待遇表現の丁寧さの計算モデル － 語尾の付加による待遇値変化 －

白 土 保[†] 井 佐 原 均[†]

待遇表現の丁寧さの計算モデルとして、待遇表現に語尾を付加した際の待遇値（待遇表現の丁寧さの度合い）の変化に関する定量的なモデルを提案した。このモデルでは（１）それぞれの待遇表現に対し、その表現が用いられるべき状況が待遇値に関する正規分布として表される、（２）それぞれの語尾に対し、その語尾が付加される待遇表現が用いられるべき状況が待遇値に関する正規分布として表される、というふたつの仮定を立て、待遇表現に語尾を付加した際に得られる情報量を定義した。そして更に、語尾の付加による待遇値の変化量は、付加の際得られる情報量に関する一次式で表すことができる、という仮定を立て、語尾の付加による待遇値の変化量を、語尾が付加される前の待遇表現に対する待遇値の関数として定義した。このモデルの妥当性を検証するため、ふたつの異なった発話状況において用いられる待遇表現のグループそれぞれに対し、語尾の付加による待遇値変化を求める心理実験を行った。その結果、いずれのグループにおいても語尾の付加による待遇値変化は、提案されたモデルによって予測された傾向に従い、モデルの妥当性が支持された。

キーワード： 待遇表現, 計算モデル

A computational model for politeness of expressions - changes of politeness with word endings addition -

TAMOTSU SHIRADO[†] and HITOSHI ISAHARA[†]

A computational model for the degree of politeness changes with the addition of word endings in polite expressions is proposed. In this model, two stochastic features are assumed as follows: (1) for each polite expression, the situation where the polite expression is likely to use can be expressed as a probability distribution of politeness value, and (2) for each ending word, the situation corresponding to the most suitable polite expression to which each ending word adds can be expressed as a probability distribution of politeness value. The degree of politeness changes with the addition of word endings is defined by the amount of information obtained from the addition of word endings. The result of the psychological experiments supports the validity of the proposed model.

KeyWords: *polite expression, computational model*

[†] 郵政省通信総合研究所 関西先端研究センター, Kansai Advanced Research Center, Communications Research Laboratory, M.P.T.

1 はじめに

従来の自然言語処理研究の多くは、言語の論理的側面に注目したものであった。しかし、計算機が人間と同じように自然言語を取り扱うことができるようになるためには、言語の論理的な取り扱いだけでなく、言語が人間の感性に及ぼす働きの実装が不可欠である。このような観点から、我々は感性を取り扱うことのできる自然言語処理システムの開発に向けた基礎研究のひとつとして、待遇表現の計算モデルに関する研究を行っている。

待遇表現とは、話し手が、聞き手及び話題に含まれる人物と自分との間に、尊卑、優劣、利害、疎遠等のような関係があるかを認識し、その認識を言語形式の上に表したものである（鈴木 1984）。本研究ではこれらの関係を総称して**待遇関係**と呼び、待遇表現に対して心理上持つ丁寧さの度合いを**待遇値**と呼ぶ。

さまざまな待遇表現を柔軟に取り扱うことができる自然言語処理システムを構築するためには、待遇表現の構成要素と待遇表現全体の待遇値の関係を記述するモデルが必要である。しかし、数学的な形式化に重点を置いた研究（水谷 1995）、心理実験による待遇表現の計量化に重点を置いた研究（荻野 1984）、あるいは丁寧さを考慮した文生成プログラムの開発（田中 1983）などの従来の待遇表現に関する研究においては、このようなモデルの提案、及び心理実験に基づくモデルの検証は行われていなかった。

本研究では、話し手及び聞き手以外の人に関する話題が含まれないような発話内容に関する待遇表現に限定した上で、待遇表現に語尾を付加した際の待遇値の変化に関する計算モデルを提案する。モデルの妥当性の検証を行うため、(1) ある事柄について**知っている**という意図を伝える際に用いられる待遇表現のグループに対し、語尾:“よ”を付加した際の待遇値変化、及び(2) 聞き手が会議などで**発言するか否か**を聞き手に質問する際に用いられる待遇表現のグループに対し、語尾:“ます?”を付加した際の待遇値変化を求める心理実験を行った。実験の結果、いずれのグループにおける待遇値変化もモデルから予測される傾向に従い、モデルの妥当性が支持された。

2 語尾の付加による待遇値変化の計算モデル

2.1 モデルが想定している発話状況

本研究では、第三者（話し手及び聞き手以外の人）に関する話題が含まれないような発話内容に関する待遇表現に限定したモデルを提案している。ただし第三者に関する話題が含まれるような待遇表現に対しても、その表現を話し手と聞き手の待遇関係、話し手と第三者の待遇関係、及び聞き手と第三者の待遇関係、のそれぞれに対応した構成要素に分けることが可能な場合は、今回提案するモデルをそれぞれの要素に対して独立に適用することにより、ある程度の

モデル化が可能であると考えられる。

2.2 待遇値の確率分布

荻野はクロス集計表に基づく待遇表現の計量化の研究において、ほとんどすべての待遇表現の待遇値は一次元の値として表現できることを示した（荻野 1986）。本研究ではこの結果をふまえ、任意の待遇表現 P は一次元の値として計量化できると考え、待遇表現 P を計量化した値を $V(P)$ と記す。

$V(P)$ は、計量化の手法に依存した尺度空間上で割り当てられた値であり、本来は待遇値（待遇表現に対して心理上持つ丁寧さの度合い）とは異なるが、ここでは $V(P)$ を単に待遇値と呼び、待遇表現 P に対して心理上持つ丁寧さの度合いと同義に用いる。

ひとつの待遇表現はいろいろな待遇関係の場面において用いることが可能であるが、これはそれぞれの待遇表現 P に対し、その表現が用いられるべき待遇関係を表す待遇値が一定の確率分布を持っており、 P の計量化によってその確率分布の平均が $V(P)$ として観測されていることを示唆する。ここでは、確率分布を正規分布（ N_P と記す）であるとし、仮定 1 を提案する。

仮定 1: 待遇表現 P (待遇値 $V(P)$) に対し、 P が用いられるべき待遇関係が、待遇値に関する平均 $V(P)$ の正規分布（ N_P ）として表される。

図 1 に待遇表現 P_1 : “知ってます” (待遇値 $V(P_1)$)、及び P_2 : “存じてます” (待遇値 $V(P_2)$) に対する確率分布の例を示す。

図 1 “知ってます” 及び “存じてます” に対する確率分布の例

2.3 整合度

我々は、待遇表現 P へ語尾 E を付加する際、文法的には間違いでなくてもその付加がしっくりとする、あるいはしっくりとしない、などの印象を持つことがある（前者の例としては、 P : “知ってる” への E : “よ” の付加、後者の例としては、 P : “存じあげております” への E : “よ”

の付加)。

これは、それぞれの語尾 **E** に対し、**E** が付加される待遇表現が用いられるべき待遇関係が、待遇値に関する確率分布 (N_E と記す) を持っており、**E** を実際の待遇表現 **P** に付加する際には、**P** に対する確率分布 N_P と N_E との間の類似性の大小に応じ、しっくりする／しっくりとしない、などの印象の違いが現れるからであると考えることができる。ここでは、 N_E が正規分布であるとし、仮定 2 を提案する。

仮定 2: 語尾 **E** に対し、**E** が付加される待遇表現が用いられるべき待遇関係が、待遇値に関する正規分布 (N_E) として表される。

いま、確率分布 N_P と N_E との間の類似性の大きさを、これらの共通面積の広さ C として表す。以下 C を、**P** と **E** との間の整合度と呼ぶ。

図 2 に、**P**: “知ってます” と **E**: “よ” との間の整合度 C の例を示す。図中 μ_E は正規分布 N_E の平均、 X は N_E を表す確率密度関数と N_P を表す確率密度関数の交点の x 座標を表す。

図 2 整合度 C の例

ここで N_P , N_E の分散をそれぞれ σ_P^2 , σ_E^2 とすると、整合度 C は次式で定義される。

$$C \triangleq \frac{1}{\sigma_P \sqrt{2\pi}} \int_{-\infty}^X e^{-\frac{(x - V(P))^2}{2\sigma_P^2}} dx + \frac{1}{\sigma_E \sqrt{2\pi}} \int_X^{+\infty} e^{-\frac{(x - \mu_E)^2}{2\sigma_E^2}} dx \quad (1)$$

式 (1) は $\mu_E \leq V(P)$ の場合を想定しているが、 $V(P) \leq \mu_E$ の場合も同様の式で定義できる。

2.4 語尾の付加によって得られる情報量と待遇値変化

いま、待遇表現が持つ情報 = 意図を伝えるのに必要最小限の情報 + 話し手と聞き手の待遇関係に応じた丁寧さ (あるいはぞんざいさ) を伝える情報、と考えると待遇値は後者の情報量に対応した値であると考えられる。

また、待遇表現が持つ情報量のうち、意図を伝えるのに必要最小限の情報量は語尾 **E** の付加によってほとんど変化しない、と考えると、**E** の付加による待遇表現の情報量の変化 (即ち、

E の付加によって新たに得られた情報量 I は、話し手と聞き手の待遇関係に応じた丁寧さを伝える情報量の変化にほぼ一致することになる。

従って、**E** の付加による待遇値変化 $\Delta(\mathbf{P}, \mathbf{E})$ と、**E** の付加によって得られた情報量 I との間には一定の関係があると考えられるが、ここではこの関係を線形であるとし、仮定 3 を提案する。

仮定 3: 待遇表現 **P** に語尾 **E** を付加した際の待遇値の変化量 $\Delta(\mathbf{P}, \mathbf{E})$ と、**E** の付加によって得られた情報量 I との間には線形の関係がある。

仮定 3 を式で表したものが、式 (2) である。

$$\Delta(\mathbf{P}, \mathbf{E}) \triangleq k_1 \cdot I + k_2 \quad (2)$$

ただし、係数 k_1, k_2 は **P**, **E** 及び待遇値の計量化方法に依存した定数である。

待遇表現 **P** への語尾 **E** の付加によって得られる情報量 I は、整合度 C を用い $I = \log_e(1/C)$ で与えることができる。なぜなら、整合度 C は **P** によって期待されるすべての待遇関係の空間の中で **E** によって期待される待遇関係が生じる確率を表している、と考えることができ、更にシャノンの情報量（例えば、Abramson 1969）によると確率 p の事象が生起したことを知ったときに得る情報量は $\log_e(1/p)$ で与えられる（ここでは対数の底を e とした単位で情報量を定義する）からである。

以上から、待遇表現 **P** に語尾 **E** を付加した際の待遇値の変化量 $\Delta(\mathbf{P}, \mathbf{E})$ の計算モデルは次式で定義される。

$$\Delta(\mathbf{P}, \mathbf{E}) \triangleq k_1 \cdot \log_e(1/C) + k_2 \quad (3)$$

ただし、 C は式 (1) で定義される整合度。

3 計算モデルから予測される、語尾の付加による待遇値変化

前章で提案した計算モデルから予測される性質として、待遇値変化 $\Delta(\mathbf{P}, \mathbf{E}) = V(\mathbf{P} + \mathbf{E}) - V(\mathbf{P})$ の待遇値 $V(\mathbf{P})$ に関する性質に注目する（ただし “ $\mathbf{P} + \mathbf{E}$ ” は、**P** へ **E** を付加して作られた待遇表現を表す）。

ここでは、個々の待遇表現における待遇値変化については議論せず、待遇表現の集まりにおける待遇値変化の普遍的な特性を調べることを目的とするが、この目的からは数学的な取扱いの簡便さのため確率分布 N_P, N_E の分散 σ_P^2, σ_E^2 に関し $\sigma_P^2 = \sigma_E^2 (= \sigma^2)$ と仮定しても差し支えないと考えられる。このとき式 (1) は、 $X = (\mu_E + V(\mathbf{P}))/2$ に関する式として、式 (4) のように書き直すことができる。

$$C(X) = \frac{2}{\sigma\sqrt{2\pi}} \int_{-\infty}^X e^{-\frac{(x-V(\mathbf{P}))^2}{2\sigma^2}} dx \quad (4)$$

式 (4) を $(x-V(\mathbf{P}))/\sigma=t$ として正規化し, 式 (5) を得る.

$$C(X') = \frac{2}{\sqrt{2\pi}} \int_{-\infty}^{X'} e^{-\frac{t^2}{2}} dt \quad (5)$$

ただし, $X'=(\mu_E - V(\mathbf{P}))/2\sigma$.

$C(X')$ の形の関数は誤差関数と呼ばれ解析的には解けないことが知られている (例えば, 森口 1957). このため $C(X')$ を, より取り扱いやすい関数で近似することを考える.

いま $V(\mathbf{P})$ と μ_E の間で $|\mu_E - V(\mathbf{P})| \leq 5\sigma$ (即ち, $|X'| \leq 2.5$) が満たされるものとする. この制約は, 確率分布 N_P (あるいは N_E) 全体の面積の中で, N_P と N_E の共通部分の面積 $C(X')$ が占める割合が ($\sigma_P = \sigma_E$ の場合には) 1.2 % 以上の場合に相当し, これは $\mathbf{P}$ と $\mathbf{E}$ の任意の組み合わせに関する多くの状況において満たされると思われる.

このとき $C(X')$ を, X' についての一次式 $p(X') = k_3 X' + k_4$ (k_3, k_4 は定数) に関する関数 $e^{-p(X')}$ で近似する. 関数 $e^{-p(X')}$ は, $|X'| \leq 2.5$ において関数 $C(X')$ を数値的によく近似する. なぜなら, $|X'| \leq 2.5$ における関数 $C(X')$ の値を表すデータ群に対し関数 $e^{-p(X')}$ で回帰すると, 回帰の当てはまりの良さを示す決定係数 R^2 (例えば, Snedecor 1972) は約 0.97 となるからである.

よって, $C(X') \simeq e^{-p(X')}$ を式 (3) に代入し, 更に $X'=(\mu_E - V(\mathbf{P}))/2\sigma$ を代入して整理すると式 (6) が得られる.

$$\Delta(\mathbf{P}, \mathbf{E}) \simeq K_1(\mu_E - V(\mathbf{P})) + K_2 \quad (6)$$

ただし, $K_1 = k_1 k_3 / 2\sigma$, $K_2 = k_2 + k_1 k_4$.

以上から, 待遇表現 $\mathbf{P}$ への語尾 $\mathbf{E}$ の付加による待遇値変化 $\Delta(\mathbf{P}, \mathbf{E})$ は, $\mu_E - V(\mathbf{P})$ に関する一次式で表されることが予測される.

4 モデルの妥当性の検証実験

前章で提案されたモデルの妥当性を検証するため, 語尾 $\mathbf{E}$ を固定 (即ち μ_E , 及び σ_E^2 を固定) した状況において, いくつかの異なった待遇表現 $\mathbf{P}$ に $\mathbf{E}$ を付加した際の待遇値変化 $\Delta(\mathbf{P}, \mathbf{E})$ が $V(\mathbf{P})$ の一次式で表されることを確かめるための心理実験を行った.

実験は, 10代~70代の関西在住の日本人男女97名を被験者とし, 一対比較法によって行った. それぞれの待遇表現 (語尾を付加する前の待遇表現, 及び待遇表現に語尾を付加して作られた待遇表現) に対する待遇値は, サーストンの比較判断の法則 (ケース V) による計量化手続きにより求めた.

実際のいろいろな待遇関係の場面において、どの待遇表現が用いられるかには個人差がある。しかし、待遇表現間の丁寧さの大小には普遍性があると考えられる。即ち、この実験で得られる待遇値 $V(P)$ は、各個人がその待遇表現に対して持つ絶対的な値ではなく、待遇表現間の丁寧さの大小の程度に関する相対的かつ普遍的な値と考えられる。実験の詳細は以下の通り。

4.1 実験に用いた表現

実験では2種類の発話状況を想定し、それぞれの状況において用いられる待遇表現のグループを実験刺激とした。

[待遇表現グループ1]

語尾を付加する前の待遇表現としては、ある事柄について**知っている**という意図を伝える際に用いる待遇表現 21 種類（表 1）を用い、各待遇表現に付加する語尾としては、終助詞“よ”を用いた。（従って、計量化の対象となる待遇表現は $21 \times 2=42$ 種類）。

この場合、語尾“よ”の付加により、概して、元の待遇表現がよりぞんざいになることが予測される（中川 1996）。

待遇表現グループ 1 は、少数の被験者を用いた予備的な実験（白土 1996）において刺激として用いられた待遇表現の種類を増やしたものに当たる。

表 1 待遇表現グループ 1

1: 分かる	8: 知ってる	15: 存じております
2: 分かります	9: 知っている	16: 存じあげてます
3: 分かってる	10: 知ってます	17: 存じあげております
4: 分かっている	11: 知っています	18: 承知してる
5: 分かっています	12: 知っております	19: 承知しています
6: 分かっています	13: 存じてます	20: 承知しています
7: 分かっております	14: 存じています	21: 承知しております

[待遇表現グループ 2]

語尾を付加する前の待遇表現としては、聞き手が会議などで**発言するか否か**を聞き手に質問する際に用いる待遇表現 19 種類（表 2）を用い、各待遇表現に付加する語尾としては、助動詞“ます”を用いた（従って、計量化の対象となる表現は $19 \times 2=38$ 種類）。

この場合、語尾が付加される待遇表現の語尾が変化する（例えば、“言う？”＋“ます”は、“言います？”になる）。また、“ます”の付加により、元の待遇表現がより丁寧になることが予測される。

表 2 待遇表現グループ 2

1: 言う？	8: 仰せになる？	14: 述べる？
2: 言われる？	9: 仰せになられる？	15: 述べられる？
3: 話す？	10: しゃべる？	16: お述べる？
4: 話される？	11: しゃべられる？	17: お述べになられる？
5: お話になる？	12: おしゃべりになる？	18: おっしゃる？
6: お話になられる？	13: おしゃべりになられる？	19: おっしゃられる？
7: お話する？		

4.2 一対比較法による心理実験

いま、計量化の対象となる n 個の待遇表現を $P_1, P_2, \dots, P_n$ と記す。一対比較法では、 $P_1, P_2, \dots, P_n$ 中の異なる全ての待遇表現の対（計 nC_2 対）を作り、一対ずつ被験者に呈示する。被験者は呈示された一対の待遇表現のうちいずれの表現がより丁寧な表現だと感じるかを回答するように求められる。また、両方の表現が同じ位丁寧であると感じた場合はその旨回答するように求められる。

4.3 サーストンの比較判断の法則に基づく計量化

一対比較法による実験の結果、待遇表現 P_i が P_j より丁寧だと判断した被験者数を全被験者数で割った値を p_{ij} とする。ただし、両者が同じ位丁寧な表現だと判断した回答に対しては、その回答をした被験者数を p_{ij} 及び p_{ji} の計算にそれぞれ半分ずつ割り振る。更に、 p_{ij} を Z 得点（例えば、田中 1977）で表した値を Z_{ij} とする。一対比較法においては、被験者にふたつの刺激 P_i, P_j の間の丁寧さの大小に関する判断を行わせることになるが、この際、各刺激 P_i, P_j に対する弁別過程がそれぞれ、平均: μ_i, μ_j , 標準偏差: σ_i, σ_j （ただし、 $\sigma_i = \sigma_j = \sigma$ ）の正規分布に従い、両弁別過程の間の相関係数: $r_{ij} = 0$ と仮定する（サーストンの比較判断の法則ケース V）。このとき、次式が成立する（例えば 田中 1977）。

$$\mu_i - \mu_j = Z_{ij} \sqrt{2} \sigma \quad (7)$$

従って最も小さい μ の値 μ_0 を 0 と置くと、式 (7) によって他の全ての μ_i を μ_0 からの相対値として決めることができる（ただし、各 μ_i の値は $\sqrt{2} \sigma$ を単位とした値）。以上によって得られた μ_i を待遇表現 P_i の待遇値とする。

5 実験結果

待遇表現グループ1, 待遇表現グループ2 それぞれに対して得られた待遇値を用い, 横軸 (X軸) に語尾 **E** が付加される前の待遇表現 P_i の待遇値 $V(P_i)$, 縦軸 (Y軸) に **E** の付加による待遇値変化 $\Delta(P_i, E) = V(P_i + E) - V(P_i)$ をプロットした図をそれぞれ図3, 図4に示す. 図の各点に添えられた番号は, 表1, 表2それぞれにおける待遇表現を示す番号である.

図3 語尾“よ”の付加による待遇値変化

図4 語尾“ます”の付加による待遇値変化

各図における点群 $(x_i, y_i) = (V(\mathbf{P}_i), \Delta(\mathbf{P}_i, \mathbf{E}))$, $i=1, 2, \dots, n$ を, 直線 $y=ax+b$ で回帰した. ここで, 回帰パラメタ a, b は次式で推定される (例えば, 肥田 1961).

$$a \triangleq \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (8)$$

$$b \triangleq \bar{y} - a\bar{x} \quad (9)$$

ここで, $\bar{x}, \bar{y}$ はそれぞれ i に関する x_i, y_i の平均, n はサンプル点数である.

回帰直線のデータへの当てはまりの良さは, 次式で定義される決定係数 R^2 により評価した (例えば, 肥田 1961).

$$R^2 \triangleq \frac{\sum_{i=1}^n (\bar{y} - (ax_i + b))^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

更に, 回帰直線の傾きの有意性を検定量 $T=a/\sigma_a$ を用い自由度 $f=n-2$, 危険率 5% で t 検定した (例えば, Snedecor 1972), ただし σ_a は回帰直線の傾きの標準偏差である. 以下, 自由度 f の t 分布における 5% 点を $T_{0.05}(f)$ と記す.

5.1 待遇表現グループ 1

図 3 は, 語尾:“よ” が付加される前の待遇表現の待遇値が大きいほど“よ”の付加による待遇値変化 (待遇値の減少量) が大きくなる傾向があることを示している.

直線 $y=ax+b$ によって回帰した結果, $a=-0.28$, $b=0.71$, $\sigma_a=0.027$, $R^2=0.84$ となった. また, 回帰直線の傾きは有意に負 (検定量 $T=-10.37 < -1.73 = -T_{0.05}(19)$) であった.

5.2 待遇表現グループ 2

図 4 は, 語尾:“ます”の付加による待遇値変化が待遇表現 $\mathbf{P}_{18}$:“おっしゃる?” (図 4 中の矢印のついた点) 付近で最大で, この点を境にして左側の (待遇値がより小さい) 領域で単調増加, 右側の (待遇値がより大きい) 領域で単調減少の傾向を示している.

左右それぞれの領域に含まれる点群に対し, 別々の直線 $y=ax+b$ によって回帰したところ, 左側の領域 ($\mathbf{P}_{18}$ の点を含め, データ点数 12 個) では, $a=0.58$, $b=1.6$, $\sigma_a=0.052$, $R^2=0.925$ となり, 回帰直線の傾きは有意に正 (検定量 $T=11.2 > 1.8 = T_{0.05}(10)$) であった. また右側の領域 ($\mathbf{P}_{18}$ の点を含め, データ点数 8 個) では, $a=-0.54$, $b=8.47$, $\sigma_a=0.194$, $R^2=0.749$ となり, 回帰直線の傾きは有意に負 (検定量 $T=-2.78 < -1.9 = -T_{0.05}(6)$) であった.

6 考察

待遇表現グループ1に対する実験結果は、語尾“よ”の付加による待遇値変化 $\Delta(\mathbf{P}, \mathbf{E})$ が $V(\mathbf{P})$ に関する傾き負の直線でよく近似されることを示唆する。この結果は、 $\Delta(\mathbf{P}, \mathbf{E}) = K_1(\mu_E - V(\mathbf{P})) + K_2$ の係数 K_1 が正定数である場合に当たる。

待遇表現グループ2に対する実験結果は、語尾:“ます”の付加による待遇値変化 $\Delta(\mathbf{P}, \mathbf{E})$ が表現:“おっしゃる”より待遇値が小さい領域においては $V(\mathbf{P})$ に関する傾き正の直線、表現:“おっしゃる”より待遇値が大きい領域においては $V(\mathbf{P})$ に関する傾き負の直線でよく近似されることを示唆する。この結果は、 $\Delta(\mathbf{P}, \mathbf{E}) = K_1(\mu_E - V(\mathbf{P})) + K_2$ の係数 K_1 が前者の領域では負定数、後者の領域では正定数である場合に当たる。

以上の結果は、 $\Delta(\mathbf{P}, \mathbf{E}) = K_1(\mu_E - V(\mathbf{P})) + K_2$ の係数 K_1 が

- (1) $\mu_E < V(\mathbf{P})$ のとき正の定数、
- (2) $\mu_E > V(\mathbf{P})$ のとき負の定数

となっている、と考えることにより以下のように説明が可能である。

[待遇表現グループ1に対する結果の説明]

グループに含まれるすべての待遇表現 $\mathbf{P}_i, i=1, \dots, 21$ に関し、語尾 $\mathbf{E}$:“よ”に対する確率分布 N_E の平均 μ_E との間で、 $\mu_E < V(\mathbf{P}_i)$ が満たされる（即ち、 $\mathbf{P}_i, i=1, \dots, 21$ のいずれの待遇表現も、“よ”との整合性が最も大きい待遇表現より待遇値が大きい領域にある）と考える。このとき係数 K_1 は常に正定数となり、従って $\Delta(\mathbf{P}, \mathbf{E})$ は $V(\mathbf{P})$ に関する傾き負の一次式となる。

[待遇表現グループ2に対する結果の説明]

語尾 $\mathbf{E}$:“ます”に対する確率分布 N_E の平均 μ_E と、 $\mathbf{P}_{18}$:“おっしゃる”に対する待遇値 $V(\mathbf{P}_{18})$ がほぼ等しいと考える。このとき $\mathbf{P}_{18}$ より待遇値が小さい待遇表現 $\mathbf{P}_i, i=1, 2, 3, 4, 5, 7, 10, 11, 12, 14, 15$ に対しては係数 K_1 は負定数となり、従って $\Delta(\mathbf{P}, \mathbf{E})$ は $V(\mathbf{P})$ に関する傾き正の一次式となる。また、 $\mathbf{P}_{18}$ より待遇値が大きい待遇表現 $\mathbf{P}_i, i=6, 8, 9, 13, 16, 17, 19$ に対しては係数 K_1 は正定数となり、従って $\Delta(\mathbf{P}, \mathbf{E})$ は $V(\mathbf{P})$ に関する傾き負の一次式となる。

7 まとめ

待遇表現の丁寧さの計算モデルとして、待遇表現に語尾を付加した際の待遇値の変化に関する定量的なモデルを提案し、心理実験によるモデルの妥当性の検証を行った。今回実験に用いた待遇表現、及び語尾については、実験によって得られた待遇値変化はモデルから予測される傾向に従うことが示され、モデルの妥当性が支持された。

今後、実用性のある定量的なモデルの構築のためには、各待遇表現や各語尾に対する確率分布の平均、分散、係数 K_1 及び K_2 の推定が必要となる。

また今回は、第三者に関する話題を含む発話内容に関する待遇表現は対象としなかったが、今後、このような表現に対する本計算モデルの拡張を検討して行く。

謝辞 心理実験にご協力頂いた、ATR人間情報通信研究所足立整治博士に深謝致します。

参考文献

- Abramson N. (宮川洋訳) (1969). 情報理論入門. 好学社.
- 肥田野直・瀬谷正敏・大川信明 (1961). 心理教育 統計学. 培風館.
- 水谷静夫 (1995). 待遇表現概要. 計量計画研究所.
- 森口繁一他 (1957). 数学公式 II. 岩波書店.
- 中川裕志・小野晋 (1996). “日本語の終助詞の機能 — 「よ」「ね」「な」を中心として —.” 自然言語処理, 3(2), 3-18.
- 荻野綱男 (1984). “敬語の丁寧さを決定するもの.” 数理科学, No. 258, 43-50.
- 荻野綱男 (1986). “待遇表現の社会言語学的研究.” 日本語学, 5(12), 55-63.
- 白土 保・井佐原 均 (1996). “待遇表現の計算モデル — 語尾の付加による待遇値変化について —.” 情処研報, 96-NL-116, 115-120.
- Snedecor G.W. (畑村又好他訳) (1972). 統計的方法. 岩波書店.
- 鈴木一彦・林巨樹編 (1984). 研究資料日本文法 9 敬語法編. 明治書院.
- 田中幸子・林四郎・荻野綱男・樺島忠夫 (1983). 朝倉書店日本語新講座 5 第 3 章. 朝倉書店.
- 田中良久 (1977). 心理学的測定法. 東大出版会.

略歴

白土 保: 1983 年電気通信大学電気通信学部計算機科学科卒業. 同年日本 IBM (株) 入社. 1986 年郵政省電波研究所 (現 通信総合研究所) 入所. 鹿島宇宙通信センター, 平磯宇宙環境センターを経て, 1992 年より関西先端研究センター知的機能研究室勤務. 主任研究官. 感性情報処理の研究に従事. 日本音響学会音楽音響研究会幹事. 電子情報通信学会, 情報処理学会, 日本音響学会各会員.

井佐原 均: 1978 年京都大学工学部電気工学第二学科卒業. 1980 年京都大学大学院工学研究科電気工学専攻修士課程修了. 同年通商産業省工業技術院電子技術総合研究所入所. 1995 年郵政省通信総合研究所入所. 現在, 同関西先端研究センター知的機能研究室長. 自然言語処理の研究に従事. 情報処理学会, 言語処理学会, 人工知能学会, 日本認知科学会, ACL 各会員.

(1997 年 3 月 11 日 受付)

(1997 年 4 月 10 日 再受付)

(1997 年 5 月 6 日 採録)