

日本語発話文に含まれる敬語の誤用を自動的に指摘するシステム —プロトタイプの開発—

白土保* 丸元聡子** 村田真樹* 井佐原均*
* 情報通信研究機構 (NICT), ** 計量計画研究所 (IBS)

1. まえがき

敬語は日本語の重要な特徴のひとつとされており、日本語の敬語は単に依頼、要求あるいは人を示す代名詞において見られるだけでなく、言語体系、及び言語行動のほぼ全般にわたって発達している。このような特徴を持つ言語は日本語以外では、韓国語、チベット語、及びジャワ語など世界中に少数しか見られない (林 1974)。

ところが現代の日本社会において、日本語の敬語に関する様々な誤用が指摘されてきている (菊地 1997; 蒲谷 1998)。日本社会における敬語の誤用は、言語によるコミュニケーションを通じた社会的人間関係の構築を妨げる恐れがある。特にビジネスの場面における敬語の誤用は、時として円滑なビジネスを進める上での障害にもなり得る。このため、一般的には敬語の誤用はできるだけ避けることが望ましい。敬語の誤用を避けるには、敬語の規範に関する正しい知識の習得が不可欠である。このような知識習得を効率的に行なうため、敬語学習を支援する計算機システムの実現が期待される。

以上の背景の下、我々は日本語発話文に含まれる2種類の敬語の誤用、即ち、語形上の誤り、及び運用上の誤りを指摘するシステムのプロトタイプを開発した。本システムは、日本語発話文、及び発話内容に関係する人物間の社会的上下関係を表すラベルを入力とし、入力された日本語発話文における誤用の有無、誤用の箇所、及び誤用の種類 (後者2つは誤用有りの場合のみ) を出力する。これに類似した機能を搭載した日本語入力支援ツールなどが最近開発されてきているが、既存のシステムは、主に語形上の誤りを対象としており、運用上の誤りについては極めて限られた表現しか扱うことができなかった。

本システムのように、発話文に含まれる敬語の誤用を指摘するシステムの構築にあたっては、

(1) 敬語の規範を何処に求めるか? 及び (2) 発話状況をどう取り扱うか? が問題になる。本研究では、以下の考え方に基づきこれらの問題に対処している。

(1) 敬語の規範

敬語 (正確には、敬語を含む言語一般) は時代の経過と共に変化する。例えば、二重敬語の語形を持つ敬語 (“お話になられる” 等) は伝統的な国語学においては誤用とされてきたが、近年では必ずしも誤用としては認識しない人が少なからずいることが報告されている (文化庁 1999)。このため、現代の日本において幅広く社会のコンセンサスが得られている敬語の規範の定義はないと考えられる。

この問題に対し、本研究では、多くの国語学者の間で共通して述べられている規範にできるだけ厳密に準拠する、という立場を取る。即ち、少数の学者のみによって主張されている規範や、現代の日本社会において敬語として許容されているが厳密に規範的かどうか不明な敬語、及びその運用法は規範としては認めないこととする。従って、本研究で提案するシステムは、やや過剰に誤用を指摘してしまう可能性がある。しかしこのことは、少しでも誤用の可能性のある表現をできるだけ漏らさずピックアップできる、という利点として考えることもできる。ここで、実用的な観点からは、問題となる表現を誤用として指摘するだけでなく、その誤用が現代の日本社会において実際にどの程度深刻なものであるかについても示すことが望ましい。しかし、このような指標化を行なうには大規模な言語学的調査が必要であるため、本研究の対象とはしていない。今回提案するシステムは、敬語の誤用を指摘するシステムのプロトタイプであり、実用的なシステムの構築に向けた拡張は今後の課題としている。

(2) 発話状況の取り扱い

ブラウンらのポライトネス理論等 (Brown and Levinson 1987) の従来の敬語研究で指摘されているように、発話状況に応じた適切な敬語 (即ち、運用上、正しい敬語) を用いるためには、発話に関わる人物間の社会的上下関係 (年齢差や社会的地位の違い)、人物間の親疎、及び各人物の体面 (Goffman 1967) に対して発話意図が及ぼすリスク、等の要因を考慮することが必要である。このうち、発話に関わる人物間の社会的上下関係に関する認識は、敬語を運用する際に考慮すべき事項として日本語の敬語に関する殆ど全ての文献において明記、あるいは暗示的に述べられている (菊地 1996, 1997; 蒲谷 1998; 国立国語研 1990, 1992)。これは、敬語の運用の規範に関わる要因としては、社会的上下関係が最も重要な要因であることを示唆する。従って本研究においては、敬語の運用上の規範は発話に関わる人物間の社会的上下関係のみに基づいて定義されている。

以下では、本研究における“敬語の誤用”の定義を述べた後、それに基づいた誤用指摘システムについて述べる。更に、様々なテストデータを用いたシステムの評価、及び評価結果に基づくシステムの改善等について考察する。

2. 敬語の誤用

敬語の誤用は、(1) 語形上の誤り、及び (2) 運用上の誤りに大別できる。前述のように、本研究で提案するシステムは両方の誤りを指摘することができる。

(1) 語形上の誤り

語形が敬語として規範的ではない表現の使用。本研究で提案するプロトタイプでは、敬語の文献 (菊地 1996, 1997; 堀川 1969; 宮地 1999) 等を参考に語形上の誤りの表現リストを定義した。

(2) 運用上の誤り

語形は正しいが、発話に関わる人物間の社会的上下関係と整合しない表現の使用。例えば、伝統的な国語学では、話者より社会的に上に位置付けられる聞き手に対する発話文の文末は丁寧にするべきとされている。従って、このような社会的上下関係における発話文として、文末が丁寧でない文は運用上の誤りと見なされる (ここでは、発話に関わる人物間の親疎、及び各人物の体面に対して発話意図が及ぼすリスクは考慮していない)。

3. 敬語誤用指摘システム

3. 1 本システムが対象とする発話文

本システムに入力される発話文は、以下の3つの制約を満たすものとする。

制約1: 発話文には述語が1つだけ含まれる。

制約2: 発話文に関係する人数は2名~4名。

ここで、2名のときは話者、聞き手 (人物名は“L”に固定)、3名のときは話者、聞き手、及び発話文中で参照される人物1名 (人物名を“A”に固定)、4名のときは話者、聞き手、及び発話文中で参照される人物2名 (人物名をそれぞれ“A”, “B”に固定) とする。

制約3: 述語の主語、及び補語に当たる人物を指す代名詞や人物名 (“私”, “L”, “A”, “B”等) は省略しない。

以上の制約は、現行の構文解析や意味解析等を用いた際の、文構造の解析 (特に、述語の主語、補語の同定等) における解析エラーが生じる可能性のあるような複雑な文を排除するために設けた。従って、高い精度の文解析手法が将来開発されれば、これらの制約はより緩やかにできると考えられる。

尚、発話文に関係する人数が5名以上の状況はまれである (ここで、“彼ら”等のように複数人からなるグループの場合は、1グループを擬人化して1名相当とみなす) と考えられるため、日常用いられている発話文の殆どは制約2を満たすと考えられる。また、制約3により、発話文中で参照する人物名は“A”, “B”に固定されているが (“L”については“あなた”等の代名詞も使用可能)、人名の情報などを事前にシステムに登録することによって人名を直接取り扱うことも可能である。

3. 2 システムの入出力

システムの入出力の例を図1に示す。図中、記号“A > B > L > S”は、話者 (S)、聞き手 (L)、人物A、及び人物Bの社会的な上下関係を表すラベル (以下、“上下関係ラベル”と記す) であり、“>”の左側に現れた記号に対応する人物は右側に現れた記号に対応する人物より社会的地位が高いものと定義する。本システムでは、上下関係ラベルが表す人物間の社会的上下関係は常に正しいと想定して処理を行なっている。上下関係ラベルと共に、聞き手 (L) に対する話者 (S) の発話文が入力される。この例では、発話意図: “A

がBに言った” に対応する発話文が入力されている。システムは入力された発話文の語形上の誤り、及び発話文と上下関係ラベルとの整合性をチェックし、誤用の箇所が見つかった場合には、その箇所、及び誤用の種類を出力する。この例では、語形上の誤りは見つからなかったが、話者（S）より社会的地位が高い聞き手（L）に対する発話文において文末が丁寧ではないため、“誤用”と判定されている。

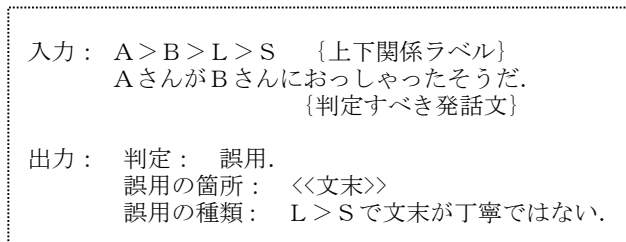


図1：入出力の例

3. 3 処理の流れ

本システムにおける処理の流れを図2に示す。システムに入力された発話文はまず形態素解析され、形態素の並びに変換される。次に、語形誤り表現リストを用いて、語形の誤りがチェックされる。例えば、形態素の並びの中に、“ご” + サ変動詞 + “なす” + “れる” が現れた場合には、語形上の誤り：“ご～なされる”と判断し、その旨を出力して処理を終了する。語形上の誤りが見つからなかった場合には、格解析を行い述語の主語、及び補語の同定を行なう。格解析では、形態素の並びの異なる箇所に、「人物を指す語 + 格助詞（“が” / “を” / “に”等）」が2箇所現れるようなパターンを始め様々なパターンのテンプレートを用意することによって、主語、補語を推定する。例えば、形態素列の中に、人物1を指す語（+敬称）+ “が”、及び、人物2を指す語（+敬称）+ “に” が現れた場合には、述語の主語は人物1、補語は人物2、と推定する（以下、述語の主語と推定された人物を *subj*、述語の補語と推定された人物を *obj* と記す）。図1の発話文に対しては、*subj* がA、*obj* がBと推定される。格解析の結果、及び、敬語タイプ辞書（表1）を用いて、発話文が持つ敬語上の特徴を表すパターン（表2。以下、“敬語特徴パターン”と呼ぶ。）を抽出する。図1の発話文に対しては、 $s=1, o=1, e=0, p=1$ と

なる。最後に、整合表（表3）を用いて、敬語特徴パターンと上下関係ラベルとの整合性がチェックされ、判定結果が出力される。図1の例では、上下関係ラベルがA(=*subj*)>B(=*obj*)>L>Sであるため、表3における $s=1, o=1, p=1$ とは整合するが、 $e=0$ とL>Sは整合しない。従って、この箇所が運用上の誤りと判定される。ここで、表3の右の欄における記号：S、Lは上下関係ラベルと同様、それぞれ話者、聞き手を指す。また、*subj, obj*がSの場合は下記のルールを適用する：

- *subj* = S のときは、 $s=0, p=0(S > obj) / p=2(S < obj)$ 。
- *obj* = S のときは、 $o=0, p=0(S > subj) / p=1(S < subj)$ 。

尚、表3は、発話に関わる人数2名～4名（主に4名）における様々な発話意図に対して、教科書的知識に基づいて作成した正例（教科書的に厳密に正しい発話文 + 上下関係ラベル）からの人手による学習で構築した。

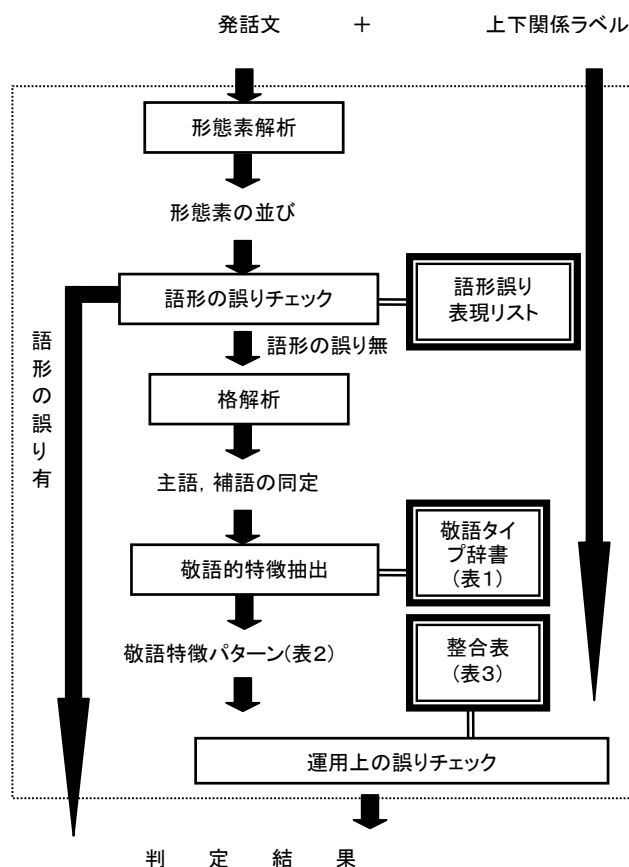


図2 処理の流れ

表1 敬語タイプ辞書 (一部)

形態素の部分的並び	敬語タイプ
A+さん	敬称
B+氏	敬称
L+様	敬称
お+～+する	謙讓語
ご+～+する	謙讓語
頂く	謙讓語
申し上げる	謙讓語
お+～+なさる	尊敬語
ご+～+に+なる	尊敬語
です.	丁寧語

表2 敬語特徴パターン

敬語特徴パターンの各要素の値	条件
s=0	<i>subj</i> の敬称なし
s=1	<i>subj</i> の敬称あり
o=0	<i>obj</i> の敬称なし
o=1	<i>obj</i> の敬称あり
e=0	文末が丁寧でない.
e=1	文末が丁寧.
p=0	述語が尊敬語でも謙讓語でもない.
p=1	述語が尊敬語.
p=2	述語が謙讓語.
p=3	述語が尊敬語かつ謙讓語 (二方面敬語).

表3 整合表

敬語特徴パターンの各要素の値	上下関係
s=0	$S > subj$
s=1	$subj > S$
o=0	$S > obj$
o=1	$obj > S$
e=0	$S > L$
e=1	$L > S$
p=0	$S > subj \wedge S > obj$
p=1	$subj > obj \wedge subj > S$
p=2	$obj > S > subj$
p=1 or 3	$obj > subj > S$

4. むすび

本研究では、日本語の発話文に含まれる誤用を指摘するシステムのプロトタイプを構築した。本システムは、語形上の誤り、及び運用上の誤りを指摘することができる。特に、述語の敬語タイプに関する運用上の誤りを詳細に指摘することができるのが特長である。

システムの妥当性をチェックするため、現在、様々な発話意図、発話に関わる人数：2名／3名／4名、及び主語、補語の全てのバリエーション、の全ての組み合わせに対して、教科書的に厳密に正しいとされる発話文（正例）を準備している。正例は、我々の研究グループとは独立した組織の言語学研究者グループに作成を依頼した。今後は正例を用いたシステムの検証、及び改善を行って行く。

参考文献

- Brown, P. and Levinson, S. 1987. *Politeness - Some universals of language usage*, Cambridge, 1987.
- Goffman, E., *Interaction ritual: essays on face to face behavior*, Garden City, 1967.
- 蒲谷宏, 川口義一, 坂本恵, 敬語表現, 大修館書店, 1998.
- 菊地康人, 敬語再入門, 丸善ライブラリー, 1996.
- 菊地康人, 敬語, 講談社, 1997.
- 国立国語研究所, 日本語教育指導参考書 17 敬語教育の基本問題 (上), 大蔵省印刷局, 1990.
- 国立国語研究所, 日本語教育指導参考書 18 敬語教育の基本問題 (下), 大蔵省印刷局, 1992.
- 林四郎, 南不二男編, 敬語講座 1 敬語の体系, 明治書院, 1974.
- 文化庁文化教育部国語課, 世論調査報告 平成 10 年度 国語に関する世論調査 - 敬語・漢字・外来語 -, 1999.
- 堀川直義, 林四郎, 敬語用例中心ガイド, 明治書院, 1969.
- 宮地裕, 敬語・慣用句表現論 - 現代語の文法と表現の研究(二) -, 明治書院, 1999.