

敬語表現への自動変換システムの構築*

明河 直樹 蒲地 俊太郎 堀 智允 韓 東力
 日本大学文理学部 情報システム解析学科

1 はじめに

敬語表現への自動変換システムはビジネスの場でもとても利便性のあるものである。既存研究の中には、文章全体から「相手」や「場」のレベルと文単位の主語の取得を行い、敬語の種類を決定する変換モデルの提案[1]や、動詞のみに着目し敬語に換言するもの[2]があるが、文中の人物の上下関係を考慮して敬語表現に正しく変換できない等の問題点がある。

本研究では係り受け解析を行うことによって、文中に複数の主語が出てきた場合でも、人物の上下関係を明らかにし、各主語に対応した動詞をそれぞれ適切な敬語表現に換言することができるようなシステムを構築した。さらに「取得できなかった主語の補充」を行うことによりシステムの性能向上を図った。

本稿ではこれらの手法を用いて構築したシステムの各モジュールと評価実験について述べる。

2 情報抽出と属性付与

本研究で構築したシステムの流れは図1のようになっている。ビジネス文書と文書の送り手、送り先の人物情報をシステムに入力すると、敬語表現に変換された文書が出力される。ここで送り手と送り先の人物情報である「名前」と「役職」を入力するのは、文中から人物間の上下関係を取得した際、より正確に人物に情報を付与する為である。

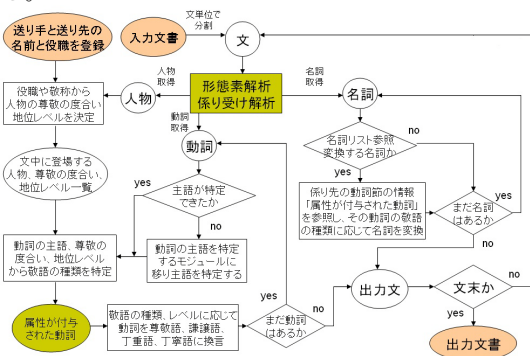


図1 システムの流れ

2.1 基本情報の抽出

まず文から動詞、名詞、係り受け関係等の基本情報を取得する為に係り受け解析器 Cabocha を用いる[3]。

ここで取得した係り受け関係を用いることにより、重文等における複数の述語のそれぞれに対し

て、主語を特定することを可能にした。以下に、ある入力文を Cabocha で解析した時、出力される文構造と係り受け関係の例を図2に挙げる。

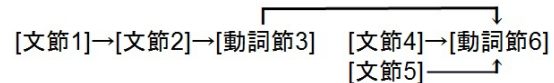


図2 入力文の係り受け関係例

文節3と文節6は動詞が含まれる為、動詞節と呼ぶ。各動詞節の主語を特定する際の候補となる文節は、「各動詞節の係り元の文節」である。つまり「動詞節3」の主語を特定する段階で考慮する文節は、係り元の文節「文節2」であり、「動詞節6」の場合は「文節4」と「文節5」である。

2.2 人物情報の取得

本システムでは以下のようなアルゴリズムを用いて文中の人物についての情報を取得する。

- ①文を文節ごとに区切る
- ②助詞の主格「は」「が」「も」、所有・目的格「を」「の」「と」「に」「へ」「まで」が付随している単語を抽出
- ③EDR 概念辞書[4]を用いて、抜き出した単語が人間を表す単語であるかどうかを判定
- ④人間と判定できた単語を我々が用意した「自分辞書」と「相手辞書」、始めに入力した送り手と送り先の情報と照合
- ⑤文中に出てきた人物を 自分 or 相手 or 第三者と表記してまとめる

ここで「自分」や「相手」等と表記をまとめるのは、人物の情報を整理し、のちに敬語の種類を決定する段階で用いる為である。

2.3 尊敬の度合いと地位レベルの取得

人物に役職や敬称が付随していた場合、人物に尊敬の度合いと地位レベルを付与する。表1で示すように、尊敬の度合いを3段階に分類し、度合いによって尊敬語、謙譲語の変換を使い分ける。

表1 地位レベル(一部)

尊敬の度合い1	地位レベル	尊敬の度合い2	地位レベル	尊敬の度合い3	地位レベル
社員	0	さん	2	副社長	12
先輩	1	主任	3	社長	13
		係長	4	会長	14
		課長	5	...	
		...		様	21
		専務	11		

* Automatic Paraphrase into Honorific Form

度合い3がもっとも丁寧な敬語へと変換されるようになっている。地位レベルとは人物間における上下関係のことであり、敬語の種類を特定する際に用いられ、人物に付与された数字が大きい方が目上の人物にあたる。

3 動詞の主語特定

動詞の主語特定においては、文の係り受け関係を用いる。動詞の係り元の節で助詞の主格「は」、「が」、「も」が付随している単語が人間であれば、その単語を動詞の主語と特定する。以下の例1で例を示す。

例1：私^は部長に/言いました。
 (“私は”は“言いました。”に係っている)
 ⇒ “言いました。”の主語は“私”と特定

日本語という言語は主語が度々省略される言語であり、この段階を終えても主語が特定できていない動詞があった場合、各種の言語的要素を用い、主語の補完を行う。

補完に係わる言語的要素として、「非人間動詞」、「推量・様子・伝聞」、「助述表現の語彙的特性」、「内部感情を表す形容詞・形容動詞」、「格に関する特性」という5つの規則を用いて主語を補完する。

3. 1 非人間動詞

主語が人間になりえない動詞は尊敬語・謙讓語に変換する必要が無い為、主語の取得ができなかった動詞が「主語が人間になりえない動詞」(以後、本稿では非人間動詞と呼ぶ)であるかどうかを判定する。

判定は文献[2]を参考に、EDR 辞書内にある日本語動詞共起パターン副辞書の表層格情報を利用し、ガ格になる名詞の意味素性が“人間”でない動詞を非人間動詞とする。非人間動詞と判定された場合、主語を補完せず動詞を丁寧語に変換する。次に非人間動詞の例を例2で示す。

例2：沸騰する、故障する

3. 2 推量・様子・伝聞

主語の取れなかった動詞に推量、様子、伝聞を意味する表現が使われている場合、一部の例外を除き、自分が主語にはなり得ないと考えられる。

推量、様子、伝聞を意味する表現が動詞の直後にある場合、予め用意してある「推量・様子・伝聞表現リスト」と照らし合わせ、動詞の“主語 = 相手 or 第三者”として補完する。以下の図3で“推量、様子、伝聞を意味する表現の一覧”を示す。

～かもしれない/～かもしれません	～まい ^ㇿ
～(する)そうだ/そうです	～(よ)う ^ㇿ
～(し)そうだ/そうです	～らしい(です) ^ㇿ
～だろう/でしょう	～ようだ/ようです ^ㇿ

図3 推量、様子、伝聞を意味する表現の一覧

例3：海外に行くらしい、開業するそうだ

3. 3 助述表現の語彙的特性

日本語のある助述表現(文末表現)には、「誰々が」「誰々に」という情報が含まれている為、省略を補完するための機能が含まれている[5]。一部の例外を除き、“主語 = 自分”と補完できるものと、“主語 = 相手 or 第三者”に補完できるものがある。こちらも予め用意してあるリストと照らし合わせ、補完を行う。以下の図4と図5で“助述表現の一覧”を示す。

願望を意味する「～(し)たい」
試行を意味する「～てみる」
勧誘を意味する「～しましょう」
「～しよう」

図4 主語 = 自分 と補完できる助述表現の一覧

忠告を意味する「～(ほう)がよい」
義務を意味する「～なければならぬ」
許可を意味する「～ても構わない」

図5 主語 = 相手 or 第三者 と補完できる助述表現の一覧

3. 4 内部感情を表す形容詞・形容動詞

日本語の内部感情を表す形容詞、形容動詞は、その主語として話し手自身、つまり自分しかとり得ない[6]。このことより、内部感情を表す形容詞、形容動詞の主語が省略されている場合、“主語 = 自分”として補完する。判定は予め用意してある“内部感情を表す形容詞・形容動詞リスト”を用いる。

内部感情を表す形容詞、形容動詞リストは日本語語彙大系[7]にある、用言意味属性体系の「心的状態」から抽出し、リストを作成した。

例4：楽しい、悲しい

3. 5 格に関する特性

1つの用言の複数の格に同じ人名が入ることは稀である為[8]、文中の表層格(ヲ格、ノ格、ト格、ニ格、へ格、マデ格)にある人物は動詞の主語として補完できないよう制限する。ただし、第一人称の場合は複数の格に使われる場合が比較的に多いので、この規則は“相手”もしくは“第三者”に限定する。

3. 6 補完の流れ

上記5つの性質を利用し、複数の補完候補を1つに絞る。ここではスロットを用意し、各々のモジュールで得られた主語補完候補によって、補完する主語のスロットに“1”、補完すべきでない主語のスロットには“0”を入力する(“1”は初期値を示す)。これを規則ごとにすべて行い、各スロットの主語候補毎にAND演算を行いスロットの

最終状態を決める。最後はそのスロットを左から探索し、最初の“1”の場所に応じた主語を補充することとする。以下の図6で例を示す。

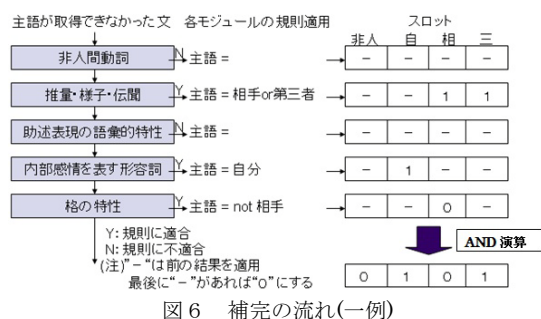


図6 補充の流れ(一例)

5つの規則を適用した結果、候補が複数あった場合、下記の優先順位で主語を補充する。もし候補が無い場合は主語を自分とする。

(非人間>) 自分>相手>第三者

準備実験を行い考察した結果、非人間動詞が正しい判定をする確度をもっとも高く、またビジネス文書では第一人称が圧倒的に省略されることが多い(約8割)ことが分かった。さらに、次いで相手、第三者という傾向が見られたので、上記のように優先順位を決定した。

4 敬語表現への換言

これまでに得た情報を用い、動詞がどの種類の敬語に変換されるかを特定した後、文を敬語表現へと変換していく。

4.1 敬語の種類とレベル判定

以下の手順で敬語の種類とレベルを決定する。

- ②□ 詞の主語を確認
- ②送り手(自分)と送り先(相手)との間の上下関係に応じて敬語の種類を特定
- ③尊敬の度合いに応じて敬語レベルを特定
- ④動詞に敬語の種類とレベルを属性として付与

本研究では敬語の種類を「尊敬語」「謙讓語」「丁寧語」「丁寧語」の4つに定めた。既存研究[1]では丁寧語を使用しなかったのに対し、本研究で新たに「丁寧語」を加えた理由は、謙讓表現をより正確にする為である。動詞がどの種類の敬語に変換されるかは、以下の条件を元に決定する。

- 尊敬語へ変換
 - ・(主語=相手)and(相手>自分)
 - ・(主語=第三者)and(第三者>(自分 and 相手))
- 謙讓語へ変換
 - ・(主語=自分)and(相手>自分)
- 丁寧語へ変換
 - ・(主語=第三者)and((自分 and 第三者)<相手)
- 丁寧語へ変換
 - ・尊敬レベル0、謙讓レベル0
 - ・主語が取得できていない

A>Bの表記は人物Aの地位レベルが人物Bより上であることを表す。尊敬の度合いに応じて尊敬語、謙讓語に敬語レベルを付与する。ここで参照する尊敬の度合いは、「相手」に付与されている尊敬の度合いである。「相手」に付与されている尊敬の度合いが1であれば、尊敬語レベル1あるいは謙讓語レベル1と決定する。丁寧語、丁寧語については、尊敬の度合いによって使い分ける必要がないので、敬語レベルを付与しない。

4.2 動詞の敬語表現への換言

動詞に付与された「敬語の種類とレベル」情報を参照し、動詞を敬語表現へと換言する。動詞の変換は二つのパターンに分かれる。

a.特殊変換

「言う」等の動詞は単語自身が「おっしゃる」等に変化する為、予めリストを用意し、リスト内の動詞であれば対応する変換を行い、活用形、活用の仕方に応じ語尾も変化させる。その際、変換後の単語の活用形を用いる。なお、丁寧語は全てこの特殊変換にあたる。以下の表2に尊敬語特殊変換リストの一部を示す。

表2 尊敬語特殊変換リスト(一部)

一般	尊敬 レベル1表現	尊敬 レベル2表現	尊敬 レベル1活用形	尊敬 レベル2活用形
する	される	なさる	一段	五段・ラ行特殊
いる	いらっしゃる	いらっしゃる	五段・ラ行特殊	五段・ラ行特殊
言う	言われる	おっしゃる	五段・ワ行促音便	五段・ラ行特殊
行く	いらっしゃる	お出かけになる	五段・ラ行特殊	五段・ラ行

b.規則変換

特殊変換を行う動詞以外は、表3のルールを用いる。特殊変換同様、活用形と変換前の動詞の活用の仕方によって変換後の動詞の語尾も修正する。

表3 規則変換のルール

敬語の種類/レベル	変換規則
尊敬語 レベル2	お+動詞(連用形)+になる
尊敬語 レベル1	動詞(未然形)+れる 動詞(未然形)+られる(母音動詞)
謙讓語 レベル2	お+動詞(連用形)+いたす
謙讓語 レベル1	お+動詞(連用形)+する

4.3 文全体の変換

文章に対して一文ずつ取り出し変換を行っていく。このとき必要な情報はCabocha解析で抽出した文の構造、動詞の場合ほどの種類の敬語・レベルに変換するかという情報である。文を構成している単語の品詞を確認していき、品詞が動詞であれば特殊変換、規則変換を用いて動詞の換言を行う。ただし、必ずしも動詞が敬語へと換言されるとは限らない。主語が「非人間」となっているものに対して換言を行わず、それ以外の動詞でも換言を行わない場合がある。以下に変換しない動詞について説明する。

a. 動詞の種類が「非自立動詞」の場合

その動詞のみで文節を構成することができない「非自立動詞」については変換を行わない。

例5：仕事してきた(きた=非自立動詞)

b. 文構造が「自立動詞」+「自立動詞」の場合

動詞が2つ続いて意味を持つ場合もあり、このような場合、前の「自立動詞」の変換は行わない。

例6：言い添える、取り急ぐ

c. 助詞「連体化」、助動詞「体言接続」が付随している動詞の場合

動詞が名詞に係り、一つの意味を持つ名詞となるので変換は行わない。

例7：励ましの言葉、動きそうな絵

動詞が変換された場合、変換前の動詞の活用形が「未然形」、「連用形」、「仮定形」と「命令形」のいずれであるかにより、語尾を変化させる。また、一文の動詞全てが変換され、文末まで処理が終わったら、文末を丁寧語表記に変換を行う。この作業を文がなくなるまで繰り返す。

4. 4 名詞の変換

ビジネスの場でよく使われる名詞をいくつか抜粋したものをリスト化し、文中にリスト中の名詞があれば、変換を行う。

表4 ビジネスでよく使われる名詞(一部)

一般表現	尊敬表現	謙譲表現
会社	御社	弊社
企業	御社	弊社
銀行	貴行	当行

名詞の変換については以下の規則を用いる。

●名詞+「は」「が」「も」の場合

(助詞が主格の場合)

係り先の動詞節が

- ・尊敬語に変換された場合→尊敬表現に変換
- ・謙譲、丁寧語に変換された場合→謙譲表現に変換

●名詞+「の」「に」「へ」等の場合

(助詞が所有・目的格の場合)

係り先の動詞節が

- ・尊敬語に変換された場合→謙譲表現に変換
- ・謙譲、丁寧語に変換された場合→尊敬表現に変換

5 評価実験

本システムの有効性を検証するため、主語補完実験とアンケートによる評価実験を行った。

5. 1 主語補完の実験

予め用意してある実験用データから、無作為に選んだ26の文章(主語省略箇所が100箇所)をシステムに入力し、システムが補完した主語の精度を測る。その結果、省略箇所100個の内77個が正しく補完できた。

5. 2 アンケートによる評価

本研究に関わっていない5人にアンケートを受けてもらった。実験データから無作為に10の文章を選び、それらをシステムに入力し、敬語表現へ換言した。変換前の文章と変換後の文章を被験者に提示し、○・×・△の3段階での評価をしてもらった。評価はそれぞれ

○→変換が正しく行われており自然な文章

×→変換が正しく行われておらず不自然な文章

△→どちらともいえないが許容範囲内の文章

となっている。○は2、△は1、×は0とした時の平均値を求めたところ、平均値は1.32という数値が出た。

5. 3 考察

主語補完の実験でもアンケートによる評価実験でも、一定の成果を出すことが出来たが、まだまだ改善の余地はあると考えられる。より細かい規則や条件を追加することにより、更なる精度の向上が期待できる。

6 おわりに

本研究では、文章を敬語表現へと変換し、ビジネスの場で文書のやり取りを支援するシステムを構築した。係り受け解析を導入することにより、動詞ごとに主語を取得できるようにした。また文中に省略されている主語の補完も試みた。しかし実際の日本語文においては主語が省略されていることが多く、現状の補完手法ではすべての場合に対応し切れない。今後は主語補完の手法をさらに拡張し、より正確で実用性の高いシステムを構築していきたい。

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,情報処理学会自然言語処理研究会報告,2005(94), pp.1-6
- [2]野口聡,南條浩輝,吉見毅彦:動詞の通常表現から敬語表現への換言,言語処理学会第13回年次大会,PC3-1.(2007)
- [3]<http://chasen.org/~taku/software/cabocha/>
- [4]http://www2.nict.go.jp/r/r312/EDR/J_index.html
- [5]工藤育男,友清睦子:日本語の述部の特性を用いた省略の補完機構について,電子情報通信学会論文誌,D-II,Vol.J76-D-II, No.3, pp.624-635.(1993)
- [6]久野暉:談話の文法,大修館書店(1978)
- [7]池原悟,宮崎正弘,白井諭,横尾昭男,中岩浩巳,小倉健太郎,大山芳史,林良彦:日本語彙典大系,岩波書店.(1999)
- [8]磯崎秀樹,賀沢秀人,平尾努:辞書式順序を持つペナルティによるゼロ代名詞解消,情報処理学会論文誌,Vol.47, No.7, pp.2279-2293.(2006)