

修士論文

日韓機械翻訳システムにおける 敬語翻訳に関する研究

指導教員

江崎 浩 教授

東京大学大学院 情報理工学系研究科 電子情報学

学籍番号・氏名

56447 金 ナヨン

提出日

平成19年2月2日

目次

第1章	はじめに	2
1.1	背景	2
1.2	本論文の流れ	2
第2章	日本語と韓国語の敬語	3
2.1	敬語とは	3
2.2	日本語の敬語	3
2.3	韓国語の敬語	4
2.4	用法の比較	4
2.4.1	上下関係	4
2.4.2	内と外の区別	5
2.5	文法の比較	6
2.5.1	尊敬語の比較	6
2.5.2	謙讓語の比較	8
2.5.3	丁寧語の比較	9
2.5.4	その他 - 助詞	10
2.5.5	本研究のスコープ	10
第3章	日韓用例ベース機械翻訳システム	11
3.1	用例ベース翻訳とは	11
3.1.1	用例構築モジュール	11
3.1.2	翻訳モジュール	11
3.2	関連研究	12
3.3	日韓用例ベース機械翻訳システムの構築	13
3.3.1	基本となる用例ベース翻訳システム	13
3.3.2	日韓翻訳用例ベース翻訳システム	15
3.3.3	関連研究	19
第4章	提案する敬語翻訳手法	20
4.1	敬語翻訳の問題点	20

4.2	敬語情報知識の解析方法	21
4.2.1	名詞の敬語情報解析	21
4.2.2	動詞の敬語情報解析	23
4.3	敬語表現変換ルール	24
4.3.1	対象とする文形と人物関係の制限	24
4.3.2	敬語表現変換ルール	26
4.4	翻訳の流れ	29
4.5	実験	30
4.5.1	実験設定	30
4.5.2	検証実験	31
第5章	おわりに	34

概要

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。このような状況のなか、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。日本語と韓国語は、他の国の言語と比べて語源的にも近く、さらに、両言語とも敬語が発達しているなど、非常に類似した言語であると言える。

このように、両言語は敬語の特徴をもっているものの、これまでの機械翻訳分野では敬語の翻訳に大きな注意を払ってこなかった。本研究では、両言語の敬語表現の違いを考慮した翻訳システムを提案する。

敬語表現を適切に翻訳するためには、まず、それぞれの言語でどのように敬語が表現されているか捉える必要がある。日本語では、敬語は主に動詞と名詞で表現されるため、名詞については、ルールとして記述し、動詞については形態素解析処理の結果を用いることで、敬語表現を捉える。

次に、日本語の敬語表現がどのように韓国語で表現されるかを整理する必要がある。提案システムは、整理した知識を敬語変換ルールという形で保持し、これを適応することで、自然な敬語表現の韓国語翻訳を生成する。

実験の結果、提案する枠組で日韓方向の敬語翻訳を扱うことを示せた。

第1章 はじめに

1.1 背景

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。そのような状況のなか、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、特に日韓翻訳に焦点をあてる。日本語と韓国語は、他の国の言語と比べて語源的にも近く、さらに、両言語とも、敬語が発達していることなど、非常に類似性が高い。

しかし、これまでの機械翻訳分野では敬語の翻訳に大きな注意を払ってこなかった。これは、敬語に厳密な韓国語においては重大な問題となりうる。そこで、本研究では、両言語の敬語表現の違いを考慮した翻訳システムを提案する。

1.2 本論文の流れ

まず、日韓両言語の敬語現象について述べる(2章)。そして、本研究のもととなる用例ベースシステムについて述べる(3章)。最後に、敬語を処理するモジュールについて述べ、実験結果と考察を行う(4章)。

第2章 日本語と韓国語の敬語

本章では、両言語の敬語の特徴を比較し、日韓翻訳において問題となりうる点について、考察する。

2.1 敬語とは

敬語の概念は、さまざまな説があるが、おおよそ「敬意を表す言葉」と規定されることが多く、人が言葉を用いて自らの意思や感情を人に伝える際に、単にその内容を表現するのではなく、聞き手や第三者(以降「話題の人物」と呼ぶ)と、自らの人間関係・社会関係について気持ちのあり方を表現するものである [12].

このような敬語の概念自体は日韓共通のものであるが、敬語の種類や出現する品詞などの面では両言語に差異がある。

2.2 日本語の敬語

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¹。ここでは、この3種類の敬語について述べる [6, 29, 34, 37].

(1) 尊敬語(「いらっしゃる・おっしゃる」型)

聞き手又は話題の人物の行為・ものごと・状態などについて、その人物を立てて述べるもの。

例) お祖父さんが御飯を 召しあがる。

(2) 謙譲語(「伺う・申し上げる」型)

話し手から聞き手又は話題の人物に向かう行為・ものごとなどについて、その向かう先の人物を立てて述べるもの。

例) 私は先生に本を 差し上げた。

¹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。

- (3) 丁寧語(「です・ます」型)
話や文の聞き手に対して丁寧に述べるもの。

例) 私は韓国人 です.

2.3 韓国語の敬語

日本語の敬語と同じく、尊敬される対象によって「主体敬語法」「客体敬語法」「聴者敬語法」の3つに分けられる。これらの各敬語法に使われる言葉を尊敬語、謙讓語、恭遜語という[30]。ここでいう恭遜語は、日本語の丁寧語に相当する。

- (1) 主体敬語法(尊敬語)
文の主体を高めて表現するもの... 尊敬語

例) 할아버지께서는 진지 를드신다. (お祖父さんが 御飯 を 召しあがる.)

- (2) 客体敬語法(謙讓語)
主語がする行動に影響される方を待遇するもの... 謙讓語
現代の韓国語では退化しつつある。

例) 나는 선생님께 책을 드렸다. (私は先生に本を 差し上げた.)

- (3) 聴者敬語法(恭遜語)
聞き手に対するもの... 丁寧語

例) 저는 한국인 입니다. (私は韓国人 です.)

2.4 用法の比較

以上のように、敬語の種類が異なるばかりでなく、敬語の用法においても、両言語では差異が存在する[32]。

2.4.1 上下関係

敬語が用いられるということは、そこに人間の上下関係が存在するということである。ここで「話し手」「聞き手」「話題の人物」という三者が存在すると、それらの上下関係と用いられる敬語の種類は両言語共通で、表2.1のようになる。

例えば、お婆さんに、「어머니께서는 말씀하셨습니다。」(お母さんがおっしゃいました。)という表現は、両言語とも適切でなく、聞き手を優先とするため(表2.2を参照)、尊敬動詞は使わない。よって、「어머니는 말했습니다。」(お母さんが言いました。)となる。

	上下関係	尊敬語	丁寧語
(1)	[題] < [聞] < [話]		
(2)	[聞] < [題] < [話]		
(3)	[題] < [話] < [聞]		
(4)	[聞] < [話] < [題]		
(5)	[話] < [聞] < [題]		
(6)	[話] < [題] < [聞]		

表 2.1: 上下関係

	娘	お母さん	お婆さん
(6)	[話]	[題]	[聞]

表 2.2: 上下関係の例

しかし，韓国語の現代語では親族内など関係なく話し手より上の場合，尊敬表現を使う場合が多く見られる。

2.4.2 内と外の区別

以上のように，両言語とも，同じ状況下では基本的に同じ敬語が用いられる。しかし，ある種の状況では，敬語の用法に差異が見られる [4]。

- 日本語 [相対敬語]

職場や家庭で，内外を区別して言語を選ぶことを社会的な親疎関係の反映としてみることになる。このような性質を相対敬語という。これは，自分と親しいものは謙譲語，遠いものは尊敬語として表現するものである。

- 韓国語 [絶対敬語]

韓国語では日本語のように内外で尊敬語の使用を変えるとということはなく，親や兄弟・先輩など身内の動作・状態に言及する場合も尊敬語を使う。このような性質を絶対敬語という。これは，日本語では謙譲語で表現する文が，韓国語の場合，尊敬語で表現するものである。相対敬語と絶対敬語をまとめると表 2.3 になる (表 2.1(5) の場合)。

以上述べたように、日韓で敬語に差異が見られる (表 2.4 を参照)。

	上下関係 内と外の区別	日本語	韓国語
		謙讓語 尊敬語 丁寧語	尊敬語 丁寧語
(5)-(内)	[話(内)] < [聞(内)] < [題(内)]	×	
(5)-(外)	[話(内)] < [聞(外)] < [題(内)]	×	

表 2.3: 外と内の区別

	日本語	韓国語
例 1	課長は今席を外しております.	과장님은 지금 자리에 안 계십니다. (課長様は今席を外していらっしゃいます.)
例 2	うちの社長が伺います.	저희 사장님께서 가실 겁니다. (われわれの社長様がお行きになります.)
例 3	只今, 父はおりませんが.	지금 아버지께서는 안 계시는데요. (只今, お父さんはいらっしゃいませんが.)

表 2.4: 外と内の区別の例

(1) 例 1 と例 2

会社の地位の高い人のことを外の人に話す時, 日本語は「外の人」を優先するため, 謙讓語を用いるが, 韓国語では「話し手」との関係を優先するため, 尊敬語を使う.

(2) 例 3

家族内の状況でも同じ現象が見られる. 例のように, 外出中の父親当ての電話を息子が取った場合, 韓国語では敬語で対応するのは普通である. このように韓国の敬語では, 特に目上に対して言語形式は絶対的なものである.

2.5 文法の比較

本節では, 両言語で, 敬語がどのように文法上表現されるかについて述べる.

2.5.1 尊敬語の比較

(1) 動詞, 形容詞

(a) 日本語

尊敬動詞の語形は助動詞である「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「(ら)れる敬語」型, 動詞の前に「お/ご」をつけ, さらに語末

に「～になる」をつけて尊敬の意味を持たせる「なる敬語」型，その単語(動詞)自体に尊敬の意味が含まれている「自体敬語(例：いらっしゃる，おっしゃる)」型等がある。

(b) 韓国語

尊敬動詞の語形は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の「시(/シ/)」を挿入させて尊敬動詞を作る「시」型と，日本語と同じく「自体敬語」型が存在する。また，「시」の場合は，その使われる範囲が広く，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。形容詞に尊敬語を使うのは日本語にはない韓国語の特色の一つである(表 2.5 を参照)。

例	普通表現	尊敬表現
日本語 1 韓国語 1	目が大きい. 눈이 크다.	目が大きくていらっしゃる. 눈이 크「시」다.
日本語 2 韓国語 2	今日は特にきれい. 오늘은 특히 예쁘다.	今日は特にきれいでいらっしゃる. 오늘은 특히 예쁘「시」다.

表 2.5: 尊敬表現の例(形容詞)

(2) 体言

名詞，接頭辞，接頭～接尾辞，接尾辞，代名詞等を用いて，対象人物及び関連(人)物を高めるために用いる。

(a) 日本語 表 2.6 を参照。

体言の種類	尊敬表現	例
接頭～接尾辞	お～さま，さん	お嬢さま，お父さん，お祖母さん等
接尾辞 1	～さま，さん	仏さま，兄さん等
接尾辞 2	～先生	山本先生，校長先生等

表 2.6: 日本語の体言

(b) 韓国語 表 2.7 を参照。

体言の種類	尊敬表現	例
名詞		진지(/ジンジ/御飯), 약주(/ヤッジュ/お酒), 치아(/チア/ご歯) 等
接尾辞 1	~님(/ニム/様)	아버님(/アボニム/御父様), 선생님(/ソンセンニム/先生様) 等
接尾辞 2	~분(/ブン/方)	여자분(/ヨザブン/女性方) 等
代名詞		이분(/イブン/このかた), 그분(/クン/そちらさま) 等

表 2.7: 韓国語の体言

2.5.2 謙讓語の比較

(1) 動詞

(a) 日本語

動詞の謙讓表現には、「お/ご ~ する」「お/ご ~ いたす」「~ ていただく」「させていただく」「お/ご ~ いただく」などの接語方式 (1) によるものと、謙讓語を用いて謙讓表現をする方法 (2) がある。

- | | |
|---------|-------|
| (1) 呼ぶ | お呼びする |
| (2) 尋ねる | 伺う |

また、日本の謙讓語はさまざまな場面、状況で用いられている謙讓語が多様に発達している。

(b) 韓国語

日本語の謙讓語に比べて韓国語の謙讓語は自体謙讓語に限られているため、その数が極めて少ない。これは、韓国語の謙讓動詞は中世では活発に用いられていたが、尊敬語の「시」のような役割をしていた「옵」という形態素が消えたことによって急激に退化したものと考えられる。現在使われている自体謙讓語は次のようなものである。

- | | |
|---------------|------------------|
| 문다(/ムッタ/尋ねる) | 여쭙다(/ヨチュッタ/伺う) |
| 보다(/ボダ/見る) | 뵈다(/ベッタ/お目にかかる) |
| 주다(/ジュダ/あげる) | 드리다(/ツリダ/差し上げる) |
| 주다(/ジュダ/あげる) | 올리다(/オルリダ/差し上げる) |
| 데리다(/デリダ/連れる) | 모시다(/モシダ/お連れする) |

(2) 名詞

(a) 日本語

ちち, はは, むすめ, わたくし, 自分, 小宅等

(b) 韓国語

자식(/ザシツク/子息), 저(/ゾ/わたくし)等

2.5.3 丁寧語の比較

● 日本語

助動詞「です」「ます」などの形があって, 比較的尊敬語や謙譲語よりは単純な形をしている. 名詞, 形容詞, 形容動詞には「です」を, 動詞には「ます」を付けるという品詞別による接続方式の区別がある.

● 韓国語

韓国語にも「です」「ます」と同じ働きをする語尾である「요」「버니다」があって, 日本語の丁寧形と同じ働きをする. ただ, 日本語とは違って, 名詞や形容詞や動詞等につけることが可能であり, 接続方式であって品詞別の区別は特にはない. 例は以下のようになる.

普通表現	丁寧表現
가다(/ガダ/行く)	갑니다/가요(/ガムニダ, ガヨ/行きます)
책이다(/チェギダ/本だ)	책입니다/책이에요(/チェギンニダ, チェギエヨ/本です)
크다(/クダ/大きい)	큽니다/커요(/クムニダ, コヨ/大きいです)

また, 「話し手」を低める表現の「저/ゾ/わたくし」と, 丁寧表現の「... 버니다/~ます」とともに使うことによってよりいっそう「聞き手」に丁寧さを表すことができる. さらに, 日本語では基本的に丁寧語とぞんざいな語の2通りしかないが, 伝統的な韓国語には大きく分けて6通りの形がある. 目上に使うもの, 目下・同等に使うもの, 大人や地位のある目下・同等に使うもの, 目下を同等あるいは目上に引き上げて使うものの4つがあり, さらにくだけた場面でのみ使われるものに目上用と目下・同等用の2通りがある. とはいえ, 実質的な標準語であるソウル方言においてはこのうち2通りの敬語系統は中年以下では用いられず, 実質4通りに数が減っている. また, 表2.8のように「極尊称」「普通尊称」「半語²」「普通卑称」「極卑称」の五つに分ける場合もある.

²ぞんざいな語と呼ばれるもの.

階称	叙述形
極尊称	~ 하시다/~ 습니다(ございます)
普通尊称	~(으)오/~ 소(ます・です)
半語	~ 아(어)/~ 지(だ・である)
普通卑称	~ 네(だ)
極卑称	~ 니다/~ 는다(だ)

表 2.8: 韓国語の丁寧語

2.5.4 その他 - 助詞

日本語と韓国語の敬語用法において最も大きな違いは、韓国語では前述した形容詞の敬語形の使用と、助詞の敬語形が存在するという点である。

例 1 아버지께서 말씀하시었다. (お父さんはおっしゃった.)

例 2 교수님 께서 강의하시고, (教授は講義なさって,)

韓国語には助詞の尊敬形を用いる主体尊待という概念がある。それは、主語の次に助詞の尊敬形の「께서(/ケソ/)」が付き、用言には「~시」が挿入された時に表現される尊敬表現である。上の例文のように、助詞の尊敬形とともに、叙述語「하다」には「~시」が挿入され、「お父さん」に尊敬の意を表す、同じく、「강의하시고」には「~시」の挿入によって、教授に対して尊敬の意を表すものである。助詞をまとめると以下ようになる。

主体助詞	助詞の尊敬表現
이(/イ/が), 가(/ガ/が), 으로부터(/ウロプト/から)	께서(/ケソ/)
어게(/エゲ/に)	께(/ケ/)

2.5.5 本研究のスコープ

以上の比較から分かるように、日韓の敬語表現は同じ点も多いが、差異も多く、そのいくつか(内と外の関係や話題の人物の上下関係の判断)は、現在の自然言語処理で扱うのは困難である。本研究では、日韓方向に翻訳方向をしばり、その際に顕著となる(1)日本語で敬語表現がない語(形容詞, 代名詞, 助詞等)の処理や(2)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。

第3章 日韓用例ベース機械翻訳システム

この章では，用例ベース日韓機械翻訳システムについて述べる。

3.1 用例ベース翻訳とは

用例ベース翻訳とは，翻訳対象となる入力文に対して，あらかじめ蓄積された用例の中から類似した文を複数探し出し，それらの組み合わせることで翻訳結果となる出力文を作り出す翻訳方式である [20, 25, 27]. 用例ベース翻訳システムの枠組は，大きく分けて二つの処理からなる。(1) 用例構築モジュール (アライメント) と，(2) その用例を用いた翻訳 (翻訳モジュール) である [1, 23, 24].

3.1.1 用例構築モジュール

まず日本語と翻訳対象の言語 (ここでは英語) の対訳コーパスが用例として用意される.[13, 38] この対訳コーパスの日本語と英語の各文は，それぞれ構文解析器によって構造化され，同じ意味内容を持つ日本語と英語の二つの木の組み合わせが出来る。この二つの木の間を，日英辞書を用いて，同じ意味を持つフレーズ同士を，対応づける。この対応付けをアライメントと呼び，これは用例を用いるための下準備である。対応する単語が見付からなかった句については，近くの句が対応している相手にまとめて対応づけする。このアライメント作業をすべての用例文について行い，対応づけされた用例の巨大なデータベースを構築する。図 3.1 の Translation Examples が構築されたデータベースにあたる。例えば「私のサイン」の語句は「私の」と「サイン」がそれぞれ「my」と「signature」に対応している。

3.1.2 翻訳モジュール

翻訳システムは対応づけされた用例のデータベースから，翻訳する日本語文と同じ語句を含む用例を収集する。そして，その同じ語句に用例文で対応している英語側の部分木を切り取り，翻訳する日本文の構造に合わせて組み合わせる。こうして，翻訳する日本語文と対応する英語文を作成することが出来る (図 3.1 を参考)[16, 26, 31].

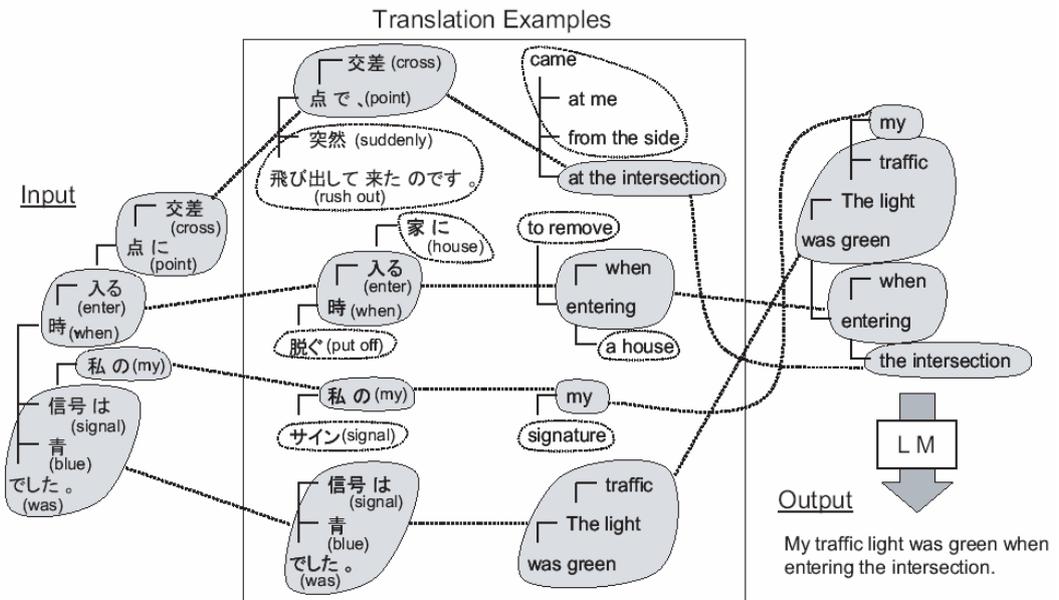


図 3.1: 翻訳の例

3.2 関連研究

日韓機械翻訳においては、これまで、Rule-Based Machine Translation(RBMT)の方式の中でも、特に「直接翻訳方式」が用いられることが多かった。これは、入力された日本語文の形態素解析を行い、それに対応する各韓国語を置き換える方式である。しかし、この直接方式は、ある程度の翻訳は実現可能なものの、様々な問題を含んだものである。例えば、日韓両言語の文節の間では語順が一致しているが、述部の形態素の間では語順の違いが多くみられることがある。また、日本語の否定表現が韓国語では否定的意味を持つ他の一つの用言に対応することがある。このように、両言語間が1対1対応が出来ないときも多く、したがって、直接翻訳方式のシステムでは自然な韓国語述部を生成するのは大変手間がかかるものになる。

このような問題を解決するため、これまで、連語パターン [35, 36] 等、例外の表現についての規則を利用したり、[3] では Neural Network を用いて訳語選択の曖昧性を解消する方法などが提案されてた。また [5, 9] では、進行形や「れる/られる」のように限られた表現を対象とした研究も行われている。

以上の関連研究から分るように、直接翻訳方式で、実用的なシステムを作るのは簡単なことではなく、精度をあげるためには、非常に複雑で巨大なシステムが必要とされる。このような複雑さを解決するために、現在では、多様な言語現象をすべて翻訳用例として扱う用例ベース翻訳方式が将来性のある翻訳手法として期待されている [2, 7, 10, 11, 39].

3.3 日韓用例ベース機械翻訳システムの構築

用例ベース翻訳方式は対象とする言語対に依存せず，基本的には，対訳コーパスと言語解析器があれば実現できる翻訳方式である。しかし，精度のよい翻訳を得るためには各言語の特徴を考慮した処理モジュールが必要となる。この章では，まず基本となる用例ベース翻訳システムについて述べ，次に日韓翻訳において修正が必要な処理について述べる。

3.3.1 基本となる用例ベース翻訳システム

用例ベース機械翻訳の基本的な流れは，まず，木構造の対訳用例文より，木構造に変換された入力文と最も似たる用例文を探しだし，それらを組み合わせて翻訳文を生成する。

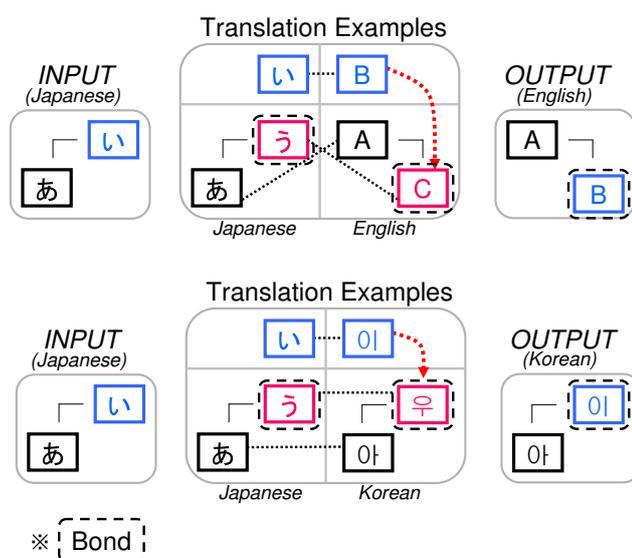


図 3.2: 言語の独立

例えば，図 3.2 のように，日英も日韓も木構造された入力文と同じ単語を探し出し，組み合わせる同じ処理で翻訳が行われている。日英の場合，文「あ い」と同じ単語を持つ木構造を対訳用例文の中から探し出し，入力文と違う「う」と対応する英語の「C」を入れ替えることによって英語文の翻訳が出来る。従って，多くのルールとコストが不要で簡単に翻訳が出来る。

用例ベース翻訳は、単に各単語の置き換えを行うルールベースシステムよりも訳語選択や語順などの複雑な処理を自然に実現できる。

[入力文] 勉強するのは 楽な ものではない。

[出力文] 공부하는 것은 편한 것은 아니다。

例えば、以上のような翻訳をする場合、ルールベースの翻訳結果では曖昧性を持つ単語「楽な」が適切に翻訳出来ない。つまり、出力文の「편한」の意味は'comfortable'にあてはまるもので翻訳されている。適切な翻訳のためには、'easy'の意味を持つ「쉬운」で翻訳しなければいけない。こういう問題は、用例ベース翻訳で解決出来る。つまり、入力文と最も似ている用例がある場合、簡単に翻訳が出来る。

翻訳結果は図 3.3 のように、「のは楽なものではない」に対応する韓国語の表現を、そのまま使うことが出来る。

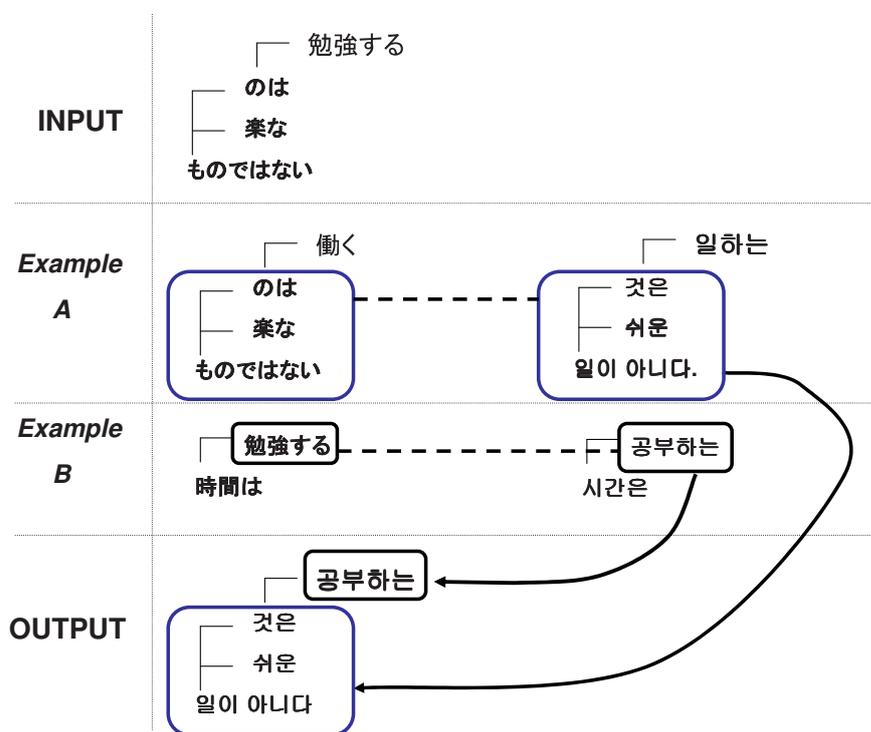


図 3.3: 用例ベース翻訳結果

3.3.2 日韓翻訳用例ベース翻訳システム

以上のような用例ベース翻訳システムも細部について対象言語独自の処理が必要とされる。例えば，韓国語を扱う際にはパッチムと機能語と内用語の処理について修正が必要となった。

(1) バッチム (받침) の処理

まず，バッチムの理解のため，ハングルの構造について述べる。ハングルは表音文字である。子音・母音を表すそれぞれの字母があり，それぞれの字母が1つの音を表すアルファベット型の文字である。しかしながら，ハングルの字母はローマ字のように一列に並べて書かれるのではなく，基本的に1音節ごとに一まとめにして1文字を成す。従って，ハングルはアルファベットであると同時に音節文字でもある。

文字の構成は子音字母は基本字母が14個，合成字母が5個の計19個，母音字母は基本字母が10個，合成字母が11個の計21個であり，合成字母を含めた字母の総数は40個である。

子音字母

基本字母	合成字母
ㄱ ㅋ ㆁ ㄷ ㅌ ㄴ ㄹ ㄷ ㅌ ㄴ ㄹ ㄷ ㅌ ㄴ ㄹ	ㄱ ㅋ ㆁ ㄷ ㅌ ㄴ ㄹ

母音字母

基本字母	合成字母
ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅝ	ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅝ

文字の組み合わせ

字母は2つ以上を組み合わせることで1文字を成す。1文字の構成は「子音字母+母音字母」あるいは「子音字母+母音字母+子音字母」のどちらかである。伝統的に先頭の子音字母を初声，母音字母を中声，末尾の子音字母を終声と呼ぶ。その終声を「バッチム」と呼ぶ。

「子音字母+母音字母」から成る文字は図3.4のように3つのタイプがある。図3.4の(1)は，母音字母が縦長の「ㅏ, ㅑ, ㅓ, ㅕ, ㅗ, ㅛ, ㅜ, ㅠ」である場合には，子音字母(初声)を左に，母音字母(中声)を右に配置する(例：가)。図3.4の(2)は，母音字母が横長の「ㅜ, ㅠ, ㅡ, ㅝ」である場合には，子音字母(初声)を上にも，母音字母(中声)を下に配置する(例：고)。図3.4の(3)の場合は，母音字母が「ㅟ, ㅠ」

「ㄱ, ㄴ, ㄷ, ㄹ, ㅁ, ㅂ, ㅅ, ㅇ」である場合には, 子音字母(初声)を左上に, 母音字母(中声)を下から右にかけて配置する(例: ㄱㅏ).

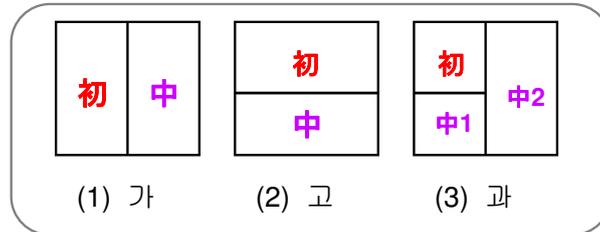


図 3.4: 字母の組み合わせ 1

終声のある場合には, 図 3.5 のようにこれらの下に終声を置く(図 3.5(4)の例: ㄱㅏ, (5)の例: ㄱㅓ, (6)の例: ㄱㅜ).

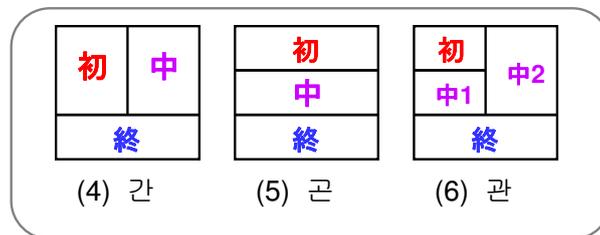


図 3.5: 字母の組み合わせ 2

なお, 終声として用いることのできる子音字母は「ㄱ, ㅋ, ㆁ」を除いた 16 個である. また, 朝鮮語の形態音素表記のために, 終声は 2 つの子音字母を左右に組み合わせることがある. 正書法で認められている組合せは図 3.6 の終声「青い文字」の 11 種類である.

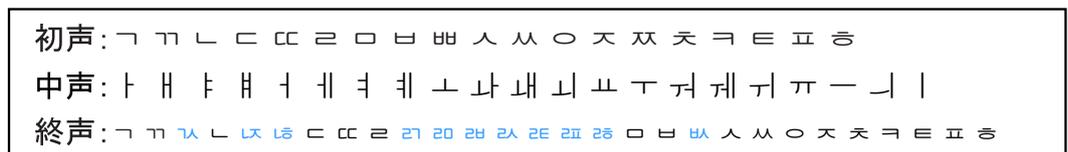


図 3.6: 初・中・終声の文字構成

バッチムの特徴は, それの有無によって次の文字に発音の影響があるため, 同じ意味で違う文字を用いて表現する(表 3.1 を参照). このような特徴を持つ韓国語であるた

バッチム	가	하	을
有	이	은	을
無	가	는	를

表 3.1: バッチムの有無

め，組み合わせで翻訳を行う用例ベースの方式では，バッチムの有無を考慮した処理が必要となる。

次は，バッチムの処理について述べる。ハングル文字には各初声，中声，終声の文字コードで構成されている。よって，その文字コードを，割り算をすることにより，3つの部分に認識するため，バッチム(終声)の有無が判断出来る。その判断によって，次に来る文字のタイプを与える。例えば，「私は」と「彼らは」の場合，両方同じ「は」であるが「私」に対応する「나」と「彼ら」に対応する「그들」からみると「나」にはバッチムがないため，次の「は」は「는(/ナ/は)」を用いる。一方「그들」の場合，バッチムの「ㄹ」があるため，次の「は」は「은(/ウン/は)」を用いる(図 3.7 の例を参照)。よって，日韓用例ベース機械翻訳システムの場合，バッチムの処理によって相応しい翻訳結果が得られる。

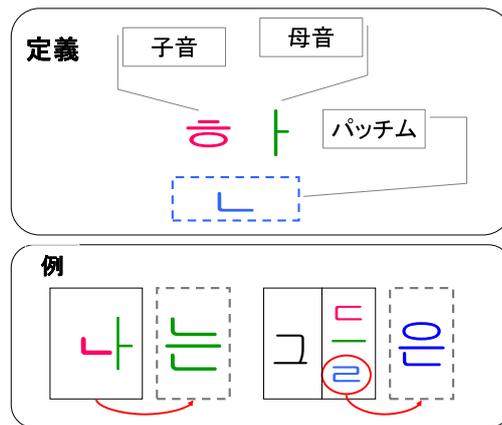


図 3.7: バッチムの定義と例

(2) 機能語と内容語の語順

もう一つの韓国語固有の特徴は，機能語と内容語の語順である。一般に，文法的な役割を持つ語を機能語 (function words)，それ以外の一般的な意味を持つ語を内容語 (content words) という。

機能語には助詞，助動詞，接続詞，形式名詞，連体詞，副詞，指示代名詞等がある。また，内容語には名詞，動詞，形容詞等がある。これらの区別は基本的に言語非依存であり，あらゆる言語に存在する文法範疇である。しかし，内容語と機能語の位置関係は，言語によって異なる。例えば，日本語と韓国語は内容語が前，機能語が後になる。英語の場合は機能語が前，内容語が後に位置する。また，中国語は基本的には，英語と同じ位置ではあるが，稀に逆になる場合もある。図 3.8 が示すように，各言語の「東京で会う」とその内容語(四角で囲まれた部分)を示す。

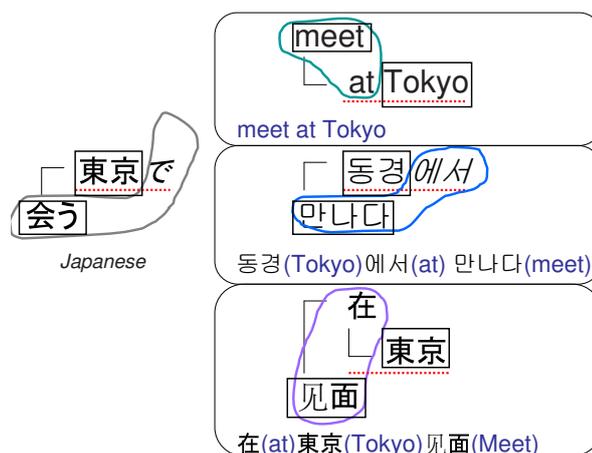


図 3.8: function words の例

機能語と内容語を区別し，翻訳することは，翻訳を行う上で必要であり，翻訳結果の語順の調整に必要な変数となる。

(3) 敬語問題

日韓用例ベース機械翻訳システム上，図 3.9 のように，適切な表現を選ぶのに問題が生じる。それは，対訳用例が，日本語側「質問しました」を韓国語側へ訳される時，主語を尊敬する場合としない場合によって敬語表現の入った表現になるかならないか異なるためである。よって，主語等による動詞の敬語表現の有無を判断する処理が必要となる。敬語処理については次の 4 章で詳しく述べる。

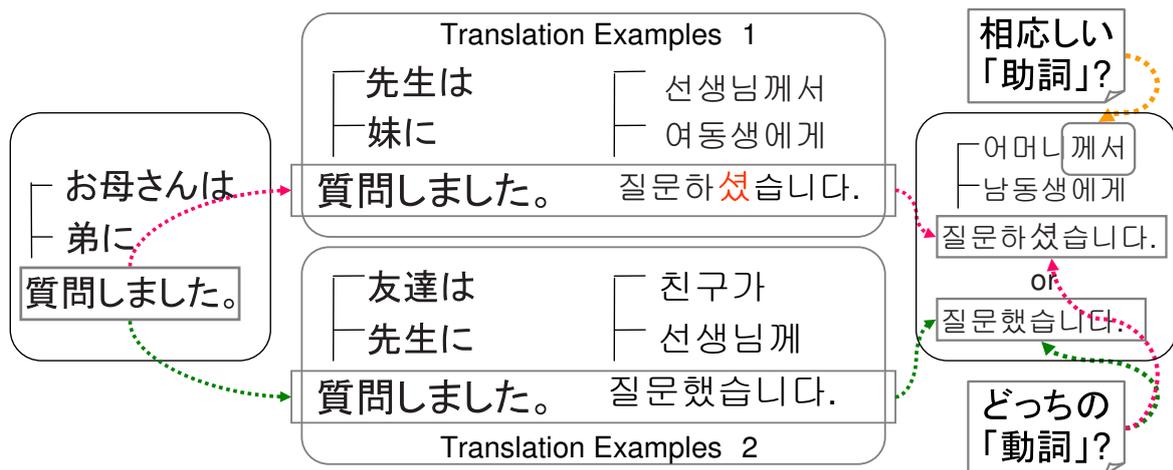


図 3.9: 用例ベース敬語翻訳の例

3.3.3 関連研究

日韓機械翻訳上，言語依存処理として重大な処理が敬語翻訳である．両言語とも敬語が他言語よりが発達している．その反面，今迄は敬語翻訳に関する研究はあまりされていない状況である．例えば，ルールベースでの敬語翻訳の結果の場合，以下のようにほぼ考慮されていない状況である．

- (1) 母は先月から神戸におります.(「に降りる」で訳される)
- (2) 東京から帰ってまいりました.(「ってまいる」が直訳される)
- (3) 高津さんは有能でいらっしゃる.(「でいらっしゃる」が直訳される)
- (4) 鈴木先生はいつもお忙しい.(「お忙しい」と「忙しい」の区別が付かない)
- (5) 鈴木先生はその問題を詳しく調べられた.(可能/尊敬の判断が出来ない)

関連研究としては，ルールを用いての「れる，られる」の処理 [5] がある．また，HPSG の敬語情報の詳細化 [8, 22] 等がある．言語学からの韓国語と日本語の敬語研究 [17, 18, 19, 21, 33, 40] はそれぞれ活発に行われてきたが，翻訳を扱った研究は存在しない．

第4章 提案する敬語翻訳手法

この章では、用例ベース翻訳で敬語翻訳をどのように行うか、その手法について述べる。

まず、4.1章では翻訳結果にどのような敬語の問題があるか、詳細に議論する。

次に、問題を解決するためには、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解析する必要がある。これは「～です/ます」のような比較的容易に捉えられる文末表現から「先生」「大臣」といった尊敬の対象となる語彙レベルの表現まで様々なレベルにわたる。4.2章では、これらの知識の解析方法について述べる。

また、解析された日本語入力文の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)用例の変換処理という2つの形で表現した。

4.3章では、これらをまとめた敬語表現変換ルールを説明する。ただし、日本語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る。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。

4.4章では実際の翻訳の流れを具体的に述べる。

4.1 敬語翻訳の問題点

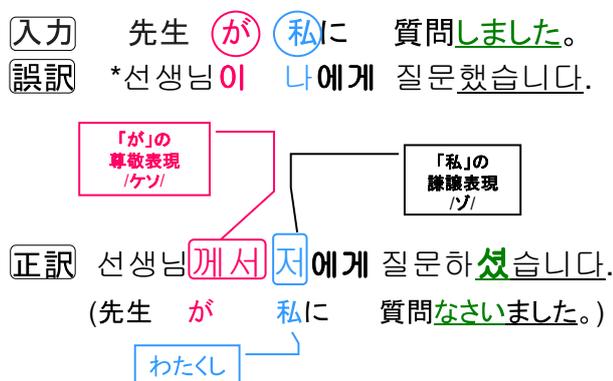


図 4.1: 敬語翻訳の誤例

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 4.1 上の入力に対して，次の表 4.1 のように 3 つの用例があった場合を考える：

	日本語	韓国語
用例 1	学生が質問しました.	학생이 질문했습니다.
用例 2	先生が....	선생님이....
用例 3	私に....	나에게....

表 4.1: 用例 1 ~ 3

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 4.1 の中段のような誤訳となってしまう。これを正訳にするためには，次のような 3 点を考慮する必要がある。

- (1) 主語が「先生」という尊敬人物であるため，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」は主語が尊敬人物であるため，「質問なさいました」にあてはまる「질문하셨습니다.」を使う。

4.2 敬語情報知識の解析方法

前章の正訳のような翻訳を実現するためには，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解析する必要がある。2 章にて述べたように，日本語の敬語は名詞と動詞にて表現されることが多く，それぞれ (4.2.1 節)，(4.2.2 節) の方法で解析を行う。

4.2.1 名詞の敬語情報解析

日本語では，名詞にて表現される敬語は，(1) 尊敬される対象の人物 (以降，尊敬人物と呼ぶ)，または，(2) 尊敬される人物に付随する物 (以降，尊敬物と呼ぶ) に分類される。

- (1) 尊敬人物

まず，尊敬人物（例えば「社長」「先生」等）の判定について述べる。これは，敬語動詞の格フレーム [28] の格要素として存在するかどうかで調べる。以下に，敬語動詞とそこから得られる尊敬人物の例を示す。

- 尊敬語：なさる，いらっしゃる，おっしゃる，くださる，おいでになる
 <ガ格> 先生が今年の夏も講演をなさいました。
- 謙譲語：いたす，うかがう，さしあげる，拝見する
 <二格> 社長にお話をうかがいました。
 <ヲ格> 新しい先生を遠くから拝見した。

以上の結果を用いて，152 単語の尊敬人物を収集した（表 4.2 を参照）。これらは格フレームより収集した 790 単語から人手で選択した。

尊敬人物

方/先生/知事/皆さん/仏様/オーナー/選任/名人/座長/同土/兄さん/
 ユーザー/管理人/弊/あなた方/先輩/達人/お姫様/お爺さん//町/
 総理/ファン/代表/博士/お嬢さん/秘書/奥さん/作者/所長/陛下/
 師/恩師/課長/弁天/王様/先客/当社/税理士/御方/職人/飼い主/ドクター/
 会長/ゲスト/助役/議長/ママ/方々/信者/首相/王/会員/医者/勇者/
 店長/長/官/霊/駅員/先祖/母/店員/兄/委員/派/学者/次長/主/教師/
 教授/神/人々/局長/官/サンタ/同志/仏/部長/老人/師匠/隊長/どなた/
 女性/夫妻/客/長/省/課/室長/年配/さま/サマ/医/持ち主/お宅/叔父/
 奥様/町長/区長/皆様/神父/宅/かあさん/かた/医師/理事/助教授/
 みなさま/イエス/社長/両親/様/親戚/年寄り/署/将軍/職員/匿名/
 おばあさん/王子/夫人/教官/強者/全員/おじいさん/その方/嬢/天/
 殿下/主人/小人/市長/お方/監督/ 猛者/弊/御挨拶/父/作家/ごあいさつ/
 お母さん/貴/当方/員/神様/天子/みなさん/大臣/大家/聖人/国/議員/

表 4.2: 尊敬人物

これを用いて，尊敬人物が存在するならば，+人 を，存在しない場合は -人を敬語情報として付与した。

また，以降の処理で，尊敬人物が主語か目的語かの区別も必要となるため，実際は，4 種類で区別を行った。主語か目的語の区別は，構文解析器 KNP[14] の結果を用いガ格なら主語，ヲ格，二格なら目的語とする（以下の表 4.3 を参照）。

	+ 人	- 人
主語 (subject)	+ 人 s	- 人 s
目的語 (object)	+ 人 o	- 人 o

表 4.3: 尊敬人物の区別

(2) 尊敬物

尊敬物 (例えば「お言葉」等) については, 人手で語彙リスト (「お宅」「お年」「お子さん」「御言葉」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.

4.2.2 動詞の敬語情報解析

動詞の敬語情報としては, 次の 3 種類のフラグを用意した. まず, (1) 聞き手を尊敬するかどうかで, + 聞 or - 聞, (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.

これは, 形態素解析器 JUMAN[15] と構文解析器 KNP[14] の結果を用いて動詞の表 4.4 の情報を調べ, 敬語情報を付与する.

	+聞/+S/+O	JUMAN/KNP の解析結果 Feature
(1)	+ 聞	デス列基本形動 or 接尾辞ます型
(2)	+S	<敬語:主体尊敬表現> or <尊敬動詞:*>
(3)	+O	<敬語:謙讓表現> or <謙讓動詞:*>

表 4.4: 動詞の敬語情報

例えば「先生がおっしゃいました。」という入力文の場合「おっしゃる<敬語:主体尊敬表現>」から「主語」を尊敬していることが「ました<動詞性接尾辞ます型>」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. よって以下のような敬語情報が付与されることとなる. 表 4.5 は動詞の敬語情報解析の例である.

先生が (+ 人 s) おっしゃいました (+S, + 聞).

JUMAN/KNP の解析結果 Feature	例
動詞性接尾辞ます型	です (+ 聞)
<尊敬動詞:行く/いく:来る/くる:居る/いる...>	いらっしゃる (+S)
<敬語:主体尊敬表現>	お食べになる (+S)
<敬語:謙遜表現>	お尋ねする (+O)
<謙讓動詞:言う/いう>	申し上げる (+O)

表 4.5: 動詞の敬語情報解析の例

4.3 敬語表現変換ルール

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) 用例の変換処理という2つに分解し，これを敬語変換ルールという形で表現する(図 4.4 に示す.)。

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。

この章ではまず，文形の制限について説明し，次に，変換ルールについて述べる。

4.3.1 対象とする文形と人物関係の制限

文				
主体部		客体部		述部
主語	助詞	目的語	助詞	動詞、 形容詞
人	ガ、ハ	人、物	ヲ、ニ	

図 4.2: 文形の制限

(1) 対象とする文形

文には多くのバリエーションがあるが，本研究では以下の図 4.2 ように SV と SOV だけを対象とする(「主体部 + 客体部 + 述部」)。そして，敬語表現に関係ある「主語」と「目的語」は以下ようになる：

(a) 主語

文		ID
主語	目的語	
[話]		(1)
[話]	Obj	(2)
Subj		(3)
Subj	[話]	(4)
Subj	Obj	(5)
[聞]		(6)
[聞]	Obj	(7)
Subj	[聞]	(8)
[話]	[聞]	(9)
[聞]	[話]	(10)

※1

図 4.3: 文の制限

- 代名詞：1 人称，3 人称¹
- 一般名詞：尊敬人物

(b) 目的語

- 代名詞：1 人称，3 人称
- 一般名詞：尊敬人物，尊敬物

(2) 人物関係の制限

SOV 文形には，最大で「話し手 ([話])」「聞き手 ([聞])」「話題の人物」という 3 者が存在することになる。また，これらが，主語となったり，目的語となったり，非常に多くの人間関係のバリエーションが存在する。これらすべてを扱うのは，困難であるため，本研究では，基本的ないくつかのパターンを扱うことにした。まず，主語を **Subj**，目的語を **Obj** と表記した場合，図 4.3 の 1 のようなパターンの文型を扱う。これは，「聞き手」が主語や目的語にいる場合，韓国語の場合は，名前や役職で呼ぶのが一般的であるためである。

¹2 人称を対象としないことについては「人物関係の制限」にて述べる。

4.3.2 敬語表現変換ルール

この章では，定義した日本語文を韓国語へ翻訳する際，必要となる敬語表現を考慮した変換ルールに関して述べる. 図 4.4 は日本語文の各部の構成要素によって韓国語文の変換を必要とするルールを表したものである (ただし，両言語間で敬語表現が変化しない場合は表から省略し，日韓翻訳において，なんらかの処理が必要なものを図示した.).

図 4.4 中の表記の定義は，以下の表 4.6 のとおりである.

表記	説明
一般名詞	一般名詞の中，尊敬人物と尊敬物のみ考慮
人	尊敬人物 (例：先生，社長，お母さ等)
物	尊敬物 尊敬関連名詞 (例：御飯，お宅等)
聞	丁寧フラグ
S	尊敬フラグ
O	謙譲フラグ
+	敬語表現
-	敬語表現を使わない (非敬語表現)
±	敬語表現と関係なし

表 4.6: 表記定義

次に名詞と助詞の変換処理について述べる. 本研究では，おおきく五つに分けた. 以下より詳しく述べる.

処理 1 は，尊敬関連名詞であるなら尊敬物に変換 (図 4.7 を参照) する.

+ 物
진지 (/ジンジ/御飯)
말씀 (/マルスム/御言葉)
연세 (/ヨンセ/お年)
자녀분 (/ザゼブン/お子さん)
댁 (/デック/お宅)
...

表 4.7: 尊敬物

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	ㄱ	+O、+聞	(b)	
					+O、-聞			処理2	ㄱ	+O、-聞	(c)	
		一般	-物	+人	+聞	処理3				+聞	(d)	
					±O、+聞	処理3		処理4	ㄱ	+O、+聞	(e)	
					±O、-聞			処理4	ㄱ	+O、-聞	(f)	
3人称 代名詞 (彼/ 彼女)	ガ or ハ			+S	処理2	ㄱ			+S	(g)		
		代名詞	1人称	ヲ or ニ	+S、+聞	処理2	ㄱ	処理3		+S、+聞	(h)	
					+S、-聞	処理2	ㄱ			+S、-聞	(i)	
		一般	3								(j)	
				+物	ヲ or ニ	+S	処理2	ㄱ	処理1		+S	(k)
				-人		+S、-O	処理2	ㄱ			+S、-O	(l)
+人	+S、±O	処理2	ㄱ	処理4		ㄱ	+S、+O	(m)				
一般名詞 +人	ガ or ハ			±S	処理4	ㄱ			+S	(n)		
		代名詞	1人称	ヲ or ニ	±S、+聞	処理4	ㄱ	処理3		+S、+聞	(o)	
			3		±S、-聞	処理4	ㄱ			+S、-聞	(p)	
		一般	+物	-人	+人	±S、+O	処理4	ㄱ	処理2	ㄱ	+S、+O	(q)
						±S	処理4	ㄱ	処理1		+S	(r)
						±S、-O	処理4	ㄱ			+S、-O	(s)
			±S、±O	処理4	ㄱ	処理4	ㄱ	+S、+O	(t)			

図 4.4: 変換ルール

処理2は、尊敬名詞の3人称の処理である。韓国では性別を問わず同じ表現（「彼」に対する敬語表現）を用いるのが特徴である（表4.8を参照）。

日本語	韓国語
主語が単数の「彼，彼女」	그분(/グブン/彼様)
主語が複数の「彼ら，彼女ら」	그분들(/グブンヅル/彼ら様)

表 4.8: 3人称の処理

処理3は、1人称の処理である。韓国語は「聞き手」を尊敬することによって自分を謙る表現を用いることが特徴である（表4.9を参照）。

日本語	韓国語
主語の助詞が「は」の場合	저(/ゾ/わたくし)
主語の助詞が「が」の場合	제(/ゼ/わたくし)
主語が「私たち」の場合	저희(/ゾヒ/われわれ)

表 4.9: 1人称の処理

処理4は、「尊敬人物」の後に「님(/ニム/様)」を付ける処理である。韓国語は目上の人に対して「様」よく用いる。よって、先生や教授、社長等の日本語の呼び方そのまま訳してしまうと、大変失礼にあたる（表4.10を参照）。

日本語	韓国語
先生	선생(先生)님(様)
教授	교수(教授)님(様)
社長	사장(社長)님(様)
...	...

表 4.10: 尊敬人物の処理

また、助詞の処理に関しては「変換ルール」に直接示すように「が，から」の場合は「께서(/ケソ/)」に、「に」の場合は「께(/ケ/に)」に変換し、用例を組み合わせの際生じる不適切な助詞表現をカバーできる。

最後に、変換ルールの各文に対応する入力文の実例を表4.11に示す。この表中のIDは図4.4のIDと対応している。

主語	助詞	目的語	助詞	動詞/形容詞	ID
私	は			行きます (+ 聞).	(a)
私	は	彼	に	伺いました (+O, + 聞).	(b)
私	は	彼	に	伺った (+O, - 聞).	(c)
私	は	御飯	を	食べました (+ 聞).	(d)
私	が	先生	に	申し上げます (+O, + 聞).	(e)
私	が	先生	に	申し上げた (+O, - 聞).	(f)
彼	が	私	に	おっしゃいました (+S, + 聞).	(h)
先生	が	御飯	を	食べています (-S, + 聞).	(r)

表 4.11: 変換ルールの入力文の例

4.4 翻訳の流れ

この章では 4.2 章と 4.3 章で述べた敬語情報の知識や変換ルールによる実際の翻訳の流れを具体的に説明する。

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。基本的には、用例の日本語側が同じ敬語情報を持つものをとってくる。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。

以下に、具体的な 2 つの入力文で、どのように翻訳が実現されるかを示す。

入力文 1: 先生が弟に言いました。

(1) 敬語情報の追加:

- 主語の「先生」は、尊敬人物であるため「+人s」を付与
- 目的語の「弟」は、尊敬人物には入らないため「-人o」を付与
- 動詞の「言いました」は、動詞性接尾辞ます型のため「+聞」を付与

よって、入力文は「先生(+人s)が弟(-人o)に言いました(-S, +聞).」になる(非尊敬動詞であるため「-S」になる.)。

(2) 用例の検索:

先のステップで得られたタイプの用例を検索する。よって、名詞に関して +人s, -人o の敬語情報を持つ動詞を探す。

用例 1: お母さん(+人s)が妹(-人o)に言いました(-S, +聞).

用例2：弟(-人o)に言いました(-S, +聞).

(3) 敬語変換ルールの参照：

組み合わせた用例の結果をもって，図 4.4 の韓国側の処理を行う. 入力文のタイプの場合は図 4.4 の ID[s] に該当する.

(4) 処理確認を終えると，韓国語の翻訳結果を出力する.

入力文2：先生はとても親切です.

(1) 敬語情報の追加：

- 主語の「先生」は，尊敬人物であるため「+人s」を付与
- 動詞の「親切です」は，聞き手を尊敬する野出「+聞」を付与

よって，入力文は「先生(+人s)は親切です(-S, +聞).」になる(非尊敬動詞であるため「-S」になる.).

(2) 用例の検索：

先のステップで得られたタイプの用例を検索する. よって，名詞に関して+人sの敬語情報を持つ動詞を探す.

用例1：お爺さん(+人s)は親切です(-S, +聞).

用例2：先生(+人s)は言いました(-S, +聞).

(3) 敬語変換ルールの参照：

組み合わせた用例の結果をもって，図 4.4 の韓国側の処理を行う. 入力文のタイプの場合は図 4.4 の ID[n] に該当する.

(4) 処理確認を終えると，韓国語の翻訳結果を出力する.

4.5 実験

4.5.1 実験設定

用例ベースの翻訳システムが実現出来たのは，精度の高い様々な言語ツールが必要となる. まず，用いた日韓用例ベース翻訳構築のため必要なツールやリソースについて紹介する.

- 形態素解析・構文解析 (韓国 KAIST)²
 - 依存構造解析 Parser
- 日韓対訳辞書 (free software)³
 - 6 万語ペア
- 日韓対訳コーパス (韓国高麗大学)⁴
 - 約 2 万文
 - ドメイン：旅行，小説，憲法等複数のドメインが存在する。

4.5.2 検証実験

前章の図 4.4 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 30 文：対訳用例 100 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 4.5 に示す (図 4.5 の ID は図 4.4 の ID に対応する)。

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，今後，実用的なシステムを構築するためには，次のような課題が残されている。

- 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。

今後，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課題である。

同時に，本研究では，韓国語側で敬語情報をほとんど扱わなかったが，韓国語側でも同時に解析することで精度をあげることが期待される。これには，具体的に次のような処理が考えられよう。

(1) 動詞

尊敬動詞の手がかりとなる「시」に敬語情報を付与する。また，丁寧動詞に関しては相手に対する敬意の度合によってさまざまな表現が存在するため，日本語と対応する「丁寧と非丁寧」の 2 種類のみで分ける。動詞のまとめは表 4.12 で示す。

²<http://bola.kaist.ac.kr>

³<http://hyam.dip.jp>

⁴<http://transkj.com>

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。)
(n)	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。)
(o)	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 4.5: 翻訳結果の例

	+ 聞/+S/+O	韓国語の形態素解析結果 Feature
(1)	+ 聞	습니다/ef, 니다/ef
(2)	+S	시/ep, 십/ep, 셉/ep 等
(3)	+O	人手でリストアップ

表 4.12: 動詞の解析

また，表 4.12(3) の謙讓動詞に関しては，退化しつつあるため，限られたものしか存在しない(4.13 を参照).

普通表現	謙讓表現
묻다(/ムッタ/尋ねる，聞く，話す)	여쭙다(/ヨチュッタ/伺う)
보다(/ボダ/見る)	뵈다(/ベッタ/お目にかかる)
주다(/ジュダ/あげる)	드리다, 올리다(/ヅリダ, オルリダ/差し上げる)
데리다(/デリダ/連れる)	모시다(/モシダ/お連れする，ご案内する)

表 4.13: 謙讓動詞

以上の解析方法を用いた敬語情報付与の例は以下ようになる.

오셨 습니다(/オシ ヨッスムニダ/いらっしやいました) 오/ncn + 션/ep(+S) + 습니다/ef(+ 聞) 오 + 션(+S) + 습니다(+ 聞)
--

図 4.6: 敬語情報付与の例

(2) 助詞・形容詞

さらに，日本語にはない助詞と形容詞に関する敬語情報の解析は以下ようになる.

	+S/+Ps/+Po	韓国語の形態素解析結果 Feature
形容詞	+S	動詞と同じく, 시/ep, 십/ep, 션/ep 等
助詞	+Ps	께서/jca
助詞	+Po	께서/jca

表 4.14: 助詞と形容詞の解析 Feature

このように、韓国語側にも敬語情報を解析することにより、両言語の敬語情報を用いたより緻密な翻訳を行うことが期待されよう.

第5章 おわりに

本論文では、日韓敬語の比較より敬語を考慮した翻訳手法について述べた。これは、基本となる日韓用例ベース機械翻訳システムと、そこの追加された敬語処理モジュールからなる。また、翻訳結果は提案する枠組の原理的妥当性を示し、今後、より大きいコーパスを用い実証的に手法の妥当性を検討する予定である。

また、今回の研究では、文に登場する人物や、文形などにおいて制限を設けたが、提案する枠組を拡張し、より複雑な状況の敬語現象を扱えるかどうか、今後の課題として望まれる。

謝辞

本研究に対してさまざまなご意見・ご指導をいただき，また，多方面からのご協力を賜わりました，黒橋禎夫教授，江崎浩教授に心より感謝いたします。また，荒牧英治氏には研究をはじめ多くのことを教えていただきました。心より感謝いたします。最後に，日ごろからお世話になっております研究室の皆さまにも感謝いたします。

参考文献

- [1] Marko Auerswald. Example-based machine translation with templates. In *Verbmobil: Foundations of Speech-to-Speech Translation*, pp. 418–427, 2000.
- [2] Key Sun Choi and Tea Wan Kim. 日韓機械翻訳システムの現状及び分析. 言語処理学会第2回年次大会発表論文集, 1996.
- [3] You-Jin Chung, Sin-Jae Kang, Kyonghi Moon, and Jong-Hyeok Lee. Word sense disambiguation in a korean-to-japanese mt system using neural networks. In *Proceedings of the 16th International Conference on Computational Linguistic. Coling 02*, 2002.
- [4] Heakyong Ga. 日韓両言語における敬語の対照研究. 白帝社, 2001.
- [5] Myoung-Su Han, Kyong-Hi Moon, and Jong-Hyeok Lee. 日韓機械翻訳の性能向上のための助動詞れる/られるの曖昧性解決. 韓国情報科学紙, 2000.
- [6] Sangonk Lee Iikseob Lee and Wan Chai. 韓国の言語. シング文化社, 1997.
- [7] Incheoul Iim. 韓日機械翻訳. 情報科学会紙, 1989.
- [8] Jong-Bok Kim, Peter Sells, and Jaehyung Yang. Deep Processing of Honorification Phenomena in a Typed Feature Structure Grammar. In *The Second International Joint Conference on Natural Language Processing (IJCNLP-05)*, 2005.
- [9] Jung-In Kim, Kyung-Hee Moon, and Jong-Hyeok Lee. 日韓機械翻訳システムにおける進行形の翻訳処理. 韓国情報科学紙, 2000.
- [10] Taeseok Kim. 日韓機械翻訳システムの研究及び開発. 情報科学会紙, 1997.
- [11] Taewhan Kim. 機械翻訳技術の概要及び動向. 電子工学会紙, 1997.
- [12] Wonho Kim. 「お～になる」型と「れる」型敬語考. 日本学報 第25号 韓国日本学会, 1990.
- [13] Kevin Knight. A statistical mt tutorial workbook. *JHI summer workshop*, 1999.

- [14] Sadao Kurohashi and Makoto Nagao. A syntactic analysis method of long Japanese sentences based on the detection of conjunctive structures. *Computational Linguistics*, Vol. 20, No. 4, pp. 507–534, 1994.
- [15] Sadao Kurohashi, Toshihisa Nakamura, Yuji Matsumoto, and Makoto Nagao. Improvements of Japanese morphological analyzer JUMAN. In *Proceedings of The International Workshop on Sharable Natural Language*, pp. 22–28, 1994.
- [16] Sadao Kurohashi and Toshiaki Nakazawa. Example-based machine translation pursuing fully structural nlp. 2005.
- [17] Dong-Young Lee. Computation of relative social status on the basis of honorification in korean. In *Proceedings of the 16th International Conference on Computational Linguistic. Coling 96*, 1996.
- [18] Jong-Bok Lee and Peter Sells. Honorification in korean as expressive meaning. In *Korean Linguistics 13*, 2006.
- [19] Hiroyuki Maeda, Susumu Kato, Kiyoshi Kogure, and Hitoshi Iida. Parsing Japanese Honorifics in Unification-Based Grammar. In *Proceedings of the 26th Meeting of the Association for Computational Linguistics (ACL'88)*, 1988.
- [20] 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.
- [21] Shigeko Nariyama, Hiromi Nakaiwa, and Melanie Siegel. Annotating Honorifics Denoting Social Ranking of Referents. In *The Second International Joint Conference on Natural Language Processing (IJCNLP-05)*, 2005.
- [22] Carl Pollard and Ivan A Sag. *Head-Driven Phrase Structure Grammar*. University of Chicago Press, 1994.
- [23] Stephen D. Richardson, William B. Dolan, Arul Menezes, and Monica Corston-Oliver. Overcoming the customization bottleneck using example-based mt. In *Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL2001) Workshop on Data-Driven Methods in Machine Translation*, pp. 9–16, 2001.
- [24] Harold Somers. Example-based machine translation. In *Machine Translation*, pp. 113–157, 1999.

- [25] Eiichiro Sumita. Example-based machine translation using dp-matching between word sequences. In *Proceedings of the Annual Conference of the Association for Computational Linguistics (ACL2001) Workshop on Data-Driven Methods in Machine Translation*, pp. 1–8, 2001.
- [26] Eiichiro Sumita, Hitoshi Iida, and Hideo Kohyama. Translating with examples: A new approach to machine translation. In *Proceedings of the 3rd TMI*, pp. 203–212, 1990.
- [27] Tony Veale and Andy Way. Gaijin: A bootstrapping, template-driven approach to example-based mt. In *Proceedings of NeMNL97, New Methods in Natural Language Processing*, 1997.
- [28] 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.
- [29] 菊地康人. 敬語. 講談社学術文庫, 1997.
- [30] 金孝子. 日本語 敬語表現. (株)時事英語社, 1991.
- [31] 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.
- [32] 森下喜一, 池景来. 日本語と韓国語の敬語. 白帝社, 1989.
- [33] 太田ふみ, 笠井俊信. 機械処理を意識した敬語判定とその応用システム. 言語処理学会 第 12 回年次大会発表論文集, 2006.
- [34] 辻村敏樹. 敬語の分類について. 論集日本語研究 9 敬語, 有精堂出版, 1978.
- [35] 朴哲済, 文敬姫, 李根培. 連語パターンによる日韓機械翻訳システムの構築とその評価. 自然言語処理, 1995.
- [36] 朴哲済, 李根培. 日韓機械翻訳における連語パターンを用いる変換手法. 情報処理学会, 1997.
- [37] 渡辺実. 敬語体系. 北原保雄編, 論集日本語研究 9 敬語, 有精堂出版, 1978.
- [38] 渡辺太郎, 奥乃博. 日英機械翻訳の統計的モデル化の事例研究. 人工知能学会論文誌, 1997.
- [39] 白京姫, 大竹清敬, 山本和英. 異なる原言語からの翻訳による同義表現の分析. 言語処理学会 第 10 回年次大会発表論文集, 2004.

- [40] 白土保, 丸元聡子, 村田真樹, 内元清貴, 井佐原均. 日本語発話文に含まれる敬語の誤用を自動的に指摘するシステム-ウチソトの考慮、及び gui の開発-. 言語処理学会 第 12 回年次大会発表論文集, 2006.