

日韓機械翻訳における意味接続関係を用いた韓国語の生成方法[†]

金 泰 錫^{††} 浦 昭 二^{††}

日本語と韓国語は語順が同じであるなど文法的な面で類似点が多い。しかし、単純に日本語と韓国語の単語を直接対応づけるだけでは不自然な訳文が生成されることが多い。単純対応で不自然な訳文になる原因のひとつとして、韓国語の助詞はその前の言葉の意味によって細分化されることが挙げられる。また、韓国語には日本語の助動詞に対応する品詞がなく、複数の助動詞が使われた日本語の述部とは、その意味の表現の仕方と配列順序が異なることも一因である。これらの問題を解決するためには、助詞の翻訳にはその前接単語の意味を考慮して対応し、用言と助動詞に対しては、後接する単語の意味を考慮して対応する必要がある。そこで、著者らは日韓機械翻訳において一番問題になる用言、助動詞、助詞に対して、各単語ごとに前後にある単語の意味(意味接続関係)を考慮した訳語をあらかじめ登録(単語翻訳テーブル)し、それを用いて韓国語を生成する方式を提案する。この方式によって、プロトタイプ的な翻訳システムを構築し、翻訳を行った結果、1)韓国語の不規則用言は、複雑な活用を行うために、一般的な生成ルールを作成しにくい問題、2)日本語の述部の表現順序との違いに対する韓国語の対応問題、3)日本語の転成名詞に対する韓国語の対応問題が解決でき、その有効性を確認した。

1. はじめに

最近、社会の高度情報化とともに国際間の情報交流が激しくなり、機械翻訳システムの研究は急速に発展している。日韓機械翻訳に関してもこの数年の間に大変活発な研究が行われている¹⁾。そして、日韓両国語間では、語順や文法的な性質が類似していることから、直接翻訳方式による研究が行われている^{2)~4)}。日韓機械翻訳の場合、両国語間の類似点を活かすことにより、他言語間でよく行われているトランスファや中間言語方式に比べ、構文解析や意味解析のかなりの部分の省略が可能である。これまで、日韓両国語間の機械翻訳システムに関する主な研究として、日本との国際協力による韓国科学技術院をはじめ¹⁾、限られた助述表現と動詞のみを対象とする翻訳方法²⁾、そして、日本語の格助詞のみを対象とする助詞の扱い方法などが発表されている^{3),4)}。しかし、両国語間の機械翻訳が実用化されるためには、両国語間の用言や助動詞の複雑な活用の問題と助詞の使い方の違いなどの解決が問題点として残されている^{2)~4)}。本論文では、日韓機械翻訳を対象として、両国語間の用言や助動詞の複雑な活用の問題と助詞の使い方の違いに対応する方法を提案し、その有効性について検討する。

日韓機械翻訳において、注目すべき類似性として次の4点がある。

(1) 語順がほとんど同じである

例外もあるが、通常、日本語の単語に対応した韓国語を選び、日本語の語順どおりに並べればよい。

(2) 主語の省略が多い

主語の省略による曖昧さについて、その主語を推論する意味解析を行う必要がない。

(3) 単複数の区分がない

名詞の表現に単数、複数の区分がないため、それを解析するための意味解析を行う必要がない。

(4) 文法活用形の性質が似ている

用言の活用形態が似ており、日本語の活用の未然形、連体形、仮定形、命令形の場合は、ほとんど1対1に対応可能である。しかし、連用形と終止形の場合、その対応は多様である⁵⁾。

著者らは、上記の4点の類似性を利用した日韓機械翻訳システムを試作した。そして、新聞記事、マニュアル、科学技術文献などから集めた474文について翻訳を行った結果、文法的には少々誤りがあっても、人間が読んで理解できる範囲を合格として、約75%の翻訳率を得ることができた^{5),6)}。

一般的に、機械翻訳システムにより完ぺきな翻訳結果を得ることは不可能であり、後編集処理に頼ることが多い。しかし、実用化のためには、可能な限り、人間の手作業を減らすことが理想である。そのため、直接翻訳方式での不合格と判定された文について分析し

[†] Generation of Korean Based on Connection Forms of the Semantics in Japanese-Korean Machine Translation by TAI-SUK KIM and SHOJI URA (Department of Administration Engineering, Faculty of Science and Technology, Keio University).

^{††} 慶應義塾大学理工学研究科管理工学専攻

た結果、多訳性によるものを除き、ほとんどが助詞と述部の扱いに関するものであった。この原因として次のものを挙げるができる。

- 1) 日本語で同じ助詞が使われていても、その助詞がついている体言の使用上の意味によって、韓国語では異なる助詞が使われる(以下では、助詞がついている体言または用言を前接単語と呼ぶ)。
- 2) 韓国語には日本語の助動詞に対応する品詞がなく、日本語の助動詞が表す意味を韓国語では用言の語尾、または補助的連結語尾と補助用言で表現する。
- 3) 韓国語の用言は、その活用形が日本語と基本的には一致するが、日本語より多様であり、しかも、語幹まで変化する不規則用言が多いため、一般的な生成ルールを作成することは難しい。

著者らは、このような問題点を解決するために、次のような翻訳テーブルを考案し、それを用いた韓国語の生成方法を提案する^{6)~8)}。

- 1) 用言翻訳テーブル：日本語の用言に対する韓国語は、日本語の用言に後接する付属語の意味、または文法情報(品詞名と活用形)によって異なる(以下では、用言に後接する付属語を後接単語と呼ぶ)。それを各用言ごとにまとめたテーブル(これには、動詞翻訳テーブル、形容詞翻訳テーブル、形容動詞翻訳テーブルの区別がある)。
- 2) 助詞翻訳テーブル：日本語の助詞に対する韓国語は、前接単語の意味、または文法情報によって異なる。それを各助詞ごとにまとめたテーブル。
- 3) 助動詞翻訳テーブル：韓国語には日本語の助動詞の機能を持つ品詞がないため、その意味によって用言の語尾、または補助的連結語尾と補助用言によって対応しなければならない。後接する単語の意味、または文法情報によって異なる韓国語に対応するが、それを各助動詞ごとにまとめたものが助動詞翻訳テーブルである。日本語の助動詞は複数の意味を表す場合があり、その意味ごとに翻訳テーブルを用意する。

韓国語生成に当たっては、日本語の解析処理で認定された単語とその意味に応じた各翻訳テーブルを辞書から読み込む。そして、1文節の中で前後の意味接続関係を参照し、各翻訳テーブルから、適切な訳語を選択する。

本論文の構成は以下のとおりである。

- ・2章では、日本語の用言、助詞、助動詞を韓国語に

翻訳するときの問題点を示し、上記の3種類の翻訳テーブルについて述べる。

- ・3章では、各翻訳テーブルを用いて韓国語を生成するために必要な日本語の解析と韓国語の生成について述べる。
- ・4章では、本論文で提案した翻訳テーブルを用いたときの翻訳結果を例示し、その有効性について検討する。

2. 意味接続関係に基づいた翻訳テーブル

日本語と韓国語の間では、助詞と述部の表現の仕方が異なる^{2)~4)}。そのため、用言、助詞、助動詞に対しては、あらかじめその違いを考慮した韓国語を用意することは日韓機械翻訳で大変有効である。

本章では、両国語間の用言、助詞、助動詞について比較し、日韓機械翻訳システムのための各単語翻訳テーブルについて述べる。

2.1 用言の翻訳テーブル

日本語の用言はその品詞別に決まった規則によって活用を行う。しかし、韓国語の用言には、一般的な韓国語の音韻規則では説明できない活用を行う用言(不規則用言)が多い¹³⁾。そして、日本語の用言を韓国語に翻訳するとき、日本語の用言の後にくる助動詞などの意味を考慮して、韓国語の用言を決めなければならない。これは、韓国語には、用言に意味をそえる日本語の助動詞に対応する品詞がなく、日本語の助動詞の意味が韓国語の用言の活用語尾として表現されるためである。また、日本語では、否定の意味を持つ単語(ない、ん、ぬ)によって否定を表すが、その否定語に対応する韓国語は前接単語によって区分される(表6参照)。また、韓国語では、主語の心の状態に関わる特定の動詞(解る、知るなど)に対しては、その否定の意味を持っている対立語によって表現する(3.3節で詳説)。

次に挙げるものは動詞の翻訳の違いを示す例文である。

例文 1) 野球を見る。

야구를 보다. (yagurul boda.)

この“見る”は終止形であって、そこで文が終了していることから、韓国語の“보다(boda)”の終止形に対応する“보다(boda)”に訳される。

例文 2) 野球を見るが、

야구를 보지만, (yagurul bojiman.)

この場合の“見る”は例文1と同じ終止形であるが、

接続助詞“が”が付いているので、その文節が確定の逆接の意味になる。この場合、接続助詞“が”は確定の逆接の意味を持つ韓国語“지만(jiman)”に訳される。そして、その文節の意味が確定の逆接を表すためには、韓国語の動詞“보다(boda)”はその語幹である“보(bo)”だけになる。

例文 3) 野球を見た。

야구를 보았다. (yagurul boattda.)

この“見”は連用形であり、過去の意味の“た”が接続されている。この場合、韓国語では過去等の時制の表現は用言の活用語尾で表現する。したがって、語幹“보(bo)”に過去時制の意味を持つ語尾“았(att)”を付け、“보았(boatt)”に訳される。

例文 4) お金がありません。

돈이 없습니다. (doni opsumnida.)

“ある”に対応する韓国語“있다(ittda)”の語幹“있(itt)”に丁寧と否定の意味の言葉を付けて訳すと“돈이 있지 않습니다(doni ittji ansumn ida)”になるが、これは不自然な表現である。この場合は、“있다(ittda)”の否定の意味を表す別の用言“없다(opda)”を用いる。そして、“없다(opda)”の語幹“없(op)”に丁寧の意味を付け“없습니다(op sumnida)”のように翻訳しなければならない。

例文 5) 誤った考えが、

잘못된 생각이, (jalmotdyon sengkaki)

“考え”は格助詞“가”が接続されている転成名詞である。これは、動詞“考える”に対応する韓国語“생각하다(sengkakhada)”の語幹“생각하(sengkakha)”の一部分である“생각(sengkak)”に訳される。一般に、日本語の転性名詞を翻訳するとき、韓国語では、語幹の一部分、またはまったく別の名詞に訳される。

上の例文 1)~5)でわかるように、日本語の用言を韓国語に翻訳するためには、日本語の用言に続く付属語の意味を考慮してから韓国語の用言を決めなければならない。韓国語の用言は、後接単語の意味によって語幹が変化したり、別の用言が必要な場合があるためである。このような問題を、すべての用言に対して一般的な翻訳ルールを作ることは難しい¹³⁾。このため、日本語の用言ごとに、後接単語の意味を考慮して、それに適切な韓国語をまとめたテーブルが必要になる。それが用言翻訳テーブルである。用言すなわち動詞、形容詞、形容動詞全体に通じるテーブルの形式がないので、それぞれ異なる翻訳テーブルを用意する。

表 1 は動詞“行く”の翻訳テーブルである。動詞翻訳テーブルには後接する単語の意味を考慮した適切な韓国語を登録する。例えば、表 1 の第 1 行は、否定を表す助動詞が付いている場合、韓国語では、対立語を用いて訳す場合がある。そのときの、対立語を登録するために、設けられている。そして、第 2 行は、動詞“行かせる”の場合、“行か”は未然形であり、その後に使役

表 1 動詞‘行く’の翻訳テーブル
Table 1 Table for translation of verb ‘iku.’

活用形	活用単語	接続情報による韓国語の対応			適用例	出現頻度 (1002)
		No.	後接単語の意味および文法情報	対応する韓国語		
未然形	行か	1	否定	가 (ga)	行かない	89
	行こ	2	否定以外	가 (ga)	行こう	6
連用形	行き	3	丁寧, 希望	가 (ga)	行きます	100
		4	丁寧+過去	갔 (gatt)	行きました	33
		5	丁寧+否定	가 (ga)	行きません	29
		6	様態	갈 (gal)	行きそうだ	0
		7	用言	가기 (gagi)	行きにくい	6
		8	連用中止形	가서 (gaso)	行き,	47
		9	転成名詞	행 (heng)	行き	21
		10	過去	갔 (gatt)	行った	257
		終止形	行く	11	推定, 伝聞	갈 (gal)
12	丁寧			가 (ga)	行くでしょう	6
13	接続助詞			가 (ga)	行くが,	76
14	終了			가다 (gada)	行く.	6
連体形	行く	15		갈 (gal)	行く人は	305
仮定形	行け	16		가 (ga)	行けば	9
命令形	行け	17		가 (ga)	行け.	0

を表す助動詞“せる”が付いていることから“가(ga)”に訳される。第2行の訳語は、その後に使役、推量、意志などの否定以外の意味を表す助動詞がくるときの訳語を示す。そして、表1の第8行は動詞“行く”が転成名詞で使われたときには、韓国語“행(heng)”がその訳語を示す。そして、例文4)のように語彙対立によって否定を表す用言の場合には、その対立語に対する翻訳テーブルを作成し、登録しておく。なお、形容詞、形容動詞の翻訳テーブルについては、金¹⁰⁾を参照されたい。

2.2 助詞の翻訳テーブル

日本語の助詞を韓国語に訳す時、それが格助詞ならば、1対1に対応させても、かなりの程度まで翻訳可能である³⁾。しかし、格助詞にも例外があり、その助詞の前接単語の意味を考慮して異なる言葉に訳さなければならないことがある^{3),4)}。また、韓国語の助詞は、その機能と意味が同じであっても音韻の衝突を避けるため、前接単語の最終文字の形態によって変化することがある。ここでは、日本語の助詞“に”を例にして述べる。

例文6) 運動場に 集合する。

운동장에 집합하다.

(undonjange jippaphada.)

体言が“場所”、“時間”、“期間”の意味を表すときは“에(e)”に訳される。

例文7) 君に 仕事を 頼む。

너에게 일을 부탁하다.

(unege ilul butakhada.)

助詞“に”の前接単語が“生物”である場合は、“君”が“対象”を表すので、“에게(ege)”に訳される。

例文8) 野球を見に行く。

야구를 보러가다.

(yagurul boroe gada.)

助詞“に”が動詞の連用形に接続され、その文節が“動作の目的”を表すときは“러(roe)”，または“으러

(uroe)”に訳される。両者の違いは音韻からくるものであり、前接文字の終子音の有無によって区分することができる。

例文9) 野球を見るにはお金がかかる。

야구를 보는데는 돈이든다.

(yagurul bonundenun donidunda.)

動詞の連体形に接続され、その文節が動作の条件の意味を表すときは、“는데(nunde)”に訳される。

上の例文6)~9)でわかるように、日本語の助詞を訳すには、その前接単語の使用上の意味を考慮しなければならない。表2は助詞“に”を韓国語に訳するための前接単語の意味を考慮した翻訳テーブルである。そして、格助詞以外の助詞についても、表2と同様に、各助詞ごとの翻訳テーブルを用意する。しかし、ほとんどの日本語の助詞に対して、その前接単語の意味によって韓国語を対応させることができるが、両国語の間に微妙な使い方の違いがあり、前接単語の意味だけでは対応しきれない場合がある²¹⁾。特に、日本語の助詞“の”の使い方は大変多様であるので、適切な韓国語を対応させるにはもっと詳細な調査が必要である^{3),19),20)}。また、日本語には、複数の単語が複合しあい、全体で1つの意味を表す複合辞¹²⁾がある。複合辞の中には、助詞と同様な動きをする表現がある。例えば、“からすると”、“にとって”などの慣用的な表現である。これらの中には各訳語の接続だけでは正しく翻訳されない場合があり、さらに検討を加える必要がある。

2.3 助動詞の翻訳テーブル

用言に意味をそえる日本語の助動詞と同じ機能を持つ品詞が韓国語にはない。受身、使役、可能などの意味は補助的連結語尾と補助用言によって表現する^{14)~16)}。そして、否定と時制などは、用言語尾、または補助用言で表現する。

次の例文は助動詞“ませ”がそれに後接する単語の意味によって別の韓国語に訳されることを示している。

表2 助詞‘に’の翻訳テーブル
Table 2 Table for translation of particle ‘ni.’

No.	前接単語	前接単語の意味 および文法情報	対応する韓国語	適用例
1	体言	場所, 期間, 時間	에 (e)	学校に行く。
2	体言	対象	에/에게 (e/ege)	君に頼む。
3	体言	結果, 理由, 方向	로/으로 (ro/uro)	失敗に終わる。
4	動詞または助動詞の連用形	動作の目的	러/으러 (roe/uroe)	見に行く。
5	動詞または助動詞の連体形	動作の条件	는데 (nunde)	見るにはお金がかかる。

表3 助動詞‘ます’の翻訳テーブル
Table 3 Table for translation of auxiliary verb ‘masu.’

活用形	活用単語	接続情報による韓国語の対応			適用例	出現頻度 (171)
		No.	後接単語の意味および文法情報	対応する韓国語		
未然形	ませ	1	否定	지 않습니 (ji ansumni)	行きません。	32
		2	否定+過去	지 않았 (ji anatt)	行きませんでした。	4
	ましよ	3	勧誘	(으)십시오 ((u)bsi)	行きましょう。	7
連用形	まし	4	接続助詞	으(서) (u(syo))	行きまして。	0
		5	過去	습니 (subni)	行きました。	38
終止形	ます	6	終了	(스) 습니다((su)bnida)	行きます。	79
連体形	ます	7	接続助詞	(스) 습니다((su)bnida)	行きますが。	11
		8	その他	(으)실 ((u)sil)	行きます時。	0
仮定形	ますれ	9		(으)시 ((u)si)	行きますれば。	0
命令形	ませ	10		(으)세요 ((u)seyo)	いらっしやいませ。	0
	まし	11		(으)세요 ((u)seyo)	いらっしやいまし。	0

例文 10) 学校に行きません。

학교에 가지 않습니다.

(hakkyoe gaji ansumnida.)

この場合、“ませ”は、その活用形が未然形であり、後接単語の意味が“否定”であるので、表3の第1行の“지 않습니(ji ansumni)”に訳すことができる。日本語の述部の意味表現順序は“丁寧、否定”であり、韓国語は“否定、丁寧”である。

例文 11) 学校に行きませんでした。

학교에 가지 않았습니다.

(hakkyoe gaji anattsumnida.)

この場合、“ませ”は、その活用形が未然形であり、後接単語の意味が“否定”、“丁寧”、“過去”であるが、“ませ”が“丁寧”を表しているため、“否定”と“過去”の意味によって、表3の第2行の“지 않았(ji anatt)”に訳すことができる。日本語の意味表現順序は“丁寧、否定、丁寧、過去”であり、韓国語では“否定、過去、丁寧”の順である。

日本語の各助動詞に対して、それに後接する付属語の意味によって適切な韓国語に対応させたものが助動詞翻訳テーブルである。表3に助動詞“ます”の翻訳テーブルを示す。

日本語の述部に複数の助動詞が使われているとき、各助動詞翻訳テーブルからあらかじめ両国語間の表現の違いが考慮された訳語を選択することができる。そして、助動詞“れる”、“られる”は複数の意味を持つが、その意味ごとに別の翻訳テーブルが必要である。なお、助動詞“です”は特別に扱う必要がある。“です”は、その前接単語の性質によって3つの言葉に訳されるので、次の3種の翻訳テーブルを用意する(詳細は金¹⁰⁾

を参照)。

- ①前接単語が体言、助詞、形容動詞型活用を行う助動詞の語幹である場合は、韓国語“입니다(imnida)”を母体にした翻訳テーブル
- ②前接単語が形容動詞の語幹の場合は韓国語“습니다(hamnida)”を母体にした翻訳テーブル
- ③形容詞の終止形と形容動詞型の活用を行う助動詞の終止形の場合は韓国語“합니다(sumnida)”を母体にした翻訳テーブル

3. 日本語の解析と韓国語の生成処理

日本語の解析では、日本語の文から最小単位である単語の認定を行う。そして、体言の意味と助動詞の意味を抽出する。このとき、多義性を持つ単語や意味判定が困難なときについては対話的にその意味を与える。

韓国語の生成処理では、認定された単語とその単語の意味によって、辞書から各翻訳テーブルを読み込む。次に各翻訳テーブルから適切な韓国語を選択し、韓国語の文法にしたがって韓国語を生成する。

3.1 日本語の形態素解析

日本語文の形態素解析を行う際には、自立語を登録した自立語辞書と助詞、助動詞を登録した付属語辞書を利用する。形態素解析では文の先頭から自立語辞書とマッチングし、成功した単語を各品詞別解析部に渡す。品詞別解析部では、活用品詞の場合、その活用を行い、成功かどうかを判定する。次に、自立語の後には付属語がくる場合が多いことから先に付属語を探索する²²⁾。そして、その自立語の文字数(先読み単語の文字数)を求める。複数の文節が存在した場合は、各文

表4 日本語の形態素解析の例
Table 4 Example for analysis of Japanese morpheme.

認定された単語	品詞名	活用形	活用文字数	意味	付属語の文字数	先読み文字数	備考
学校	名詞			場所	1(に)	2(行き)	
に	助詞						1文節目 学校に
行き	動詞	連用形	1(き)		6(ませんでした)	1(.)	
ませ	助動詞	未然形	1(せ)	丁寧	1(ん)		
ん	助動詞	終止形		否定	2(でし)		
でし	助動詞	連用形	1(し)	丁寧	1(た)		
た	助動詞	終止形		過去			2文節目 行きませんでした
.				終了			終了(.)

節の自立語の文字数、活用文字数、付属語の文字数、先読み単語の文字数の総和が最も大きい文節を選ぶ。ただし、その総和が同じである場合は、活用文字数、付属語の文字数が大きい文節を優先する。そして、次の自立語を探し出し、文が終了するまで繰り返す⁷⁾。

表4はその解析例であり、用言と助動詞の場合はその活用形も結果として出力する。

3.2 単語の意味抽出

翻訳テーブルから適切な訳語を選択するためには、体言と助動詞の意味を抽出する必要がある。ここでは、“意味”という言葉それぞれの場合で、異なる意味合いで用いている。すなわち、体言の意味は、体言に助動詞が付いて表す使用上の意味のことであり、助動詞の意味は、用言または体言などに意味をそえる文法上の意味である。

(1) 体言の意味抽出

体言の場合の意味とは、例えば、行き先を示す“場所”、建てる“目的”物など、その体言が使われる目的のことであり、日本語の助動詞を翻訳するとき、その意味に応じた訳語を選ぶために使われる。これは、韓国語の助動詞は前接体言の意味によって、異なる言葉が使われるためである。例えば、表2の1, 2, 3番は、日本語の助動詞“に”が前接体言に付いて表す使用上の意味によって異なる韓国語に訳される例である。

体言の意味分類については、意味解析処理のために、数千種もの分類が必要とされている²⁵⁾。しかし、本論文では、体言の意味分類の目的は、その意味に応じた韓国語の助動詞に訳すことであり、一般的に行われている意味分類方法^{18), 25)}とは異なる分類を行う。つまり、韓国語の助動詞が使われた名詞句が表す使用上の意味によって体言を分類することである。

次の例文で、IPAL辞書¹⁸⁾の体言の意味分類と本論

表5 体言の意味分類

Table 5 Semantic classifications of uninflection word.

No.	意味分類	体言の例
1	理由	風で倒れる。
2	場所	運動場に行く。
3	由来, 起点	日本から来た人。
4	方向, 進路	右に回る。
5	材料, 道具, 手段, 方法	紙で作る。
6	資格	班長に選ばれる。
7	変性, 結果	失敗に終わる。
8	付与, 対象	太郎にあげる。

文で用いられる体言の使用上の意味分類の違いについて述べる。

例文 12) 学校に行く。

この場合の“学校”は、前者では“場所”に分類されている。そして、後者でも、“学校”はその行き先を示しているために、前者と同じく、“場所”に分類する。

例文 13) 学校を建てる。

この場合の“学校”は、前者では“建物”に分類されている。しかし、後者での“学校”は、建てる目的物を示しているために“目的”に分類する。この場合の助動詞“を”の訳語は“学校”に目的格の資格をそえる韓国語の助動詞“을(rul)”に対応する。

例文 14) 学校を休む。

この場合の“学校”は、前者では“動作”に分類されている。しかし、後者での“学校”は、休む動作の目的物を示しているために、例文 13)と同じく、“目的”に分類する。

このように、体言の意味は、韓国語の助動詞が使われた名詞句の意味によって分類したものであり、“高校用韓国語文法”²⁶⁾の韓国語の助動詞の使い方に基づいて、表5のように分類する。この分類にしたがって、体言が一

表6 韓国語の否定語の選択表
Table 6 Selection of negative expression in Korean.

日本語の否定語	否定語の前接情報	韓国語の否定語	日本語の例文	韓国語の例文
ない	動詞(居る) 体言+助詞 可能動詞 (可能の意味の助動詞)	없다 (opda)	友達がいない。 お金がない。 行けない。	친구가 없다 (chinguga opda) 돈이 없다 (doni opda) 갈 수 없다 (gal su opda)
	形容詞 体言+助詞(では, じゃ) 形容動詞 動詞(その他)	지 않다 (ji anta)	美しくない。 先生ではない。 重要でない。 行かない。	아름답지 않다 (arumdapji anta) 선생이지 않다 (sunsengiji anta) 중요하지 않다 (jungyohaji anta) 가지 않다 (gaji anta)
ん(ぬ)	動詞(有る, 居る) 可能動詞 (可能の意味の助動詞)	없다 (opda)	お金がありません。 友達がいません。 食べれぬ。	돈이 없습니다 (doni opsumnida) 친구가 없습니다 (chinguga opsumnida) 먹을 수 없다 (mokuisu opda)
	動詞(その他)	지 않다 (ji anta)	雨が降らぬ。	비가 오지 않다 (biga oji anda)
	助動詞(ます)	다 (da)	行きません。	가지 않습니다 (gaji ansumnida)

一般的に使われる使用上の意味を体言の辞書に登録する。

この分類については、日本語の助詞 42 個を対象として検討を行い、十分に対応可能であることを確認した¹⁰⁾。しかし、体言の意味は共起する述語によってその意味が異なる²³⁾など、その意味判別を自動的に行うことは、まだ問題が残っている。このような体言の意味判別の問題は本論文の主題ではないため、ここでは、辞書に登録している意味と異なって使われた場合は、韓国語の生成過程で、その意味を対話的に与えることにする。

(2) 助動詞の意味抽出

日本語の助動詞は複数の意味を持つことがあるが、その意味がわかれば、その意味を表す韓国語に対応させることができる²⁾。つまり、複数の意味を持つ日本語の助動詞には、その意味別に異なる韓国語に対応させればよい。助動詞の意味は“くわしい国文法”²²⁾に基づいて分類し、それを辞書に登録しておく。しかし、日本語の複数の意味を持つ助動詞“れる”、“られる”の意味判別には、まだ問題が残っている²⁾。ここでは、韓国語の生成過程で、その意味を対話的に与えることにする。

3.3 韓国語の生成処理

韓国語の生成処理では、まず、日本語の形態素解析で得られた各単語とその意味によって、各単語翻訳テーブルを辞書から読み込む。次に、単語の意味接続関係によって単語翻訳テーブルから適切な韓国語を選択する。

ここでは、表 4 での“学校に行きませんでした”を例

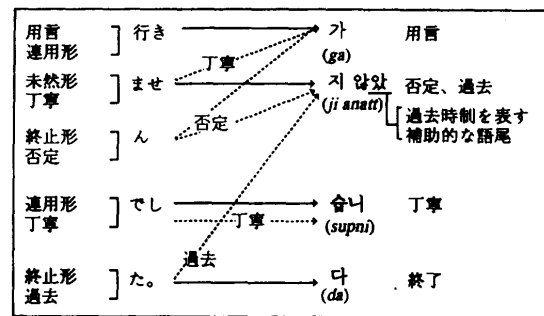


図1 日本語と韓国語の述語の表現の比較
Fig. 1 Comparison of Japanese and Korean predicative expression.

にして説明する。まず、助詞“に”は前接单語(名詞“学校”)の意味(場所)から“에(e)”と翻訳される(表 2 の第 1 行)。そして、“行きませんでした”は、その文節の中に現れている助動詞の意味は“丁寧”、“否定”、“丁寧”、“過去”の順である。動詞“行き”の活用形は連用形であり、その後“ませ(丁寧)”と“ん(否定)”が接続されているために、“行き”は“가(ga)”(表 1 の第 5 行)に翻訳される。次に、丁寧の意味を持つ助動詞“ませ”はその活用形が未然形であり、後接单語として自分の意味とは異なる“否定”、“過去”が続くことから、“지 않았(ji anatt)”(表 3 の第 2 行)に翻訳される。このときの“ませ”に対応する韓国語“지 않았(ji anatt)”は否定と過去の意味を同時に表す表現である。次に、“ん(否定)”は 3 種類の韓国語に対応しているため、表 6 からその前接单語が助動詞“ませ”であることで、“다(da)”の翻訳テーブルを選択する。そして、“다(da)”の翻訳テーブルの中から後接单語の意味が丁寧である欄

表7 'です'に対応する3つの韓国語
Table 7 Korean expressions of corresponding to 'desu.'

No.	'です'の前接単語	対応する韓国語	例 文
1	体言 助詞(の, からなど) 助動詞(形容動詞型活用)	입니다 (imnida)	先生です 東京からです 行くようです
2	形容動詞の語幹	합니다 (hamnida)	健康です
3	形容詞(連体形, 終止形) 助動詞(その他)	습니다 (sumnida)	美しいです 行きたいです

の訳語を選択する。この場合の“ん(否定)”については、その前の単語“ませ”に否定の意味が移されたため、その訳語は必要ない(図1を参照)。次に、“でし(丁寧)”も3種類の韓国語に対応しているため、表7からその前接単語が助動詞であることで、“습니다(sumnida)”の翻訳テーブルを選択する。そして、“습니다(sumnida)”の翻訳テーブルの中から、後接単語“た”の意味が過去であるため、“습니(sumni)”に訳す。次に、“た(過去)”の過去の意味は“ませ(丁寧)”の訳語に移されたため、文の終了を表す叙述表現“다(da)”に訳す(図1参照)。次に、韓国語の文法にしたがって、各選択された韓国語を並べると、“학교에 가지 않았습니 다.(hakkyoe gaji anattsumnida)”が得られる。韓国語の生成アルゴリズムは次のとおりである。

[アルゴリズム] 韓国語の生成手順

ステップ1. 日本語解析の結果から1文節の各単語
翻訳テーブルを読み込む

ステップ2. 自立語の韓国語の選択

2-1. 体言の場合

日本語の体言に対応する韓国語を選択する

2-2. 用言の場合

用言の活用形と後接する意味情報、または文法情報によって用言翻訳テーブルからその韓国語を選択する(否定によって別の用言に変換する場合は、その処理を行い、2-2へ)

ステップ3. 助詞の韓国語の選択

3-1. 体言に接続された場合

体言の意味によって助詞翻訳テーブルから韓国語を選択する

3-2. 用言、または助動詞に接続された場合

前接単語の品詞とその活用形によって助詞翻訳テーブルから韓国語を選択する

ステップ4. 助動詞の韓国語の選択

後接単語の意味によって助動詞翻訳テーブルから

韓国語を選択する

ステップ5. 文節が終了でなければ3へ(接続付属語の処理)

ステップ6. 韓国語の文法による再合成

6-1. 体言の最終文字形態による助詞の音便処理

6-2. 空白処理

ステップ7. 文が終了でなければ1へ(次の文節処理)

前に述べたとおり、“解る”、“知る”などの動詞に否定の意味を持つ助動詞が接続されたときは、その対立語に変換する必要がある。例えば、図2の“知りませんでした”の場合、後接する“ん”が否定の意味を持っているために“知る”は“모르다(moruda):知らないの意味”の翻訳テーブルに切り替える。次に、重複する意味を整理し、付属語の活用形を合わせる。そして、各単語の翻訳テーブルから整理された情報によって韓国語を選択し、韓国語を生成する。図2は否定による韓国語の用言変換処理を示す。このような変換処理は、変換された用言の後接単語の意味によって、適切な訳語に訳すことが可能である。しかし、日本語の否定文は、韓国語の否定補助用言を用いた否定文、または対立語を用いた肯定文に訳されるなど⁹⁾、今後、さらに、検討を加える必要がある。

4. 翻訳テーブルの有効性と翻訳結果の考察

まず、ここで提案した単語翻訳テーブルの有効性を調べるため、P. W. Andersonの論文¹⁷⁾を対象にして、そこに現れた動詞1002個と助動詞“ます”171個を用いて、表1と表3の有効性を確認するために、プロトタイプ的な翻訳システムを構築し、各意味接続別の頻度を調べた。1002個の動詞の活用形とその後接情報が一致するものの頻度を表1に示す。その結果、動詞の多義性による誤訳を除いて、すべての動詞が自然な韓国語に訳されたことから、動詞翻訳テーブルが有効であると思われる。そして、171個の助動詞“ます”の後接情報別の頻度を表3に示す。助動詞“ます”が自然な韓国語の丁寧な表現に訳されたことから、表3が有効であると思われる。そして、表3のような意味接続別翻訳テーブルを各助動詞ごとに作成することで、助動詞の翻訳に十分対応できるものと思われる。なお、出現頻度が零である型があるが、それらは科学技術文献ではほとんど用いられない表現である。

さらに、単純な直接翻訳方式⁹⁾で生成された不自然さがどの程度解消されるかを調べるため、提案した各

単語翻訳テーブルを用いて行った主な翻訳結果を表 8 に示す。表 8 の 1 番は特定の用言が否定されたとき、韓国語にはその対立語彙で否定を表す例であり、2 番は両国語間の表現順序の違いが解消された例である。

表 8 の 3 番は助動詞“られる”が“受け身”、“可能”、“自発”、“尊敬”のそれぞれの意味によって異なる訳語が用いられる例である。表 8 の 4, 5 番は助動詞“です”が前接単語によって異なる訳語が用いられた例であ

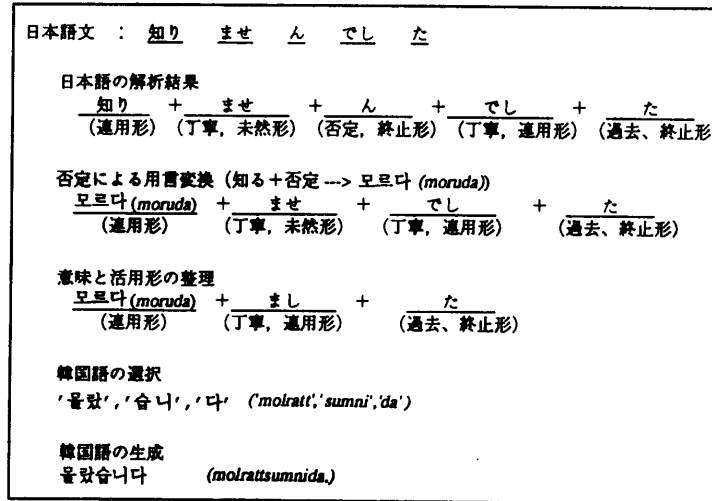


図 2 否定による韓国語の用言変換

Fig. 2 Convert of Korean declinable-word for negative expression.

表 8 翻訳結果の比較

Table 8 Comparison of translation results.

No.	日本語文	直接対応による結果	直接対応による誤訳の原因	翻訳テーブルによる結果
1	～つもりはありません。	예정은 있지 않습니다 (yejungun ittji ansumnida.)	否定の意味によって‘ある’は別の用言‘있다(opda)’に翻訳しなければならない。	예정은 없습니다 (yejungun opsumnida.)
2	～訳ではありませんでした。	이유는 있지 않습니다 (iyunun ittji ansumni ottda.)	1番と同じ。しかも、過去時制の表現に誤りがある。	이유는 없었습니다 (iyunun opsotsumnida.)
3	～広く知られています。	넓게 알게 되어 있습니다 (nulge algedyeoe ittsumnida.)	助動詞‘られる’が受け身に訳されていない。	넓게 알려져 있습니다 (nulge alryojoe ittsumnida.)
4	～違うのですが、彼は～	다를 것이지만, 그는 (darulgot ijiman, gunun)	‘です’は前接単語の品詞によって対応する韓国語が異なる (表 7 参照)。	다른 것입니지만, 그는 (darungotipnidaman, gunun)
5	～美しいです。	아름답습니다 (arumdapipnida)	4番と同じ。	아름답습니다 (arumdapsumnida)
6	～見るにはお金がかかる。	보게는 돈이 든다 (bogenun doni dulda)	‘に’は動詞‘見る’の連体形についてその文節が動作の条件を表すため、‘는데(nunde)’に翻訳しなければならない。	보는데는 돈이 든다 (bonundenun doni dulda)
7	～見に行く。	보게 가다 (boge gada)	‘に’は動詞‘見る’の連体形についてその文節が動作の目的を表すため、‘러(roe)’に翻訳しなければならない。	보러 가다 (boroe gada)
8	～書いた。	썼다 (suottda)	不規則用言の語幹に誤りがある。	썼다 (suttda)
9	～考えます。	생각합니다 (senggakhamnida)	連用形‘考え’の場合は正しく訳されている。	생각합니다 (senggakhamnida)
10	～考えをする。	생각하기를 하다 (senggakhagiul hada)	転成名詞‘考え’の訳語に誤りがある。	생각을 하다 (senggakul hada)

る。表8の6, 7番は助詞“に”がその前接単語である動詞“見る”の活用形によって、異なる韓国語で翻訳された例である。表8の8番は不規則用言の語幹が変化した翻訳例であり、9番と10番は動詞の連用形と転成名詞の翻訳の違いを表す例である。

このように、本論文で提案した翻訳テーブルを用いて、述部の翻訳処理を試みた結果、韓国語の述部がより自然な韓国語に生成されることを確認できた。

意味接続関係に基づく翻訳テーブルを用いる日韓機械翻訳システムにおいて、体言や用言の多義性の問題などがあるが、これらについて検討を加えながら、今後、さらに翻訳システムを充実させていきたい。

また、両国語間の類似点のみを利用した翻訳システムに用いられた474文についての翻訳結果など全体的な評価については、否定文の扱いを含むシステムに完了させた後、総合的に評価・検討を加え、別途、報告する予定である。

5. おわりに

日韓両国語間の類似点を考慮した直接翻訳方式での不自然な訳文を避けるため、述部中の意味接続関係を考慮した翻訳テーブルを提案し、それを利用する翻訳方式を提案した。

その結果、直接翻訳方式での不自然な訳文になる原因であった韓国語の不規則用言に対する活用問題、述部の表現順序の違いの問題、転成名詞に対する韓国語の対応問題、対立語による否定表現の問題、意味によって細分化される韓国語の助詞の問題などが解決でき、より自然な表現に翻訳可能であることがわかった。しかし、解決を要する若干の問題が残っている。そのひとつは、体言の使用上の意味判別の問題である。これは文型による解析処理を取り入れれば、かなりの部分が解消されると思われる¹⁸⁾。また、日本語の助詞“の”のように、その使い方が多様である助詞をより精密に処理するためには、名詞の意味素性と名詞間の意味関係を明確にする必要がある^{23)~25)}。

今後、これらの問題点を解決しつつ、両国語間の類似点を活かした日韓翻訳システムの実現に向け、さらに改良、拡張を図りたい。

参考文献

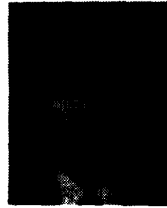
- 1) 野村浩郷, 田中穂積(編): 機械翻訳, Bit 別冊, 共立出版(1988.9).
- 2) 李 義東, 中嶋正之, 安居院猛: 助述表現の意味対

- 応による日韓述部機械翻訳システム, 情報処理学会論文誌, Vol. 31, No. 6, pp. 801-809(1990).
- 3) Choi, K. S. ほか: 日本語翻訳システムの環境下での韓国翻訳システム開発のための一考察, 情報処理学会自然言語処理研究報告, NL-68-4(1988).
- 4) 渡辺正己: ATLAS-Iにおける日韓機械翻訳システムの一考察, 第38回情報処理学会全国大会論文集, pp. 341-342(1989).
- 5) Kim, T. S. and Ura, S.: A Study on “Japanese-Korean” Machine Translation by Similarity between Two Languages, Technical Report No. 91001, Dep. of Administration Eng. Faculty of Science, Keio Univ.(1991. 2).
- 6) Kim, T. S. and Ura, S.: A “Japanese-Korean” Machine Translation Based on Conjugated Words Analysis, *The 1991 International Conference on Electronics, Information and Communications (ICEIC'91)*, P. R. China, pp. 199-203(1991. 8).
- 7) 金 泰錫, 金 政仁, 大駒誠一, 浦 昭二: 意味接続関係に基づく翻訳テーブルを用いた日韓機械翻訳における日本語の形態素解析, 第43回情報処理学会全国大会論文集(3), pp. 201-202(1991).
- 8) 金 泰錫, 浦 昭二: 意味接続関係に基づいた翻訳テーブルを用いた日韓機械翻訳システム, 第43回情報処理学会全国大会論文集(3), pp. 203-204(1991).
- 9) 金 泰錫, 浦 昭二: 日韓機械翻訳における述部の否定語の処理, 第44回情報処理学会全国大会論文集(3), pp. 141-142(1992).
- 10) 金 泰錫, 浦 昭二: 日韓機械翻訳における意味接続関係を考慮した翻訳テーブル, Technical Report, No. 92002, Dep. of Administration Eng., Faculty of Science, Keio Univ.(1992. 3).
- 11) 洪 思満: 国語特殊助詞論一意味分析, 学文社(韓国)(1987).
- 12) 森田良行, 松木正恵: 日本語表現文型一用例中心・複合辞の意味と用法, アルク(1990).
- 13) 張 S. J.: 活用論研究, 塔出版社(韓国)(1989).
- 14) 玉村文郎: 講座日本語と日本語の教育, 第7巻日本語の語彙・意味(下), pp. 81-105, 明治書院(1990).
- 15) 宋 敏: 韓日両国語の比較について(日本語古層への道), 月刊言語, Vol. 16, No. 7, pp. 118-128(1989).
- 16) 李 M. J.: 現代朝鮮語の受身表現について(日本語との対照), 言語研究, 第71号, pp. 70-72(1987. 3).
- 17) Anderson, P. W. (森, 米沢(訳)): 科学の美学について, ヒューマン・ハーモニー(科学と社会の調和をもとめて), pp. 17-43, 三田出版会(1989).
- 18) 情報処理振興事業協会: 計算機用日本語基本動詞辞書 IPAL(解説編), pp. 30-54(1987).
- 19) 島津 明, 内藤昭三, 野村浩郷: 助詞「の」が結ぶ名詞の意味関係の subcategorization, 情報処理学会自然言語処理研究報告, 53-1, pp. 1-8(1986).

- 20) 平井 誠, 北橋忠宏: 日本語文における「の」の連体修飾の分類と解析, 情報処理学会自然言語処理研究報告, 58-1, pp. 1-8(1986).
- 21) 田中康仁, 吉田 将: 語と語の関係について—“に”について—, 情報処理学会自然言語処理研究報告, 73-5, pp. 33-42(1989).
- 22) 吉村賢治, 日高 達, 吉田 将: 文節数最小法を用いたべた書き日本語文の形態素解析, 情報処理学会論文誌, Vol. 29, No. 1, pp. 40-46(1983).
- 23) 橋本三奈子, 桑畑和佳子, 村田賢一: 計算機用日本語基本名詞辞書の試作, 第43回情報処理学会全国大会論文集(3), pp. 213-214(1991).
- 24) 島津 明, 内藤昭三, 野村浩郷: 日本語文意味構造の分類—名詞句構造を中心に—, 情報処理学会自然言語処理研究報告, 47-4, pp. 25-32(1985).
- 25) 池原 悟, 宮崎正弘, 横尾昭男: 日英機械翻訳のための意味解析辞書, NLC 91-19, pp. 23-30(1991).
- 26) Cho, K. B.: ハイライト高校国語文法, pp. 65-72, 志学社(韓国)(1989).
- 27) 田近洵一: くわしい国文法, pp. 116-164, 文英堂(1981).

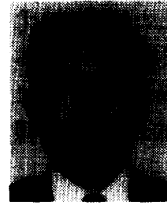
(平成4年4月1日受付)

(平成4年9月10日採録)



金 泰錫 (正会員)

1981年韓国国立慶北大学工学部電子工学科卒業, 同年来日, 日本(株)FMCでSE研修. 1983年韓国(株)国際コンピュータ Eng.でシステム開発に従事. 1984年韓国国立慶北大学助手. 1988年日本慶應義塾大学大学院理工学研究科管理工学専攻修士課程修了. 現在同大学院博士課程在学中. 自然言語処理, 自動プログラミングの研究に従事.



浦 昭二 (正会員)

1952年東京大学工学部応用数学科卒業. 現在, 慶應義塾大学理工学部教授. 理学博士. 情報システム, ソフトウェア工学, 日本語教育システムなどの研究に従事. 「情報システムハンドブック」(培風館)を編著.