

対話文における誤入力訂正処理

柿ヶ原 康二 相沢 輝昭

(株)ATR自動翻訳電話研究所

自動翻訳電話システムを実現するには音声認識、機械翻訳、音声合成の3つの技術が必要であり、それらの技術のインターフェースを効率よく実現することはシステムにとって重要な課題である。本報告では音声認識で発生する認識誤りを言語情報を用いて訂正する手法を提案する。第一段階として、文節に区切って発話された文の自立語が正しく認識されているものとして、自立語列から最適な付属語を推定し、認識結果の付属語と比較することで発話を決定する手法を提案する。今回の報告では自立語列から付属語を推定する手法についてその概略を示す。

Error-Correcting Method in Spoken Dialogue

Koji KAKIGAHARA, Teruaki AIZAWA

ATR Interpreting Telephony Research Laboratories

Twin 21 Bldg. MID Tower, 2-1-61 Shiromi, Higashi-ku, Osaka 540 Japan

The Interpreting Telephone System is composed of three main technology, speech recognition, machine translation and speech synthesis. The interface technology among them is a matter of importance for the system. This paper presents a method for correcting errors on speech recognition for Japanese conversation by using linguistic information. As the first step, we assume that the jiritsugo-words such as noun, verb, adjective and so on, are recognized correctly on isolated word block recognition. Then our method 1) generates the fuzokugo-words such as auxiliary verbs and particles so that they, with the recognized jiritsugo-words, constitute a meaningful sentence, 2) compares them to the candidates of speech recognition, and 3) decides the whole sentence. This paper mainly discusses the second step of our method.

1. はじめに

異なる言語を話す人の電話でのコミュニケーションをはかるための自動翻訳電話システムを実現するには、音声認識技術、機械翻訳技術、音声合成技術の確立が必要である。

従来から機械翻訳は文法的にもまた意味的にも正しい文を翻訳対象にしており、誤りが存在すると、正しく翻訳することができない。しかも音声認識には認識誤りは避けられないものであり、機械翻訳部と音声認識部の効果的なインターフェースは自動翻訳電話システムを実現するための非常に大きな課題である。

本稿では、そのようなインターフェースの問題を機械翻訳の立場から検討した中間結果を報告する。

2. 問題の設定

2.1 音声認識結果としての文節ラティス

音声認識としては比較的实现可能性の高い文節発話を前提とする。その場合、それぞれの文節に対する認識結果の候補がラティスの形で与えられる。図1は「原稿の締め切りはいつですか」という質問文を文節に区切って発話した場合の認識結果の例で、自立語部分とその自立語に付属する付属語列の候補がラティスを形成している。図中の数値は、その候補の認識の確からしさを表す尤度である。

機械翻訳部への入力としては、このような音声認識結果としての文節ラティスを仮定する。このとき、インターフェースの問題は、この文節ラティスから文節ごとの最適候補を選択する問題になる。

発話文	文節1 原稿の		文節2 締め切りは		文節3 いつですか	
認識結果	原稿 4	を 4 も 3 の 1	締め切り 4	は 3 が 2	いつ 3	ですか 4 でした 2 べしか 1
	検討 2	を 4 も 3 の 1	刺激 2	には 4 にが 1	椅子 2	ですか 4 でした 2 べしか 1
	念頭 1	を 4 も 3 の 1			行く 1	ですか 4 でした 2 べしか 1

図1 音声認識結果の例

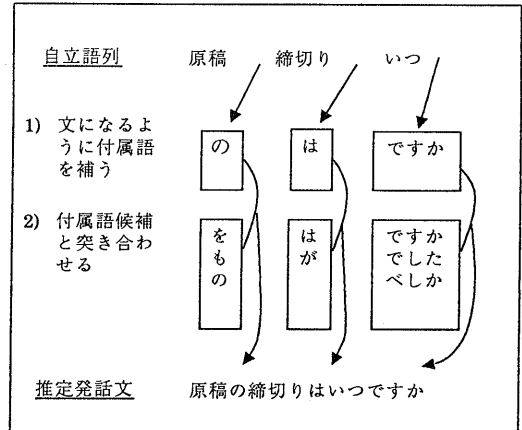


図2 付属語推定の流れ

2.2 文生成による付属語推定

前述の認識結果の文節ラティスに対して、本稿では第一段階として、何らかの方法で自立語の部分の選択が終わったとして、それ以降の付属語選択の問題を考える。現実の状況としては、文節ラティスの各文節の自立語候補から、尤度等を用いて自立語が選択されたとし、その自立語に対する付属語候補から最適な付属語を選択する部分問題を考えることになる。図1の例では、

原稿 締め切り いつ (1)

という自立語が選択されたとして、それに対する付属語候補

を は ですか
も が でした
の が べしか (2)

から最適な付属語を選択する問題になる。あるいは端的に、自立語の部分は正しく認識されたとし

て、付属語の部分の候補を絞る問題として捕らえることもできる。このような問題を考えることは、付属語の多い日本語の場合、十分に意味があると考えられる。

さて、自立語(1)に対する付属語候補(2)から最適な付属語を選択する方法として、以下の2段階の手順を考える。

- 1) 自立語(1)のそれぞれにうまく付属語を補って、まとまりのある文を生成する。
- 2) 補った付属語と、認識結果としての付属語候補(2)とを突き合わせ、最適なものを選択する。

全体の流れを図2に示す。

文末に対しては、その表現が任意に変化するため、付属語の推定は非常に困難である。しかしながら、対話文の場合には、表現の種類をいくつかに分類し、特徴的な表現に関わる付属語部分を準自立語として扱うことにより推定が可能になると考える。この点については今後の検討課題である。そこで、今回は文末の表現が固定されている質問文に対して検討を行った。

音声認識の文節候補から最適なものを選択する手法として堀ほか(1)の一連の研究がある。

2.3 付属語推定のレベル

自立語列から付属語部分を推定する場合にはどのような情報を用いるかが重要な問題となる。推定には以下のレベルが考えられる。

- 1) 構文レベル；文法情報のみで推定
- 2) 意味レベル；意味情報を用いて推定
- 3) 文脈レベル；文脈情報を用いて推定

1) 構文レベル

自立語の持つ文法情報のみで接続する付属語が推定できる場合であり、自立語の品詞、活用のある自立語の時はその活用型、活用形から接続する付属語を推定する。

2) 意味レベル

自立語の持つ文法情報だけでは付属語が推定出来ない場合であり、名詞の持つ意味素性と述語がどのような名詞と共起するかを考慮して付属語部分を推定する。

3) 文脈レベル

付属語部分がそれ以前の発話内容を参照しないと推定できない場合であり、現在話されている話題や発話者の意図などを知る必要がある。

文脈レベルでの付属語の推定は難しい問題を多く含んでおり、話題、意図の参照などは機械翻訳

部からフィードバックされる情報を用いることが必要になると思われる。そこで本報告では意味レベルまでを検討の対象としている。

3. 対象とする対話文

対話文の分析のために収集している端末間対話文に対して付属語推定の可能性を調査した。端末間対話文は電話対話の基本的性質を保持しており、それぞれの文の乱れが少なく、電話対話に対する処理を検討するための前段階として適当である(2)。この端末間対話は『国際会議の問い合わせ』という状況設定で収録されており、それを構成している発話文は大部分が質問文とその応答文になっている。質問文は相手の返答を要求するものであるが、一般に図3に示されるように分類される(3)。

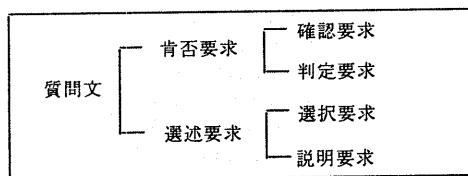


図3 質問文の分類

以下に端末間対話中の質問文の例を示す。

- 確認要求； 三日間通しで100ドルですね。
 判定要求； 同時通訳のサービスは有るのでしょうか。
 選択要求； 会議に参加するのですか、それとも聴講するのですか。
 説明要求； 会議での公式言語は何ですか。

これらの分類のほかに判定と説明を同時に要求しているものとして判定説明文が考えられる。

判定説明要求； 外に何か質問は有りますか。

国際会議への問い合わせの6つの対話(合計369発話)のうち質問文は157文ある。これらの質問文の内、それぞれの要求文の数を表1に示す。確認要求については他の質問文と文末の表現が異なるので本報告の付属語推定処理では対象外とした。

表1 質問文中の各要求文の数 (単位 文)

	確認	判定	選択	説明	判定説明
数量	7	80	2	52	16

4. 結合価パターンを用いた付属語推定

4.1 名詞意味素性と結合価パターン

意味レベルで付属語を推定する場合には、述語がどのような名詞と共起するかという情報が非常に有効である。述語と名詞との意味的な関係により名詞が述語に対してどのような格要素となるかが分かればその名詞に接続する付属語が決まる。

表2 設定した名詞の意味素性

素性名	下位分類と例
抽象概念 abstract (abs)	論理(logic,log);事、理論、方法、内容 状態(state,sta);様子、場合 言語(language,lan);日本語、何語 学術(learning,lea);分野、専門 意向(intention,int);興味、希望 価値(value,val);必要 記号(sign,sig);名前 労役(labor,lab);用件 才能(ability,abi);技術 境遇(situation,sit);資格 その他(others,oth);一緒、泊まり
行為 action (act)	予定、制限、割引、要約、通訳、 参加、発表、募集、
具象物 concrete (con)	文書(document,doc);要約、原稿 交通(transport,tra);バス、地下鉄 物品(article,art);写真、用紙、電話 貨財(money,mon);お金、カード 任意(various,var);これ、それ、どれ
人間 human (hum)	人間(human,hum);発表者、学生 任意(various,var);だれ、そちら
場所 location (loc)	施設(institution,ins);ホテル、会場 地域(region,reg);京都、海外 位置(position,pos);間、方、そちら 任意(various,var);どこ
時間 time (tim)	時間(time,tim);間、後、今、時間 任意(various,var);何時、いつ
数 number (num)	数量(amount,amo);字数、何名 単位(unit,uni);円、ドル 費用(cost,cos);登録費、参加料、費用 価格(price,pri);無料、いくら
任意 diverse (div)	ほか、なに

しかしながら、このような格支配の関係からは、深層の格が把握され、それがかならずしも実際の発話文とは一致しない。そこで、ここでは表層での名詞と述語の結合を求めるために結合価情報を利用する[4]。述語と名詞との共起関係を調べるためには先ず名詞の持つ意味を分類する必要がある。本報告では表2に示されるように上位の分類としては一般的な分類を用い、下位の分類にタスクを考慮した意味素性を用いている。行為を表す名詞は「する」を付加してサ変動詞になるものである。

次に質問文中の述語に対してそれがどのような意味素性を持った名詞とどのような付属語(格助詞)を伴って表現されるかを調べる。表3は質問文中に出てきた述語の例である。これらの述語に対して名詞との結合関係を調べた。表4はデータの中に出てきた質問文に対する結合価パターンの例である。表4の結合価パターンの欄のNは名詞であり、[]のなかの略記号は表2の名詞意味素性の略記号に対応している。Vは述語である。これら実際に質問文に出てきた表現をもとに一般的な結合価パターンを求めるのであるが、収録データの中にその述語が持つすべての表現パターンは出現していないので、一般に表現される述語と名詞の結合価パターンとタスクを考慮して決定している。例えば「持つ」の場合は、基本的な結合価パターンは

「持つ」

$N[\text{hum}]^{\text{g}} + N[\text{abs}/(-\text{sta}, -\text{lan}, -\text{val}, -\text{lab})] + \text{V}$

$N[\text{hum}]^{\text{g}} + N[\text{con}/-\text{tra}] + \text{V}$

$N[\text{hum}]^{\text{g}} + N[\text{con}] + N[\text{loc}] + \text{V}$

のような表現である。実際の対話文では、主格に相当する名詞はほとんど省略されており、また主題化によって助詞「は」をとめない、語順がさま

表3 質問文中の述語

品詞	例
動詞	有る、出す、出来る、出る、 はいる、始まる、払い戻す、開く、 行く、要る、返す、かかる、 聞く、決まる、持つ、 送る、支払う、頼む、
形容名詞	駄目、必要、可能、
形容詞	はやい、いい、宜しい、

ざまに変化する。表4の例の「カードはお持ちですか」の表現に対して基本形は

<基本形>

「あなたがカードを持つ」

N[hum]が+N[con/mon]を+V

である。この基本形を変化させると、

<変化形>

- (1) N[con]は+V
- (2) N[con]を+V
- (3) N[hum]が+N[con]を+V
- (4) N[hum]は+N[con]を+V

のような表現パターンが考えられる。同様に「お手元に登録用紙はお持ちですか」の表現に対しては以下のような変化形が考えられる。変化形の結合価パターンはその出現確率が高いと思われるものを先に記述してある。

<基本形>

「あなたが手元に登録用紙を持つ」

N[hum]が+N[loc/pos]に+N[con/doc]を+V

<変化形>

- (1) N[loc]に+N[con]を+V

表4 結合価パターンの例

述語	結合価パターン
持つ	カードはお持ちですか。 N[con/mon]は+V お手元に説明書をお持ちですか。 N[loc/pos]に+N[con/doc]を+V お手元に登録用紙をお持ちですか。 N[loc/pos]に+N[con/doc]は+V 何に興味をお持ちですか。 N[div]に+N[abs/int]を+V
入る	同時通訳は入りますか。 N[act]は+V 通訳が入るのでしょうか。 N[act]が+V
始まる	会議は何時から始まりますか。 N[act]は+N[tim/var]から+V
開く	会議はどこで開かれるのですか。 N[act]は+N[loc/var]で+V
ある	同時通訳のサービスはありますか。 N[act]は+V 字数に制限はありますか。 N[num/amo]に+N[act]は+V ほかに質問はありますか。 N[div]に+N[act]は+V

- (2) N[loc]に+N[con]は+V
- (3) N[hum]が+N[loc]に+N[con]を+V
- (4) N[hum]は+N[loc]に+N[con]を+V

それぞれの述語についてその変化形を調べ、述語辞書に記述しておくことによって名詞の付属語を推定することができる。特に説明要求に対してはその述語と共起しやすい疑問詞とともに結合価パターンを記述している。図4に述語の辞書記述例を示す。

送る

読み おくる
品詞 動詞
活用型 ラ行五段
基本形
N[hum]が+N[con]を+N[hum]に+V
N[hum]が+N[con]を+N[loc]に+V
変化形
N[con]を+V
N[con]は+N[loc]に+V
N[hum]にN[con]を+V
N[hum]がN[con]を+V
N[hum]が+N[con]を+N[hum]に+V
いつ
N[con]は+N[tim]までに+V
N[tim]までに+N[con]を+V
どこ
N[con]は+N[loc]に+V
N[con]は+N[loc]へ+V
N[loc]に+N[con]を+V

録音する

読み ろくおんする
品詞 動詞
活用形 サ変
基本形
N[hum]が+N[act,abs]を+N[loc]で+V
N[hum]が+N[act,abs]を+N[con]で+V
変化形
N[act,abs]を+V
N[loc]で+N[act,abs]を+V
N[con]で+N[act,abs]を+V
N[act,abs]は+N[loc]で+V
N[act,abs]は+N[con]で+V
どこ
N[act,abs]は+N[loc]で+V
N[con]で+N[act,abs]を+V

図4 述語の辞書記述例

4.2 「名詞+の+名詞」の扱い

名詞Aと名詞Bが格助詞「の」で結合されると、「A+の+B」はそれぞれの名詞の持つ意味によってさまざまな意味内容をあらわす。今回調査した質問文にもいくつかの「A+の+B」が現れている。そこで、二つの名詞の間に格助詞「の」を挿入した場合に「A+の+B」が矛盾のない意味内容を表現しているかどうかを判定することが必要になる。「A+の+B」の表現に対する研究が以前からなされており、ここでは島津⁶⁾に従い、以下の様に分類した。

<「AのB」の分類>

- (分類1) AとBに格関係が存在する。このとき、Aが述語になる場合とBが述語になる場合がある。
- (分類2) Aを基準として場所、時間などを示している。
- (分類3) BがAの性質、属性、数量などを示している。
- (分類4) AがBにかかり、所有、場所、時、数量、限定などを示す。

表5に実際に質問文に現れた「AのB」の例を示す。それぞれの分類において表2の意味素性で表現すると、分類1の場合は格関係が存在するので

- 「開発の技術」：act+の+abs/abi
「技術で開発する」
- 「支払の期日」：act+の+tim
「期日に支払う」
- 「同時通訳のサービス」：act+の+act
「通訳(すること)をサービスする」
- 「通訳電話の開発」：con/art+の+act
「通訳電話を開発する」

のように述語にならない方の名詞にはさまざまな意味素性の名詞がくる。分類2の場合は

- 「会議の間」：act+の+tim
- 「会議の後」：act+の+tim
- 「会場の近く」：loc/ins+の+loc/pos

のように意味素性がactの名詞に対してtimの名詞が結合し、locの名詞はloc,conなどの名詞に結合する。分類3、分類4の場合はデータの数が少なく、意味素性での結合規則の分類はできないが、現段階では個々の分類の視点となる名詞の辞書にどのような意味素性の名詞と助詞「の」で結合されるかを記述しておくことにより、助詞「の」が挿入

されるかどうかを判断する。ここで視点とはそれぞれの分類に対して基準となる名詞である。

- <視点> 分類1；述語になる名詞が視点
- 分類2；Aが視点
- 分類3；Bが視点
- 分類2；Bが視点

図5に名詞の辞書記述の例を示す。

4.3 付属語推定処理

次に構文レベル、意味レベルで付属語を推定する手法を具体的に述べる。推定処理は次の5つのステップで行われる。

- step1 自立語に対して辞書を参照して、品詞、活用、名詞に対しては意味素性を付与する。
- step2 構文レベルで推定できる部分を処理する。
- step3 述語と名詞の共起情報を用いて名詞に付属する助詞を決定する。
- step4 step3において付属語が付加されていない自立語に対し、場所、時などの意味マーカを持つ名詞について助詞を決定する。

表5 「名詞+の+名詞」の分類

分類	質問文中の例
分類1	Aが述語の場合 開発の技術、会議の公式言語、 参加の費用、支払の期日、 発表の内容、 Bが述語の場合 同時通訳のサービス、 ホテルの手配、宿泊の手配、 通訳電話の開発、技術の会議、 会議の参加、
分類2	時間 会議の間、会議の後 場所 会場の近く
分類3	会議の様子、自分の名前、 ホテルの名前、タクシーの場合、
分類4	所有 会議の要綱集、 場所 どのカード、ほかの発表者 ちかくのホテル 時 今年の会議 数量 何名の会議 限定 発表者の方々、参加者の人たち

step5 [名詞]+[名詞]の繋がりにおいて助詞「の」が挿入可能であれば「の」を挿入する。

構文レベルで推定できる部分には以下のような規則が適用される。

<構文レベル規則>

- 接続詞には付属語は付かない
- 連体詞には付属語は付かない
- 副詞には付属語は付かない。
- 動詞,連用形+付属語+動詞→接続助詞「て」
- 動詞,終止形+付属語+動詞→接続助詞「と」
- 動詞,連用形+付属語+形容詞→接続助詞「て」
接続助詞「ても」
- 動詞,仮定形+付属語+形容詞→接続助詞「ば」

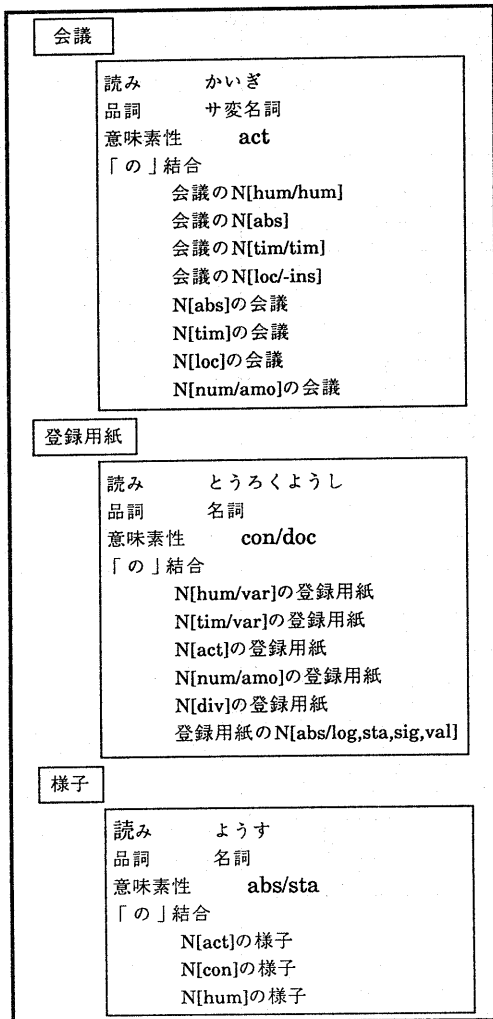


図5 名詞の辞書記述例

動詞,連体形+付属語+名詞→付属語は付かない
 文末の形容詞,終止形には「ですか」が付く
 文末の動詞,連体形には「のですか」が付く

次に例を用いて付属語の推定手順を示す。(図6)
 「登録用紙はいつまでに送ればよろしいですか(例1)」という質問文に対して処理を行う場合、自立語は正しく認識されているという仮定を設けているので入力される自立語列は「登録用紙、いつ、送れ、よろしい」である。辞書をもとに各自立語の品詞、活用、名詞であれば意味素性を調べると、

- 登録用紙 → 名詞、con/doc
- いつ → 名詞、tim/var
- 送れ → 動詞、仮定形
- よろしい → 形容詞、終止形

となる(step1)。「送れ」+付属語+「よろしい」に対しては構文レベルでの推定が可能で、付属語は「ば」である。文末は形容詞の終止形なので「ですか」が付加される(step2)。次に動詞「送る」の結合価パターンから

N[con]は+N[tim]までに+V

が採用される(step3)。次に名詞の連続する部分に助詞「の」が挿入可能かどうか調べる。「登録用紙+いつ」に対して意味素性から「いつの登録用紙」は可能であるが「登録用紙のいつ」は不可能なので助詞「の」は挿入されない。以上で付属語がすべて推定され、「登録用紙はいつまでに送ればよろしいですか」が導かれる。他の例として「会場で会議の様子を録音してもいいですか(例2)」の場合は、入力される自立語列は「会場、会議、録音し、いい」である。辞書から、

- 会場 → 名詞、loc/ins
- 会議 → 名詞、act
- 様子 → 名詞、abs/sta
- 録音し → サ変動詞、未然、連用形
- いい → 形容詞、終止形

となる(step1)。「録音し」+付属語+「いい」は構文レベルでの推定が可能で、付属語は「て、ても」である。文末は形容詞の終止形であるので「ですか」が付加される(step2)。次に動詞「録音する」の結合価パターンから

N[loc]で+N[act,abs]を+V

が採用されるが、「会議」と「様子」のどちらも「録音する」の目的語になりうる(step3)。「会場」は意味素性がlocであるので助詞「で」が付加される(step4)。次に名詞の連続する部分に助詞

「の」が挿入可能かどうか調べる。「会場+会議」「会議+様子」の二つがあり、どちらも辞書からは「の」が挿入できる(step5)。すなわち「録音する」の目的語は「様子」となり、「会議」と「様子」は「の」で結合される。以上で付属語がすべて推定され、「会場(で,の)会議の様子を録音して(て,ても)いいですか」が導かれる。

5. おわりに

本報告では音声認識結果に対して自立語部分が正しく認識されているという仮定のもとに、自立語列からその自立語間の付属語を推定する手法の概略を述べた。特に国際会議の問い合わせという限定されたタスクのもとで、質問文に対して検討を行った。調査した質問文の数がまだ少ないので詳細の検討には至っていないが、付属語推定の可能性は導けたと考える。現在まだ簡単な質問文しか扱っていないが、今後更に複雑な文を調査し、結

合価パターンを精度を上げるとともに、意味素性の分類を更に細かくする必要があると思われる。また、文脈レベルでの推定方法については具体的な手法はまだなく、話題、意図の抽出方法やその話題や意図によって文の表現がどの様に変化するかなど、検討項目が多く残されている。これらについても多くのデータに対して分析を行い最適な手法を確立して行きたい。

謝辞 本研究の機会を与えて下さるとともに適切な助言を述べられたATR自動翻訳電話研究所 樽松明社長に感謝します。また熱心に討論して下さいた言語処理研究室の諸氏に感謝します。

参考文献

- [1] 堀ほか:「音声理解システムのための連想パーザASP」電子情報通信学会研究会資料NLC87-11(1987)
- [2] 有田ほか:「メティアに依存する会話の様式」情報処理学会研究会資料NL61-5(1987)
- [3] 南:「文法と意味II」朝倉書店 1985
- [4] 荻野:「文法と意味I」朝倉書店 1983
- [5] 島津ほか:「日本語文意味構造の分類-名詞句構造を中心に-」情報処理学会研究会資料NL47-4(1985)

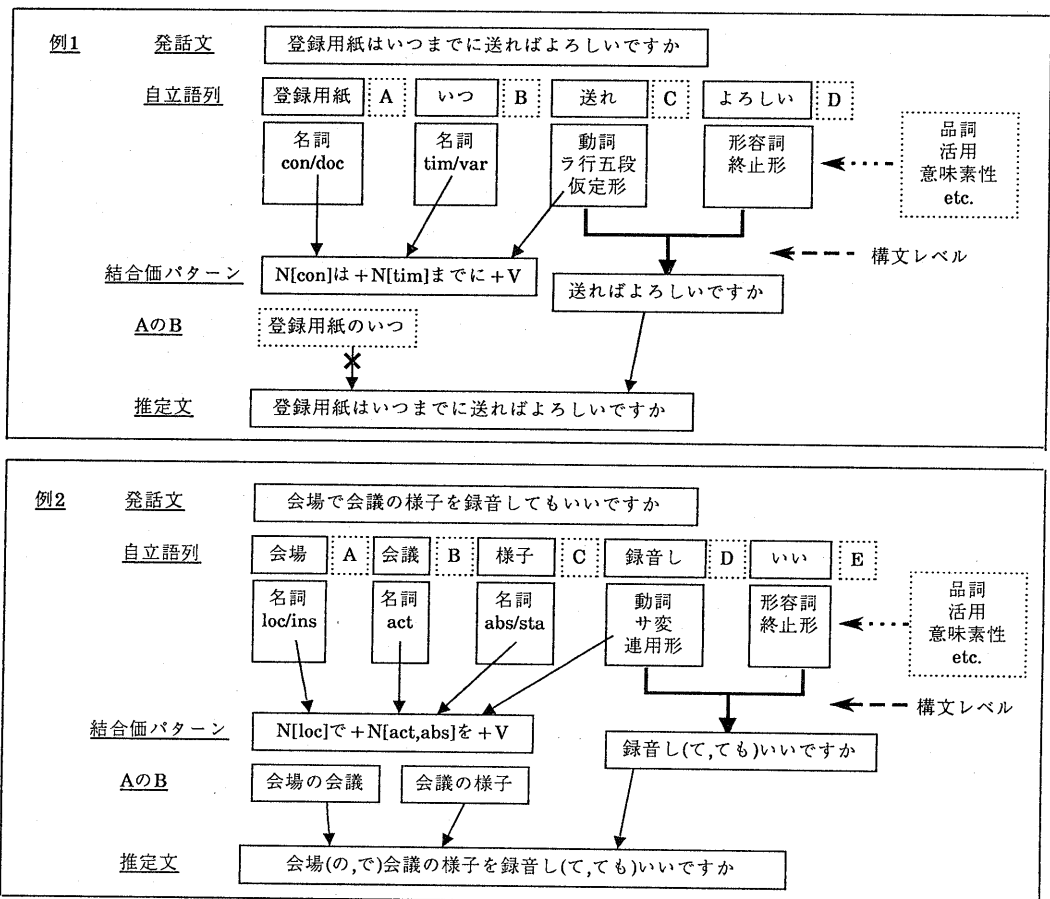


図6 付属語推定処理