

外国人向け敬語文理解支援システムの提案

李 国慶[†]吉野 孝[†][†] 和歌山大学

1 はじめに

2012年の国際交流基金の調査^{*1}によると、全世界の日本語学習者は約399万人である。また、2009年より9.1%増加しており、日本語を勉強する外国人は増加している。外国人日本語学習者の敬語使用に関する結果によると、外国人日本語学習者は日本人と同じように敬語は難しいと感じている[1]。そこで本研究では、外国人向け敬語文理解支援システムを提案する。本システムでは、「変換」「解釈」「翻訳」の3つの機能を段階的に利用することで、外国人の日本語における敬語文理解支援を目指す。

2 関連研究

大野らは、独自の変換ルールにもとづく日本語敬語表現の変換・解析システムを開発した[2]。変換には、形態素解析や名詞の意味属性などが用いられる。しかし、この研究の目的は、敬語文から普通語文への自動変換であり、システム利用者の理解については考慮されていない。

日本語における読解支援に関する研究としては、日本語読解学習支援システムであるリーディングチュウ太があげられる[3]。このシステムは、入力文に含まれる全ての単語の解釈を提示する。しかし、単純に単語の解釈の支援だけでは、外国人が文の意味を理解するには不十分である。また、利用者は長い入力文と解析結果を照らし合わせながら読む必要があるため、これらの対応付けが難しい。

本研究では、1文ごとの解析結果を「変換」「解釈」「翻訳」の三段階で提示することで、外国人の敬語文理解支援を目指す。

3 敬語文理解支援システム

3.1 システムの概要

図1にシステムの流れを示す。本稿では、「変換」「解釈」「翻訳」の3つの機能で、段階的に結果を提示する。その理由は、結果全てを提示すると情報量が多く、文の理解に時間がかかる可能性があるためである。図2にシステムの実行例を示す。図2-(1)より第一段階では、まず、外国人が文章を読む負担を軽減させるために、文章を1文ずつに分ける。次に、形態素解析 Mecab^{*2}と敬語変換ルールを用いて「変換」を行った結果を提示する。図2-(2)より第二段階では、外国人が第一段階の結果で意味の理解が困難な場合、第一段階の「変換」に加えて「解釈」を行った結果を提示する。図2-(3)より第三段階では、第二段階の結果で意味の理解が困難な場合、第二段階の「変換」および「解釈」に加えて、敬語原文の「翻訳」を行った結果を提示する。以下に各機能について述べる。

3.2 変換

本稿では、敬語文を分かりやすい文に変更することを「変換」と定義する。また、「分かりやすい文」とは外国人が理解しやすいと考えられる文のことである。例

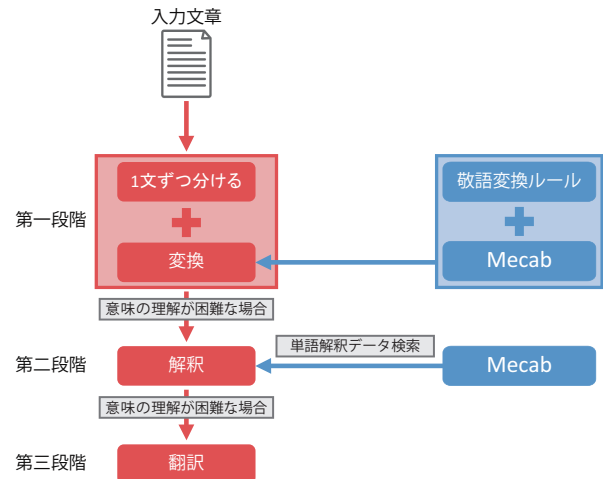


図1: システムの流れ

(1) 第一段階 一行ずつ分け、分かりやすい文に変換

このたび喜父様の急速の移転に接し、あまりにも突然のことで驚いて(い)ます
 ここに(正確な)意味の意を表するとともに、真意を申し上げます
 解題してから、急に具合が悪くなったことか
 調査を共にした貴父さまが世界の悲しみはいかほどかと察します

(2) 第二段階 単語の解釈

このたび喜父様の急速の移転に接し、あまりにも突然のことで驚いて(い)ます
 喜父→父
 急速→急に(なった)
 移転→移入報告

(3) 第三段階 翻訳

このたび喜父様の急速の移転に接し、あまりにも突然のことで驚いて(い)ます
 喜父→父
 急速→急に(なった)
 移転→移入報告
 01번에 조부님의 급사의 비보에 받고 매우 놀란것으로 보고 있습니다

図2: システムの実行例

例えば、「お願い申し上げます」という敬語文は「願います」という分かりやすい文に変更される。「変換」された文章は日本語としては正しくない場合があるが、外国人にとっては敬語文より理解が容易である可能性が高い。

3.3 解釈

本稿では、難しい単語を簡単な文章で説明することを「解釈」と定義する。Mecabの辞書をカスタマイズし、単語の「解釈」を辞書に入力した。「解釈」は文をMecabで解析し、単語の「解釈」がある単語を抽出する。難しい単語と簡単な文章を判断する基準は日本語能力試験N2(中級)を基準値にした。以下に、難しい単語と簡単な文章の定義を示す。

難しい単語: 中級以上の単語と敬語文によく使う単語。

簡単な文章: 初級と中級以上の単語を使った文章。

例えば、「寝食」という単語は「日常生活(食べることと寝ること)」と説明される。

Proposal of the honoric sentence comprehension support system

Guoqing Li[†] Takashi Yoshino[†]

[†]Graduate School of System Engineering, Wakayama University

^{*1} 日本語学習者数調査結果: http://www.jpfr.go.jp/j/japanese/survey/result/dl/survey_2012/2012_s_excerpt_j.pdf

^{*2} Mecab: <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

表 1: 協力者の属性

協力者	国籍	英語	日本語レベル	日本滞在歴
A	韓国	不得意	N2	1年～2年
B	韓国	不得意	N2	1年～2年
C	韓国	不得意	N2	3ヶ月
D	ベトナム	不得意	N2	6ヶ月
E	ベトナム	得意	N2	6ヶ月
F	ベトナム	不得意	N3～N2	3ヶ月
G	ラオス	不得意	N2	3年以上
H	オーストラリア	得意	N3～N2	3ヶ月
I	フランス	得意	N3～N2	3ヶ月

表 2: 文章の属性

文章	文数	1文を分割した文数	評価値の満点
1	6	2部分: 6	60
2	8	2部分: 8	80
3	7	2部分: 6 3部分: 1	75

このたびは貴社社長様ご逝去の報に接し、驚愕のほかございません。ご家族はもとより、貴社ご同様のご愁傷のほどいかにばかりかと拝察致します。ここに謹んで哀悼の意を表しますとともに、ご冥福をお祈り申し上げます。ご生前中はひとかたならぬご高配を賜わり、弊社が今日まで社業発展を遂げることができたのも、ひとえに社長様のお陰と深く感謝しております。

図 3: 評価実験に利用した文章

3.4 翻訳

本稿では、敬語文を対応言語に変更することを「翻訳」と定義する。現在対応言語は、英語、中国語、韓国語、ベトナム語である。

4 評価実験

4.1 実験概要

2014年12月26日に本システムを使用した時の敬語文理解度について調査するため、評価実験を行った。本実験では、中国人以外の外国人を対象とした。その理由は、中国人は漢字で文章の意味を理解できたためである。実験協力者は外国人学生9名である。表1に協力者の属性を示す。日本滞在歴が3ヶ月の協力者は交換留学生である。

本実験では、本システムと2つの既存システム(Google翻訳³、リーディングチュウ太)の比較を行った。協力者は、各システムを用いて、敬語文3文章⁴の意味の記述を行った。また、比較のため各システム使用前にも敬語文3文章の意味の記述を依頼した。図3に評価実験に利用した敬語文章の一部を示す。各システムによる実験結果への影響を考慮してラテン方格法を用いた。

また、協力者が書いた文章を日本人大学生3名に評価してもらった。協力者が書いた敬語文章の意味を評価するため、各文章の1文を、文の意味のまとまりで2～3部分に分け、各部分に対して5段階評価を行った。評価値は、「5: 同じ意味」「4: 文法などに多少問題があるが、大体同じ意味」「3: 意味は何となくつかめる」「2: 雰囲気は残っているが、元の意味は分からない」「1: 全く違う意味、または意味が分からない」である。表2に各文章の文数と1文の分割状況、また文章の評価値の満点を示す。評価後に、各システムの使用前および使用時の得点平均率を計算した。以下に、得点平均率の計算式を示す。

$$(1人当たりの各システムの平均評価値) = \frac{(各部分の評価値合計)}{(文章評価満点値) \times (評価した人数)}$$

³ Google 翻訳: <https://translate.google.co.jp>

⁴ 今回、実験に使用した敬語文3文章はそれぞれ下記 url より参照した。

文章 1: <http://www.bunnrei.net/oiwaijyou/kannreki/201301101896.html>

文章 2: <http://www.bunnrei.net/oiwaijyou/kannreki/201301101896.html>

文章 3: <http://www.printmate.co.jp/postcard/model.701.php>

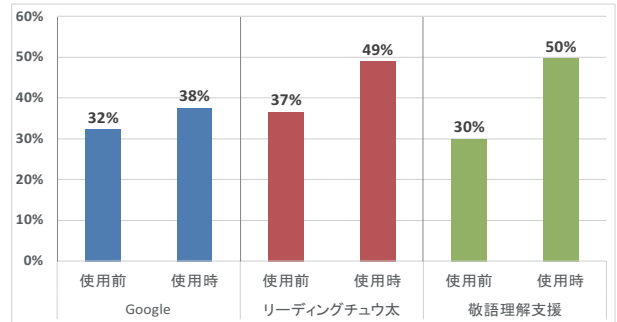


図 4: 各システム使用前および使用時の得点平均率

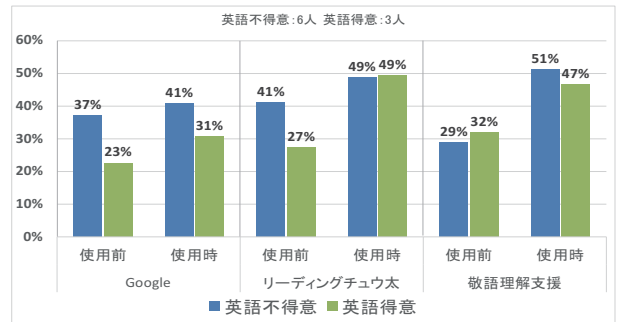


図 5: 英語得意協力者および英語不得意協力者の得点平均率

$$(各システムの得点平均率) = \frac{(1人当たりの各システムの平均評価値合計)}{(協力者の人数)}$$

4.2 実験結果と考察

図4に各システムの使用前および使用時の得点平均率を示す。図4より、本システム使用時の得点平均率が最も高く、本システムを使用することで、文章の理解度が高くなったと考えられる。しかし、リーディングチュウ太使用時の得点平均率と比較すると、理解度の差は見られなかった。

図5に英語得意協力者および英語不得意協力者の得点平均率を示す。図5より、英語得意協力者の本システム使用時の得点平均率がリーディングチュウ太使用時の得点平均率より低かった。リーディングチュウ太では、英語の解釈が提示されるため、協力者は英語の説明で文の意味を理解したと考えられる。このことから、他言語得意協力者に対応するため、単語の解釈を翻訳する機能が必要であることが分かった。

5 おわりに

本稿では、外国人向け敬語文理解支援システムを提案した。評価実験を行った結果、下記のことが分かった。

- (1) 本システム使用時における文の理解度が一番高かったが、リーディングチュウ太使用時における文の理解度と比較すると、差が見られなかった。
- (2) 他言語得意協力者に対応するため、単語の解釈を翻訳する機能が必要である。

今後は、単語解釈の翻訳機能の実装と、解釈用データの追加を行う。

謝辞

本研究の一部は、JSPS 科研費 24220002 の助成を受けた。

参考文献

- [1] 魏春娥: 日本人と外国人日本語学習者の敬語使用に関する考察: アンケート調査結果の分析を中心に, 山口大学文学会誌 63, pp.125-150(2013)。
- [2] 大野満, 横山晶一, 西原典孝: 日本語敬語表現の変換・解析システム, 言語処理学会 第9回年次大会 発表論文集, pp.218-221 (2003)。
- [3] 川村よしこ: 日本語読解支援システム「リーディング・チュウ太」の取り組み, ウェブマガジン「留学交流」, 2012年10月号, pp.1-11(2013)。