

日本語敬語表現の変換・解析システム

大野満†, 横山晶一‡, 西原典孝‡

†山形大学大学院 理工学研究科, ‡山形大学 工学部

1.はじめに

日本語は尊敬語、謙譲語、丁寧語をもちいた敬語表現が高度に発達した言語である。その場の状況や人間関係そして心理的要素などを瞬時に判断して使い分けるこの表現は、日本語を円滑に使う上で非常に重要な役割を果たしている。普段、親しい間柄で用いる言葉とは異なり、敬語表現を正確に使い分けることは日本人にとっても容易なことではない[1]。

情報化社会の急速な発展に伴い、機械を用いて敬語表現を扱うことの意義は増大している。敬語体系が話し言葉を軸に培われてきたことから、敬語表現を正確に扱うことができるならば、自然な日本語を扱う対話システムを構築することができると考えられる[2]。

敬語表現を扱う処理には、常体文から敬体文への変換と、敬体文から常体文への変換が考えられる。前者を実現するためには、対話者間の相対的地位やその場の状況などを考慮する必要がある。高度な意味解析や構文解析を必要とするためシステム化は難しい。それに比べ、逆の処理つまり、敬体文から常体文への変換処理では、対話者間の相対的な地位などを考慮せずに比較的一意に変換する語彙を決定できる。

本研究では、敬語の形式的な分類に基づいた変換規則を設定することによって、敬語文を普通語文へ機械的に変換するシステムを構築した。敬語の特徴でもある、多義性の敬語動詞については、その意味を形態素、名詞の意味属性、用言の文型パターンなど、比較的簡素な要素を用いて同定した。

2.背景

2.1.従来の研究

敬語表現は、日本語として重要な言語表現であるため、昔から様々な観点から研究が行われてきた。しかし、そのほとんどが言語学の観点から扱ったもので歴史的な背景や文化的側面などに重きを置いたものであった[1,3]。

一方、機械処理の観点から敬語表現を扱ったものとしては、杉村の研究がほとんど唯一である。この研究では、対話者間の上下や親疎の関係を数値化し、単語検索、パターンマッチング、単一化などを行うことにより文章を解析し敬語表現を扱おうと試みている[4]。

前研究[5]では、「である」調で書かれた常体文を「です、ます」調で書かれた敬体文に変換するシステムを構築した。この場合、ほとんどが文末での変換を行えば良く簡素な処理で変換が行えた。

本研究[6,7]では、もう一つ踏み込んだ領域である、敬体文から常体文への変換を取り扱う。常体文から敬体文への変換は、相手との相対的地位などを考慮するために、一般的に一对多の関係になるが、逆の場合には一对一または、

一对多でも場合の数が少なくなる。そのため、システム化が比較的容易に行えると考えられる。

2.2.敬語の特徴

2.2.1.敬語語彙の分類

敬語は、その形式から付加形式と置き換え形式の二つに分類することができる[8]。

1)付加形式

- ・接頭辞、接尾辞の付加
「お美しい」、「村長どの」、「お地藏さま」
- ・補助動詞、助動詞の付加
「書きなさる」、「読まらる」
- ・接頭辞、補助動詞の付加
「お連れいただく」、「お礼申し上げる」

2)置き換え形式

- 普通語とは別の敬語語彙を用いる。
「言う」に対する「おっしゃる」
「考え」に対する「所存」

2.2.2.多義語の存在

「いらっしゃる」、「いただく」などそれ一語で複数の意味を持つ語彙が敬語には存在する。これらは状況別に使い分けられるが、機械的にそれを同定することは非常に難しい。本研究では、日本語語彙大系[9]の意味体系と構文体系を用いて意味の同定を試みた。

2.2.3.特殊な言いまわし

敬語では、「お書きいただく」や「お礼申し上げます」などの「動詞+動詞」「名詞+動詞」といった特殊な言いまわしが許されている。どちらも対応する普通語にそのまま変換してしまうと不自然な日本語になる。したがって、変換する際に何かしら補助的な処理をしなければならない。前者の場合、敬語動詞が補助動詞の役割を果たしていると思なせるので、「助詞+動詞」の形の補助動詞と同等であるとして扱う。後者の場合、話し言葉の特徴である助詞の省略が行われていると見なし、助詞の補間を行う。

2.2.4.現代ではあまり使われない敬語

「御(ぎよ)、皇、宸、聖、勅、天、宝、鳳、竜」はどれも敬意を示す接頭辞であるが天皇のみに使う敬語であるため普通の敬語文には現れない。

また「お前」や「貴様」などのように昔は最上級の敬語であったものが時とともにその敬度を落としていき現在では同等かそれ以下の相手に使う言葉となり敬語とは見なせないものなど、現代ではあまり使われていない敬語については敬語として考えない。

2.2.4.慣用表現

「お気に召す」「いらっしゃいませ」など慣用的な使い方をする言葉。挨拶語などとも呼ばれ、それに対応する普通語が無いものもあるものに分けられる。前者は敬語、普通語どちらでも使う表現として、後者は例外表現として扱う。

3.変換処理

変換処理の大まかな流れ図を図1に示す。入力として、敬語を含む文章を想定する。

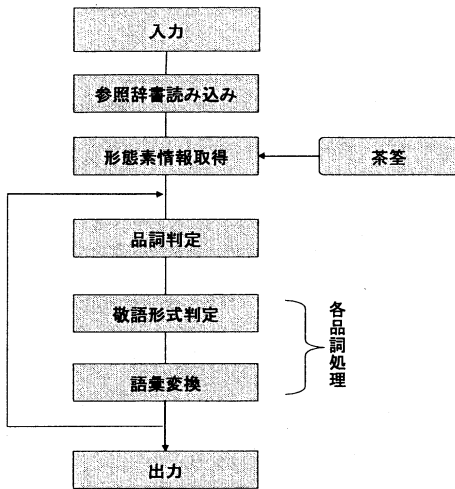


図1.処理の流れ図

3.1.参照辞書

置き換え形式敬語語彙に対する普通語語彙の相関関係を記した対訳辞書ファイルを作成した。名詞を体言辞書に、形容詞・動詞を用言辞書にそれぞれまとめた。その記述形式は以下ようになる。

- (普通語語彙¥パラメータ
 (尊敬語語彙¥パラメータ)
 (謙讓語語彙¥パラメータ)
 (丁寧語語彙¥パラメータ)
)

ここで用いられているパラメータは、体言、用言で異なる。体言辞書では、その語彙に対する品詞情報をパラメータとし、用言辞書では、その語彙に対する活用型情報をパラメータとする。

現在の登録語彙数は、体言辞書に約200語、用言辞書に約100語である。

3.2.形態素解析

汎用形態素解析ソフト「茶釜」[10]を用いて形態素解析を行う。その際、より正確な解析結果を得るために、「茶釜」の辞書ファイルを敬語表現用にカスタマイズした。行

う操作は茶釜でうまく解析されない、主に余計に細かく形態素分けされる語彙を繋げて一つの単語として追加登録すること、複合語として解析される語彙を削除することの二つである。表1に追加・削除語彙例を示す。

表1.茶釜辞書、追加・削除語彙例

追加語彙	削除語彙
ごらん(名詞-一般)	お願い(名詞-サ変)
おでまし(名詞-一般)	お礼(名詞-サ変)

3.3.語彙変換

3.3.1.基本変換規則

2.2.1節で述べた敬語表現の分類に従った基本変換規則を以下に記述する。

1)付加形式

基本的に付加された語彙を削除することで普通語化が可能だが、補助動詞を付加した形の場合、削除するとその補助動詞の持つ意味情報が抜け落ちてしまうので、単純に削除するのではなく、その補助動詞が持つ意味を保ったまま普通語に変換する。

- ・接頭辞+名詞・形容詞+接尾辞

付加された接頭辞、接尾辞を削除することで普通語化を行う。

「お地藏さま」→「地藏」

「おいしい」→「美しい」

- ・接頭辞+名詞+補助動詞

接頭辞は削除し、補助動詞に関しては名詞の種類に依存する。

i)サ変名詞

動詞「する」+接続助詞「て」を補間する。

「ご相談戴く」→「相談してもらう」

ただし、補助動詞が「する」の意味を持つ敬語動詞である場合補間を行わない。

「ご配慮いたす」→「配慮する」

ii)一般名詞

格助詞「を」を補間する。

「お礼申し上げます」→「礼を言う」

- ・接頭辞+動詞+補助動詞

接頭辞は削除し、補助動詞に関しては次の規則に従う。

i)補助動詞が「する」の意味を持つか、「になる」である。

補助動詞を削除する。このとき本動詞の活用形は補助動詞に依存する。

「お書きになった」→「書いた」

「お読みなさい」→「読め」

ii)補助動詞がi)以外である。

接続助詞「て、で」を補間。どちらを補間するかは本動詞に依存する。

「お書きいただく」→「書いてもらう」

「お読みくださる」→「読んでくれる」

- ・動詞+補助動詞 or 助動詞

補助動詞、助動詞を削除することで普通語化を行う。変換後の活用形は削除動詞に依存する。

「書き換え」→「書け」
「読まれた」→「読んだ」

2)置き換え形式

・名詞、形容詞

対訳辞書の登録語彙とマッチングする語彙に変換する。

「逝去」→「死去」
「宜しい」→「良い」

・動詞

基本的には対訳辞書にマッチングする語彙に変換する。

「差し上げる」→「与える」
「存じ上げる」→「知る」

多義の敬語語彙の場合、意味を同定してから対応普通語に変換する。意味同定に関しては次節で述べる。

「ゆっくりとお話を伺いましょう」
→「ゆっくりと話を聞きましょう」

なお、変換後の語彙の活用形は変換前の動詞に依存する。

3.3.2.補助動詞による意味同定

多義語のうち、その意味を補助動詞として扱われるか否かで判定できる語彙が存在する。「いただく」の普通語語彙は、「もらう」、「たべる・飲む」で、前者は補助動詞として扱えるが後者は補助動詞として扱うことができない。従って、「いただく」が補助動詞として使われているときは、その意味を「もらう」に同定することができる。このように意味を同定できる語彙は、「あがる」「差し上げる」などがある。

3.3.3.日本語語彙大系[9]を用いた意味同定

3.3.2.節の条件に当てはまらない場合、つまり補助動詞として意味を同定できないとき、または自立した動詞として扱われるとき、意味を同定するために日本語語彙大系から意味体系、構文体系を利用する。意味体系から名詞の意味属性を、構文体系から動詞の文型パターンを利用する。名詞の意味属性については、「シソーラス自動獲得システムの構築に関する研究」で作成されたシステム[11]を用いる。動詞の文型パターンは、該当箇所を人手で抽出したものをを用いる。以下使用例を示す。

変換前の文：

- 1)「小学校の運動会でメダルをいただいた」
- 2)「小学校の運動会で林檎をいただいた」
- 3)「小学校の運動会でジュースをいただいた」
- 4)「小学校の運動会で林檎を友達からいただいた」

上記 1)~4)で用いられる文型パターンと名詞の意味属性階層構造をそれぞれ表 2、3 に示す。

表 2.「いただく」文型パターン

文型パターン	普通語
N1 を N2 から/より/に~, N1[*] N2[3 主体]	貰う
N1 を~, N1[-856 飲物・たばこ 838 食料 534 生物]	食べる
N1 を~, N1[857 飲物]	飲む

表 3.名詞の意味属性階層構造情報

名詞	意味属性階層構造
メダル	[1 名詞/2 具体/533 具体物/706 無生物/707 自然物/893 道具/925 目印・象徴物/927 記章] [1 名詞/1000 抽象/1001 抽象物/1002 抽象物(精神)/1062 言語/1080 言語/1099 記号/1102 紋章]
林檎	[1 名詞/2 具体/533 具体物/534 生物/671 植物/672 植物(個体)/673 樹木/674 果樹] [1 名詞/2 具体/533 具体物/706 無生物/707 自然物/838 食料/853 嗜好品等/854 果物]
ジュース	[1 名詞/2 具体/533 具体物/706 無生物/707 自然物/838 食料/853 嗜好品等/856 飲物・たばこ/857 飲物/859 コーヒー・ジュース] [1 名詞/1000 抽象/1235 事/1236 人間活動/1560 行為/1657 娯楽/1680 スポーツ]
友達	[1 名詞/2 具体/3 主体/4 人/5 人間/111 人間 (対人関係)/112 人間 (交際関係) /124 友・なじみ/125 友人]

1)表 3 に示した、名詞「メダル」の意味属性から[838 食料 534 生物]でないことがわかるので、表 2 の「貰う」文型パターンに合致するため、意味が「貰う」と見なされる。

2)表 3 に示した、名詞「林檎」の意味属性から[838 食料 534 生物]に該当することがわかるので、表 2 の「食べる」文型パターンに合致するため、意味が「食べる」と見なされる。

3)表 3 に示した、名詞「ジュース」の意味属性から[838 食料]と[857 飲物]に該当するので、「食べる」か「飲む」に意味を絞れるが、「食べる」文型パターンに、[-856 飲物・たばこ]が存在する。一が付く意味属性は、その文型パターンでは使われないことを意味するので、「貰う」は除外され、意味が「飲む」と見なされる。

4)表 3 に示した、名詞「林檎」の意味属性から「食べる」の文型パターンに該当するが、名詞「友達」の意味属性が[3 主体]であることから、最長一致法により、表 2 の「貰う」文型パターンが適用され、意味が「貰う」と見なされる。

従って 1)~4)の文は、以下のように変換される。

変換後の文：

- 1)「小学校の運動会でメダルを貰った」
- 2)「小学校の運動会で林檎を食べた」
- 3)「小学校の運動会でジュースを飲んだ」
- 4)「小学校の運動会で林檎を友達から貰った」

3.4.敬語構造

変換処理の過程で得られる敬語の構造は以下の通りである。

- 1)敬語の始点と終点
- 2)敬語の種類(置き換え or 付加形式)
- 3)敬語の種類(尊敬 or 謙譲 or 丁寧)

これらは、システムの内部データであるが、表示することで、敬語の扱われ方を示すことができる。本システムでは、敬語の始点に開括弧を、終点に閉括弧を充て、開括弧の直前にその敬語の分類情報 3)を付加したものを敬語構造情報とした。表 4 に実際の出力例を示す。

表 4.敬語構造出力例

敬語文	敬語構造
私が部長に伺ってきます	私が部長に(伺っ[謙])てきます
そこにおられる人は、校長です	そこに((おら[寧])れる[尊])人は、校長です。

4.実行例

前述の変換処理に基づいたシステムを用いて手紙で使われる文章を普通語文へ変換した。その実行結果を以下に示す。出力文では、下線部が変換適用箇所である。

入力文：

このたびご尊父様のご急逝の悲報に接し、あまりにも突然のことで驚いております。ここに謹んで哀悼の意を表するとともに、ご冥福をお祈り申し上げます。ご帰宅なさってから急にお具合が悪くなられたとか。寝食を共になさったご尊父様ご他界のお悲しみはいかほどかとお察し申し上げます。お慰める言葉もございませんが、お力落としなくなお一層のご健闘をお祈りいたします。それがご尊父様へのなによりのご供養と存じ上げます。同封のものは誠に少額にて恐縮ですが、ご尊父様のお好きだったお菓子でもお供えくたされれば幸いです。本来ならばすぐ参上いたしましてご焼香いたすところでございますが、私、先月より高血圧にて東大寺病院に入院中の身でございますのでご容赦賜りますようお願い申し上げます。末筆ながらご自愛のほどお祈りいたします。

敬語構造出力：

このたび(ご[尊父][尊]様[尊])の(ご急逝[尊])の悲報に接し、あまりにも突然のことで驚いて(お[寧])ります。ここに謹んで哀悼の意を表するとともに、(ご冥福[尊])を(お祈り申し上げ[謙])ます。(ご帰宅なさつ[尊])てから急にお(お具合[尊])が悪く(なられ[尊])たとか。寝食を共に(なさつ[尊])た(ご[尊父][尊]様[尊])(ご他界[尊])の(お悲しみ[尊])はいかほどかとお(お察し申し上げ[謙])ます。(お慰める[謙])言葉も(ご[寧])ませんが、(お力落とし[尊])なくなお一層の(ご健闘[尊])を(お祈りいたし[寧])ます。それが(ご[尊父][尊]様[尊])への(ご供養[尊])と(存じ上げ[謙])ます。同封のものは誠に少額にて恐縮ですが、(ご[尊父][尊]様[尊])の(お好きだっ[寧])たお菓子でも(お供えくたされ[尊])ば幸いです(ご[寧])ります。本来ならばすぐ(参上いたし[謙])まして(ご焼香いたす[寧])ところで(ご[寧])りますが、私、先月より高血圧にて東大寺病院に入院中の身で(ご[寧])りますので(ご容赦[尊])(賜わり[謙])ますよう(お願い申し上げ[謙])ます。末筆ながら(ご自愛[尊])の(ご[寧])お祈りいたし[寧])ます。

出力文：

このたび父の急逝の悲報に接し、あまりにも突然のことで驚いています。ここに謹んで哀悼の意を表するとともに、冥福を祈りを言います。帰宅してから急に具合が悪くなったとか。寝食を共にした父世界の悲しみはいかほどかと察して言います。慰めをする言葉もありますが、力落としなくなお一層の健闘を祈ります。それが父へのなにより

の供養と知ります。同封のものは誠に少額にて恐縮ですが、父の好きだったお菓子でも供えてくれれば幸いです。本来ならばすぐ訪問しまして焼香するところでありすが、私、先月より高血圧にて東大寺病院に入院中の身でありますので容赦を貰いますようお願いいたします。末筆ながら自愛のほど祈ります。

概ね良好な結果が出力されたが、網掛け部分は不自然に変換された箇所である。以下に網掛け部分についての考察を述べる。

1)「父」

「ご尊父様」は、相手の父親のことを述べている個所だが、変換適用後は「父」だけになり相手の父という情報が抜け落ちている。しかしながら、敬語構造情報と照らし合わせれば、尊敬語という情報が得られるので「父」が自分側ではなく相手側の人物を指していることが分かる。

2)「冥福を祈りを言います。」

格助詞「を」が二回続く不自然な日本語になっている。これは、「お祈り申し上げる」の茶釜での解析結果で「祈り」が動詞ではなく「名詞」として取れているために生じている。仮に「祈り」が動詞として取れていたとすれば、変換結果は「冥福を祈って言います」となる。

5.おわりに

本研究では、形態素解析、名詞の意味属性など、比較的簡素な要素を用いて敬語文を普通語文へ自動的に変換するシステムを構築した。4章で示したように、ある程度良好な変換結果を得た。

本システムは、敬語を解析し普通語へ変換することが目的であったが、敬語を解析することを利用して、多重敬語に関する誤用判定なども今後の展望としては一考の価値があると考えられる。また、敬語構造情報の表示が少々見づらいので、改善する必要がある。

参考文献

- [1]菊地康人:敬語、談話社会学文庫(1997)
- [2]横山晶一:敬語表現を常体に直すシステム、筑波大学プロジェクト報告書(2001)
- [3]林四郎:敬語の体系、明治書院(1974)
- [4]杉村領一:日本語待遇表現の解析と状況意味論、コンピュータソフトウェア vol.3 No.4 (1986) pp.350-358
- [5]羽根木規:常体から敬体への変換システムの作成、山形大学卒業論文(2000)
- [6]大野満、横山晶一、西原典孝:敬語表現変換システムの試み、情報処理学会東北支部 2000 年度第 4 回研究会 (2001)
- [7]大野満:敬語表現変換システムに関する研究、山形大学卒業論文(2001)
- [8]奥山益郎:状況分類敬語用法辞典、東京堂(1999)
- [9]日本語語彙大系、NTT コミュニケーション科学研究所 (1997)
- [10]形態素解析システム「茶釜」、奈良先端科学技術大学院大学
- [11]阿部亮介:シソーラス自動獲得システムの構築に関する研究、山形大学卒業論文(2001)