して複合名詞を構成することはない (例:「十人二つ」) のに対し、時詞は「十二時十分一秒」のように複合名詞を構成するのが普通だからである。

そこで、時間の表現については [[reln]] の下に [[hour][[minute]] などのラベルを設け、その下に数詞を表現する。

例: 二十時三分

```
[[parm !X01]
 [restr [[reln time]
         [entity !X01]
         [hour [[reln number]
                [iden [[place-10 二-1]
                       [place-1 []]]]]]
         [minute [[reln number]
                  [iden [[place-1 三]]]]]]]]
```

6.2.1.4 取立助詞が後続する名詞句

取立助詞「くらい」「ぐらい」「ほど」「など」「ばかり」などに対しては、[[contr]] というラベル名を与えて SEM 部に表現する。

例: 二百人ぐらい

```
[[parm !X01]
 [restr [[reln number]
         [entity !X01]
         [unit 人]
         [contr ぐらい-p-1]        ← 取立助詞
         [iden [[place-100 二-1]
                [place-10 []]
                [place-1 []]]]]]]
```

また、「二百人も参加する」のように、数量詞に後続して概数を表す「も」、および、「誰か知りませんか」のように、疑問詞に後続して不定であることを表す「か」は、[[suffix]] というラベル名で [[SEM]] に表す。

例: 誰か

```
[[parm !X01]
 [restr [[reln 誰-1]
         [entity !X01]
         [suffix か-indefinite]]]]
```

「誰か」を解析する規則を作成した。

```
ADV -> (N-com POSTP-indef)        ; 誰か
```

6.2.1.5 人名、地名など固有名詞の名詞句

人名、地名など固有名詞から構成される名詞句には、[[reln]] の下にラベル名を与え、それぞれの語の [[iden]] の値を代入する。

例: 渡辺明

```
[[parm !X01]
 [restr [[reln name]
         [entity !X01]
         [family-name 渡辺-1]
         [last-name 明-1]]]]
```

人名に接尾辞が接続した場合

```
[[parm !X01]
 [restr [[reln name]
         [entity !X01]
         [family-name 渡辺-1]
         [last-name 明-1]
         [honorific-title 様]]]]
```

例: 京都府相楽郡精華町光台

```
[[parm !X01]
 [restr [[reln address]
         [entity !X01]
         [pref 京都]
         [gun 相楽]
         [machi 精華]
         [town 光台]]]]
```

人名の複合名詞に関して以下の規則を追加した。

```
N-com => (N-family)                    ;(長尾先生) は
N-name -> (N-name N-suffix)            ; 渡辺明様
N-family -> (N-family N-suffix)        ; 長尾先生
```

地名の複合名詞に関して以下の規則を追加した。

```
N-ward -> (N-ward N-chome)          ; 北区 / 茶屋町二十三
```

6.2.1.6 接頭辞、接尾辞類が後続する名詞句

接頭辞には [[prefix]] というラベル名を、また、単位辞ではない接尾辞には [[suffix]] というラベル名を与えて名詞句を構成する。

```
例: 約十分おき
[[parm !X01]
 [restr [[reln name]
         [entity !X01]
         [prefix 約]
         [iden [[place-10 一-1]
                [place-1 []]]]
         [unit 分-1]            ←この「分」は数量を表す
         [suffix おき-1]]]]
```

6.2.1.7 形式名詞型の意味表現

形式名詞型の語の補語は、[[reln [[iden]]]] の下のレベルに、[[parm [[restr]]]] 構造のまま表現される。

この場合の形式名詞には、単独では現れず、必ず補語を伴う語も含まれている。「(参加する) 方」「(プリンスホテルの) ほう」「(参加する) ため」「(準備の) ため」などが、その例である。

```
例: プリンスホテルのほう
[[parm !X01]
 [restr [[reln ほう-1]
         [entity !X01]
         [iden [[parm !X02]
                [restr [[reln named]
                        [iden プリンスホテル-1]
                        [entity !X02]]]]]]]]
```

6.2.1.8 形式名詞型の意味表現をとる接尾辞

接尾辞でも意味的に形式名詞であると判断された語に関しては、形式名詞と同じ意味表現を与える。

p-206 今月中に登録料をお支払いください

「今月中」の意味表現

```
[[parm !X01]
```

```
[restr [[reln 中-fn-time]
        [entity !X01]
        [iden [[parm !X02]
               [restr [[reln 今月-1]
                       [entity !X02]]]]]]]]
```

数量詞や時詞に接尾辞が接続して生成される名詞句に対応して以下の規則を追加した。

```
N-temp -> (N-quant N-suffix-temp)     ; 一時間後
N-adv -> (N-temp N-adv-suffix)        ; 今月中
```

6.2.1.9 副詞修飾の副詞

E セットでは副詞 (句) を修飾する副詞を扱った。

p-203 これはなかなかよく書けている論文です

副詞は通例は述語を修飾するが、上の例では、副詞を修飾している。このような文を解析するために、次の規則を作成した。

```
ADV-kernel -> (ADV-degr ADV-kernel)     ; なかなか / よく
N-com -> (ADV-degr N-com)               ; ちょうど / ホテルの前から
ADV-fadv -> (ADV-cond ADV-fadv)         ; もし / それが無理ならば
```

意味表現は次のようになる

```
[[mann [[parm !X01]
        [restr [[reln よく-1]
                [entity !X01]
                [degr [[parm !X02]
                       [restr [[reln なかなか-2]
                               [entity !X02]]]]]]]]]]
```

6.3 PRAG 部の変更

取立助詞は [[SEM]] 部に表しているが、取立の対象および取立の範囲を明確にするため、それぞれ [[focus]][[scope]] として PRAG 部にも表す。

例: 参加者は二百人ほどです

```
[[sem !X08[[reln だ-identical]
           [obje !X10[[parm !X01]
                      [restr [[reln 参加者-1]
                              [entity !X01]]]]]
           [iden !X07[[parm !X05]
                      [restr [[reln number]
```

```
                                [iden [[place-100 二-1]
                                       [place-10 []]
                                       [place-1 []]
                                       [unit 人]
                                       [contr ほど-p-1]
                                       [entity !X05]]]]]]]
      [aspt stat]]]
[prag [[restr [[in [[first [[reln polite]
                            [agen !X02 [[label *speaker*]]]
                            [recp !X03 [[label *hearer*]]]]
                     [rest !X13]]]
              [out !X13]]]
       [speaker !X02]
       [hearer !X03]
       [aspe [[in !X19]]
             [[out !X20]]]
       [prsp-terms [[in [[first [[focus !X07]
                                 [prsp-mode hodo]
                                 [scope !X08]]]
                           [rest !X14]]]
                     [out !X14]]]
       [topic [[in [[first [[focus !X10]
                            [topic-mod wa]
                            [scope !X08]]]]
                     [rest !X15]]]
              [out !X15]]]]]]
```

第 7 章

M セット (モデル会話文) の追加作業

7.1 音声認識誤り検出用の意味情報の付加

音声認識部から渡された解析結果について、意味的に誤った文を除外するため、解析文法に意味情報を付加した。

なお、この意味情報は、M セットの範囲で誤った文が制限されるように付加したものである。

意味情報の内容は以下のとおり

7.1.1 述語のとる必須格補語の意味素性の制限

「お取り」の補語に対する意味素性の制限が緩いため、音声認識の結果である「お願いをお取りできます」(正しくは、「お部屋をお取りできます」) という文を解析している。

そこで、「お取り」の補語の意味素性を 2 種類記述する。

```
;; 一般的な意味制約
#-SPREC(<!m syn head grfs obj semf> == [[anim -]])
;; 音声認識用の意味制約
#+SPREC(<!m syn head grfs obj semf> == [[anim -][hum -][conc +]])
```

音声認識部と接続しない場合には、-SPREC の行が読み込まれる。一方、音声認識部と接続した場合には +SPREC の行が読み込まれる。その結果、「お取り」は、[[conc -]] の素性をもつ「お願い」を補語として取ることができなくなり、この文は失敗する。

このような意味制約を加えた語は他に、「書く」、「送る」、「分かる」、「申し込む」がある。

7.1.2 格助詞のとる補語の意味素性の制限

「会員について詳しい事も教えてください」(正しくは「会議について詳しい事を教えてください」) を排除するため、「について」の補語の意味素性を制限する。

「について」は本来、どのような素性の名詞句もとることができるが、M セットの中では、[[hum +][anim +]] の素性をもつ名詞句をとることはないため、以下のような制限を加える

```
;; 音声認識用の意味制約
#+SPREC(<!m syn subcat first semf> == [[anim -][hum -]])
```

このような意味制約を加えた語は他に、「と」、「など」、「ほど」がある

7.1.3 任意格要素の修飾先の制限

任意格補語 (「〜に」および「〜へ」) の誤った修飾先を排除するために、「〜に」「〜へ」の修飾しうる述語を限定する素性 [[tran]] を設定し、誤った文を排除する。

例: 意味へは早速お送りいたします

助詞「へ」に以下の行を追加する

```
(<!m syn head coh syn head tran> == +)
```

これに対し、「お送り」には以下の行を追加する。

```
(<!m syn head tran> == -)
```

これにより、[[syn [[head [[coh [[syn [[head [[tran +]]]]]]]]]]]] である「意味へ」が、[[syn [[head [[tran -]]]]]] である「お送りいたし」を修飾することに失敗する。

[[tran -]] を設定した語は他に「載る」「かかる」「振り込む」がある。

7.1.4 形式副詞の修飾先の制限

「すれば失礼します」(正しくは「それでは失礼します」) などの文を排除するため、形式副詞の修飾先を制限する。

その形式副詞の修飾する述語の種類を以下のように記述する

形式副詞「ば」の場合

```
(<!m syn head coh syn head morph ctype> == desu)
```

(「今会議に申し込めば参加料はいくらですか」のために)

形式副詞「と」の場合

```
(<!m syn head coh syn head morph ctype> == (:or desu cons-uv))
```

(「京都駅からですとおよそ六千円かかります」のために)

7.1.5 感動詞を修飾する副詞の制限

また、感動詞にかかる副詞の種類を制限するため、感動詞の品詞を V から V-interj に変更した。そのために、スタートルールの下に SENT というカテゴリを設け、V, V-sfp, V-interj がそこを経由するようにする。それに伴って、以下の規則を追加した

```
V-interj -> (ADV-state V-interj)
```

仕様書では、E セットに対して意味情報を付加するということだったが、M セットに対して行った。対象となった文法のバージョンは Ver3.2.1, Ver3.2.2 である。

参考文献

[衛藤・永田, 91] 衛藤純司, 永田昌明: 日本語単一化文法における語順変化と主題化の効率的な解析手法, 情報処理学会第43回全国大会 3G-1, 1991.

[江原ほか, 90] 江原暉将, 井ノ上直己, 幸山秀雄, 長谷川敏郎, 庄山富美, 森元逞: ATR対話データベースの内容, ATRテクニカルレポート TR-I-0186, 1990.

[奥津ほか, 86] 奥津敬一郎, 沼田善子, 杉本武: いわゆる日本語助詞の研究, にほんごの凡人社, 1986.

[小倉ほか, 89] 小倉健太郎, 坂野俊哉, 保坂順子, 森元逞: 音声言語日英翻訳実験システム (SL-TRANS), ATRテクニカルレポート TR-I-0102, 1989.

[竹沢ほか, 93] 竹沢寿幸ほか (14名): ATR音声翻訳システム ASURA の実装, ATRテクニカルレポート TR-I-0303, 1993.

[寺村, 82] 寺村秀夫: 日本語のシスタクスと意味 I, くろしお出版, 1982.

[寺村, 84] 寺村秀夫: 日本語のシスタクスと意味 II, くろしお出版, 1984.

[永田・久米, 90] 永田昌明, 久米雅子: SL-TRANS における日本語文法の概要, ATRテクニカルレポート TR-I-156, 1990.

[永田, 91] 永田昌明: 単一化文法の効率的な解析手法, 情報処理学会第42回全国大会 5C-8, 1991.

[永田・衛藤, 92a] 永田昌明, 衛藤純司: 日本語基本辞書の概要, ATRテクニカルレポート TR-I-0250, 1992.

[永田・衛藤, 92b] 永田昌明, 衛藤純司: 日本語基本辞書の意味情報に関する検討, ATRテクニカルレポート TR-I-0251, 1992.

[永田・衛藤, 92c] 永田昌明, 衛藤純司: SL-TRANS の日本語標準表記に関する検討, ATRテクニカルレポート TR-I-0278, 1992.

[永田 ほか, 93] 永田昌明, 田代敏久, 衛藤純司, 坂口明子: 日本語解析文法 解説書, ATRテクニカルレポート TR-I-0335, 1992.

[益岡・田窪, 89] 益岡隆志, 田窪行則: 基礎日本語文法, くろしお出版, 1989.

[吉本, 87] 吉本啓: 日本語品詞の分類, ATRテクニカルレポート TR-I-0008, 1987.

[吉本, 88] 吉本啓: 句構造文法に基づく日本語文の解析, ATR テクニカルレポート, TR-I-0049, 1988.

[Kogure et. al, 90] K. Kogure, H. Iida, T. Hasegawa, and K. Ogura: "NADINE: An Experimental Dialogue Translation System from Japanese to English", Proc. Info Japan'90, 1990.