

メッセージにおける 言語表現の分析とその生成

堀井統之 今村賢治 加藤恒昭 大山芳史
NTT情報通信処理研究所

メッセージでは、内容が同じであっても、送り手と受け手の社会的上下関係、親疎関係、送り手の性別などによって、様々な言語表現が用いられる。我々は、送り手と受け手の関係や送り手の属性からメッセージの表層を生成するために、これらの言語表現について検討を続けている。

本稿では、電報文を対象として行なった、言語表現の分析結果について報告する。今回は、特に電報文で顕著に現れる「敬語表現」、「男性表現・女性表現」に着目し、その特徴を抽出した。さらにその結果を取り込んだ簡単なシステムを例に、その生成方法を述べる。

A Method for Generating 'Messages' Using the Analysis of Telegrams

Motoyuki HORII Kenzi IMAMURA Tsuneaki KATO Yoshifumi OUYAMA

NTT Communications and Information Processing Laboratories

1-2356, Take, Yokosuka-shi, Kanagawa 238-03, Japan

We present the analysis of telegrams and propose a method for generating 'Messages'.

'Messages', if their contents are the same, vary their forms according to the social and interpersonal relationships between the sender and the receiver, and the gender of the sender. The analysis of 'Messages' shows that many honorific and male/female expressions are used in telegrams. We obtain their morphological patterns from the analysis and construct an algorithm for generating 'Messages' using the patterns. We also describe experimental systems based on this algorithm, and show that it is effective for generating various messages.

1.はじめに

現在我々は、メッセージ、特に電報文を対象として、その蓄積、検索、生成（合成）について検討を行なっている。

メッセージは書き言葉でありながら、特定の送り手（話し手）から特定の受け手（聞き手）への通信手段である。したがって、送り手と受け手の社会的上下関係、親疎関係、送り手の性別などにより、内容が同じであってもその表層に用いる言語表現が異なってくる。

大量のメッセージを蓄積する場合、内容が同じであるメッセージをその表層の違いによってすべて蓄積しておくのは、効率が悪い。どれか1文だけを蓄積しておき、送り手と受け手の関係や送り手の属性に応じて、リアルタイムにその表層を変換する方法が好ましいと考えられる。

本稿では、電報文に現れる言語表現の分析結果について述べ、さらに簡単なシステムを例にその生成方法について報告する。

2.言語表現の位相

言語は、話し手、書き手といった言語主体の属する社会集団の違いや、言語を使用する場面の違いによって、異なる表現を用いる。このような現象を言語表現の位相と言う[1]（図2-1）。

メッセージにおいても、送り手と受け手の上下関係による丁寧さの違いや親疎関係による親密さの違い、送り手の性別による表現の違い、さらには地域差による言葉の違い（方言）など、様々な位相が現れる。

そこで今回は、電報文の中で顕著に現れる、

- ・送り手・受け手の上下関係による丁寧さの違い
- ・送り手の性別による表現の違い

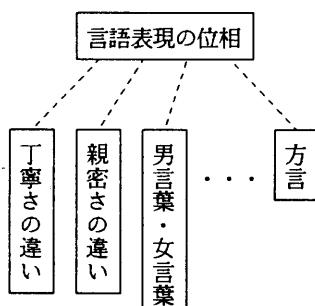


図2-1 言語表現の位相

に着目し、その分析を行なった。それについて、3章、4章で述べる。

3.丁寧さの違いを表す言語表現

文全体の丁寧さ、さらにはメッセージ全体の丁寧さを決定する要因は様々であると考えられるが、ここではその表層的な要因となる敬語表現について分析した。

3.1.敬語表現の現れる単位

敬語表現の形式は、大きく次の2つに分けられる[2]。

- ・敬語的成分を付けるもの

「祝う」に対して、

お祝いします、お祝い申し上げます など

- ・特定の語（敬語表現専用語）を用いるもの

「来る」に対して、

いらっしゃる、おいでになる など

丁寧さの違いは、これらの敬語表現が用いられているかどうか、またどんな敬語表現が用いられているかにより決定される。ここで、敬語表現の現れる単位として、

お二人／の／幸せ／を／お祈りしております／。

のような単位を設定し、これを『パート』と呼ぶ。パートは、丁寧さの違いにより表現を変え得る最小単位であり、「お二人」、「幸せ」、「お祈りしております」は、それぞれ丁寧さの違いによって、

二人-お二人

幸せ-お幸せ

祈っている-祈っています-お祈りしておりますなどの表現をとる。「の」、「を」のように、丁寧さの違いにより表現が変わらない部分については、パートは「単語」に等しい。

また、敬語的成分を付けたり、特定の語を用いることにより、敬語表現を生成することを『敬体化する』と呼び、上記の「二人」、「幸せ」、「祈っている」などのように、敬体化されていないパートを『原形』と呼ぶ。

3.2.用言パートと体言パートの分析

ここでは、敬語表現がよく用いられている用言パート（用言を含んだパート）と体言パート（体言を含んだパート）について、その分析結果を述べる。

(1) 用言パート

用言パートの原形は、原則として、用言とそれに付く助動詞によって構成されており、その敬語表現は用言またはパート末の単語を敬体化することにより生成される。

用言パート内の敬体化について以下に述べる。

用言に現れる敬語表現

動詞においては、その動作主体により尊敬語、謙譲語が使い分けられ、それぞれにいくつかの表現がある。表3-1にそれらの各表現と、数量化手法を用いて得られた丁寧さの違い[3]を示す。その他に、動詞自体が変わる敬語表現専用語がある。

来る → いらっしゃる、おいでになる

言う → おっしゃる、申し上げる

形容詞の場合、接頭語「お」が付く（お美しい）。

パート末に現れる敬語表現

基本的には、「ます」、「です」、「ございます」といった『丁寧語』が付く。

祈ります、お祈りします

美しいです、美しゅうございます

持ってくるそうです

しかし、助動詞がパート末の場合には、

だろう→でしょう→ございましょう

た→ました（でした→ございました）

などのように、特殊な形をとる場合がある。

その他の敬語表現

例外的に、用言及びパート末以外を敬体化する場合がある。

知らないかった（ない+た）→知りませんでした

聞いていた（「聞く」が謙譲語の時）

→うかがっておりました

以上に示したのは、用言パートを単独でみた場合の敬体化であるが、実際の利用においては、その文中の役割

表3-1 尊敬語、謙譲語の体系

	尊敬語の体系	謙譲語の体系
大	お～になる	お～させていただく
↑		お～申し上げる
丁寧さ	～（ら）れる	お～いたす
↓		お～する
小	常体	常体

に応じて、現れ得る敬語表現が異なってくる。

これは以下のいくつかの場合に分けられる。

文末にある場合

文末にある用言パートにおいては、

持ってくるそうだ

→持ってこられるそうです

（→持っておいでになるそうです）

喜んでいるだろう

→お喜びになっているでしょう

などのように、上記で述べたとおりに敬語表現が用いられる。

連用中止の場合

連用中止の用言パートは文末の用言パートと同様に用言を中心としたパートであるが、文末の用言パートと異なり、敬体化されてもパート末に丁寧語が現れない。

ご子息様のご結婚をお喜び申し上げ（×お喜び申し上げまし）、お二人の前途に、幸多かれとお祈りいたします。

接続助詞を伴う場合

接続助詞を伴う用言パートとは、

これから的人生、いろいろとあるでしょうが、雨のあとには虹も出ます。

において、接続助詞「が」を伴っている用言パート「あるでしょう」などのことである。

接続助詞の結び付きの強さに関して、その分類が南[5]によって行なわれているが、主節の主用言パートに丁寧語が現れている時に接続助詞を伴う用言パートに丁寧語が現れるかどうかもこの南の分類に準じている。

すなわち、接続助詞「が」を含んだ従属節のように独立性が高い場合には、上記例のようにその用言パートには主節の主用言パートと同様に丁寧語が付いたほうが自然である。

これから的人生、いろいろとあるだろう（？）が、雨のあとには虹も出ます。

また、接続助詞「と」を含んだ従属節のようにあまり独立性が高くなき場合は、接続助詞を伴う用言パートに丁寧語が付かない方が自然になる。

春が来て、夏が来て、冬を越える（○）と、また春が来ます。

春が来て、夏が来て、冬を越えます（？）と、また春が来ます。

(2) 体言パート

体言パートに現れる敬語表現には、

接頭語「お」「ご」

必ず付く語（ご多幸、お陰）

付け得るもの（ご家庭、お二人）

付かないもの（門出、笑顔、ハネムーン）

敬称（接頭語、接尾語）

～さん、～様、お～さん

敬語表現専用語

君・おまえ→あなた、人→かた

などがある。

4. 言語表現における男女差

- (1) 山のもみじが君たちの燃える愛で、真っ赤に紅葉し始めた。きれいです。おめでとう（あまり燃えすぎて、山火事をおこすなよ）。
- (2) 山のもみじがあなたたちの燃える愛で、真っ赤に紅葉し始めました。きれいです。おめでとう（あまり燃えすぎて、山火事をおこさないでよ）。

上記(1)(2)のメッセージは、同じ内容のメッセージを別の言語表現で表したものであるが、(1)は男性が、(2)は女性が発信したものと感じができる。このように、発信者の性別を感じさせることができ表现を本稿では男性表現、女性表現と呼ぶ。言語表現における男女差の生成では、これらの男性表現、女性表現を生成することを目的とする。

もちろん、男性表現は男性が、女性表現は女性が必ず使うという絶対的なものではなく、男性表現は男性が、女性表現は女性が好んで使うことが多いという相対的なものである[6]。

また、男性表現、女性表現の利用は、受け手や場面によっても変化する。例えば、大衆の前でのスピーチ、演説で、女性が講演者である場合には、男性とほとんど変わらない話し方をする[6]。電報というメッセージに限ってみても、子供に送る誕生日の祝電では、男性が女性表現に近い言語表現を用いる場合がある。しかしながら今回は、言語表現の性差は送り手の性別にのみ左右されるものとして扱い、受け手および場面の影響は考慮しなかった。

4. 1. 男性表現・女性表現の特徴

言語位相の男女差は主として社会言語学の立場から研究されてきた[7][8]。彼らは、歴史的、社会的立場からその差と原因を整理しているが、計算機による表層の生成を考える場合、工学的な立場から再整理を行なう必要がある。

社会言語学等で知られている男性表現と女性表現の差異のうち、形態素調整で表層を生成することが可能と思われるものには以下に示すものがある[6][7][9]。

- ①同じ内容、同じ心的距離の相手に発話する場合、男性に比べ女性の方が丁寧な表現を用いることが多い。
- ②終助詞によって代表される文末表現は、性差ニュアンスを持つものが多い。その際、男性は主張を強める表現、女性は主張を和らげる表現を多く使う。
- ③人称代名詞が異なる。
- ④男性は、「でけー」「うるせー」等の標準形からの逸脱形を使うことがある。俗語、卑語の使用も多い。
- ⑤女性は、「お入口」「おコーヒー」などの美化語、「おひや」「むらさき」などの隠語女房詞を使うことがある。
- ⑥感嘆詞は、男性・女性で異なることがある。
- ⑦女性は男性に比べ、「すごーく」等の強意語や、「すてき」等の甘いムードをかもし出す形容詞を多く使う。

この他、構文的な特徴として、女性に多く主語の省略がみられる等がある。

上記の男女差のうち、③、④、⑥、⑦は単語自体の置き換えで、⑤は接辞の付与・削除及び単語の置き換えで、②は用言文節の助動詞・終助詞の変更により、表層の生成が可能であると考えられる。①については、3章で論じた特徴があるため、丁寧さの違いを表す敬語表現の問題となる。

4. 2. 男性表現→女性表現変換用分析

4. 1. で述べた特徴が実際のメッセージにどのように現れるか確認するために、書き換え実験により分析を行なった。表層構造の男女差の生成を行なうためには、男性表現→女性表現の相互変換または中間言語→男女性表現生成を検討すべきであるが、今回は男性表現→女性表現の一方向変換に限り分析を行なった。作業に当たっては、例文集より収集した電報文約1700メッセージを女性3名に分担して、「あなたならどう書くか」とい

表4-1 男性表現→女性表現書き換え分析結果

分類	出現回数	書換パターン例	例
体言文節	丁寧さの変更	6 7	名詞→「お」または「ご」+名詞 名詞→名詞+「さん」 金→お金 おでん屋→おでん屋さん
	人称代名詞等	1 5 1	人称代名詞→人称代名詞 固有名詞+「君」→固有名詞+「さん」 我々→私たち ○○くん→○○さん
	俗語	1 8	ボイン→大きな胸 お袋→お母さん
	卑語	2	悪ガキ→悪い子
	文語的名詞	4	当方→私ども
	男女名詞	3	妻→奥様
用言文節	丁寧さの変更	2 9 0	名詞+「だ」→名詞+「です」 未然形+「よう」→未然形+「ましょう」 甘い涙だ→甘い涙です 満喫しよう→満喫しましょう
	文末表現	1 4 5	終止形+「な(禁止)」→未然形+「ないで」 命令形+「よ」→連用形+「ね」 メソメソするな→メソメソしないで 聞かせろよ→聞かせてね
	標準形からの逸脱形	9	標準形からの逸脱形は、標準形にした 上で書き換える 一杯やっか→一杯どう しっかりせいよ→しっかりしなさいよ
	俗語	2 2	俗語は標準形に書き換える 足を洗った→抜け出した うまい→おいしい
	古語・文語	9	古語・文語は、現代語・口語にする 奮闘すべし→奮闘しなさい 叶わぬ→叶いません
その他	感嘆詞	1	やったぜベイビー→やったね
	接続詞	8	だが→けれど
	未分類	3 4	文の持つ意味が変わったもの等 広げることです→広げましょう

う観点から書き換えを行なってもらい、その書き換わった部分を中心に分析を行なった。従って、書き換わった部分は、男性表現のものが女性表現、または女性でも男性でも使用する中立表現になっている。

なお、書き換えの際には、メッセージの話題、構造等を変化させず、形態を変化させるのみにとどめるよう留意した。

表4-1は、メッセージ書き換え結果を、書き換え元（男性表現）の観点から整理したものである。出現回数は、全メッセージ中、何回書き換わったかをカウントしたもので、「一杯やっか」→「一杯どう」等、標準形からの逸脱形と俗語が混在しているものは、両方にカウントしてある。以下、分析結果を列挙する。

(1) 出現回数

出現回数を比較すると、丁寧さの変更、人称代名詞等の変更、文末表現の変更が多く、全体の約80%を占め

ている。未分類を全体に含めない場合、約90%を占めている。従って、この3種の変換をサポートすれば、男性表現→女性表現へはかなりの部分を変換できることになる。

(2) 丁寧さの変更

丁寧さの変更は、体言文節、用言文節の両方に現れているが、どちらも丁寧さが上がる方に書き換えられており、下がる方に書き換えられたものはなかった。

(3) 体言文節

体言文節に関する書き換えは245回あったが、そのうち、220回までが人を表す名詞が書き換わっていた。今回の書き換えは、男性が送ったメッセージを、同じ状況下で女性が送った場合の書き換えであるが、送り手の性別が変わったことにより、人間関係が変化したため、送り手・受け手自身、または送り手・受け手と関係のあ

る人の呼び方が変わったものと思われる。

また、「妻」→「奥様」など、性別が含まれる名詞の書き換えで、俗語・卑語・人称代名詞等に分類できなかったもの（表4-1では男女名詞として分類）があったことも、人間関係の変化に起因しているものと思われる。

従って、男性表現→女性表現変換を行なう場合、人間関係の正確な把握を行なわないと誤った変換を行なう可能性がある。

（4）用言文節・俗語

用言文節の俗語として分類したものの中には、比喩表現に近いものがいくつかあった。例えば、結婚の祝電のメッセージで、

真面目一方の君が、あっという間に、わが社のミス〇〇をおとすとは

↓

真面目一方のあなたが、あっという間に、わが社のミス〇〇をお嫁にするとは

などである。比喩表現に関する変換は未検討であるが、「おとす」→「お嫁にする」に関しては、変換対象分野を結婚の祝電に限定しないと不可能のように思われる。

5. 簡単なシステムによる評価

ここでは、3章、4章で述べた、我々の得た知見をもとに作成した簡単なシステムによって行なった実験の評価結果を示す。

5.1. 丁寧さ変換システム

5.1.1. システムの構成

本システムの構成を図5-1に示す。

入力される電報文は1文単位であり、あらかじめパートごとに区切られている。各パートは、3章で述べたような、パート末に丁寧語を付けるかどうかなどの情報ももっており、各パート単位に辞書・テーブルを用いて、指定された丁寧さの表現に変換する。

丁寧さは10段階（最も丁寧な表現が10、最もぞんざいな表現が0）であり、入力される電報文にはその文の取り得る丁寧さの範囲の情報も付与されている。これは、たとえば、

ご両親様のお喜びもひとしおでございましょう。
という文のような、ぞんざいな表現には出来ないよう

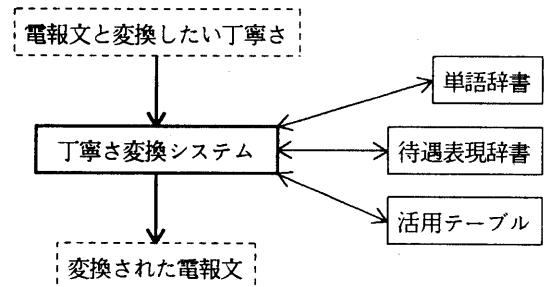


図5-1 丁寧さ変換システムの構成

文に対して、5≤取り得る丁寧さ≤10といった範囲を与えることにより、不自然な文の生成を防いでいる。

(g) 両親の喜びもひとしおだろう。(x)

逆に、丁寧になると不自然になるような文もあり、この場合にも適用できる。

単語辞書は、表記、品詞、活用などといった情報の他に、丁寧さに応じた敬語表現のパターンが記述されている待遇表現辞書へのポインタをもっており、これらの辞書と活用テーブルを利用して変換を行なう。

5.1.2. 評価

本システムを用いて丁寧さが変換された様々な文について、その評価を行なった。

入力電報文は22文で、その22文から合計127文が生成された（1文につき2~8のバラエティ、次ページの図5-2に一例を示す）。被験者にはその127文の自然さと、それぞれの入力電報文から生成されたバラエティの丁寧さの順番が正しいかどうかについて評価してもらった。

その結果を以下に示す。

自然な文が生成された割合

・・・127文中108文（約86%）

丁寧さの順番が正しかった割合

（上記で不自然とされた文は除去している）

・・・22文中20文（約91%）

いずれもかなり高い数値を示しており、本システムを用いた丁寧さ変換方法の有効性が示されている。

不自然な文が生成された例としては、敬語表現が重複して過剰になり過ぎた場合が多い。

お若いお二人のご結婚を祝します。（？）

これはパートレベルではなく、文レベルでの調整が必要な部分である。

また、丁寧さの順番がシステムと被験者で違った例は、

入力電報文：二人の幸せと、ますますの繁栄を、祈る。

丁寧さ	丁寧さの変換された文
1～3	二人の幸せと、ますますの繁栄を、祈る。
4	二人の幸せと、ますますの繁栄を、祈ります。
5	お二人のお幸せと、ますますのご繁栄を、心から祈ります。
6～7	お二人のお幸せと、ますますのご繁栄を、心からお祈りします。
8	お二人のお幸せと、ますますのご繁栄を、心からお祈りいたします。
9～10	お二人のお幸せと、ますますのご繁栄を、心からお祈り申し上げます。

図5-2 システムによる丁寧さの変換例

いずれも「～させていただく」と「～申し上げる」の丁寧さの順番である。被験者は「～申し上げる」の方を丁寧だとしている。「～させていただく」に関しては、もともと関西方言から入ってきた謙譲表現形式なので[10]、これを最も丁寧とするかどうかや、謙譲表現として認めかどかは、個人によって見解が異なる。この個人差も、メッセージを扱う場合、大きな検討課題と言える。

5.2. 男女表現変換システム

4.2. の分析結果から、男性表現→女性表現変換ルールを作成し、図5-3のシステムで評価を行なった。

図5-3において、男性表現→女性表現変換ルールは、男性表現を構成する形態素列とそれに対応する女性表現の形態素列のペアであり、表4-1における書き換えパターンとはほぼ同じものである。丁寧さの変更に関しては、丁寧さ変換システムを用いるべきであるが、今回は男性表現→女性表現変換ルールの中に含めて評価を行なった。

4.2. で述べた分析結果のうち、約950メッセージ分の分析結果から75ルールを作成し、ルール作成に用いた電報文202メッセージ（結婚:202）と、他の電報文177メッセージ（結婚:135、誕生日:28、母

の日:6、父の日:8）について出力文の品質を女性にチェックしてもらったところ、表5-1の結果が得られた。

表5-1 男性表現→女性表現変換評価結果

		分析電報文	分析外電報文		
成功	無変換成功	115(57%)	80%	37(21%)	50%
	変換成功	47(23%)		51(29%)	
失敗	無変換失敗	6(3%)		0(0%)	
	変換失敗	18(9%)	20%	79(45%)	50%
	品質不良	16(8%)		10(6%)	

表5-1における品質不良とは、日本語として間違ってはいないが、読んだときに奇異な感じを受けるもので、変換したために七五調が崩れたメッセージ等が含まれる。

この結果によると、女性表現に変換すべきであるにも関わらず変換しなかった失敗（無変換失敗）は非常に少ない。このことは、ルールにマッチするものが多く、変換しきっていることを示している。従って、変換を抑制するルールの充実が必要不可欠である。

変換を失敗したもの（変換失敗および品質不良）についてその原因を分析したところ、ルールの充実で対処可能なものとして、

- ①文末のみ変換可能な文節を、文末以外でも変換した。
- ②方言／古語を誤って変換した。
- ③「ください」は命令形であるため、命令形の変換ルールを適用し、「くださってください」に変換した。
- ④既に女性表現となっている文を変換した。

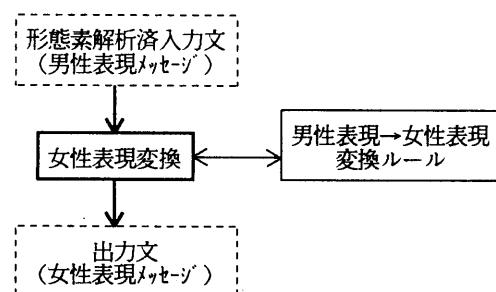


図5-3 男性表現→女性表現変換評価システム

がある。また、ルールの充実では対処不可能なものとして、

⑤引用句／決まり文句を変換した。

⑥七五調の文を変換したために、読んだときのリズムが狂った。

⑦主体が通常とは異なる文を変換した。

結婚はおれが先輩。わからないこと教えます。

↓

ご結婚はわたしが先輩。わからないこと教えます。

⑧単体では「お」「ご」をつけるが、複合語となり、

「お」「ご」を付けられなくなった名詞に「お」「ご」を付加した。

⑨文節の丁寧さをあげたが、前後とのバランスが悪くなつた。

があった。変換ルールの充実では対処不可能なものの中には、文節の丁寧さをバランスよく上げる等、丁寧さ変換システムで既に実現済みのものもあり、男女表現変換システムと丁寧さ変換システムの融合は必要不可欠である。

6. 位相変換システムへの拡張

今回は、丁寧さの違いを表す言語表現及び男性表現・女性表現について分析し、丁寧さの変換、男性表現から女性表現への変換を、簡単なシステムにより実現した。

内容が同じメッセージの言語表現を、様々な場面、状況に応じて生成するためには、今回検討していない親密さの違いを表す言語表現や方言などについても分析を行ない、言語表現の各位相の変換を1つのシステムで実現しなくてはならない。

このような位相変換システムを考える場合、各位相間で共通に扱える部分と独立に対処する部分との切り分けが重要である。すなわち、各位相間の関係を明らかにすることにより、共通に処理できる部分は共通化する必要があるということである。5.2.でも触れたが、今回扱った丁寧さの変換と男性表現から女性表現への変換に関しては、ある程度共通化可能であることがわかった。また親密さの変換についても、丁寧さの変換や男性表現・女性表現の変換と関係があることは容易に予想できる。

したがって、今回のような各位相ごとの変換システムから、共通に扱える部分を1つの同じ枠組みで捉えていくことにより、効率的なメッセージ変換処理を行なうことができると考えられる。

7. おわりに

本稿では、電報文を対象として、メッセージに現れる言語表現について分析を行なった。特に丁寧さの違いを表す言語表現、男性表現・女性表現に関しては、その特徴を抽出し、簡単なシステムによりその生成方法を検討した。

今後は、今回明らかになった課題を克服するとともに、メッセージ全般を扱える汎用的なシステムの実現に向けて、検討を続けていく予定である。

[謝辞]

本研究の機会を与えて下さいましたNTT情報通信処理研究所メッセージシステム研究部坂井陽一部長に深く感謝いたします。

また、日頃ご指導頂いている東田正信グループリーダーをはじめ、同研究部の皆様にも感謝いたします。

[参考文献]

- [1] 小川、林他：「日本語教育事典」、大修館書店、1982
- [2] 文化庁：日本語教育指導参考書2「待遇表現」、大蔵省印刷局、1971
- [3] 堀井、加藤、大山：「メッセージにおける待遇表現の数量化」、情報処理学会第40回全国大会(5F-6)、1990
- [4] 水谷静夫他：朝倉日本語新講座5「運用Ⅰ」、朝倉書店、1983
- [5] 南：「現代日本語の構造」、大修館書店、1974
- [6] 益岡、田窪：「基礎日本語文法」(p.201～204)、くろしお出版、1989
- [7] 國廣哲彌編：「日英語比較講座・第5巻・文化と社会」第3版(p.151～169)、大修館書店、1986
- [8] ジェニファー・コーン：「女と男のことば」、研究社出版、1990
- [9] 荻野綱男：「社会が言語に与える影響」、信学会ソフトウェア科学会「言語とその環境シンポジウム」、1990
- [10] 林大他：「図説日本語」、角川書店、1982