

敬語文章変換システムの作成

飛鳥井 元晴[†] 岸 義樹[‡][†] 茨城大学大学院理工学研究科情報工学専攻[‡] 茨城大学工学部情報工学科

1 はじめに

我々が他者との人間関係を良好に保つためにはコミュニケーションが重要である。コミュニケーションには様々なものがあるが、代表的なものとして言語があげられる。日本語には敬語表現をはじめとする言葉遣いが存在し、適切に用いることで他者と円滑にコミュニケーションを取る助けとなる。しかし、敬語表現には様々なルールが存在し、適切な敬語を扱うのは用意ではない。本研究では、敬語表現への変換ルールを分かrazずとも敬語表現を用いた文章を使えるように補佐するべく、ユーザが入力した日本語文章を敬語表現を用いた文章に変換するシステムを作成し、変換を行った。

2 提案手法

一般的な文章において、主語が自分側の場合は、相手に自分の行為などをへりくだって伝える謙譲語が使われ、相手側の場合は、対象となる相手を高める尊敬語が使われる [1]。本システムでは、自然言語で書かれた文章に対して形態素解析を行い主語となる語を特定し、主語の属性が自分側か相手側かに関する属性判定を行うことにより、文章中に現れる形態素に対して謙譲語と尊敬語どちらの変換を行うかを決定する。その後、決定された表現のパターンへと変換処理を行う。今回の研究では特に動詞の変換について重点的に考察を行った。

本システムの流れは以下の通りである。それぞれの実装については次節で述べる。

1. 入力文への形態素解析
2. 主語となる単語の決定
3. 得られた主語の属性判定
4. 各形態素への変換情報の付与
5. 敬語変換処理
6. 変換された文章の出力

表 1: 主語となる単語

品詞分類名	品詞 ID	助詞
代名詞	42,43,44	「は」・「が」
名詞接尾, 人名	55	
代名詞, 一般	59	

3 実装内容

システムの実装内容と流れについて記述する。

3.1 入力文への形態素解析

入力文に対し形態素解析を行う。形態素解析エンジンには Mecab を用いた。品詞を数値化した品詞 ID と素性情報を取得する。

3.2 主語となる単語の決定

本研究において主語として扱うものを表 1 にまとめた。品詞が特定の名詞で続く形態素が助詞「は」・「が」であるものを主語とした。候補が 2 つある場合は両方を保持することとする。主語の前に名詞が続いていた場合はそれらを繋げたブロックを主語とする。

3.3 得られた主語の属性判定

得られた主語の属性判定を行う。ここでの属性とは「自分側」、「相手側」、「その他」のいずれかである。先ほど取得した主語に対してパターンマッチングを行い、「自分側辞書」に登録されていれば自分側、「相手側辞書」に登録されているか、「敬称辞書」に登録されている単語を含む場合は相手側、どちらでもない場合はその他とする。「自分辞書」、「相手辞書」、「敬称辞書」は『日本語語彙体系』 [2] のカテゴリをもとに作成した。

3.4 各形態素への変換情報の付与

主語が自分側となる場合には謙譲語変換情報を付与し、相手側となる場合には尊敬語変換情報を付与する。属性がその他及び主語未発見の場合には謙譲語変換情報を付与することとした。主語となる語が 2 つあり、それぞれが自分側と相手側の属性を持つ場合は、最初に現れた主語候補+助詞「は」・「が」までとその文章で

Creation of honorific sentence conversion system

Motoharu Asukai[†], Yoshiki Kishi[‡]

^{†‡}Ibaraki University

4-12-1 Nakanarusawa, Hitachi, Ibaraki, 316-8511, Japan

最後に登場する用言となる語である動詞・形容詞・形容動詞語幹以降からなる部分と、それ以外の部分で分ける。2つのブロックに対し主語の属性に応じた変換情報を付与する。

3.5 敬語変換処理

形態素ごとに、付与された情報と品詞 ID から行う変換を決定し、敬語表現に変換したものと置き換えを行う。自立動詞の変換処理を次節で述べる。

3.6 変換された文章の出力

変換が行われた文章を出力する。

4 自立動詞の変換処理

敬語変換処理における動詞の変換について説明する。敬語の動詞の変換には、「食べる→召し上がる」のような単語自体が変化する「特殊形変換」と、「歩く→お歩きになる」のような単語に「お～になる」といった語を付け足す「付け足し型変換」に分けられる [3]。以下に付け足し型変換である動詞を含む「歩いて」の尊敬語変換の流れを示す。特殊形かどうかを判定する辞書は『敬語言い換え辞典：日常語から一発変換!』[4]の項目を参考に作成した。

1. 基本形である「歩く」が尊敬語特殊形辞書に登録されているかを確認する。
2. 登録されていないので付け足し型変換処理を行う。
3. 「歩く」の活用パターンは、五段カ行イ音便活用なので、基本形から1字削除して語幹「歩」を取得する。
4. 助詞「に」に接続するので「き」を付け加え連用形の形にする。
5. 「歩い」は、連用タ接続なので、後ろに付け足す「になる」を同じく連用タ接続である「になっ」の形に変換する。
6. これらを結合して「お歩きになっ」を得ることができるので、「歩い」と入れ替えて「お歩きになっ」という変換が完了する。

5 実験

判定文をシステムに入力し、実際に敬語変換を行い動作を確認する。

5.1 対象判定文

判定文には全文検索システム「ひまわり」[5]を用いて『青空文庫』パッケージ [6] から抽出した文を用いた。

5.2 結果および考察

主語が「私」や「～君」などが使われており、動詞の数が少ない文章では、違和感のない敬語文章に変換を行うことができた。また、うまく変換できなかった文には以下のようなパターンがあった。

1. 主語の取得ミス。「母」といった名詞は一般名詞に分類されており、人物であっても主語として取得できなかった。しかし、一般名詞も主語と含めてしまうと人物でないものも主語として取得してしまうため、専用の辞書を作り対処する必要がある。
2. 物体の動作に対する変換。形態素解析だけでは、判断できないため係り受け解析等を用いて人物についてのみ変換処理を行うようにする必要がある。
3. 連用タ接続の変換ミス。“死んだ→お亡くなりになった”となるべきだが、“お亡くなりになった”となってしまった。これは動詞が置き換わることにより続く助詞との接続があわなくなったものと考えられる。

6 まとめ

本研究では、形態素解析による主語決定及び形態素素性情報を用いた敬語文章変換システムを作成し、判定文を用いて実行することで、システムの動作を検証した。簡単な文章では違和感なく敬語変換を行うことができたが、うまく変換できない文もみられた。

今後は変換のルールの見直しや係り受けへの対応、変換時の助詞接続の修正などを行うことで敬語変換の精度の向上を目指したい。

参考文献

- [1] 田添丈博, 渡辺千亜季, 椎野努: “敬語表現への言い換えに関するコンピュータモデルの構築”, 情報処理学科研究報告. 自然言語処理 169-1, pp.1-6, 2005.
- [2] 池原悟: “日本語語彙大系”, 岩波書店, 1997.
- [3] 岩下志乃, 岩切智希: “状況に応じた対話による敬語学習システム”, 知能と情報, 日本知能情報フェジ学会誌, journal of Japan Society for Fuzzy Theory and Intelligent Informatics 20(5), pp.709-719, 2008.
- [4] 学研辞典編集部: “敬語言い換え辞典: 日常語から一発変換!”, 学習研究社, 2005.
- [5] 大学共同利用機関法人 人間文化研究機構 国立国語研究所, “全文検索システム「ひまわり」”
- [6] 山口昌也: “『青空文庫』パッケージ”