

用例に基づく類似文生成による 日本語学習のための対話型ドリル作成システム

塩谷教子[†] Harold L. SOMERS

Centre for Computational Linguistics
Department of Language and Linguistics
University of Manchester Institute of Science and Technology

本稿では、日本語のためのCALL(Computer Assisted Language Learning)に関する簡単な教材作成システムについて述べる。本システムは教師の入力した一例文から複数の類似文を生成する。システムは、まず教師が入力した一例文を解析する。そして、入力例文内の単語と教師が示唆する用例単語より、辞書および意味階層木(semantic feature hierarchy)から代替候補となる意味特徴(semantic feature)を類推し、類似文を生成する。意味特徴を類推する過程で動詞と主語・目的語の名詞との共起に関する意味的制限を導き、それを次回の単語抽出の過程に反映させる。さらに生成された複数文から日本語学習用ドリルを出力する。本稿では、この過程を、日本語学習システム開発を効果的に支援する一手法として提案する。

Example-Based Drill Generation for Japanese CALL

Atsuko SHIOYA[†] Harold L. SOMERS

Centre for Computational Linguistics
Department of Language and Linguistics
University of Manchester Institute of Science and Technology

We introduce a simple authoring system for Japanese Computer-Assisted Language Learning (CALL). The system generates a number of new sentences from a teacher's input sentence. According to the input sentence structure, the system extracts alternative words using methods of retrieval of example words and inferring their semantic features from the lexicon and the semantic feature hierarchy. Semantic restrictions on the subject and the object of the verb in the input are found and used for the next extraction. In addition, drills are produced according to the generated sentences. This process can provide very effective support for producing CALL courseware.

[†] 現在、株式会社プロテック

Presently with PROTECH Corporation

1. はじめに

本稿では、日本語CALL(Computer Assisted Language Learning)¹³⁾においてドリル問題生成過程で教師を支援するために作成したシステム*について述べる。本研究の目的は、ある種のコンピュータ化されたドリルの情報源となりうる自動生成文を、用例に基づく手法を用いて提供することにより、教師の仕事を支援することである。

システムは教師が入力した一文から複数の類似文を生成する。生成される文は入力文の構文に基づいて構成される。これは単に、文生成過程を支援するだけでなく、教師がドリル問題について明確なイメージを持たない場合でも、一入力文からどんなドリル問題を作成すべきかの意思決定支援の役割も提供することが出来ると考える。即ち、システムは教師の創造活動支援の役割を担うことも可能である。この役割は以下の考えに基づく。

- (1) 文生成の有効範囲は、適用される辞書が包括する範囲に依存する。従って生成文の最終極限は辞書の拡大により無限にもなりうる。
- (2) 異なったレベルでの文生成が、文法規則や辞書を調整することにより実現可能である。言語学習のための生成文の語彙の範囲やレベルは、文法や辞書により定義することができる。従って、本システムは、言語学習においてどんなレベルにも対応が可能となる。
- (3) 多くの言語学習用ドリルや練習問題は類似または同一構文から構成されている場合が多い。従って、生成された類似文を利用してさまざまなドリル問題への適用が考えられる。

2. システム構成

システムの実現には、SICStus Prologを用い、日本語文の解析には構文解析システムSAX⁹⁾を利用した。システムの構成およびその動作概要を以下に述べる。各モジュールとその機能構成は図1に示すとおりである。本稿では実線で囲まれたモジュールを現段階における成果として述べる。太波線で示された部分は現段階で簡単にその動作を担うが、今後の課題としてより充実したものを考えている。点線部分は現時点では実現されていないが、より統合的なCALLの実現に対して課題とする部分である。

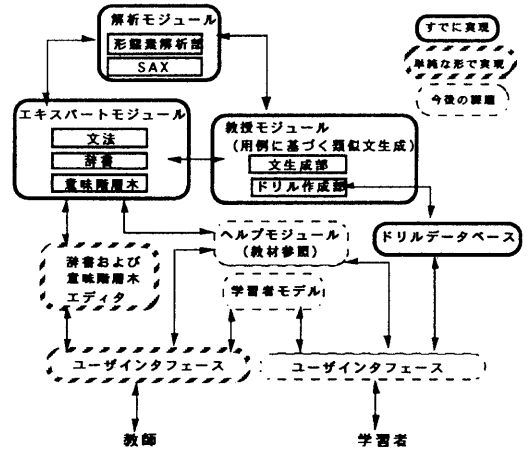


図1 システム構成

Fig.1 Configuration of the system

現段階において実現されている各モジュールの概要を以下に述べる。他のモジュールについてはシステムの拡張性として後述する。

エキスパートモジュール

本モジュールは日本語文法ルール、日本語辞書、日本語の意味特徴(semantic feature)における階層木(semantic feature hierarchy) (以下、意味階層木)といった3つの言語学上の知識から構成される。

解析モジュール

本モジュールは、形態素解析部⁹⁾およびSAXによる構文解析部から成る。

教授モジュール

本モジュールで用例に基づく類似文生成を行う。主なタスクは、教師による入力文から複数の文を生成する文生成と教師が生成文を基に日本語学習のための問題作成を行うためのドリル作成である。本稿では、ドリル作成部として穴埋め問題の作成を実現したが、これは他の種類のドリルや演習問題作成部との置き換えが可能である。

ドリルデータベース

システムによって生成されたドリルは解答と共にファイルに蓄積される。

意味階層木エディタ

現システムでは、意味階層木の修正のために意味特徴ノードの追加、削除、移動、保存といった、簡単なオペレーションを用意した⁹⁾。

3. 用例に基づく類似文生成

システムは、入力文に対して代替可能な単語を辞

*本システムの詳細はShioya⁹⁾を参照のこと。

書から抽出し、構文に従って置き換えを行うことによって複数の文を生成することができる。代替可能な単語を抽出する過程では用例に基づくアプローチを適用した。

3.1 代替可能単語の抽出

入力文に対して代替可能な単語は、その意味特徴に関する情報を用いて、辞書を検索し抽出される。抽出の過程は以下の通りである。

I) まず教師は入力した文の中でどの単語を代替可能な単語と置き換えたいかを選択する*。

II) 次に代替可能な全単語が以下の条件とそのプロセスに従って抽出される。

(1) 入力文内の動詞によって主語または目的語の名詞が共起する場合、システムはその共起の意味的制限に従って単語の抽出を行う。

(2) 名詞、形容詞または副詞の場合、次の2つのアプローチを準備した。

- システムが適切な意味特徴を自動探索し、教師との対話による確認の結果、複数の代替可能な単語を辞書から選び出す。

- 教師がいくつかの代替可能な用例単語を示すことにより、システムは単語が使われている条件での代表的な意味特徴を類推する。そして、その意味特徴から複数の代替可能単語を辞書から選び出す。この過程で、動詞に対する主語または目的語名詞に関して、共起する意味的制限が導かれたなら、その情報を知識として蓄える。

(3) 動詞の場合、入力文中の置き換えられる動詞が辞書に持つその属性により、代替可能な他の動詞を選び出す。そして、主語または目的語に対する共起の意味的制限の知識を獲得する。

以下に、(1)(2)(3)の各場合における代替可能単語の探索による抽出プロセスの詳細を述べる。

(1) 共起における意味的制限による単語抽出

入力文内で、教師が選んだ置き換え単語が、

i) 主語または目的語と解析され

かつ

ii) 入力文内の動詞がその単語に共起するという意味的制限がすでに存在する

という時、システムは代替可能単語の抽出に、その共起のための意味的制限を適用する。この時適用される意味的制限とは、動詞の抽出過程(3)動詞の抽出と共起の意味的制限の獲得を参照)

* 本稿において、「置き換え単語」とは教師が入力文内で類似文生成のために置き換えを行いたいと選択した単語とする。また、「代替可能単語」とは、置き換え単語に対して、その代替の可能な候補となり、辞書から抽出される単語とする。

において類推され、獲得される知識である。

例えば、教師の入力文が、「あなたは本を買う」であったとする。教師が「あなた」を置き換え単語に選んだとし、システムがすでに、「動詞「買う」の主語の意味特徴は[human]である」という意味的制限の知識を持っている時「あなた」に対する代替可能単語は意味特徴[human]に従って選び出される。

(2) 名詞、形容詞、副詞の抽出

意味階層木は、名詞、形容詞、副詞について用意した。従って、名詞、形容詞、副詞が代替可能な単語として抽出される時には、辞書だけでなく意味階層木も参照される。教師がこれらの品詞の単語を置き換えたい時、システムは次の2つの方法を用意している。

システムによる探索

このプロセスでは、システムは代替可能単語を、意味特徴に関する情報とユーザである教師との対話から検索する。まず、入力文内の置き換え単語の意味特徴を辞書から得る。そして、代替可能案としての単語探索のために、意味階層木を教師とシステムとの対話によってたどり適切な意味特徴を取得する。代替可能単語の適切な意味特徴として、意味階層木での入力単語の意味特徴の親ノードが選び出される。その親ノードの意味特徴が入力単語の代替案として当てはまるかどうかを教師に尋ねる。もし、当てはまらなければ、同階層である兄弟ノードについて当てはまるかを訪ねる。対話により当てはまると判断されたノードの意味特徴に従って、辞書内から全単語が検索される。この時、代替案としてあてはまる意味特徴のノードは複数でも構わない。

「人は本を読む」という文が入力されその中で「人」を置き換え単語として選んだ時の探索例を図2に基づいて示す。「人」は[human]という意味特徴を持つので、システムはその親である[animate]を意味特徴にもつ全単語が「人」に対して代替可能であるかを教師に尋ねる。親ノードが適切でなければ、兄弟ノードを順にたどり探索する。

この時、教師は意味階層木や意味特徴を特に意識する必要はない。単に「単語「人」が全ての[animate]単語で置き換えが可能であるか」といった問い合わせが行われる。

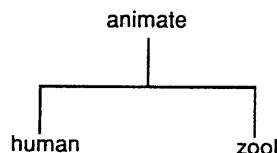


図2 「人」の置き換え過程における意味階層木の該当部分

Fig. 2 The related partial semantic feature hierarchy

探索結果として意味特徴[human]に絞り込まれ、例えば「学生」「少年」「先生」といった単語が「人」に対する代替案として抽出される。

用例に基づく探索

システムによる探索の他に、教師が候補としていくつか代替可能な単語の例を示唆することにより探索を行うこともできる。教師が例として挙げた候補を基に、システムは意味階層木を検索し、代替可能単語の代表的な意味特徴を導き出す。そしてその意味特徴に従って、辞書から複数の代替可能な単語を引いてくる。教師が示唆した単語の例（以下、代替用例）と入力文内の置き換え単語（以下、入力単語）との意味特徴の関係による探索プロセスを以下に述べる。

(a) 教師が示唆した代替用例の意味特徴と入力単語の意味特徴が同じ場合

その意味特徴をそのまま代替可能単語を代表する意味特徴として抽出を行う。例えば、教師が入力した文が「犬は食べる」であり、教師は「犬」を置き換えて文を生成したいとする。教師は、「犬」の代替として「ねこ」と「うさぎ」を挙げたとすると、システムはこれらの単語の意味特徴が全て[zool]であることから、これを代表的な代替可能単語の意味特徴とみなす（図3）。

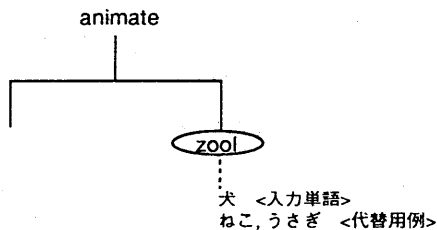


図3 入力単語と代替用例が同じ意味特徴を持つ場合

Fig. 3 Input and examples have same semantic feature

(b) 代替用例の意味特徴が入力単語の意味特徴の親ノードである場合

代替用例の意味特徴である親ノードが代表的なものとなされる。例えば、入力文が「犬は食べる」で、教師は「犬」の代替用例として「動物」「獣」といった単語を示唆したとすると、代表となる意味特徴は[animate]とされる（図4）。

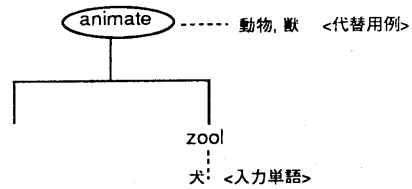


図4 代替用例と入力単語の意味特徴が親子関係の場合
Fig. 4 Examples and Input's semantic features are mother and daughter

(c) 代替用例の意味特徴は全て等しいが、入力単語の意味特徴が用例単語の意味特徴の親ノードとなっている場合

システムはどちらの意味特徴が、より限定された代替可能単語の代表的な意味特徴を示しているかを、教師との対話により導き出す。例えば、入力文が「動物は食べる」で「動物」に対する代替用例の単語に「犬」「ねこ」が挙げられた時、システムは「犬」「ねこ」の意味特徴である[zool]が代替可能単語の代表的な意味特徴となるかを教師に尋ねる。もし"yes"が返されたなら、代替用例の意味特徴が代表とみなされる（図5）。しかしもし"no"であるなら、システムは入力単語と代替用例の共通の親ノードで最も限定されたものを採り出す。

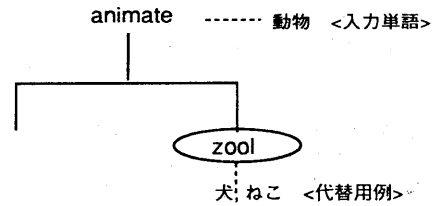


図5 入力単語と代替用例の意味特徴が親子関係の場合

Fig. 5 Example and Input's features are daughter

(d) 代替用例の意味特徴が全て同じであるが、入力単語と代替用例との間に直接の親子関係がない場合

(d) 代替用例の意味特徴が全て同じであるが、入力単語と代替用例との間に直接の親子関係がない場合

代替用例の意味特徴である親ノードが代表的なものとなされる。例えば、入力文が「私は家を買った」であり、置き換え単語を「家」として、それに対して教師が与える代替用例を「草」「本」とする。システムは、意味階層木を検索し、最も限定された両者の共通の親ノードとして[inanimate]を見つけ出す（図6）。

システムは両者に共通で最も限定された親(祖先)ノードを探し、その意味特徴が代表となるかどうかを教師に尋ねる。例えば、入力文が「私は家を買った」であり、置き換え単語を「家」として、それに対して教師が与える代替用例を「草」「本」とする。システムは、意味階層木を検索し、最も限定された両者の共通の親ノードとして[inanimate]を見つけ出す（図6）。

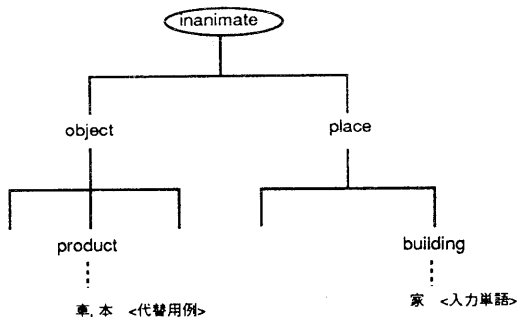


図6 入力単語と代替用例の意味特徴が直接の親子関係にない場合

Fig. 6 Examples and input's semantic features are not mother-daughter

(c)いくつか挙げた代替用例の意味特徴が異なる場合
システムは、入力単語も含め各単語の意味特徴を調べ、それらを2つのグループに分類して、その2つのうちの最も限定された親ノードを導く。そして、導き出されたものが適切かどうかを、教師との対話により確定する。例えば、入力文が「私は車を買った」であり、「車」に対する代替用例が「犬」「本」「ねこ」と挙げられた時、共通の祖先であり最も限定された親ノード[concrete]が導き出される(図7)。

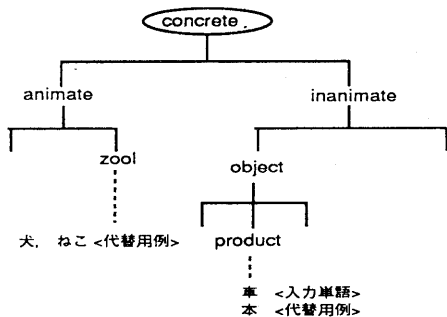


図7 入力単語と代替用例の意味特徴が分散している場合

Fig. 7 Examples and input's semantic features are mixed

(f) その他の場合

現段階ではシステムは代表的な意味特徴を導くことができない。この場合は、教師への代替用例の問い合わせフェーズに戻り、代替用例となる単語を入力し直す。

概してシステムは、入力単語と代替用例の共通で最も限定された意味特徴を導き出し、それを代表的な意味特徴として単語の検索を行うことを基本的なアプローチとしている。これは、意味階層木内で入力単語と代替用例との、最短距離を求めることである。この時、代表的な意味特徴が見つからない場合、

意味階層木を修正することが可能である。従って、教師は独自の教育戦略に則した階層木を構築することもできると考えている。現システムでは、階層木内の意味特徴ノードの追加、削除、移動、保存といった簡単なオペレーションを用意した。

さらに、このプロセスにより導かれた意味特徴は、入力文の述語である動詞の主語・目的語の名詞の共起における意味特徴の情報ともなり得る。入力文内の動詞に関する情報と文の構文解析の結果から、動詞が共起する主語・目的語名詞の意味的制限が類推され、その動詞の制限知識として、システム内に登録される。そしてその情報は、次回の名詞の抽出および動詞の抽出時で動詞の共起に関する意味的制限を決定する段階に利用される。

用例に基づく探索による抽出のプロセスの例を以下に示す。入力文は「かれは本を読みました」である。「わたし」「先生」が「かれ」に対する代替用例として与えられる。システムは、構文解析の結果、「かれ」を入力文の主語とみなす。入力文内の置き換え単語「かれ」と代替用例「わたし」「先生」の代表的な意味特徴が[human]と導き出され、その意味特徴は入力文の述語動詞である「読む」が共起する主語の意味特徴であることが類推される。そして、次回「読む」の主語となる単語が抽出される過程で、この意味特徴の情報が利用される。

また、置き換え単語が「本」である時には、その代替用例として、「てがみ」が挙げられたので、代表される意味特徴[product]が、「読む」の共起する目的語の意味特徴であると類推される。このプロセスで、これらは「読む」の共起における意味的制限知識として登録される。

```

Input a sentence:かれは 本を 読みました*
The input sentence is being analysed.....

[v(part(n(かれ), は), part(n(本), を), 読む)]

How many drill sentences do you need?: 3
Generating new sentences.....

Which words in your input sentence can be replaced?
The input words in dictionary forms are :
1: かれ 2: は 3: 本 4: を 5: 読む
Choose numbers (Type no to finish): 1 3
Any more?
(If you do not need any more alternative words, type no)
1: _no

Choose method of alternative word search for *かれ*
1 System seeks alternatives
2 User gives examples for inference of restrictions
3 Next word
4 Help
Choose a number : 2

Which words can be alternative for かれ*? :わたし 先生
[human] is inferred as a semantic restriction on the subject of the verb *読む*
読む subject = [human] save? (y/n) Y
This restriction will be saved in the [sentree/semres_subobj.dat] file? (y/n) Y

```

*下線は教師の入力部分を示す。

```
Choose method of alternative word search for *本*
  1 System seeks alternatives
  2 User gives examples for inference of
restrictions
  3 Next word
  4 Help
Choose a number : 2
```

```
Which words can be alternative for *本*? : てがみ
[product] is inferred as a semantic restriction on the
object of the verb *読む*
```

```
読む object = [product]save? (y/n) Y
This restriction will be saved in the
[semtree/semres_subobj.dat] file? (y/n) Y
```

```
3 sentences are generated.
1: 先生 は 本 を 読みました
2: 少年 は 本 を 読みました
3: 彼女たち は てがみ を 読みました
```

システムは新しい文の生成に用例検索を行うが、これは用例に基づく機械翻訳⁷⁾において原語文と用例との間の意味の類似度または距離を計算するということの代わりに、教師との対話によって入力文内の単語と代替用例の単語との類似度を導き出すことであると言える。

(3) 動詞の抽出と共起の意味的制限の獲得

動詞の代替可能単語は、その属性に基づいて辞書から抽出する。また、入力文内の動詞が持つ共起における意味的制限知識に従って、代替可能の候補となる動詞の制限知識も類推する。この過程は次の2つの手順から構成される。

まず、入力文内の動詞を辞書から引き、辞書内の属性を基に検索して属性がマッチした他の動詞を抽出する。この属性は、動作動詞か状態動詞か、移動を表すか否か、他動詞か自動詞かといった情報から構成される。抽出される動詞は入力文内の動詞と同じ属性を持ち、生成される文に対しても同じ性質を持つことが要求される。従って、これらの動詞は入力文内の代替可能動詞とみなすことができる。例えば、「見る」という動詞は、動作動詞で移動を表さず、自動詞である。同じ属性を持つ動詞には「待つ」「書く」などがある。

次に、システムは入力文内の動詞の持つ共起における意味的制限知識もチェックする。そして、その制限知識は抽出された他の動詞に対しても適用できるかもしれない。ここでシステムは、抽出された代替可能動詞に対しても同じ制限知識を持つことができるかどうかを対話的に教師に確認する。持つのであれば次回の抽出過程での利用のため、システム内に登録する。

また抽出された代替可能動詞の活用は文生成過程で、入力文内の動詞の活用に従う。

入力文「わたしはてがみを書きました」に対する代替可能単語の抽出過程の例を以下に示す。教師が「わたし」と「書く」を置き換え単語として選ぶ。「わたし」に対する代替用例として「あなた」と「かれ」を挙げると、動詞「書く」が共起する主語の意

味特徴は[human]であることをシステムは類推する。動詞「書く」に対する代替可能単語には、その辞書内の属性により「待つ」「あげる」「見る」が抽出される。これらの動詞に対しても「書く」と同じように主語に対して [human] という共起における意味的制限を持つと推論できる。この例では、次の3つの文が生成され⁸⁾、動詞も入力文内の動詞と同じように活用されて、過去、肯定、「ですます」形で生成される。

```
Input a sentence: わたしは てがみ を 書きました
The input sentence is being analysed.....
```

```
[v(part(nわたし), は), part(n(てがみ), を), 書く)]
```

```
How many drill sentences do you need?: 3
Generating new sentences.....
```

```
Which words in your input sentence can be replaced?
```

```
The input words in dictionary forms are :
```

```
1: わたし 2: は 3: てがみ 4: を 5: 書く
```

```
Choose numbers (Type no to finish): 1 5
```

```
Any more?
```

```
(If you do not need any more alternative
words, type no)
```

```
1: no
```

```
Choose method of alternative word search for
*わたし*
```

```
1 System seeks alternatives
  2 User gives examples for inference of
restrictions
  3 Next word
  4 Help
Choose a number : 2
```

```
Which words can be alternative for わたし*? :あなた
かれ
```

```
[human] is inferred as a semantic restriction on the
subject of the verb *書く*
```

```
書く subject = [human]save? (y/n) Y
This restriction will be saved in the
[semtree/semres_subobj.dat] file?
save? (y/n) Y
```

```
The word *書く* has a restriction on the subject
[human]
```

```
The word *あげる* was found as an alternative for
*書く*
```

```
Must the subject of *あげる* be [human]? (y/n) Y
```

```
あげる subject = [human] save? (y/n) Y
```

```
This restriction will be saved in the
```

```
[semtree/semres_subobj.dat] file?
save? (y/n) Y
```

```
The word *待つ* was found as an alternative for *書く*
```

```
Must the subject of *待つ* be [human]? (y/n) Y
```

```
待つ subject = [human]
```

```
save? (y/n) Y
```

```
This restriction will be saved in the
```

```
[semtree/semres_subobj.dat] file?
save? (y/n) Y
```

```
The word *見る* was found as an alternative for *書く*
```

```
Must the subject of *見る* be [human]? (y/n) Y
```

```
見る subject = [human]save? (y/n) Y
```

```
This restriction will be saved in the
```

```
[semtree/semres_subobj.dat] file?
save? (y/n) Y
```

```
3 sentences are generated.
```

```
1: かのじよらは てがみ を 待ちました
```

```
2: あなた は てがみ を あげました
```

⁸⁾本稿は、簡略化のため生成文を3つ要求する例を示している。実際は、教師が何十、何百の生成文を要求すると考えている。

3: かれ は てがみ を 見ました

3.2 類似文生成

入力文内の単語は、抽出された代替可能単語の候補に置き換えられ、その組み合わせにより多数の類似文が生成される。教師は、いくつ新たに生成文が必要であるかを、システムに対話的に要求する。辞書の中の代替可能である全候補から、要求された数だけ任意に選ばれ、入力文の構文に基づいて置換されて新しい文が構成される。新しい文を構成する代替可能単語の動詞、助動詞の活用は、その置換対象である入力文内の置き換え単語の形態素処理で得られた情報に従う。生成された文は教師との対話により、意味的に妥当なものおよび教師にとって必要なもののみが残され、ドリル生成の対象となる。

4. 日本語学習のためのドリル作成への応用

生成された文の構文で、各単語をギャップに置き換えることにより、穴埋め問題用のドリルを作成する。教師は、対話的に生成文中の品詞もしくは直接単語を指定してドリル作成を行う。品詞が指定された場合、教師はシステムが示す生成文内の解析された品詞の中からギャップ対象としたい品詞を選ぶことができる。また、ドリル全文をその品詞に対してのギャップにするか、各文ごとにギャップ対象の品詞を選ぶかといったことも、対話的に選択することができる。ドリルは選ばれた品詞の単語をギャップにすることにより作成される。一方、各文ごとに品詞または単語を指定して、ドリル作成を行う場合、教師は生成文ごとにギャップ対象とする品詞または単語を選ぶことができる。作成されたドリルはギャップにされた単語をドリルの解答として保管する。

5. システムの拡張性

所期の目的である言語学習システム作成支援のための文生成およびドリル作成の基本的な部分は実現することができた。しかし、広い意味でのCALL(コンピュータ支援言語学習)に対して、システムの拡張、改良が期待できると考えている。以下にシステムの拡張の可能性について述べる。

5.1 構造的な拡張

本システムはCALLに利用できるデータを生産するが、CALLの直接的な実現までは到達していない。本システムの最終的な設計イメージは、システムの単純な出力結果を柔軟に利用し、教師の意思決定も支援するCALL作成パッケージの理想像と考えている。次に挙げるモジュールは、本システムの最終イメージの中で実現すべき点である。

辞書、意味階層木作成モジュール

辞書および意味階層木はシステムのプロセス

中更新できるモジュールが必要である。本モジュールの実現は、教師にとってのCALL拡張に大いに助けとなると考えている。

ヘルプモジュール

教師がドリル作成する上で参照でき、また学習者が学習する上でも参考にできるような、言語学上の情報の提供は期待される場所である。モジュールの実現は電子化辞書および電子化テキストなどによるインタラクティブなアクセスを可能とすること、さらにはマルチメディア化なども考えられる。

学習者モデル

教師が、本システムから知的CALL¹⁾作成を期待するのであれば、学習者モデルが必要となる。それはシステムプロセス中に学習の知識がその段階に従って保管され、各学習者の不正解であった状況も含めた学習履歴が蓄えられること、またそれが学習過程に反映できることであると考える。

5.2 他の練習問題作成システムとしての拡張

教師モジュールのタスクは、他種の練習問題を提供するタスクと置き換えることによって、その拡張が期待できる。特にドリル作成部は、生成文を基にした他のタイプのドリルや練習問題への置換が容易である。なぜなら、言語学習ドリルや練習問題の多くは、同一または類似構文のさまざまな文に従って作成されているからである。従って、本システムの文生成部は他種の練習問題にとっても、基になる文を提供していると考える。

例えば、システムは動詞および助動詞の活用を認識処理している。これを利用して動詞の活用練習に関する問題作成もできる。特に日本語は、肯定または否定、ですます形などにより、その動詞活用が異なるので、動詞の語形変化を学習する上で、この種の練習問題は効果的な教材となろう。

5.3 他の言語への拡張

SAX解析器は、文脈自由文法で記述できる言語において解析が可能である。システムで取り扱っている意味階層木もまた、特定の言語に依存していない。従って、エキスパートモジュールは他言語への置き換えができる。また、システムでの代替用例から意味特徴を類推することによって、単語を置き換え、類似文を作成するという方法も、他言語に容易に適用できると考える。文生成およびドリル作成は、構文解析によって得られた文法構造に基づいているため、教師モジュールもまた特定の言語によるものではない。

従って、システムは言語に依存する文法、辞書および意味階層木からなるエキスパートモジュールの部分の置き換えることで、他言語への拡張が容易で

あると言える。

6. おわりに

日本語学習システム開発支援として、用例に基づく類似文生成の手法を提案し、簡単なドリル問題生成システムに応用した。最後に本システムでの要点をまとめる。システムは、教師が入力した1つの日本語文を解析し、その解析結果である構文に基づいて類似文を生成して、日本語学習のためのドリル作成を行う。計算言語学的視点からすると、システムの主要な点は用例から意味特徴を類推することによって、新たな類似文を生成する点であると考えられる。

本研究で挙げる文生成の用例方式では、入力と用例との間の最小距離を、機械翻訳の用例方式における差異計算を行う代わりに、教師の対話による類推によって求めている。代替候補の他語の置き換えの可能性は意味特徴の関係を考慮する上での対話処理によって導かれる。また、用例は教師が候補となるものを与えるため、そこですでにある程度の絞り込みがなされていることになり、無駄な検索は行われない。さらに、システムは入力文内の動詞と主語・目的語の名詞の共起における意味的制限も獲得し、文生成の際にその情報を利用する。

教育へのコンピュータ利用の視点からの主張点は、モジュール化された拡張性を持つCALL実現の可能性を示した点である。現段階では、主要ないくつかのモジュールの基本的な動作を実現した。これらのモジュールは、その拡張によりさらに教師の仕事を創造的な活動として支援することができると考えている。

謝辞

本稿をまとめるにあたり、貴重なコメントをいただきました三重大学の椎野勢教授に感謝いたします。また、寄稿および発表の機会を与えてくださったプロテック社長の久保江氏に感謝いたします。

参考文献

- 1) Jeremy Fox, Bev Labbett, Clive Matthews, Carmit Romano-Hvid and John Schostak: *New Perspectives in Modern Language Learning*, Report OL147, Learning Methods Branch, Employment Department, Sheffield (1992).
- 2) Mary-Louise Craven, Roberta Sinyor and Dana Paramskas(eds): *CALL: Papers and Reports*, La Jolla, CA: Athelstan Publications (1990).
- 3) A.Shioya: *Example-Based Drill Generation for Japanese CALL*. MSc dissertation, Department of Language and Linguistics, UMIST (1993).
- 4) 松本裕治,伝康晴,宇津呂武仁: 構文解析システム SAX 使用説明書 version 2.0, 京都大学工学部 長尾研究室, 奈良先端科学技術大学院大学 松本研究室 (1993).
- 5) Annie Gal, Guy Lalalme, Patrick Saint-Dizier and Harold Somers: *Prolog for Natural Language Processing*, Chichester: John Wiley (1991).
- 6) G.A.Gupta: *Semantic features for a dependency grammar parser of German*. MSc dissertation, Department of Language and Linguistics, UMIST (1992).
- 7) Harold L. Somers. *Example-based and Corpus-based Approaches to Machine Translation*, Proceedings of the International Symposium on Natural Language Understanding and AI(NLU & AI)(Iizuka, Japan),pp.87-101 (1992).
- 8) Makoto Nagao: *A framework of a mechanical translation between Japanese and English by analogy principle*. In A. Elithorn and R. Banerji(eds.) *Artificial and Human Intelligence*, Amsterdam: North-Holland. pp.173-180 (1984).
- 9) Everett F. Bleiler: *Masic Japanese Grammar*, Tokyo: Charles E. Tuttle (1967).
- 10) Teiji Furugori and Akiko Takeda: *An Example-based System of writing English Sentences for Japanese English Users*, *Literary and Linguistic Computing* 8, pp.85-90 (1993).
- 11) 池田尚志, 兵頭安昭: テキストベースを中心とした支援型の自然言語システムについて, 電子情報通信学会技術研究報告 Vol.91 No.297(NLC 91 27-36), pp.33-38 (1991).
- 12) 大深悦子, 坂谷内勝, 吉岡亮衛, 藤田正春, アルド トリニ, 及川昭文: 日本語教育支援システムの開発, 情報処理学会研究報告, Vol.91 No.72(CH-10) pp.1-8 (1991).
- 13) 高木清, 吉岡亮衛, 坂谷内勝, 及川昭文: 日本語教育・学習支援システムの機能構成とその操作性について, 情報処理学会研究報告, Vol.92 No.72(CH-15) pp.33-40 (1992).
- 14) 林敏浩, 矢野米雄: 外国人のための漢字学習 C A I システム - 部分構造に着目した漢字辞書の構築, 電子情報通信学会技術研究報告, Vol.92 No.23(AI 92 1-20), pp.37-44 (1992).