

アノテーションを利用した日本語表現法獲得支援システム

A System Supporting Skill Acquisition for Reading and Writing of Japanese As Second Language Using Phrasing Annotations

掛川 淳一*¹ 石川 賢太郎*¹ 海野 俊介*¹ 藤井 雅弘*¹ 伊丹 誠*¹ 伊藤 紘二*¹
 Jun-ichi KAKEGAWA Kentaro ISHIKAWA Shunsuke UNNO Masahiro FUJII Makoto ITAMI Kohji ITOH

*¹東京理科大学基礎工学部電子応用工学科

Department of Applied Electronics, Tokyo University of Science

In this paper we propose a text-based system that assists learner's acquisition of the ability of using expressions of Japanese as a second language. The system allows learners to refer to the teacher's notes of phrasing attached to the positions where any useful phrasing is made use of in the texts. The note comprises the semantic label and the form of the expression it designates as well as interpretive comments. The notes are used to retrieve text positions where the phrasing concerned or other related expressions are used and allows the learners to compare usage of the expressions in different contexts.

A preliminary evaluation experiment suggests the effectiveness of the system compared with a system similar to the proposed except for the capability of mutual referencing of the notes

1. はじめに

近年、第二言語獲得 (SLA : Second Language Acquisition) における研究を取り入れたコンピュータによる言語学習支援 CALL (Computer-Assisted Language Learning) の研究が盛んになってきている [1][2] .

我々は、これまでに日本語を学習する外国人留学生を対象として、作文のための表現法獲得を支援するシステムについて提案と試作を行ってきた .

一般に、表現は分野や文脈に依存して使い分けがなされ、同じ表層を持つ表現であっても異なる意味、意図で使用されることが多い . そのような表現を獲得させるためには、ある表現がどのような状況、場面で使用可能かを認識させるために、表現の出現する他のテキストを提示し、そこでの比較を行わせることが重要である .

そこで、我々は、表現を辞書等により意味的な把握を行ったうえで、比較事例を見ることができれば、学習者にとって、そこでの表現事例間の格や項の違い、その表現の使用される場面や状況の違いを意識しやすく、一般化を行いやすくなると考えた .

2. 提案システム

提案システムでは、教師によって教材として設定されたテキストについて、学習者に読解を行ってもらう .

教師は、設定したテキストについて、予めアノテーションを各単語 / 表現に付加しておく . 教師はその際に、それらのアノテーション (以下、「先生のノート」と記す) にはその表現に関するコメントを記述する . とくにイデオムやコロケーションのような表現に関する「表現ノート」には表現検索のためのインデックスをつける . (インデックスについては、それまでに登録されたインデックスのメニューの中から選択しても良いし、これから教えるべき表現については、新規に登録することもできるものとする .)

学習者は、教師による単語や表現に関するノートづけがされたテキストを読解していくが、未習得 / 曖昧な単語や表現に関しては、電子辞書や「先生のノート」を閲覧しながら、内容把握を

連絡先: 〒 278-8510 野田市山崎 2641, 東京理科大学基礎工学部, Tel. 04-7124-1501(ex. 4226), Fax. 04-7122-9195, 掛川 淳一 (kakegawa@itlb.te.noda.tus.ac.jp)

行うことができる . 「先生のノート」が付加されている表現について、他の例文が見たければ、ノートに記載された検索用インデックスを利用し、テキスト群からその表現が使用されているテキストを検索することができる .

提示される例文が使用されているテキストが、学習者にとって、これまでの授業において既習のものであれば、内容に関する理解がされており、その表現の形式と文脈に関する検討により表現に関する知識の確認と発見を期待できる .

また、学習者は自分のノートを作成することができ、「表現ノート」の場合は、「先生のノート」に記載されたインデックスをコピーさせる .

3. 表現ノート

「表現ノート」は、読解テキスト中に出現する表現に対して、つけられるアノテーションであり、以下のような形式をとる .

1. 表現が出現したテキストファイル名と表現の位置
2. 表現の
 - (a) 意味的カテゴリ
 - (b) 表現形式
 - (c) 構文的特徴
3. 参考にした / すべき教師、もしくは、他学習者の「表現ノート」へのリンク
4. 訳、およびメモ

なお、1. に関しては、システムが自動的に付加し、2(c) については、ハイライト箇所の指定と構文情報を利用した表現検索のために使用される .

「表現ノート」における 2(a) と 2(b) をキーとするブリーアン検索により、同一、類似、関連表現にアクセスできる .

本システムで採用する電子テキストは、構文解析等の自然言語処理が施されたものとする . システムは、ユーザがテキストから選択した箇所に含まれる構文情報を認識し、要求に応じて、その情報を可視化したり、電子辞書にリンクする . 「表現ノート」作成時には、選択された文の構文解析結果を表示し、検索された際にハイライトする位置の指定を支援する .

「表現ノート」の登録された「構文的特徴」は、登録テキスト群の中から、同一の構文的な特徴をもつ表現を検索するために使用することを想定して設けられた .

たとえば教師が学生に対する読解課題テキストを新たに設定する場合、これまでに登録された「表現ノート」の「構文的特徴」を利用することで、システムから表現ノートをつけるべき箇所の提案を受けながら、「表現ノート」を作成していくことができるようにするためである。

4. 評価実験

4.1 試作システム

試作したシステムは、構文解析 KNP[3] を施した電子テキストと、XML (eXtensible Markup Language) によって記述された表現ノート群を用いる Java アプリケーションである。

評価実験で用いた試作システムでは「表現ノート」における他の表現ノートへのリンクの機能と「構文的特徴」を用いた表現検索の機能、及び、電子辞書へのリンクの機能は実装されていない。

4.2 実験

システムを 2 種類用意し、一方は、単語や表現について簡単な英語と日本語による訳が記載されたノートが見れるもの (以後、 O_1 とする)、他方は、 O_1 の機能に加え、表現検索機能があり、比較用例文が見れるもの (以後、 O_2 とする) とし、例文参照の効果について評価を行った。

被験者は留学生二名 (それぞれ S_1 、 S_2 とする) であり、 S_1 は中国出身の理系大学生であり、第二言語として日本語、英語、また、 S_2 はウイグル出身の理系大学院生であり、媒介語として中国語、第二言語は日本語、英語である。

評価実験は、事前テスト (T_1)、システム A 使用、中間テスト (T_2)、システム B 使用、事後テスト (T_3) の手順で行った。事前、中間、事後のテストは同じ内容であり、ペーパーテストとして、テキストの中に穴を設け、候補の中から選び適切な形に語尾を変えろという形の、穴埋め作文問題を設定した。

なお、システムでの読解課題テキストと、ペーパーテストでの穴埋め作文課題テキストは [4][5] の同じ単元から抜粋したものである。テキスト中の各表現について、システム開発者が事前に教師の「表現ノート」の登録を行った。また、参照される比較用のテキスト ([4][5] より抜粋) についても同様の登録を行い、被験者に読解させるテキストでは、「意味的カテゴリ」と「表現形式」のインデックス検索で原則として 2 以上、最低 1 の比較用の例文を学習者に提示できるように設定した。

また、テキスト中の各単語については、簡単な日本語と英語による注釈を、システムではノートとして付加し、ペーパーテストでは単語帳として用意し、被験者に使用させた。

4.3 結果と検討

各システム使用後の達成度を、(T_k での得点 - T_{k-1} での得点) / (満点 - T_{k-1} での得点) とすると、以下の表 1 のようになった。

表 1: システム A, B 使用後の達成度の推移

	O_1 使用后	O_2 使用后
S_1	0.071	0.19
S_2	0.27	0.24

また、各ペーパーテストでは、問題数に対し十分な時間をとったため、実際に両被験者とも時間を余す形となり、回数を経るにつれて余らせる時間が増えていたようであった。

被験者 S_1 は、システム O_2 使用時には、例文を参照するスピードが早く、内容把握についても問題が無いようで、参照先の

テキストについての読解も早かった。表 1 より察するに、被験者 S_1 は、システム O_1 では獲得ができず、システム O_2 の使用により、獲得が進んだと考えられる。

対して、被験者 S_2 は、システム使用時には、読解のスピードは遅いが、時間をかけ、参照例文との見比べを繰り返していた。また、事前テストから中間テストの間の得点の伸びを見るに、事前テストのペーパーテストで使用されていた表現と、システムで読解したテキストで使用されている表現との間の類似性を発見し、効率的に学習してしまったものと考えられる。

しかしながら、被験者 S_1 、 S_2 とともに提案システムであるシステム O_2 使用後も得点の大きな伸びを見せた。

評価実験終了後、簡単なインタビューを行った。システム O_2 のように、「ある表現についてテキストの中に埋め込まれた例文を、数多く参照するということは有用か?」という質問に対し、両被験者から「Yes」という回答が得られた。

被験者 S_1 は、「システムに登録されていたノートの情報についてより詳細な説明がほしい」ということであった。

被験者 S_2 の方は、「予め説明を受けたが、システム O_2 を使用していくにつれて、どのようなシステムであるかをあらためて認識できた」ということであった。

今回用意したテキストに付けられたノートの情報は簡単な日本語と、簡単な英語で記述されたものであったが、被験者の英語力を測定していないので、あるいは意味的な内容把握の支援が十分でなかったかもしれない。

5. まとめと今後の課題

本システムは、第二言語における単語、表現についての知識の獲得を支援する。

学習者は、電子辞書を利用して、テキストの読解を行うことで文脈の把握を行い、その上で、そこに現れる表現法について、他の事例と比較を行うことができる。文脈に関連づけられた表現法を意識させることで獲得された知識は、作文の際の文脈に応じた表現の使い分けの知識に結び付くであろうと考えた。

提案したシステムについて部分的な試作を行い、コンテキスト付きの比較事例を提示する機能の有効性について、小規模ながら評価実験を行った。各被験者は提案システム使用後のテストにおいて、比較事例の見れないシステム使用後のテストよりも大きな達成度の上昇を確認できた。

今後の課題として、被験者の数を増やし、評価実験によりシステムの検証を行って行く予定である。また、「教師ノート」のオーサリングに関する調査と検討も同時に行っていく予定である。

参考文献

- [1] C.A.Chapelle: "Multimedia CALL: Lessons to be Learned from Research on Instructed SLA", Language Learning & Technology, Vol.2, No.1, pp.22 - 34 (July 1998).
- [2] Johann Gamper, Judith Knapp: "Review of CALL System in Foreign Language Instruction", Artificial Intelligence in Education 2001, pp.377 - pp.388.
- [3] 黒橋 禎夫: 日本語構文解析システム KNP version 2.0 b6, 1998.
- [4] 筑波大学日本語教育研究室: "日本語表現文型中級 I", 凡人社, 東京 (1993)
- [5] 筑波大学日本語教育研究室: "日本語表現文型中級 II", 凡人社, 東京 (1994)