

# 日韓機械翻訳システムにおける敬語翻訳に関する研究

金 ナヨン\*      中澤 敏明\*      荒牧 英治†      黒橋 禎夫‡  
東京大学大学院\*      東京大学医学部附属病院†      京都大学大学院‡

{nayoung, nakazawa}@kc.t.u-tokyo.ac.jp\*  
aramaki@hcc.h.u-tokyo.ac.jp†, kuro@nlp.kuee.kyoto-u.ac.jp‡

## 1 はじめに

近年インターネットの爆発的な普及により機械翻訳の必要性が高まっている。これを受けて、これまで様々な言語対で機械翻訳研究がなされてきた。本研究は日韓翻訳に焦点をあてる。日韓両言語は、語順や語彙など他言語と比べると類似性の高い言語であり、特に、両言語とも敬語表現が重要視されるのが大きな特徴である。しかし、これまでの機械翻訳は敬語の翻訳に大きな注意を払ってこなかった。このため、現在の機械翻訳システムは不自然な敬語表現を含んだ翻訳文を生成してしまうことがあり、敬語表現の厳密な韓国語においては、重要な問題となる。本研究は、両言語の敬語表現の特徴を考慮した翻訳システムを提案する。これは、用例ベース翻訳システムをもとに、敬語を処理するモジュールを追加することによって実現される。

## 2 日韓敬語の比較

まず、日韓両言語の違いを述べ、敬語翻訳に必要な知識を考察する。日本語の敬語は大きく「尊敬語」「謙譲語」「丁寧語」の3つに分けられる<sup>1</sup>。韓国語にも「尊敬語」「謙譲語」「恭遜語」が存在し、恭遜語は日本語の丁寧語に対応する。ただし、韓国語の「謙譲語」は退化しつつある[4]。

日本語の敬語は、主に動詞、名詞によって表現されるのに対し、韓国語の敬語は、動詞、名詞に加え、形容詞や助詞においても表現される。韓国語の敬語と日本語の敬語の特徴は以下のようになる：

- (1) 動詞：日本語では、接尾辞「れる、られる」を動詞の語末につけて尊敬の意味を持たせることが

<sup>1</sup>平成18年に文化審議会国語分科会より発表された「敬語の指針」では従来の「尊敬語・謙譲語・丁寧語」を5分類(尊敬語、謙譲語、丁寧語、丁寧語、美化語)した。本研究では従来通りの3分類を扱った。

できる。同じく韓国語でも、一般動詞の語幹の後に尊敬の意味を持っている尊称補助語幹「시(/シ/)」を語末につけて尊敬の意味を持たせることができる。

- (2) 形容詞：尊称補助語幹の「시(/シ/)」は、動詞だけではなく、形容詞にも取り入れられて尊敬語としての働きをする。

目が大きい。      目が大きくていらっしゃる。  
눈이 크다.      눈이 크시다.

- (3) 助詞：「に、が、は」は以下のように敬語表現が2種類存在する。

에게(/エゲ/(に))      께(/ケ/(に))  
이,는(/イ,ガ/(が,は))      께서(/ケソ/(が,は))

- (4) 名詞：日本語の「御言葉」「お車」のように、韓国語でも以下のように、一部の名詞は敬語の意味が含まれている。

- (5) 代名詞：「私」の場合、丁寧語では、「나(/ナ/私)」ではなく、「저(/ゾ/わたくし)」を用いる。また、三人称の彼や彼女の場合、「그(/グ/彼)・그녀(/グニョ/彼女)」ではなく、「그분(/グブン/彼様)」で敬意を表す<sup>2</sup>。

このように、日韓の敬語は表現される場所が異なる。このため、複数の用例を組み合わせる翻訳した場合、不自然な敬語表現を含んでしまう恐れがある。

本研究では、日韓方向に翻訳方向をしぼり、その際に顕著となる日本語で敬語表現がない語(助詞、形容詞)の処理や日本語の名詞と動詞の尊敬表現を手がかりにし、適切な用例の組み合わせを調整することにより、尊敬を考慮した日韓翻訳を実現する。

<sup>2</sup>三人称の場合、性別関係なく그분(/グブン/彼様)で表現する。

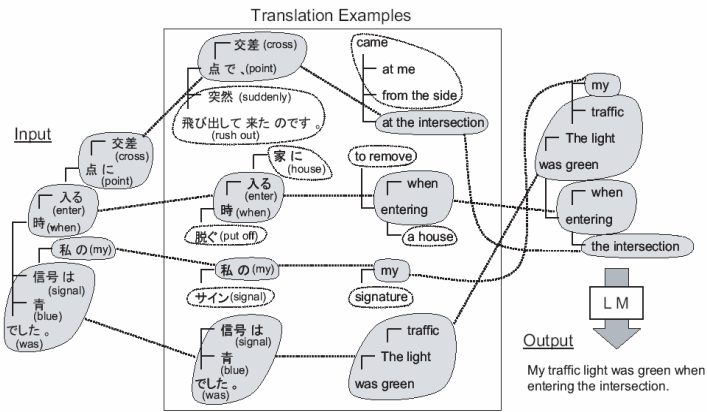


図 1: 日英の翻訳例

### 3 日韓用例ベース機械翻訳システム

まず最初に、本研究のもととなる用例ベース機械翻訳について述べる。用例ベース翻訳とは、入力文と類似した翻訳用例を探し出し、それらを組み合わせることで出力文を作り出す翻訳方式である [1, 3]。図 1 に日英の翻訳例を示す。図 1 に示されるように、入力文を部分的にカバーする複数の翻訳用例が選択され、それらの英語側の表現を組み合わせることで英語構文木を得る。最後に用例間の語順を言語モデルにより調整する。

このように、用例ベース翻訳の基本的な枠組は対象とする言語対に依存しない<sup>3</sup>。

しかし、日韓翻訳の敬語現象は、両言語で異った個所で表現されるため、複数の用例を組み合わせた際に、韓国語側で敬語表現が正しく表現されないことがある。これは、敬語表現に厳密な韓国語においては、重要な問題となりうる。例えば、図 2 上の入力に対して、次のような 3 つの用例があった場合を考える：

- 用例 1: 学生が質問しました / 학생이 질문했습니다.
- 用例 2: 先生が...。 / 선생님어....
- 用例 3: 私に...。 / 나에게....

この場合、これらの用例を組み合わせても、韓国語での敬語表現の整合性がとれない図 2 の中段のような誤訳になってしまう。これを正訳にするためには、次の 3 点を考慮する必要がある。

- (1) 主語が「先生」という尊敬名詞であるため、尊敬助詞「께서」を使う。
- (2) 動詞が「ます形」の丁寧語であるため、「私」の

<sup>3</sup>ただし、機能語の処理や韓国語のパッチムの処理等、いくつか言語別に対応する点は必要である。

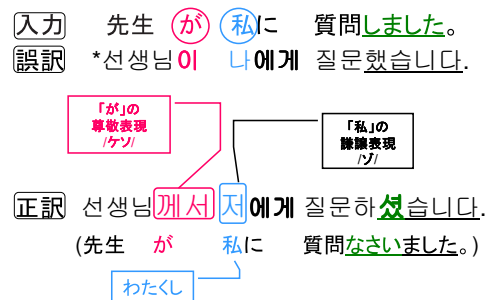


図 2: 敬語翻訳の誤訳と正訳

謙讓表現である「저」を使う。

- (3) 動詞の「質問しました」は主語が尊敬名詞であるため、「質問なさいました」にあてはまる「질문하셨습니다。」を使う。

### 4 敬語翻訳の提案手法

前章の正訳のような翻訳を実現するために、必要な処理は次の二つである。まず、入力となる日本語文にどのような敬語表現が埋め込まれているか捉える必要がある。次に、入力文に存在する敬語表現がどのように韓国語で表現されるのか知る必要がある。本研究では、この知識を、(1) 必要とされる用例のタイプ、(2) 用例の変換ルールという 2 つの形で表現した (図 3)。ただし、入力文にはさまざまな文形が存在し、また文に出現する人物の上下関係にもさまざまなバリエーションがあり、これらの組み合わせは膨大な数になる。そこで、本研究では基本的な文形である SV、SOV を対象とし、上下関係も一部制限した<sup>4</sup>。

<sup>4</sup>「主語」と「目的語」は話者より目上の人物とした。

#### 4.1 敬語情報知識の獲得方法

まず、入力文を解析し、敬語情報をとらえる。日本語の敬語情報は動詞 (4.1.1 節) と名詞 (4.1.2 節) に表現される。

##### 4.1.1 動詞の敬語情報

動詞の敬語情報としては、(1) 聞き手を尊敬するかどうかで、+聞 or - 聞を、(2) 主語を尊敬するかどうかで、+S or - S を、(3) 目的語を尊敬するかどうかで、+O or - O を動詞に付与する。

例えば、「(先生が) おっしゃいました」の場合には、「おっしゃる」から主語を尊敬していることが<sup>5</sup>、「ました」から聞き手を尊敬していることが分かり、動詞の敬語情報は「+S、+聞」となる。

##### 4.1.2 名詞の敬語情報

一般名詞から尊敬される対象の人物(「社長」「先生」など)または、尊敬される人物に付随する物(「お言葉」など)を抽出する。

尊敬される対象の人物かどうかの判定は、敬語動詞の格フレーム [2] の格要素として存在するかどうかを調べ、存在するならば、+人 を付与した。存在しない場合は - 人 を付与した。例えば、「先生」は、敬語動詞「おっしゃる」の格フレームのガ格に存在するため、+人 となる。

また、尊敬される人物に付随する物については、人手で語彙リスト (例: お宅、年、お子さん、御飯等) を作成し、マッチするものについて+物 を付与した。マッチしない場合は - 物を付与した。

#### 4.2 敬語表現変換ルール

提案する翻訳システムは、先の処理で付与した敬語情報と一致する用例を用いる。ただし、動詞の用例検索の際には、その格要素となっている名詞の敬語情報の一致も調べる。この結果、「先生が言った」の翻訳において、「言った」を「말씀하셨습니다(おっしゃいました)」と正しく翻訳できる。

また、人称代名詞や助詞等の処理も必要である。どのようにこれらを修正すればよいかは図 3 の処理に示す。

<sup>5</sup>主語の尊敬判定には形態素解析器 JUMAN の、<敬語: 主体尊敬表現> <尊敬動詞> 情報を使った。

ID		翻訳例文
(d)	入力文	私は家に行きます。(+聞)。
	出力文	저(処理3)는 집에 가겠습니다. (+聞) (わたくしは家に行きます。)
(g)	入力文	彼がいらっしゃいました。(+S)
	出力文	그분(処理2)께서 오셨습니다. (+S) (彼様がいらっしゃいました。)
(m)	入力文	先生はとても親切です。(+聞)。
	出力文	선생님(処理4)께서는 매우 친절하십니다. (+S) (先生様はとても 親切でいらっしゃいます。)
(n)	入力文	先生は私に厳しいです。(+S+聞)。
	出力文	선생님(処理4)께서는 저(処理3)에게 엄하십니다. (+S/+聞) (先生様はわたくしに 厳しくいらっしゃいます。)

図 4: 翻訳結果の例

## 5 実験と考察

図 3 に示す変換ルールに含まれるすべての入力文のタイプを用意し、提案手法の妥当性を調査した(入力文 20 文; 対訳用例 50 対)。その結果、すべてにおいて、適切な翻訳結果を得ることができ、提案する敬語翻訳の原理は妥当であると考えられる。

翻訳結果の例を図 4 に示す(図 4 の ID は図 3 の ID に対応する)。例えば、図 4 最上段の例 (ID(d)) では、日本語「行きます」が丁寧語であり、聞き手を尊敬している(+聞)。この場合、「私」に相当する韓国語を聞き手に敬意を払う「저(/ゾ/わたくし)」にする必要があるが、翻訳システムは変換ルールを参照し、処理 3 を適用することによって、正しい翻訳結果を得ることが出来ている。

以上のように、提案する枠組は原理的には妥当であるものの、実用的なシステムを構築するためには、次のような課題がある。

- 尊敬名詞の問題: 本稿では、格フレームを用いて尊敬される対象の人物を判別したが、「管理人」や「ゲスト」など、尊敬される対象の人物かどうか曖昧な場合も確認された。
- 「ノ格」等の尊敬名詞の処理: 主語と目的語の主辞の名詞のみを用いて敬語情報を付与しているた

日本語文					韓国語文					ID	
主体部		客体部		述部	主体部		客体部		述部		
主語	助詞	目的語	助詞	動/形等	主語	助詞	目的語	助詞	動/形等		
1人称 代名詞 (私)	ガ or ハ			+聞	処理3				+聞	(a)	
		代名詞	3人称	ヲ or ニ	+O、+聞	処理3		処理2	刪	+O、+聞	(b)
					+O、-聞			処理2	刪	+O、-聞	(c)
		一般	-物	ヲ or ニ	+聞	処理3				+聞	(d)
			+人		± O、+聞	処理3		処理4	刪	+O、+聞	(e)
				± O、-聞			処理4	刪	+O、-聞	(f)	
3人称 代名詞 (彼/ 彼女)	ガ or ハ			+S	処理2	刪			+S	(g)	
		代名詞	1人称	ヲ or ニ	+S、+聞	処理2	刪	処理3		+S、+聞	(h)
					+S、-聞	処理2	刪			+S、-聞	(i)
		一般	3								(j)
			+物	ヲ or ニ	+S	処理2	刪	処理1		+S	(k)
		+人	+S、± O		処理2	刪	処理4	刪	+S、+O	(l)	
一般名詞 +人	ガ or ハ			± S	処理4	刪			+S	(m)	
		代名詞	1人称	ヲ or ニ	± S、+聞	処理4	刪	処理3		+S、+聞	(n)
					± S、-聞	処理4	刪			+S、-聞	(o)
		一般	3	ヲ or ニ	± S、+O	処理4	刪	処理4	刪	+S、+O	(p)
			+物		± S	処理4	刪	処理1		+S	(q)
		+人	± S、± O	処理4	刪	処理4	刪	+S、+O	(r)		

名詞の処理			
処理1	尊敬関連名詞に変換	処理3	主語助詞[は]→지、主語助詞[が]→가、主語[私達]→저희
処理2	主語[単数]→그는、主語[複数]→그들은	処理4	[尊敬名詞]+[님(様)]

図 3: 変換ルール

め「先生の妹が」といった場合、単に「妹」として扱われてしまう。

提案手法を拡張し、これらの現象を扱えるようにすることが今後の課題である。

## 6 おわりに

本稿では、日韓の敬語を比較し、さらに敬語を考慮した用例ベース翻訳手法について述べた。

今後は提案手法を拡張し、本研究で扱わなかった現象を取り込む必要がある、また、より大きなコーパスを用いた実験を行い、実証的に手法の妥当性を検討する予定である。最後に、本研究は日韓方向に焦点を当てたが、韓日方向についての研究も今後必要となるであろう。

## 参考文献

- [1] Makoto Nagao. A framework of a mechanical translation between Japanese and English by analogy principle. In *Proceedings of the international NATO symposium on Artificial and human intelligence*, pp. 173–180, 1984.
- [2] 河原大輔, 黒橋禎夫. 頑健な格解析を実現する格フレーム辞書の自動構築. 言語処理学会 第8回年次大会発表論文集, 2002.
- [3] 荒牧英治, 黒橋禎夫, 柏岡秀昭, 加藤直人. 確率的用例ベース翻訳の実現. 言語処理学会, 2005.
- [4] 森下喜一, 池景来. 日本語と韓国語の敬語. 白帝社, 1989.