

# 対話文の英日機械翻訳における日本語待遇表現の生成

6B-2

石原 佳典 亀井 真一郎\* 赤峯 享\*

NEC技術情報システム開発

NEC C&C情報研究所\*

## 1. はじめに

マニュアル文ではなくて対話文を対象とする場合は話し手の願慮が現われる待遇表現を翻訳することが非常に重要となる。

言語は使用される社会や文化に大きく依存しているため、待遇表現を決定する要因や原則は言語毎に異なっている [1-6]。したがって、単語や句の単純な対応規則だけでは適切な翻訳結果が得られない場合がある。

我々は既に社会・文化的要因をパラメータとして抽出し、適切な丁寧度の待遇表現を生成するモデルを提案している [3,4]。今回は前回パラメータとした要因を2つに分類して、モデルの改良をおこない、それを翻訳システムに実現した。

本論文では、待遇表現の丁寧度を決定する要因は命題の内容および発話の状況であると考え、丁寧度を決定するモデルを提案する。そして、そのモデルを英日対話文翻訳システムに適用する手法について述べる。

## 2. 待遇表現の言語による違い

待遇表現が言語によって大きく異なることは良く知られている。例えば、英語の "Tell me your name." という文を常に命令形で「名前を教えろ」と翻訳しては適切でない場合が多い。状況によっては「名前を教えてください」とか、さらに丁寧度の高い「お名前をお教えてくださいませんか」とする必要がある。

日本語は独自の待遇表現の体系を持っており、待遇表現専用の語彙が発達している。具体的な待遇表現の選択は話し手と聞き手の上下関係などによって大きく影響を受ける。それに対して、英語は待遇表現専用の語彙が少なく、表現選択の原則も日本語とは異なる。したがって、英語と日本語の待遇表現を単純に単語や句のレベルで対応させることはできない [1-4]。

## 3. 待遇表現の丁寧度の決定モデル

文を命題とモダリティで構成されていると考え、待遇表現はモダリティの一種であると位置づけられる。

Generating Japanese Polite Expressions in Dialogue-Discourse Machine Translation

ISHIHARA Yoshinori, \*KAMEI Shinichirou, \*AKAMINE Susumu

NEC Scientific Information System Development, Ltd.

\*NEC C&C Information Technology Research Laboratories

同じ命題を述べる文でも待遇表現部分の違いによって具体的な表現にバリエーションが生じる。待遇表現は丁寧度の高低によって並べることができる。しかし、言語によってバリエーションの数は異なるし、英語で一番丁寧度の高い表現が日本語の一番高い表現に常に対応しているわけではないので、日本語と英語の待遇表現を丁寧度順に対応させて処理することはできない (図1)。

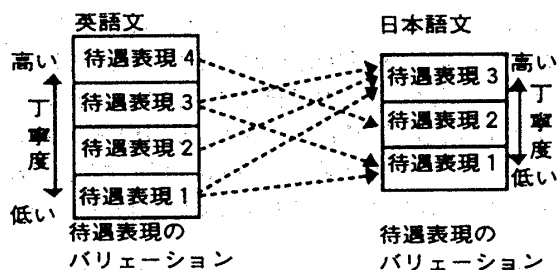


図1

この問題を解決するため、我々は待遇表現の丁寧度を決定するプロセスをモデル化した。そのモデルでは丁寧度を決定する要因を次の二つに分類した。一つは命題の内容で、もう一方は発話の状況である。

命題内容とは文においてモダリティを除いた事実関係だけを表す部分のことである。話し手は命題の内容が聞き手にたいしてどれだけ負担を与えるかを推し量り、負担が大きければ丁寧度を高く、小さければ低くしようとする。すなわち、命題の内容は丁寧度の決定に影響する。命題内容を評価することによって得られる丁寧度はどの言語においても共通であり、命題内容が同じであればどの言語においても同じ丁寧度が得られると考える。本論文ではこれを絶対丁寧度とする。

発話状況とは発話に関与する者の間の上下関係・内外関係などや発話が行なわれる場面 (会議での会話や友達同士での会話など) のことである。命題内容が同じであっても相手が自分よりも上であれば丁寧度を高く、下であれば低くしたり、また、友達同士であっても公的な場では互いに丁寧度を高くしたりする。絶対丁寧度を基準として発話状況の評価し、最終的な丁寧度が計算される。本論文ではこれを表層丁寧度とする。命題内容の評価とちがひ、発話状況の評価が表層丁寧度に与える影響は言語によって異なる。表層丁寧度の決定に際して、日本語はこの発話状況の評価の影響が大きい、日本語と比べ

ると英語は影響が小さい。

待遇表現の丁寧度を決定するモデルは図2のようになる。このモデルの特徴は命題内容の評価によって得られる絶対丁寧度を言語に依存しない普遍的ものと考え、発話状況の評価が表層丁寧度の決定に与える度合いが言語によって異なるとしたことである。

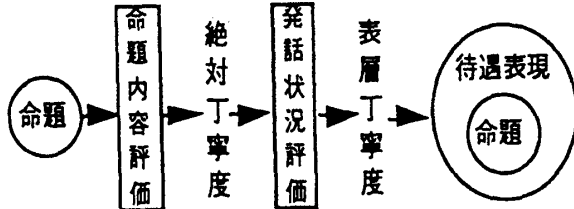


図2

絶対丁寧度と表層丁寧度の関係を図で表現すると図3のようになる。表層丁寧度の範囲は発話状況によって変化する。例えば図3中の発話状況Aでは表層丁寧度の範囲はa-c間になり、発話状況Bではb-d間になる。同じ絶対丁寧度xが、発話状況Aではyとして、発話状況Bではzとして表現される。発話状況によって限定された範囲内での相対位置 (a-c間のy、b-d間のz) は絶対丁寧度におけるxの相対位置を保存していると仮定する。それについて、a-d間でのyとzの位置とxの位置は相対的に異なる。

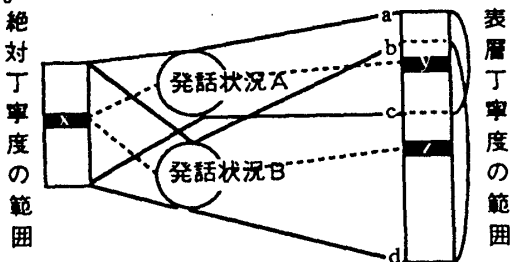


図3

#### 4. モデルのシステムへの適用

図2のモデルの有効性を確認するために実際にインプリメントをおこない、コーパスを用いて、実験をおこなった。このコーパスは会議参加間い合わせにおける申込者と事務局の二者間の対話文例集である[7]。

システムの構成は図4のようになる。英文が入力されると、解析部の待遇表現処理部はその英文の待遇表現の表層丁寧度を計算する。つぎに発話状況の評価し、絶対丁寧度を得る。そして、中間言語に絶対丁寧度を付与して生成部に渡す。生成部は中間言語に付与された絶対丁寧度と発話状況の評価結果から日本語の表層丁寧度を決定し、待遇表現を生成する。

今回の実験の発話状況は「会議参加間い合わせの場面」であり、その発話者は「申込者」と「事務局」に限定されている。発話状況の場面が固定されているので、英語の表層丁寧度の相対的な位置から絶対丁寧度を導きだすようにした。日本語を生成するときには発話者の違いを評価し、日本語の表層丁寧度を決定した。例えば、「I'll

send it to you"が入力され、発話状況の発話者が申込者の場合は「それを送ります。」を生成し、事務局の場合は申込者より丁寧度の高い「それをお送りいたします。」を生成する。

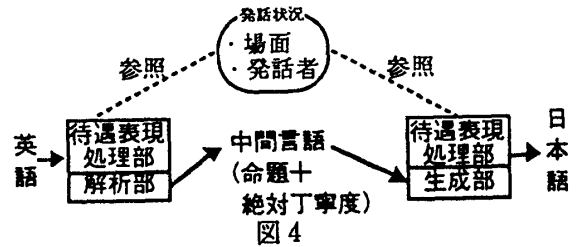


図4

#### 翻訳結果例

入力 Tell me your name.  
事務局 お名前を教えてください。  
申込者 名前を教えてください。

入力 Will you kindly tell me your name?  
事務局 お名前をお教え下さいませんか。  
申込者 名前を教えてくださいませんか。

#### 5. おわりに

本論文では待遇表現の生成要因をモデル化し、命題内容による要因と発話状況による要因とに分けた。会議参加間い合わせに状況を限定したコーパスによる評価では良好な翻訳結果を得ることが出来た。ただし、今回は発話状況を限定したコーパスにおけるインプリメンテーションであったので、今後はコーパスを拡張し、提案したモデルの検証を続けていく予定である。

#### 謝辞

今回のモデルの実験にあたり、ATR自動通訳電話研究所の飯田室長、隅田研究員には対話コーパスを提供していただきました。深く感謝致します。

#### 参考文献

- [1]南、辻村、大石(1977)「日本語4 敬語」岩波書店
- [2]鶴田、ロジャー、クリトン(1988)「英語のソーシャルスキル」大修館書店
- [3]石原、奥村、赤峯(1991)「多言語機械翻訳における「ていねい」表現の生成」人工知能学会第5回全国大会
- [4]Akamine, Ishihara, Okumura, Muraki(1991): Generation of Polite Expression in Multilingual Translation, NLPRS91
- [5]前田、小暮、久米、飯田(1987)「待遇性を考慮した対話文生成」電子情報通信学会技術研究報告NLC86-11
- [6]前田、小暮、久米、飯田(1987)「日本語対話文における敬語表現の解析」情報処理学会第36回全国大会
- [7]江原、小倉、森元(1990)「電話対話データベースの構築」情報処理学会第40回全国大会