

日本語並立表現の句構造表示の提案

— 句構造表示コーパスによる日本語学習支援のために —

Proposal of Phrase Structural View of Japanese Junctive Expressions
--- for Use in Assisting Learning Japanese by Phrase Structured Corpus ---

中村 宏
Hiroshi NAKAMURA
東京通信大学
Tokyo Online
University

掛川 淳一
Jun'ichi KAKEGAWA
兵庫教育大学
Hyogo University of
Teacher Education

伊藤 紘二
Kohji ITOH
東京理科大学
Tokyo University
of Science

<あらまし> 意味制約関係を表わす階層的句構造の作り方の獲得を支援する、第二言語としての日本語学習支援において、これまで構築してきた構文解析器の解析結果を基に作成した入れ子構造の表示に加えて、今回日本語の並立構造の作成規則から導いた「並立表現」の構造表示について提案する。学習させたい言語の表現型のそれぞれについて、文脈と意図において、典型的に類似あるいは異なる複数の使用事例が含まれるテキスト群の句構造表示をコンテンツとして用意し、それを材料として、再現的な生成と理解の課題を課し、その解決をコンテンツにおける表現型使用事例の検索による句構造表示の比較によって支援する過程において、表現の使い分けを獲得させる。句構造表示として、学習対象言語の表示と並べて、対訳された媒介語のテキストについても並立句を含む階層的句構造を示すことで、意味的な対比を支援する。

<キーワード> 第二言語学習支援システム 日本語教育 句構造表示 表現獲得 コーパス

1. はじめに

言語の獲得は、音声を含めて理解と生成のいずれについても、意味制約関係を表わす階層的句構造の作り方を習得することで行なわれるという仮説に基づき、句構造表示に導かれた第二言語学習支援の環境を提案してきた(1)(2)(3)。本稿では構文解析器の解析結果を基に作成した入れ子構造の表示に加えて、日本語の並立構造の作成規則から導いた「並立表現」の構造表示について提案する。また学習言語の表現型について、文脈と意図においてそれぞれ複数の使用事例が含まれるテキスト群の句構造表示をコンテンツとして用意し、それを基に再現的な生成と理解の課題を課し、その解決を表現型使用事例の句構造表示の比較によって表現の使い分けを獲得させる課題解決学習支援について述べる。

2. 階層的句構造

句に主辞が置かれ、それを修飾、説明する複数の句を主辞の前あるいは後に置いて句を作るこ

とを階層的に繰り返すことで、あらゆる言語における文は作られている(4)(5)(6)。主辞には、かかれて完結する主辞と、かかる句とかかられる主辞の間の関係を指示する機能的な主辞がある。また、語の単位が複合したものを主辞とすることも多い。句構造を表示するには、主辞を載せたパネルの上にそれにかかる句のパネルを載せることを階層的に繰り返す(2)(3)。

日本語の場合には、日本語係り受け解析器 CaboCha(7)による解析の仕方を基に、Japanese Phrase Structure Grammar(4)も参照して、自己完結的な主辞を「head 形態素複合」、結合機能的な主辞を「functional 形態素複合」として、二者の接続からなる chunk を句構造表現の単位とみなす。機能的な主辞が空で、完結的主辞の活用でかかりの機能を構成することもある。構造を視覚的にパネルで表わす方法としては、かかられる完結的主辞を載せた chunk パネルの前の方に、かかるほうの chunk を並べて乗せて表現する。機能的な主辞は、完結的主辞を載せたパネルの後ろの部分に付加する形で示す。

英語では、Stanford Parser(8)による解析を基に Head-Driven Phrase Structure Grammar (5) も参照して、動詞や形容詞にかかる主語、目的語、副詞句や、名詞にかかる形容詞句のように主辞の前後の位置関係だけがかかり方が指示されている句と、前置詞や接続詞という機能的な主辞が置かれた句が完結的主辞を置いた句を従えて、かかり方を指示している場合を識別する。機能的な主辞は必ず完結的主辞を置いた句がかかる先として置かれて、独自の句を構成する。パネル表示は日本語と同様、全て主辞を載せたパネルの上にそれにかかる句のパネルを載せて表示する(1)。

3. 日本語並立句の構造表示

これまで係り受け解析器による解析結果を用いた句構造表示においては、chunk の構成も含めて必ずかかれる完結的主辞を載せたパネルの前に、かかるほうの chunk を並べて乗せて表示を作成してきた。しかし並立となることばの構成においては、どちらかにかかるという表示では、表現意図が汲み取りにくくなっていた。そこで今回我々は、日本語並立表現の学習用の構造表示の提案する。図1にその表示例を示す。

並立句は動詞「V たり」の繰り返しや、時間の経過に合わせた連続的な動作の表現、同列の名詞等において、係り受けではなく「並立」していると捉えて構造を表す。また並立句の chunk では「head 形態素複合」と「functional 形態素複合」を階層表示ではなく 1 つのパネル上に仕切り線を挟んで表示していく。現行では並列句表示を人手によって作成したが、CaboCha による解析結果から自動化して並立構造の表示を作り出す方法を検討している。

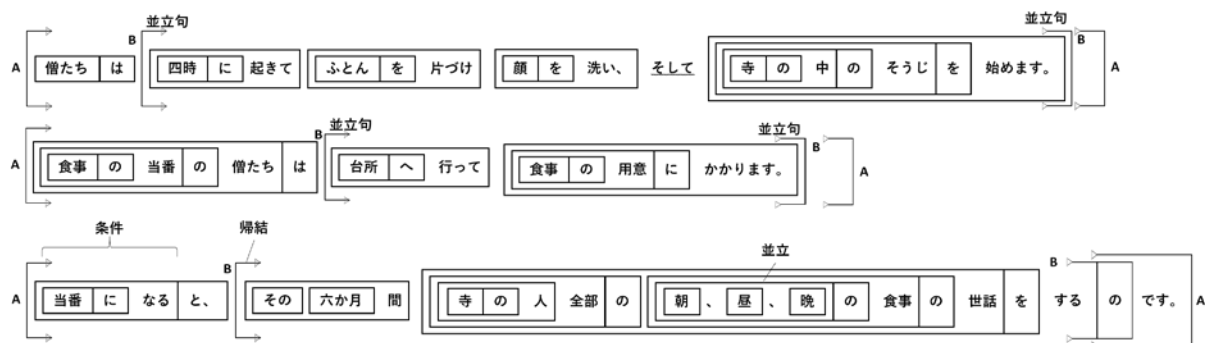


図1 並立表現の表示例

4. 対象言語表現と媒介言語表現の句構造を表示する学習支援

言語教育は、表現意図と表現の形の使い方の間のマッピングを訓練することであるが、対象言語の句構造と媒介語訳文の句構造を並べて、対象言語の句構造に対応する媒介語の句構造を対応表示すれば、(多少の訓練により)媒介語の階層的句構造が持つ意図へのマッピングができる学習者には、対象言語のその部分の意図と表現の形の対応が伝わることなる。

システムでは、学習対象言語を日本語、媒介語を英語として、学習対象の日本語文に対し、英語の対訳文を予め用意し、その構文解析結果の単純化によって生成した XML ファイルを基にして、テキストの形態素/単語をその並びのままに取り込んだパネルの階層データを作り、その表示によって、両言語における句構造と言語間での対応を視覚的に提示する。

5. 表現型の句構造マッチングに基づく句構造コーパス

対象言語のテキスト内に登場する表現型は先行研究(9)(10)を基にして、使用例の句構造に関連付けて、キーとなる表層と品詞指定をふくむ句構造として分類定義し、XML 形式のライブラリを作る。そして、句構造表現されたテキストへのマッチングにより、各表現型の使用例のテキスト句構造上の位置と構造を XML 形式で登録した「表現型使用例句構造コーパス」を用意する。なお、表現型の仕分けに、主辞の意味カテゴリが必要な場合は表現型とコーパスに、手作業で付加する。

```

<expression id="3">
  <phrase id="0" on="1,2" pos="動詞">
    <phrase id="1" pos="動詞">
      <text />
      <func pos="助詞">
        たり
      </func>
    </phrase>
    <phrase id="2" pos="動詞">
      <text />
      <func pos="助詞">
        たり
      </func>
    </phrase>
    <text>する</text>
  </phrase>
</expression>
</expressions>

```

図2 表現型ライブラリ XML

```

<phrase id="1" sid="1">西北部<func funcSpecified="true">に</func>
</phrase>なり</phrase>
</expression>
<expression id="2">
  <phrase headSpecified="true" id="3" sid="0">
    <phrase id="1" sid="0">町<func funcSpecified="true">へ</func>
    </phrase>
    <phrase id="2" sid="0">托鉢<func funcSpecified="true">に</func>
    </phrase>出かける</phrase>
  </expression>
  <expression id="3">
    <phrase headSpecified="true" id="7" sid="0">
      <phrase id="3" sid="0">出かけ<func funcSpecified="true">たり</func>
      </phrase>
      <phrase id="6" sid="0">し<func funcSpecified="true">たり</func>
      </phrase>する</phrase>
    </expression>
  </expressions>

```

図3 句構造コーパス XML



図4 コンテンツの階層的句構造の表示例

6. 句構造表示コンテンツを用いた言語学習メソッド

学習者には句構造表示コンテンツを材料として、再現的な生成と理解の課題を課し、その解決を、コンテンツにおける表現型使用事例の検索による句構造表示の比較によって支援する過程において、表現の使い分けを獲得させる。

6.1 再現作文課題学習

課題学習では、学校や市井における日常生活、職業あるいは研究など、限定した場面で用いることができるように、表現使用の汎化を行なわせる。教材テキストにおける表現型使用に関する穴埋め課題、テキストの文についてその形態素を含んでランダムに並べて提示される一覧から、形態素を選んで並び替えさせる再現作文課題、解読支援としてテキストの文を与えて意味制約関係としての階層句構造を作らせる表現解読課題を用意する。

再現作文課題では、図5に示すように、句構造表示を行なった教材テキストの中の一文を形態素に分解し、場合によっては余分な機能形態素も加え、それをランダムに並べて提示し、学習者は一覧から選んで正しく文章になるように並び替える。チェックボタンを押すと、学習者が並べた形態素の列の中で、句構造が原文に一致する部分について合体させて構造表示を行なう。こうして出来たマクロな部品の並び替えをおこなってはチェックを受けることを、階層的に行なわせることにより、再現作文を支援する。

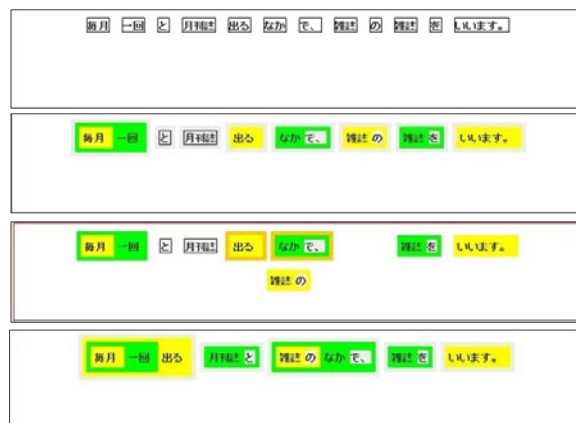


図5 再現作文の流れ

学習者は、課題学習を解決する過程において、表現型のメニューから、あるいはテキスト中の句をクリックしてそれを含む表現型のメニューから選んで、テキスト中にある同一または関連した表現型の使用例を検索し、比較して、帰納的に表現の作り方と使い分けを学習することができる。

例えば図5のような作文課題において「いいます」に着目して、それを含む表現型を検索して、XML形式の表現型ライブラリから「[用言句終止形]という」、「[体言句]を[体言句]という」、「[体言句]が[体言句]に[体言句]をいう」の3つの表現型が見つかったとする。それぞれの例として「ベトナムに行くという。」、「親の兄弟姉妹の子供を「いとこ」という。」、「私が先生にお礼をいった」が

表現型使用例句構造コーパスから見いだされたとすると、出題の形態素の部品の中から助詞が見つかる2番目の型が、階層的句構造で当てはまる組み合わせ候補として「雑誌を雑誌といいます」「月刊誌を雑誌といいます」「雑誌を月刊誌といいます」が得られ、並び替え再現作文を解いていく上での足掛かりとなる。

表現型の間には関連性のリンクが張られ、類似や対立その他の関係にある表現型を含む部分構造を検索し、英語訳文の句構造を参照しながら日本語文の句構造を比較することによって帰納的に表現の作り方と使い分けを学習することができる。

6. 2 構造解読課題学習

構造解読問題では、学習者に、テキスト上で句構造を再現させる。そのための支援として、完結的主辞を指摘させて正しければ、そこに係る句の完結的主辞の全てを指摘させ、正しければその階層の句構造を再現表示する。以下同様に主辞のかけり関係を指摘させながら各階層の句構造を再現表示していく。各階層において、完結的主辞に対応した媒介語の単語を表示し、意味の側面で支援する。

7. まとめと今後の課題

第二言語としての日本語学習において、入れ子構造の階層構造表示に加えて、日本語の並立構造の作成規則から導いた「並立表現」の構造表示について提案した。

また表現型の使用事例が含まれるテキスト群の句構造表示コンテンツを用いて、再現的な生成と理解の課題を課し、その解決を表現型使用事例の検索による句構造表示の比較によって支援する過程において、表現の使い分けを帰納的に獲得させる言語学習メソッドを提案し、課題解決支援と、表現型使用例句構造コーパスにより表現型の使用例を検索させる仕組みを試作した。

今後は並列構造表示の自動化を行ない、表現型使用例句構造コーパスを充実させ、作文・読解支援の仕組みを完成して、提案した支援に対する評価実験を行なうことを予定している。

参考文献

- (1) 伊藤紘二, 中村宏, 掛川淳一, 佐々木諒: “句構造の提示に基づく第二言語習得支援”, 日本第二言語習得学会第 15 回年次大会, pp2-4, 2015.
- (2) 中村宏, 掛川淳一, 伊藤紘二: “対象言語の表現の使用例をその媒介言語表現と共に句構造で提示して比較対応を行わせる第二言語学習支援” 日本教育工学会第 33 回全国大会, 2a-501-06, 2017.
- (3) 中村宏, 掛川淳一, 伊藤紘二: “媒介言語による表現を伴わせて句構造を表示した表現使用例の検索を手掛かりに読解と作文を行なわせる第二言語学習支援” 言語処理学会第 24 回年次大会発表論文集, pp.484-487, 2018.
- (4) T.Gunji: “Japanese Phrase Structure Grammar: A Unification-based Approach,” Springer Netherlands, 1987.
- (5) C.Pollard, I.A.Sag: “Head-Driven Phrase Structure Grammar,” University of Chicago Press, Chicago, 1994.
- (6) 時枝誠記: 国語学原論, 岩波書店, 1941.
- (7) 工藤拓, 松本裕治: “チャンキングの段階適用による日本語係り受け解析” 情報処理学会論文誌, 43 (6), pp.1834-1842, 2002
- (8) D.Klein, C.D. Manning: “Accurate Unlexicalized Parsing,” Proc. ACL, 2003
- (9) 掛川淳一, 中村宏, 関谷政則, 伊丹誠, 伊藤紘二: “自然言語処理を用いて日本語教育のための例文検索を支援するシステム” 日本教育工学会論文誌, 25 (2), pp.85-94, 2001.
- (10) 哈布日, “第二言語としての日本語の表現教材作成支援に向けた例文検索ルールデータベースの構築”, 兵庫教育大学大学院学校教育研究科修士論文, 2012.