

英会話教育用 ICAI システムにおける 2C-2 入力文解析のための文法規則

澤山ゆかり* 加藤正明* 山本秀樹**

* (株)沖テクノシステムズラボラトリ ** 沖電気工業(株)

1 はじめに

自然言語処理システムを構築するためには構文解析が重要であり、これまでに多くの英語のための文法規則が提案されてきている。これらの目的は主に文法的に正しい文の構造を解析し、文法的に誤った文を排除することである。しかし、文法的に正しい全ての文を受理できる文法規則を作成することは困難である。そのため、システムの目的に合わせて、受理する文の範囲をある程度限定した文法規則が提案されている[1]。例えば、高度な個別教育を目的とした ICAI システムが受理すべき文は、システムの教育の目的や対象としている教材に応じて決められる。現在、著者らは、学習者とシステムとの英語による対話をシミュレートすることによって、学習者の英語力を高めることを目的とした英会話教育用 ICAI システムを開発している[2]。本システムは、文脈自由文法で書かれた文法規則を用いて、学習者からの入力文を構文解析して解析木を作成し、解析木から意味表現を作成する。また、構文解析と同時に、学習者の入力文中の文法的な誤りを検出する。次に、意味表現を参照してシステムの次の発話内容を決定し、これを基に文を生成して出力する。さらに、検出された学習者の誤りによって学習者のレベルを測定し、そのレベルを教育的制御に使用する。本稿では、英会話教育用 ICAI システムが入力文を構文解析する際、どのような文を受理すべきか、及びどのような特徴を持った文法規則が適切かについて述べ、これらのことを考慮した文法規則について述べる。

2 受理すべき入力文及び文法規則の特徴

英会話教育用 ICAI システムは、英会話の教育が目的であり、文法書(例えば、文献[3]など)にある用法を用いた文を教材とするため、学習者の入力文として以下のような文を受理する必要がある。

- ・文法書に掲載されている文型、語の用法、慣用表現を含む文。特に、会話には感嘆文や省略文が頻出する傾向があるため、感嘆文及びある程度の省略文を含む。
- ・学習者の入力文中の文法的な誤りを検出できるように、学習者が犯し易い誤りを含む文。

また、対話であるという点と教育を行うという目的とを考慮すると、文法規則には次のような特徴が要求される。

- (1) 文がどの文型に分類されるか、文中の各語はどのような用法で使用されているか、どのような慣用表現が使用されているかを判別すること。また、意味解析などの構文解析後の処理を容易にするために、文型、用法、慣用表現の相違の大きさによって、解析木の形の異同を決めること。例えば感嘆文と間接疑問文のように、表層構造が似ていても用法が大きく異なるものは、異なった形の解析木を作るようにする。
- (2) 学習者の能力を把握するため、学習者がどのような文法誤りを犯したかを判別すること。
- (3) システムと学習者の間の会話の臨場感を高めるため、応答速度、即ち処理速度が速いこと。

(4) 係り受け関係などの構造の間違った木を生成しないこと。

上記の特徴のうち、(1)の学習者の使用した文型、語の用法、慣用表現の判別は、解析木の形とカテゴリの持つ素性によって行うことができる。(2)の誤りの判別も素性で行うことができるが、余分な語句が入っていたり、語句が抜けていたり、語句の順序が入れ替わっていたりする誤りで、正しい規則との素性の違いだけでは判別できないものがある。これらの誤りは、対応する規則を加えることにより検出できる[4]。

(3)の処理速度は、一般に、文法規則の数を減らすことで減少する。少ない文法規則で様々な文を受理しようとすれば、規則中のカテゴリの種類を少なくし、各カテゴリで表す語句の範囲は広くする必要はある。一方、(4)のように、間違った構造の木をなくすには、カテゴリを厳密に区別することによって各カテゴリの表す範囲を狭くして、厳密に木を作るという方法と、カテゴリを区別する代りに、語句の持つ素性で木の正しさを検査し、正しくない候補木を削除する方法がある。(1)の解析木の形については、厳密にカテゴリの区別を行うことによって、用法が大きく異なる部分からは異なる形の解析木を作成することができる。

そこで、あるカテゴリでどの範囲の語句を表せば、上記の要求を満たせるかを考えて、文法規則を作成する。

3 カテゴリの区別方法

3.1 文のカテゴリ

(a) 文のカテゴリ

文を表すカテゴリは、下に示すように平叙文、YN疑問文、WH疑問文、感嘆文といった文の種類毎に区別し、さらに省略文というカテゴリを設ける。文が節となる場合も、これらのカテゴリを使用する。文が節となる時の文法的な正しさは、例えば、「平叙文の形の節だけが、ある動詞の目的語に成り得る」など、節となる文の種類によって決まることが多い。省略文については、節となる場合がほとんどない。このため、省略のない文と省略文のカテゴリを区別することにより、節が省略文の形でないことを検査するための時間を短縮できる。

<文のカテゴリ>

[平叙文]	sdec → np vp
[YN疑問文]	sq → auxp np vp
[WH疑問文]	shwq → whp vp
[感嘆文]	excls → *wh *adj np *be
[省略文]	sind → np

<文が節となる場合>

[動詞句]	vp → *v sdec
[WH疑問文]	shwq → whp sq
[重文、複文]	sentence → sdec *conj sdec
[複文]	sentence → *conj sdec *comma sdec

(注: sentenceは節にならない文を表す)

(b) 複文と重文

副詞節が後ろに来る複文は、重文と同じ文法規則を使用して解析する。副詞節が後ろに来る複文と重文との相違は、表1に示すように接続詞の素性によって決まる節と節の関係にあり、節となる文の意味はほとんど変わらない。そこで、意味解析処理の行い易さを考

特性No. が指定されている。フィラー特性は、テンプレートに適したフィラー部分の構文的な形式を指定するもので、No. に対する値はフィラー特性テーブルに記述されている。

表2に、「望み(欲しいもの・こと)を言う(答える)」という言語機能を言語形式に変換するための発話データの例を示す。ここで、変数HOLEはフィラーが埋め込まれる部分を示している。また、表3にはフィラー特性テーブルの例を示す。

言語機能	対人関係	言語形式テンプレート	フィラー特性
望みを言う (依頼)	Neutral	I'd like HOLE.	1, 2, 6
		I need HOLE.	1, 2, 6
		May I HOLE?	0, 3
		What I need is HOLE.	1, 6
	Informal	I'm dying for HOLE.	2, 6
		I want HOLE.	1, 2, 6
		HOLE would go down well.	15
	Formal	I'd love HOLE	1, 2, 6
		I should very much like HOLE.	1, 2, 6
		What I should really like is HOLE.	1, 6
		I should really love HOLE.	1, 2, 4, 6
		I'm really longing HOLE.	1

表2 発話データの構造

No.	フィラー特性(=形式変換ルール)
0	HOLEなしとみなすことも可
1	+ TO-infinitive
2	+ N/pronoun + TO-infinitive
3	+ infinitive
4	-ING form
5	+ N/pronoun/possessive ADJ -ING form
6	+ N/pronoun/NP
:	:
15	+ N/pronoun/NP/-ING form/N-clause

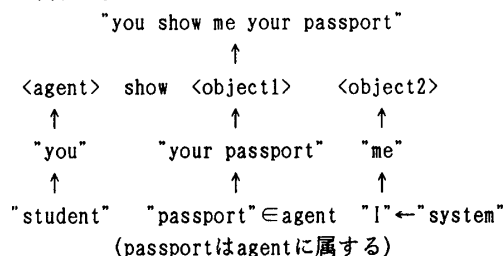
表3 フィラー特性テーブル

4. 2 文生成のプロセス

文生成のプロセスは、フィラー部分の文生成、発話データを利用した言語形式の選択、選択した言語形式(テンプレート)へのフィラーの埋め込み、文の整備という大きく4つのstepからなる。以下にその概要を示す。

- step1: 会話制御部から渡される命題内容を、フィラーとしての表層形である平叙文の形式に変換する。
- (1) 命題内容に従って動詞の選択を行う。
 - (2) 行為者に関しては、状況設定を参照して表現を決定する。
 - (3) 動詞中心の構文ルールを適用して文を生成する。例えば、テンプレート"Would you HOLE?"を用いて

文"Would you show me your passport, please?"を生成する場合、まず例1のような方法でフィラー部分に相当する文"you show me your passport"を生成する。



- 例1 "show" を中心にした構文ルールを用いた文生成の例
- step2: (1) 会話制御部から指定された言語機能に対応する言語形式の集合である発話データを得る。言語形式はテンプレートの形をしている。
- (2) 場面・状況設定により、formalityを決定する。
- (3) 制御部から言語形式(構文など)の指定があれば従い、なければ適当な言語形式のテンプレートを選択する。
- step3: step2の(3)で選択したテンプレートに、step1で生成した文をフィラーとして埋め込む。このとき、テンプレートに記述されているフィラー特性No. をキーにフィラー特性テーブルを参照し、そのNo. に対応するルールに従ってフィラーの形式を整える。
- step4: step3までで生成した文を、談話構造を考慮した自然な表現に変換する。以下に変換の種類を示す。
- (1) 旧情報に対する指示、代用、繰り返しを行う。
 - (2) 省略を行う。
 - (3) 焦点に強勢を付け、音調を整える。

5 おわりに

学習者とシステムの外国語による自由な会話のシミュレーションを通して教育的指導を行う Communicative Approach に基づいた語学訓練用知的CAIシステムの実現法を検討した。そして、文生成方式に言語機能を利用することが有効であることを述べた。

謝辞

熱心に討論に参加して下さった三重大学工学部情報工学科の椎野努教授、ならびに宮阪信次氏をはじめとする大阪ガス(株)のメンバーに深謝する。

参考文献

[1] Stratton, Florence : "Putting the Communicative Syllabus in Its Place," TESOL Quarterly, 11, 2, Jun., 131-141, 1977.

[2] 垣田編 : 英語教育学研究ハンドブック, 大修館書店, 1982.

[3] 長澤 : コミュニカティブ・アプローチとは何か その理論と展開, 三友社出版, 1988.

[4] 山本他 : 会話シミュレーションを基にした語学訓練用知的CAIシステムの構成, 情処学論, Vol. 30, No. 7, 1989.

[5] 甲斐他 : Communicative Approachに基づく知的CAIシステムの実現法(1), 第41回 情処全大, 1990.

[6] 浅野他 : 語学訓練用知的CAIシステムにおける文の生成方式, 第40回情処全大, 1990.