

# ポライトネス及び配慮表現コーパス作成と分析手法の一考察 ～対人関係を考慮した対話システムの適用に向けて～

## A Study on Politeness and Consideration Expression Corpus Creation and Analysis Method ～ Toward application of dialogue system considering interpersonal relations ~

太田 博三<sup>†</sup>

Hiromitsu Ota

<sup>†</sup>放送大学

The Open university of Japan

9924658973@campus.ouj.ac.jp

### 概要

ディープラーニングが定着し、対話システムや機械翻訳でも配慮表現が取り込まれようとしている。対人関係を配慮した表現にフォーカスしたものである。ここで、本稿は配慮表現に着目し、対話文生成の基となる小規模なコーパスを意識し、コーパス利用と分析手法について考察したものである。ポライトネス・ストラテジーを簡易なベイズ論的アプローチを用いることで、対話システムの質的向上の一助につなげたい。

キーワード：配慮表現、機械翻訳、対話システム、ポライトネス、ベイズ統計学

### 1. はじめに

ディープラーニングの進展は、スポーツ分野での緻密なデータ解析への適用やコスモロジーなどの宇宙分野や天文学への専門分野への応用が試みられている段階にあり、成熟段階にあると言える。

自然言語処理の機械翻訳の分野では、語用論の中の一つのポライトネスが必要との認識の下で、機械学習時の学習データの準備の手間暇がネックとなっている。

もともと、ポライトネスと待遇表現は文化庁の国語審議会等を見る限り、火遊びにもなりかねないため、ここ数年はあまり、言語学や言語社会学等の視点では、研究における新たな進展は見受けられない。

直感的に、コーパスを紙媒体から電子媒体への移行は国レベルのプロジェクトとして、国立国語研究所を中心に進められているものの、面白みに欠けることは、否めないのではないと思われる。

そこで、筆者自ら、ポライトネスや待遇表現、または配偶表現に関するコーパスを収集し、ユーザの運用を想定した、有効な分析手法を検討したものである。

### 2. 意味論から語用論へのブレークスルー

ポライトネスの輸入とグライスなどの概念に、統語論の故ノーム・チョムスキーの束縛理論のように、縛られてしまい、多岐にわたる学説が唱えられているが結論として皆が納得出来ているようでもない、ゆきずまった分野である。しかし、同時にディープラーニングの適用時に味のある潜在性は、誰も否定はしないものでもあると言える。

本稿では、あまり既存の議論には縛られず、対人関係を良好な状態にするものとして、大きく、曖昧に捉え、コーパスの母数や切り口となる項目を、できる限り多く用意できることを優先した。

### 3. 研究のフロー

既存の代表的なコーパスをもとに、代表的な既存のコーパスを用いて、小規模な仮説検証を行ったものである。

### 4. ベイズ論的アプローチの会話ストラテジーの検討

配慮表現に該当する語句や品詞の組み合わせの出現確率を計算し、語句とストラテジーの流れをモデル化し、その傾向を検討したものである。

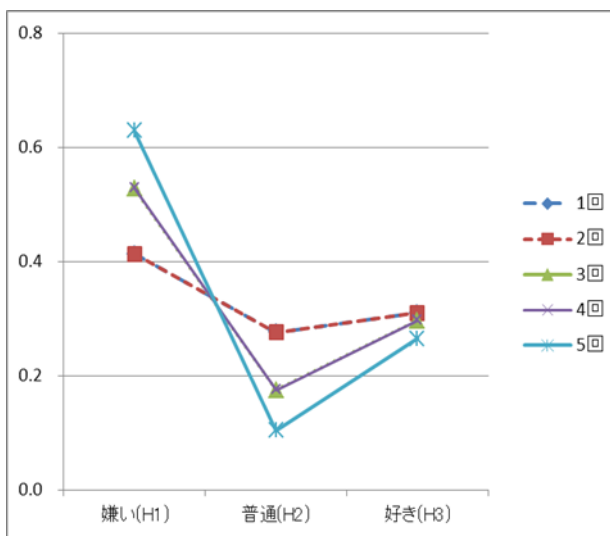
まず、水準を次の3つに定めて、モデル化する。

- 1) ポジティブ
- 2) ゼロ
- 3) ネガティブ

次に、尤度を算出し事前確率を設定する。ここでは、理由不十分の原理から事前確率を等しく設定した。

「はい」と「笑」とをベイズで推定				
(1) 「はいはいはい」と「笑」のモデル化				
	ポジティブ(H <sub>1</sub> )	ゼロ(H <sub>2</sub> )	ネガティブ(H <sub>3</sub> )	
憎因子の数	3	3	3	
愛因子の数	4	2	3	
(2) 尤度の算出				
データ(D)	$P(D H_1)$	$P(D H_2)$	$P(D H_3)$	
悪印象(憎因子選択)	0.429	0.600	0.500	
良印象(愛因子選択)	0.571	0.400	0.500	
(3) 事前確率の設定				
	嫌い	普通	好き	
最初の事前確率	0.3	0.4	0.3	
(4) データ入力と事後確率の算出				
データ回数	データ(D)	嫌い(H <sub>1</sub> )	普通(H <sub>2</sub> )	好き(H <sub>3</sub> )
		0.300	0.400	0.300
1回	良	0.414	0.276	0.310
2回	悪	0.414	0.276	0.310
3回	良	0.527	0.176	0.297
4回	悪	0.527	0.176	0.297
5回	良	0.630	0.105	0.266
6回				
7回				
8回				
9回				
10回				

図 「はい」と「笑い」のポライトネス的ベイズ更新



## 5. 結論

本稿では、1回目のOPI試験を踏まえて、省略と反復・繰り返しの視点から、フィラーや笑いをポライトネス理論の視点も意識して考察したものである。テスターによる「はいはいはい」などは一見すると、強い同意の応答句と捉えがちだが、含意されることは、決して好意的ではなく、OPI特有の「突き上げ

(Probes)」になっている。この突き上げにより、非母国語話者である受験者は、「あっ、その一、えーと(笑い)」などのように、あらゆる非言語的側面が生成されている行為が読み取れた。この部分でも、何とか答えようとしながらも、フィラーや笑いなどの非言語的側面が顕著に出ている。ここでの評価もあり得ると思われる。それは、会話のつながりを進める機能であり、発話と共に、重要な日本語能力であるとも考えられるからである。

1回目より、2回目、2回目より3回目と、人間同士の接触回数が進むにつれて、非言語的側面や語用論的な側面が重要になってくると思われる。

## 6. 今後の展望

今後の展望として、省略と反復・繰り返しは、語用論のポライトネス理論に深く関係しており、前回の接触を踏まえて生成されるものである。非言語的のフィラーや笑い、応答などは、自然な会話では、単体ではなく、ポライトネスの含意とのセットで考慮されるものであると思われる。

## 文献

- 1) 杉山ら(2018)「文脈に沿った発話理解・生成を行うドメイン特化型雑談対話システムの実験的検討」, SIG-SLUD-B802-33, 人工知能学会
- 2) 国立国語研究所「日本語学習者会話データベース 縦断調査編」  
[https://db3.ninjal.ac.jp/judan\\_db/](https://db3.ninjal.ac.jp/judan_db/)
- 3) 久野(1978)『談話の文法』大修館書店
- 4) 堂下・白井・溝口・新美・田中(1998)『音声による人間と機械の対話』「対話過程の基本的特徴と対話における省略の処理(第2章)」オーム社
- 5) ペネロピ・ブラウン・スティーヴン・C・レヴィンソン(1987)「ポライトネス 言語使用における、ある普遍現象」 “Politeness: Some Universals in Language Usage”
- 6) 松原望(2008)「入門ベイズ統計—意思決定の理論と発展」東京書籍
- 7) 福田 一雄(2013)「対人関係の言語学」 開拓社