正用コーパスを利用した日本語文法練習問題作成支援*

Computer-assisted Authoring of Japanese Grammar Exercises utilizing Corpora of Correct Japanese Usages

黒川 和也，宇津呂 武仁
Kazuya Kurokawa and Takehito Utsuro

豊橋技術科学大学 工学部 情報工学系
Dept. Information and Computer Sciences, Toyohashi University of Technology, JAPAN

あらまし 本論文では，計算機上の自然言語コーパス (正用コーパス) を利用して，日本語学習者のための教材開発を支援する方法を提案する．特に，日本語の文法習得において重要な役割を持つ機能表現に焦点をあて，機能表現の練習問題を，日本語コーパスから半自動的に作成する方法を提案する．

キーワード 日本語学習，文法問題，機能表現，語学教材，コーパス，用例検索

1 はじめに

本研究では，計算機上の自然言語コーパスを利用して，日本語学習者のための教材開発を支援する方法についての研究を行い，とくに，日本語の文法習得において重要な役割を持つ機能表現の練習問題を，日本語コーパスから半自動的に生成する方法についての研究を行う．具体的には，日本語能力試験 [日本国際教育協会 94] で出題の対象となる接続・修飾・文末などの機能表現を対象とする [黒川 01]．本研究では，計算機を用いた日本語文法学習支援の過程を，

1. 日本語機能表現の分類体系の構築，分類体系を用いた練習問題作成支援の研究
2. 学習者の誤り分析，および，分類体系を用いた練習問題作成支援技術の運用の統合による，実用的練習問題作成支援システム・機能表現学習用 CALL システムの研究

に分けてとらえ，特に本研究では 1 に焦点をあてる．

2 日本語機能表現の分類体系の構築

機能表現練習問題を作成する際は，機能表現の意味分類だけでなく，意味は異なるが統語的接続形式がほぼ同等の機能表現の組が必要となる．そこで，日本語機能表現に対して，意味による分類と統語的接続パターンによる分類を行う．

2.1 階層的意味分類体系

本稿では，[森田 89] の複合辞の階層的分類 (助詞と同様の働きをする表現として 273 種類，助動詞と同様の働きをする表現として 150 種類) を参考にして分類を行う．表 1 に [森田 89] の分類の抜粋を示す．表 1 に示すように，機能表現は，助詞と同様の働きをするものと，助動詞の働きをするものとに大きく分類され，助詞と同様の働きをするものの中で，格助詞タイプや，係助詞タイプのように文法機能によって細かく分類される．さらに，それぞれのタイプの中で，資格・立場と対象・関連のように意味による分類がなされている．

2.2 統語的接続パターンによる分類

統語的接続パターンを用いた機能表現の分類としては，以下の 3 種類を考える．

- 文末型動詞接続 · · · 終助詞タイプ，助動詞タイプに限定する．
  例： 終助詞タイプ · · · する ものだ
  助動詞タイプ · · · する ことはいけない
- 非文末型動詞接続
  例： 格助詞タイプ · · · する にあたって
  係助詞タイプ · · · する となると
  副助詞タイプ · · · する ことは
  接続助詞タイプ · · · する や否や
- 名詞接続
  例： 格助詞タイプ · · · 私 として
  係助詞タイプ · · · 私 といえども
  副助詞タイプ · · · 私 という
  接続助詞タイプ · · · 私 とすれば

3 機能表現の組の分類

次に，機能表現の階層的分類体系をもちいて練習問題を作成するために，機能表現の組を以下の 4 タイプに分類した．

- タイプ 1： 置き替えによって文の意味が不変の組
  例：病期を治すために静養し (なければならない/しなくてはだめだ)．
  この例の場合は，“なければならない” を，他の機能表現 “しなくてはだめだ” などに置き換えても文の意味が変わらない．

*本研究は，一部，文部科学省 科学研究費 特定領域研究 (A)「高等教育改革に資するマルチメディアの高度利用に関する研究」研究課題「音声言語処理技術と学習者モデルを用いた語学学習システムの研究」(2000∼2002 年度，研究代表者: 豊橋技術科学大学 工学部 中川 聖一 教授) の援助を受けている．

表 1: 機能表現の分類 (機能表現例)

助詞の働きをする表現 (273 種類)				
格助詞タイプ	係助詞タイプ	副助詞タイプ	接続助詞タイプ	終助詞タイプ
資格･立場 (をもって , でもって , として , etc.), 対象･関連 (について , にかんして , etc.) , 全 9 分類 , 63 種類	定義 (とは , というのは , といったら , etc.) , 主題化 (といえば , といって , etc.) , 全 2 分類 , 30 種類	強調 (という , といった , etc.) , 限定・非限定 (に限って , にかぎり ,etc.) , 全 6 分類 , 28 種類	同時性 (や否や , が早いが , そばから , etc.) , 仮定 (としても , にしても , にしたって , etc.) , 全 10 分類 , 96 種類	伝聞 (とのこと , ということ , etc.) , 願望・勧誘 (かな , ないかな , etc.) , 全 9 分類 , 46 種類

助動詞の働きをする表現 (150 種類)
義務・当然 (なければならない , なくてはだめだ) , アスペクト (ずにいる , てしまう , ずにしまう) 全 13 分類

- タイプ 2 : 置き替えによって文の意味が変わるが日本語として許容できる組

  例 : 病期を治すために静養 (しなければならない/するわけではない) .

  この例の場合は , “なければならない” を , 他の機能表現 “わけではない” と置き換えた場合 , 日本語として許容できるが , 文の意味が変化する .
- タイプ 3 : 置き替えによって日本語として許容できなくなる組

  例: 先生すみません . 明日 , 父の友人が日本に来るんです . その人は日本語が全然分からないので (迎えに行かなければならない/*迎えに行くに違いない) んです .

  この例の場合は , “なければならない” という機能表現を , “に違いない” に置き換えると , 日本語として許容できなくなってしまう .
- タイプ 4 : タイプ 1~3 以外で , 言い替えによって , 練習問題の素材にすることが適当でない組
  このタイプについては , 作業過程でタイプ 1~3 以外の例があれば認定する .

4 機能表現の組の分類を利用した練習問題作成

前節の機能表現の組の分類を用いれば , タイプ 1 とタイプ 2 の違いを利用した同義表現選択問題と , 正用とタイプ 3 の違いを利用した空欄補充問題の二種類の形式の練習問題が作成できる . [黒川 01] では , 空欄補充問題のみ扱ったが , 問題文の文脈によって , 誤用の選択肢でも文として意味が通じてしまうという問題点があった . そのために , 問題作成の対象となる機能表現が限定されていたが , 同義表現選択問題を用いれば , 問題作成の対象となる機能表現の範囲が広がる . 以下にそれぞれについて述べる .

4.1 タイプ 1 とタイプ 2 の違いを利用した練習問題の形式

問題例 **1** 病期を治すために 静養しなければならない .

1. 静養しなくてはだめだ
2. 静養するわけではない

上記の問題例 1 は , 問題文中の下線部 “静養しなければならない” と同義となるものを選択肢から選択する問題である . この問題では , 問題文の “なければならない” とタイプ 1 の関係がある “なくてはだめだ” を正解選択肢として , タイプ 2 の関係にある “わけではない” を誤り選択肢として用いている .

4.2 正用とタイプ 3 の違いを利用した練習問題の形式

問題例 **2** 先生すみません . 明日 , 父の友人が日本に来るんです . その人は日本語が全然分からないので ( ) んです .

1. 迎えに行かなければならない
2. 迎えに行くに違いない

上記の問題例 2 は , 空欄に適当な表現を選択肢から選択する問題である . 正解選択肢については , 元々の文の機能表現 “なければならない” をそのまま用いて , 誤りの選択肢は , 問題文の機能表現とタイプ 3 の関係にある機能表現 “にちがいない” を用いている .

5 機能表現の組のタイプの認定

前節までの方法によって , 実際に練習問題を作成するためには , 各機能表現の組がどのタイプに該当するのかを認定する必要がある . ここで , タイプ 1 に関しては , 階層的分類体系中の最小グループ (例えば , 表 1 において , 助詞と同格の働きをするグループでは , 格助詞タイプ , 資格・立場を表す “をもって” , “でもって” , “として” のようなグループ) の範囲に限定して行い , タイプ 2 , 3 に関しては , 統語的接続パターンが同一のグループ内に限定して行う . また , タイプ 2 , 3 は , (a) 文脈独立にタイプ 2 , あるいは , 3 が成り立つ組 , と , (b) 文脈に依存してタイプ 2 , あるいは , 3 が成り立つ場合がある組 , とに分けて文脈の影響を考慮した二段階の認定をする . こうすることによって , (a) は文脈による影響が無いので全自動問題作成用として , (b) は , 文脈による影響があるので半自動問題作成用として用いることが出来る .

6 おわりに

本稿では , 日本語機能表現練習問題作成支援に必要な日本語機能表現の分類体系を構築する方法 , 日本語機能表現の分類体系を用いた機能表現練習問題作成の方法について述べた .

参考文献

[黒川 01] 黒川和也, 宇津呂武仁 : コーパスを利用した日本語機能表現学習教材作成支援, 言語処理学会第 7 回年次大会論文集, pp. 367–370, 言語処理学会 (2001).

[森田 89] 森田良行, 松木正恵 : 日本語表現文型, NAFL 選書, 第 5 巻, アルク (1989).

[日本国際教育協会 94] 日本国際教育協会 : 日本語能力試験出題基準, 凡人社 (1994).