派生文法に基づく日本語動詞句のウイグル語への翻訳

小川 泰弘[†] ムフタル・マフスット[†] 杉野 花津江[†]
外山 勝彦^{†,††} 稲垣 康善[†]

日本語とウイグル語は共に膠着語であり、構文的に類似した点が多い。したがって、日本語からウイグル語への機械翻訳においては、形態素解析によって得られた各単語を逐語翻訳することにより、ある程度の翻訳が可能となる。しかし、従来の日本語文法は動詞が活用することを前提としていたため、ウイグル語への翻訳の前に、動詞の活用処理が必要であった。本論文では、日本語、ウイグル語を共に派生文法で記述することにより、日本語の活用処理を不要とすると同時に、両言語間の形態論的類似点を明確にし、単純でかつ体系的な機械翻訳が可能になることを示す。しかし、日本語とウイグル語との間の文法的差異から、単純な逐語翻訳では不自然な翻訳となる場合がある。本論文では、単語間の接続関係を考慮した訳語置換表を用いることによりこの問題を解決し、より自然な翻訳を実現した。さらに、この手法に基づく日本語-ウイグル語機械翻訳システムを作成した。このシステムでは、日本語形態素解析システムとウイグル語整形システムを、それぞれ独立のモジュールとして構成している。この設計は、他の膠着語間における翻訳にも応用可能であると考えられる。また、実験によりその翻訳精度を評価した。本論文では、特に両言語において文の中心的役割を果たす動詞句の翻訳について述べる。

キーワード: 派生文法, 機械翻訳, 日本語, ウイグル語, 膠着語

Verbal Phrase Generation based on Derivational Grammar in Japanese-Uighur Machine Translation

OGAWA, YASUHIRO[†], MUHTAR, MUHSUT[†], SUGINO, KAZUE[†],
TOYAMA, KATSUHIKO^{†,††} and INAGAKI, YASUYOSHI[†]

Since both Japanese and Uighur languages are *agglutinative languages*, they have a lot of syntactic similarities. Thus we can translate from Japanese into Uighur by replacing Japanese words with corresponding Uighur words after morphological analysis of Japanese sentences. However, Japanese verbs have been said to conjugate but Uighur verbs do not, so that we have to analyze Japanese conjugation. In this paper, by using the derivational grammar which formalizes Japanese verbs syntax without conjugation, we propose a simple and systematical translation system. Since there are some differences between Japanese and Uighur, this simple translation sometimes makes output sentences unnatural. To solve this problem, we use a word replacement table which is based on word connection relations on Uighur. The system consists of three independent modules and so we can apply this approach to the translation between other agglutinative languages. We also show the performance evaluation of the system. In this paper, we will focus on the translation method of verbal clauses which play an important role in sentences.

Keywords: *Derivational grammar, Machine translation, Japanese, Uighur, Agglutinative language*

1 はじめに

日本語とウイグル語は言語学上の区分において、共に膠着語に分類され、両言語の間には語順がほぼ同じであるなどの様々な構文的類似点が見られる。そのため、日本語-ウイグル語機械翻訳では、形態素解析が終了した段階で各単語を対応するウイグル語に置き換える、いわゆる逐語翻訳によって、ある程度の翻訳が可能となる(ムフタル, 外山, 稲垣 1994)。

ところで、学校文法をはじめとする多くの日本語文法では、文の中心的役割を果たす動詞が活用することを前提としている。しかし、ウイグル語の動詞は活用しないと考えられてきたため、両言語間の翻訳の際には、活用の有無の違いを考慮する必要があった。それに対して、(ムフタル他 1994) は推移グラフの利用を提案したが、実際の処理の際には扱いにくいという問題がある。

一方、Bloch(Bloch 1975)を源流とする音韻論に基づく文法は、活用を用いることなく日本語の動詞の語形変化を表現することが可能である。本論文では、それらの中でも、動詞の語形変化を体系的に記述することに成功している派生文法(清瀬 1991)(清瀬 1989)を使用する。派生文法は、日本語の膠着語としての性質に着目した文法であり、動詞の語形変化を語幹への接尾辞の接続として表現する。さらに、ウイグル語も同じ膠着語であるので、その語形変化も派生文法で記述可能であると考えられる。原言語である日本語と目標言語であるウイグル語の双方を共に派生文法で記述することができれば、その結果、両言語間の形態論的類似性がより明確になり、単純でかつ精度の高い機械翻訳の実現が期待できる。特に、本論文で扱う動詞句の翻訳においては、複雑な活用処理をすることなく、語幹と接尾辞をそれぞれ対応する訳語に置き換えることにより、翻訳が可能になると考えられる。

そこで、本論文ではウイグル語の動詞句も派生文法に基づいて記述することにより、活用処理を行うことなく、簡潔にかつ体系的に日本語からウイグル語への動詞句の機械翻訳を実現する手法を提案する。

膠着語間の機械翻訳に関する研究としては、日本語と韓国語との間の研究(李, 中嶋, 安居院 1990)(李, 小沢 1992)(金, 大駒 1996)(金, 李, 李 1998)が多くなされている。それらでは、日本語および韓国語の動詞がともに活用することを前提に翻訳が行われているが、両言語において活用変化の仕方が異なる点が問題とされている。例えば、日本語の学校文法においては、活用形が未然形、連用形、終止形、連体形、假定形、命令形の6つに分類されるが、これは日本語独自の分類であり、韓国語の活用形の分類とは一致しない。そのため、両言語の活用形の間で対応

† 名古屋大学大学院工学研究科情報工学専攻, Department of Information Engineering, Nagoya University

†† 名古屋大学統合音響情報研究拠点, Center for Integrated Acoustic Information Research(CIAIR), Nagoya University

をとる必要があるが、日本語の連用形は文中における機能が多岐に渡るため、韓国語の活用形と1対1に対応させることは困難である。また、日本語の学校文法が用言の活用を五段活用および上下一段活用の2種類の規則活用とカ変およびサ変の不規則活用に分類しているのに対して、韓国語には種々の不規則動詞が存在し、その変化の仕方は日本語と異なる。そうした日本語と韓国語の比較については、文献(金・大駒 1996)が詳しい。そのため、これまでの日本語-韓国語機械翻訳の研究においては、日本語の語形変化の処理と韓国語の語形変化の処理を別々に行っている。それに対して本研究では、日本語およびウイグル語の動詞は共に活用しないとしているため、活用形の不一致は問題とならない。また、動詞句の形成には派生文法に基づく同一の規則を用いるため、日本語とウイグル語の語形生成を同じ規則で扱うことが可能である。

また、日本語と韓国語との間の翻訳においては、もう一つの問題として様相表現の違いが指摘されてきた。これは、様相表現を表わす接尾辞の接続順序が日本語と韓国語で異なるために生じる問題であり、この問題を解決するために、意味接続関係によって記述された翻訳テーブルを使用する方式(金・大駒 1996)や、様相情報の意味をテーブル化し、PIVOTとして用いる方式(金他 1998)などが提案されている。日本語とウイグル語では、様相表現を表す接尾辞の接続順序は同じであるため、そうした点も問題とはならない。

しかし、日本語とウイグル語には、同じ意味役割を果していても、互いに品詞の異なる単語が存在する。そのため、それらの単語の翻訳においては、単純に置き換えただけでは不自然な翻訳文が生成される。本論文では、この問題はウイグル語の語形成の性質を利用することによって解決できることを示す。具体的には、日本語形態素解析の結果を逐語翻訳した後、ウイグル語単語の接続情報を用い、不自然な並びとなる単語列を他の訳語に置き換えることによって、より自然なウイグル語文を生成する。

さらに、本研究では形態素解析システム MAJO(小川, ムフタル, 外山, 稲垣 1999)を利用して日本語-ウイグル語機械翻訳システムを作成した。MAJO は派生文法に基づいて日本語の形態素解析を行うシステムである。MAJO の辞書は、本来、日本語単語とその品詞および意味情報の3項組で構成されているが、この機械翻訳システムでは、意味情報の代わりにウイグル語訳語を与え、日本語-ウイグル語対訳辞書として利用した。その結果、MAJO の出力結果は、そのまま日本語からウイグル語への逐語翻訳となっている。さらに、この MAJO の出力結果に前述の訳語置換を適用するモジュール、および、ウイグル語特有の性質に合わせて、最終的な出力文を整形するモジュールをそれぞれ作成した。このように、機械翻訳システムを独立のモジュールから構成する設計としたが、これにより派生文法で記述された他の膠着語との間の機械翻訳システムの実現にも応用可能であると考えられる。

なお、本論文で使用する派生文法は音韻論的手法の一種であり、入力文を音素単位で解析するため、日本語の表記の一部にローマ字を用いる。また、ウイグル語の表記においても、計算機上で扱うときの簡便さから、本来のウイグル文字ではなく、そのローマ字表記を用いる。そ

図 1 日本語-ウイグル語逐語翻訳

ここで、日本語とウイグル語との混同を避けるため、以下では、日本語の単語は「」, ウイグル語の単語は“ ”で囲んで区別する。

本論文の構成は以下の通りである。まず2章では、学校文法に基づく日本語-ウイグル語逐語翻訳の例とその問題点を指摘する。3章と4章では、派生文法に基づいて日本語とウイグル語の動詞句をそれぞれ記述し、5章で派生文法に基づく日本語-ウイグル語逐語翻訳手法を示す。6章では、単純な逐語翻訳だけでは不自然な翻訳文が生成される問題を取り上げ、7章で、その問題に対する解決法である訳語置換表を提示する。また、8章で日本語-ウイグル語機械翻訳システムの実現について述べ、9章では、実験によるそのシステムの性能評価について述べる。10章は本論文のまとめである。

2 日本語-ウイグル語逐語翻訳

日本語とウイグル語は互いに良く似た言語であり、語順もほぼ同じ¹である。そのため日本語-ウイグル語機械翻訳においては、日本語入力文の形態素解析を行った段階で、各単語を対応するウイグル語の単語に置き換えれば、構文解析を行うことなく、ある程度の翻訳が可能となる。図1は、日本語の入力文「肉ヲタクサン食ベタ」に対して単純な逐語翻訳を行った例であり、出力文“Goxni jik yédi”は自然なウイグル語文となっている。

しかし、学校文法を始めとする従来の日本語文法では、動詞が活用することを前提としていたため、翻訳の際には活用処理が問題となる。例えば、日本語の動詞語幹と活用語尾を分離しそれぞれを辞書に登録して形態素解析した場合は、活用語尾の翻訳が問題となる。図2は動詞句「作ラレル」および「作ル」を逐語翻訳によって翻訳した例であるが、ここで「作ラレル」の「ラ」および「作ル」の「ル」は、それぞれ未然形、終止形を表わす活用語尾である。翻訳されたウイグル語文においては、活用語尾「ラ」に対応する単語は存在しないが、「ル」に対しては、終止形を表わすウイグル語の接尾辞“-ydu”が対応している。また、「作ラレル」の「ル」および

1 本論文では触れないが、形容詞の比較表現などにおいて日本語と語順が異なる場合が見られる。

図 2 動詞句の翻訳

表 1 日本語の活用形とウイグル語の対応

活用形	日本語	用例	ウイグル語	用例
語幹	作		yasa	
未然形	作ラ	作ラナイ	yasa	yasamaydu
	作口	作口ウ	yasa	yasay
連用形	作り	作りナガラ	yasa	yasagaq
	作ッ	作ッタ	yasa	yasagan
終止形	作ル	作ル.	yasaydu	yasaydu.
連体形	作ル	作ル人	yasaydiġan	yasaydiġan adém
仮定形	作レ	作レバ	yasa	yasasa
命令形	作レ	作レ	yasagin	yasagin

「作ル」の「ル」は、同じ活用語尾であるが、前者は“-idu”に、後者は“-ydu”にそれぞれ翻訳されている。このように、日本語の動詞は活用すると考え、動詞語幹と活用語尾を分離した場合、活用語尾の訳語を決定するために、きめ細かな処理が必要である。

一方、日本語の動詞の活用形ごとに、対応するウイグル語の訳語を登録する手法も考えられる。これは、ウイグル語も活用していると考えた手法と言える。表 1 は日本語の動詞「作ル」と、それに対応するウイグル語動詞“yasaydu”に関して、活用形ごとに対応を示したものである。

(ムフタル他 1994)では、そのような考え方に基づいた推移グラフを導入することにより、活用形を処理していた。このグラフでは、各辺に日本語の助動詞が、また、各節点にウイグル語の訳語が割当てられており、グラフの辺を開始節点から順にたどることによって、複雑な語形変化を含む動詞や助動詞の接続を処理する。しかし、この推移グラフは開始節点が動詞の活用形ごとに異なるため、1つの動詞に対して活用形の数だけ開始節点が必要である。また、1つの助動詞に対して複数の辺が対応しているため、実際の処理の際には扱いにくいという問題がある。

本論文では、日本語動詞の活用を前提としない派生文法を利用することにより、そうした活用処理を必要としない、簡潔かつ体系的な日本語-ウイグル語動詞句翻訳手法を提案する。

表 2 動詞と接尾辞の接続例

活用形	子音幹の例	母音幹の例	接尾辞
未然形	kak-ana-i	tabe-na-i	-(a)na-i
	kak-are-ru	tabe-rare-ru	-(r)are-(r)u
	kak-ase-ru	tabe-sase-ru	-(s)ase-(r)u
	kak-ou	tabe-you	-(y)ou
連用形	kak-imas-u	tabe-mas-u	-(i)mas-(r)u
	ka ϕ -ita	tabe-ta	-(i)ta
終止形	kak-u	tabe-ru	-(r)u
連体形	kak-u	tabe-ru	-(r)u
仮定形	kak-eba	tabe-reba	-(r)eba
命令形	kak-e	tabe-ro	-e / -ro,-yo
	kak-una	tabe-run-a	-(r)una

3 派生文法による日本語動詞句の記述

従来の日本語文法は、用言の活用を前提にしており、学校文法では、その活用形は未然形、連用形、終止形、連体形、仮定形、命令形の 6 つに分類されている。

一方、日本語は言語学上の分類において**膠着語**であるとされている。膠着語とは、文法的機能を表す接辞が、実質的観念を表す語幹に結合することによって単語を形成するという性質をもつ言語の総称である。ゆえに、膠着語である日本語が活用することは不合理であると捉えられ、活用を前提としない**派生文法** (清瀬 1991)(清瀬 1989) が提案されている。本章では、派生文法による日本語の記述方法を、その特徴が現れる動詞句の形成についてまとめる。

3.1 連結子音と連結母音

動詞の不変化部分を語幹と呼ぶ。学校文法における一段活用動詞「見ル」「食ベル」などの場合は、不変化の部分「見」「食べ」がそれぞれ語幹であり、その際、語幹は母音 i か e のいずれかで終わる。また、五段活用動詞の「書ク」の場合、学校文法では「書カ」「書キ」「書ク」「書ケ」「書コ」のように末尾が変化するとされるが、これは音韻論的に考えれば「kak-a」「kak-i」「kak-u」「kak-e」「kak-o」であり、「kak」を語幹として取り出すことができ、語幹は子音で終わる。そこで、派生文法では、一段活用動詞のように母音で終わる語幹を**母音幹**と呼び、五段活用動詞のように子音で終わる語幹を**子音幹**と呼ぶ。

派生文法では、動詞の変形は動詞の語幹に接尾辞が接続したものとして考える。そのため、学校文法でいう活用形の語尾や助詞、助動詞をいずれも接尾辞として扱う。それらを学校文法における活用の形に対応させると表 2 のようになる。なお、表 2 における記号 ϕ は、音便により対応する子音が消失したことを表している。

ここで、終止形「kak-u(書ク)」「tabe-ru(食ベル)」の場合、接尾辞はそれぞれ「-u」「-ru」

である。派生文法ではこれをまとめて「-(r)u」と表記する。子音 r の有無は動詞語幹の末尾に依存して決まる。例えば「kak-」に「-(r)u」が接続した場合、子音が連続することになるので、接尾辞の先頭の r が欠落する。そのような子音を**連結子音**と呼ぶ。

一方、否定の助動詞「ナイ」が接続する場合を考えてみる。これは未然形に接続する接尾辞である。派生文法では「書カナイ」「食べナイ」を表 2 に示すように、それぞれ「kak-ana-i」「tabe-na-i」と解析する。ここで、否定を表す接尾辞は「-(a)na-」の形で表され、母音 a は母音が連続する場合に欠落する。そのような母音を**連結母音**と呼ぶ。

以上より、派生文法では、語幹と接尾辞の接続を以下の 2 つの規則で記述できる。

接続規則 1: 子音幹に連結子音を持つ接尾辞が接続する場合、連結子音を削除する。

接続規則 2: 母音幹に連結母音を持つ接尾辞が接続する場合、連結母音を削除する。

3.2 統語接尾辞と派生接尾辞

3.1 節で否定の接尾辞を「-(a)nai」ではなく「-(a)na-i」と表記した。これは「kak-ana-katta(書カナカッタ)」のように、「-(a)na-」の後にさらに他の接尾辞が接続することが可能だからである。これは、動詞語幹に接尾辞「-(a)na-」が接続することにより、新たな語幹が派生したと見ることができる。そのような語幹を**二次語幹**と呼び、二次語幹を派生する接尾辞を**派生接尾辞**と呼ぶ。日本語の派生接尾辞には他に「-(s)ase-」「-(r)are-」「-(r)e-」「-(i)mas-」「-(i)ta-」があり、それぞれ使役、受身・可能・尊敬、可能、丁寧、希望の意義を表す。二次語幹に対する接尾辞の接続に関しても接続規則 1 および 2 は適用される。例えば表 2 では、「kak-(i)mas-(r)u」における連結子音 r が削除されている。

一方、派生接尾辞に対して、「-(r)u」のように新たな語幹を派生しない接尾辞を**統語接尾辞**と呼ぶ。統語接尾辞は動詞形を形成する役割を果たす。ここで、動詞形とは終止形、連体形、連用形、命令形の四形のことである。動詞に複数の接尾辞が接続する場合には、統語接尾辞が最後に接続する。

4 派生文法によるウイグル語動詞句の記述

3 章で述べた動詞句形成の特徴は日本語だけでなく、多くの膠着語にも現れる現象であり、ウイグル語にも同様の特徴がある。我々は、ウイグル語の記述に派生文法を用いることにより、その共通点を明確にした。

例えば、日本語の動詞「書 k-」に相当するウイグル語の動詞は“yaz-”である。使役の意味を表す場合、日本語では派生接尾辞「-(s)ase-」が接続して「書 kase-」となる。同様にウイグル語では、“-ğuz-”という派生接尾辞が接続して“yazğuz-”となる。両言語間の派生接尾辞の対応を表 3 に示す。ここで、敬語表現の違いから、丁寧を表す日本語の接尾辞「-(i)mas-」に相当す

表 3 日本語とウイグル語の派生接尾辞の対応

役割	日本語	ウイグル語	日本語例	ウイグル語例
使役	-(s)ase-	-ğuz-	kak-ase-	yaz-ğuz-
受身	-(r)are-	-(i)l-	kak-are-	yaz-il-
可能	-(r)e-	-(y)ala-	kak-e-	yaz-ala-
丁寧	-(i)mas-	-	kak-imas-	yaz-
否定	-(a)na-	-ma-	kak-ana-	yaz-ma-
希望	-(i)ta	-ğū-	kak-ita-	yaz-ğū-

表 4 日本語とウイグル語の統語接尾辞の対応

動詞形	役割	日本語	ウイグル語
終止形	非完了	-(r)u	-[i]du
	完了	-(i)ta	-di
	前望	-(y)ou	-(a)y
	否定前望	-(u)mai	-maydu
連体形	非完了	-(r)u	-[i]diğan
	完了	-(i)ta	-ğan
連用形	順接	-(i)	-(i)p
	完了	-(i)te	-(i)p
	仮定条件	-(r)eba	-sa
	開放条件	-(r)uto	-sa
	却下条件	-(i)teha	-sa
	否定	-(a)zu	-mastin
	同時	-(i)nagara	-ğaq
	目的	-(i)ni	-gili
命令形	肯定命令	-e, -ro	-gin
	否定命令	-(r)una	-mağın

るウイグル語の接尾辞は存在しない。また、「-(r)are-」は受身・可能・尊敬の意味があるが、ここでは受身の意味に限定している。

また、両言語とも最後に統語接尾辞が接続することによって動詞句が形成される。上述の例では、日本語の「-(i)ta」に相当する“-di”が“yazğuz-”に接続することで、動詞句“yazğuzdi”が形成される。両言語間の統語接尾辞の対応を表4に示す。

また、日本語と同様にウイグル語にも連結母音、連結子音が存在する。例えば、ウイグル語で受身を表す派生接尾辞は“-(i)l-”であり、括弧内のiが連結母音である。よって、日本語の「作r-」に相当する動詞“yasa-”に“-(i)l-”が接続する場合、語幹末尾が母音であることから、iが欠落して“yasal-”となる。ところが、ウイグル語には連結母音、連結子音とは異なり、欠落する代わりに変化する音素も存在する。非完了の連体形を表す統語接尾辞“-[i]diğan”は、“yaz-”のような子音幹動詞に接続する場合には、[i]がそのまま表記され“yazidiğan”となるが、“yasa-”のような母音幹動詞に接続する場合には、iがyに変化して“yasaydiğan”となる。そのような

図 3 派生文法を用いた動詞句の翻訳

音素を連結半母音と呼び, [i] と表記する. このことから, ウイグル語には次の動詞接続規則もあることが判る.

接続規則 3: 連結半母音 [i] は, 子音幹に接続する場合は i に, 母音幹に接続する場合は y にそれぞれ変化する.

5 派生文法を用いた逐語翻訳

日本語-ウイグル語翻訳において, 派生文法を用いた逐語翻訳を行った例を動詞句「作 rareru」および「作 ru」について示すと, 図 3 のようになる. 図 2 と比較した場合, 単語が日本語とウイグル語の間で 1 対 1 に対応していることが判る. 図 2 の例では, 日本語の動詞語幹と活用語尾を分離していたが, 「作ラレル」における「ラ」に対応するウイグル語の訳語が存在しなかった. しかし, 派生文法では「-(r)are-」を 1 つの接尾辞とみなすことにより, ウイグル語の「-(i)l-」と対応させることが可能となる. また, 接続規則 1~3 により, ウイグル語の翻訳文も簡単に生成できる. 図 2 の例では「-(r)u」に対応するウイグル語の翻訳語が下線部のように「-idu」と「-ydu」の 2 種類あるが, 派生文法では連結半母音を利用して「-[i]du」とまとめて表記できる.

一方, 表 1 に示すように, 日本語の動詞の活用形ごとに対応するウイグル語の訳語を登録する手法もある. この手法はすでに示したように, 1 つの動詞に対し活用形の数だけ訳語を辞書に登録する必要がある. また, それを避けるため, 辞書に語幹だけを登録し, 活用形に応じて活用語尾に相当する接尾辞を付加する方法も考えられる. しかし, その場合, 例えば終止形を形成する語尾が「-idu」になるのか「-ydu」になるのかを決定する処理が必要となる. そうした処理は, 派生文法を用いてウイグル語を形成する場合に必要な処理と同じであり, 派生文法を用いた手法と比較した場合, 日本語における活用処理が必要となる点で劣っている.

なお, 日本語は語順の自由度が高いといわれるが, 動詞語幹に接続する接尾辞の順序には明らかに制約がある. 例えば, 使役を表す派生接尾辞「-(s)ase-」と受身を表す接尾辞「-(r)are-」の二つの接尾辞が動詞語幹に接続する場合, 必ず「-(s)are」「-(r)are-」の順序で接続

図 4 逐語翻訳によるウイグル語-日本語翻訳

する. すなわち, 「書 k-」に接続する場合は「書 k-ase-rare-」となり, 接尾辞の順序が入れ替わって「書 k-are-sase-」となることはない.

そうした接尾辞の接続の順序も日本語とウイグル語で同じであると考えられる. このため, 複数の接尾辞を含む複雑な動詞句も, 図 4 に示すように, 日本語入力文の形態素解析が終わった段階で, 各単語を対応するウイグル語に置き換えれば, 翻訳は基本的に可能となる.

6 逐語翻訳における問題点

これまでに示したように, 日本語とウイグル語の間の構文的小よび形態論的類似性は高いが, 異なる部分もあり, 日本語単語とウイグル語訳語を 1 対 1 に対応付けできない場合がある. その場合, 単純な逐語翻訳では不自然な翻訳となる. 本章では, そのような問題点を例を挙げて説明する.

6.1 終止形と連体形の区別

表 4 において, 日本語では終止形と連体形に同じ形の接尾辞がある. 例えば, 完了を表す統語接尾辞は終止形でも連体形でも共に「-(i)ta」である. しかし, ウイグル語では同じ役割を果たす接尾辞が終止形と連体形ではそれぞれ別の単語になる. 例えば以下の例を考える.

終止形: 彼 ga 書 ita. U yazdi.
 連体形: 彼 ga 書 ita 本. U yazğan kitap.

ここで, 日本語では共に「-(i)ta」で表されている部分が, ウイグル語では, 終止形の場合は“-di”に, 連体形の場合は“-ğan”になっている. したがって, 「-(i)ta」の翻訳においては, その動詞形に応じて, “-di”と“-ğan”のいずれが訳語として適切であるかの選択が必要となる.

6.2 派生語幹の不一致

否定の派生接尾辞は日本語では「-(a)na-」であり, ウイグル語では“-ma-”である. 「書 kanakatta」を単純に逐語翻訳すると, 以下のようになる.

書 k-	-ana-	-katta
↓	↓	↓
動詞	派生接尾辞	形容詞接続 統語接尾辞
↓	↓	↓
yaz-	-ma-	-kén

ここで, “-kén” は形容詞の語幹に接続して完了の意味を表すウイグル語であり, 日本語の「-katta」に相当する. しかし, 実際のウイグル語では, “-ma-” の後に接続する単語は “-kén” ではなく, 日本語の動詞接尾辞「-(i)ta」に相当する “-di” で, 「書 kanakatta」の翻訳としては “yazmadi” が自然である. これは日本語の「-(a)na-」が形容詞の語幹を派生するのに対して, ウイグル語の “-ma-” は動詞語幹を派生するからである.

希望を表す派生接尾辞についても, 同様の問題がある. 日本語で希望を表す派生接尾辞「-(i)ta-」は形容詞語幹を派生する接尾辞であるが, これに対してウイグル語で希望を表す動詞接尾辞は “-ğum” である. 例えば, 「私ハ書キタカッタ」の自然なウイグル語訳は “Mén yazğum bar idi.” であるが, これを日本語に直訳すると「私ハ書キタイコトガアッタ」となる. すなわち, “-ğum” は「～シタイコト」の意味で動詞を名詞化する接尾辞であり, その後に「アッタ」に相当する “bar idi” が接続して, “yazğum bar idi” と表現される. このことから, “-ğum” は動詞語幹に接続し, 名詞語幹を派生する接尾辞と考えられる.

このように派生接尾辞「-(a)na-」と “-ma-”, および「-(i)ta-」と “-ğum” ではそれぞれ派生する語幹が異なるため, 単純な逐語翻訳では不自然な翻訳となる.

6.3 サ変動詞の対応

日本語のサ変動詞とは, その単語の基本となる形が名詞であるが, 接尾辞「スル」が接続することによって動詞化する単語のことである. 例えば「開発」, 「登録」がその例であり, 日本語には数多く存在する.

ウイグル語にも, 「スル」に相当する単語として “qilmaq” がある. 例えば, 「開発」に相当するウイグル語名詞は “kéxip” であるが, これが動詞化して「開発スル」となる場合, ウイグル語では “kéxip qilmaq” となる. よって, 「スル」の訳語として “qilmaq” を対応させれば, 逐語翻訳による翻訳が可能となると考えられる.

しかし, この手法では不自然な翻訳となる例が存在する. 例えば, 「登録」に相当するウイグル語は “tizimlax” であるが, 「登録スル」に相当するウイグル語は “tizimlax qilmaq” ではな

表 5 訳語置換表

日本語	基本訳語	前接ウイグル語	後接ウイグル語	新訳語	新品詞
-(r)u	-[i]diǵan	*	文末	-[i]du	終止接尾辞
		*	句読点	-[i]du	終止接尾辞
		*	終助辞	-[i]du	終止接尾辞
-(i)ta	-ǵan	*	文末	-di	終止接尾辞
		*	句読点	-di	終止接尾辞
		*	終助辞	-di	終止接尾辞
-katta	-kén	-ma-	*	-ǵan	連体接尾辞
		-ǵum	*	bar idi	名詞接尾辞
登録	tizimlax	*	qil-	tizimla-	サ変動詞
減少	azayix	*	qil-	azay-	サ変動詞
si-,su-,se-	qil-	サ変動詞	*	-	母音幹動詞

く, “tizimlamak” である. ここで, “tizimla-maq” の “-maq” は動詞の辞書見出し形をつくる接尾辞であり, 語幹は “tizimla-” である. つまり, “tizimlax” は, 動詞語幹 “tizimla-” に名詞化接尾辞 “-(i)x” が接続することによって形成される名詞であり, “tizimlax qilmaq” は日本語の「登録スルコトヲスル」に相当する冗長な表現となる. そのため, 「登録スル」の翻訳に際しては, 逐語翻訳 “tizimlax qilmaq” ではなく, “tizimlamak” と翻訳するのが望ましい.

7 訳語置換表の導入

本章では, 6 章で挙げた問題への解決方法を示す. それらの問題は, いずれもウイグル語の訳語を, その前後に現れる単語に応じて訳し分けることで解決できる. そこで, 逐語翻訳したウイグル語単語とその前後に現れるウイグル語単語との関係から, より自然な訳語に置き換える訳語置換法を導入する.

前後に現れる単語の情報を利用する場合, 翻訳前の原言語の単語接続関係を利用する方法と, 翻訳後の目的言語の単語接続関係を利用する方法とが考えられる. 前者は, 原言語の接続関係から訳語を選択する手法であり, (金・大駒 1996) などで用いられている. それに対して, 後者はいったん各単語を翻訳し, 後処理として不自然な訳語を適切なものに置換する手法である. 前者の訳語選択手法では, 原言語と目標言語の双方に関する知識がないと訳語選択の規則を記述することは困難であるが, 後者の訳語置換手法では, 目標言語に関する知識があれば, 置換規則が記述できる. また, 別の原言語に対しても, 元の置換規則を再利用できる可能性がある. そうした点を考慮して, 本手法では, 後者の手法を採用した.

7.1 訳語置換表

本手法では, 形態素解析が終了した段階で, いったん各日本語単語をウイグル語の基本訳語に翻訳する. その後, 文の先頭から順に各訳語とその前後の訳語を調べ, 他の訳語が適切であ

る場合はその訳語に置き換える。この置換規則を記した表を訳語置換表 (表 5) と呼ぶ。

表 5 では、一番左の欄に日本語の単語が記述してあるが、これは表の理解を助けるためのものであり、実際の翻訳システムが用いる訳語置換表では省略される。次の列のウイグル語は、その日本語に対応する基本訳語である。前接ウイグル語および後接ウイグル語の欄は、基本訳語を置換する場合の条件を示しており、それらのウイグル語が前後に現れる場合、基本訳語を新訳語で置換する。前接ウイグル語および後接ウイグル語の欄には、基本的にウイグル語の単語を記述するが、規則の記述を簡潔にするため、単語の代わりに品詞を記述することも可能とする。なお、前接ウイグル語または後接ウイグル語に依存しない置換規則の場合は、条件の欄に* (don't care) を記述しておく。新訳語の欄には、条件を満たした場合に基本訳語を置き換えるための新しい訳語を記述しておく。また、今回の手法では、訳語置換を行うかどうかを文の先頭から順に検査するため、訳語を置換した場合には、置換後の訳語の品詞が必要になる。そこで、それを新品詞の欄に記述しておく。

以下では、6 章で取り上げた各問題を訳語置換表を利用して解決する方法を示す。

7.2 終止形と連体形の区別の処理

6.1 節で取り上げた終止形と連体形の区別は、後接する単語の品詞に依存して決まる。そこで、例えば、「-(i)ta」については、それに対する基本訳語を連体形の “-ğan” とし、後接ウイグル語の品詞が「文末」, 「句読点」, 「終助辞」のいずれかである場合にのみ、終止形の “-di” に置き換えるという規則を訳語置換表に記述しておく。なお、基本訳語として “-ğan” の方を選んだ理由は、置換規則の数を少なくするためである。

7.3 派生語幹の不一致の処理

6.2 節で言及した派生語幹の不一致の問題は、「-katta」の訳語を、その前に現れる単語によって訳し分けることで解決できる。そこで、「-katta」の基本訳語を形容詞の語幹に接続する “-kén” とし、前接ウイグル語が “-ma-” である場合には “-ğan” に、前接ウイグル語が “-ğum” である場合には “bar idi” に置き換えるという規則を導入する。

ところがここで、“-ma-” に後接する接尾辞の候補として、連体形を表す “-ğan” の他に、終止形を表す “-di” も考えられる。そうした区別を考えた場合、後接ウイグル語の条件も考慮し、置換規則を書く必要が生じるが、それは前接ウイグル語と後接ウイグル語の組合せに応じて置換規則が増えることになり、そのメンテナンスが困難となる。そこで、“-kén” を “-ğan” に置換するという規則だけを記述し、終止形の場合は、7.2 節で述べた置換規則で対処する。つまり、一つの訳語を置き換えた場合、すぐに新訳語を出力するのではなく、その新訳語が更に訳語置換表の別の条件を満たしていないかを検査し、もし満たしている場合は、再び訳語を置き換えるのである。その結果、置換規則の簡略化と、規則のメンテナンスの容易さが実現できる。た

図 5 提案手法における翻訳の例

だし、置換規則の書き方によっては、際限なく規則が適用され、置き換えが終了しない場合が生じるため、置換規則の記述には注意が必要となる。

7.4 サ変動詞の対応の処理

6.3 節で取り上げたサ変動詞の対応の問題は、逐語翻訳 “tizimlax qilmaq” に訳語置換を施すことで対処できる。“tizimlax” に動詞 “qil-” が後接する場合、“tizimlax” を “tizimla-” と置き換える。ただし、その後で “qil-” を消去する必要がある。そこで、置換した “tizimla-” に、特別な品詞「サ変動詞」を与え、“qil-” の前接ウイグル語の品詞が「サ変動詞」である場合に、“qil-” を何も出力しないことを表す “-” に置き換える。この「サ変動詞」という品詞は形態素解析の結果には出現せず、訳語置換の場合だけに利用する品詞である。なお、この手法では、そうした変化をする全ての名詞に対して規則が必要となるが、これは動詞語幹に名詞化接尾辞 “-(i)x” を付加することで自動的に生成できる。

7.5 訳語置換の例

図 6 翻訳システムとその動作例

この手法による例文「登録 sinakatta(登録シナカッタ).」の翻訳過程を図 5 に示す。「登録」の基本訳語は“tizimlax”であるが、日本語のサ変動詞「si-」の訳語“ḳil-”が後接するため、訳語置換表を用いて“tizimla-”に置き換えられる。次の“ḳil-”は、前の単語が“tizimla-”に置き換えられたため、前接単語が「サ変動詞」とあるという条件を満たす。したがって、何も語を訳出しないことを意味する“-”に置き換えられる。次に、「-katta」の基本訳語“-kén”は前接のウイグル語単語が“-ma-”であるため、訳語置換表の条件に合致し“-ğan”に置き換えられる。また、この置き換えられた“-ğan”を訳語置換表を用いて再検査すると、後接のウイグル語が句読点であるという条件を満たすため、終止形を表す“-di”に置き換えられる。この結果、最終的に入力文「登録シナカッタ。」に対する自然な翻訳文“tizimlamadi.”が得られる。

8 機械翻訳システムの実現

本章では、本論文で提案した手法に基づく日本語-ウイグル語機械翻訳システムの実現について述べる。我々は、日本語-ウイグル語機械翻訳システムを日本語形態素解析システム、訳語置換システム、ウイグル語整形システムの三つのモジュールで構成した。図 6 に本翻訳システムと、その動作例を示す。

日本語形態素解析システムには、我々がこれまでに開発した **MAJO** (**M**orphological **A**nalyzer of **J**apanese based **O**n derivational grammar) (小川他 1999) を使用した。MAJO は派生文法に基づいて日本語の形態素を解析するシステムであり、辞書に各単語の情報を(日本語単語、品詞名、意味)の 3 項組の形で登録している。今回作成した機械翻訳システムでは MAJO の辞書を、(日本語単語、品詞名、ウイグル語訳語)の 3 項組で表される日本語-ウイグル語辞書に置き換えて使用した。この結果、MAJO の出力結果は、そのまま逐語翻訳の結果となる。なお、変更部

分は辞書のみであり、文法やシステム自体には何ら変更を加えずに使用した。

なお、本翻訳システムでは、入力として漢字仮名混じりの日本語文を扱うが、派生文法に基づく解析を行うため、形態素解析の前処理の段階で、入力文中の平仮名の部分を日本式ローマ字表記²に変換する。形態素解析システム MAJO の出力結果は、そのまま入力文に対する逐語翻訳となり、訳語置換システムに引き渡される。

訳語置換システムは、7 章で述べた訳語置換表を実現するモジュールである。すなわち、訳語置換システムでは、訳語置換表と逐語翻訳の結果を比較し、置換表に記入された規則に該当する語の並びが出現した場合、適切な訳語に置き換える。図 6 の例では、2 回出現している動詞接尾辞「-ta」は MAJO の出力の段階では共に“-ğan”と翻訳されているが、訳語置換システムにより、文末に出現する“-ğan”は“-di”に置換されている。

ウイグル語整形システムはウイグル語の音韻規則を処理し、最終的なウイグル語翻訳文を出力するシステムである。連結母音・連結子音の削除も音韻規則の一つであり、それらはウイグル語整形システムで処理する。例えば、図 6 の例では、訳語置換システムの出力に、連結母音を含む派生接尾辞“-(i)l-”が存在するが、子音幹に接続しているため、この連結母音(i)はそのままiとして出力される。

また、ウイグル語には連結母音・連結子音以外にも母音調和や母音の弱化・同化などの音韻規則が存在する(竹内 1991)。これは文の発音を容易にするために接尾辞や語幹が変化する現象である。例えば母音調和については、ウイグル語の母音に前母音é, ü, öと後母音a, u, oの区別があり、前母音を含む動詞には前母音を含んだ接尾辞が、後母音を含む動詞には後母音を含んだ接尾辞がそれぞれ接続する。ただし、iとeは母音調和に関しては中立である。現在では、ウイグル語整形システムの内部処理においてこれらの音韻規則を処理している。図 6 の例においては、連体形を形成する完了の統語接尾辞“-ğan”が派生接尾辞“-(i)l-”を狭んで前母音を含む動詞“ét-”に接続しているため、ウイグル語整形システムによって“-ğan”が“-gén”に置き換えられ、最終的なウイグル語文では、“étilgén”となっている。

9 翻訳実験の評価と検討

本章では、我々が作成した翻訳システムを用いた翻訳実験の結果と、その評価について述べる。環境問題を扱った新聞社説など3編の日本語文138文を本システムを用いて翻訳し、生成された282個のウイグル語動詞句について翻訳精度の評価を行った。なお、このウイグル語動詞句には同じものがいくつか含まれており、異なる動詞句の数は250種であった。

また、実験に使用した辞書であるが、十分な規模の日本語-ウイグル語辞書がそもそも存在しないため、ウイグル語-日本語辞書(飯沼 1982)を電子化し、その逆辞書として日本語-ウイ

² 日本式ローマ字表記は仮名に従った表記をする。訓令式やヘボン式では「ぢ」「づ」「を」を音に従ってそれぞれ「zi」「zu」「o」と表記するが、日本式ローマ字表記ではそれぞれ「di」「du」「wo」となる。

表 6 翻訳実験の結果

	動詞句数	正解翻訳数	正解率
単純な逐語翻訳	282	104	36.9%
置換表を利用	282	197	69.9%

表 7 翻訳誤り箇所数の内訳

誤り原因	単純な逐語翻訳の場合	置換表を利用した場合
終止形と連体形の区別	52	2
派生語幹の不一致	7	0
サ変動詞の対応	60	0
直訳不能	21	21
訳語の多義性	49	49
人称接尾辞の選択	5	5
音韻規則の適用	7	7
形態素解析	2	2
合計	203	88

グル語電子化辞書を自動的に作成して(小川, 杉野, ムフタル, 稲垣 1998), 実験に利用した。ただし, この辞書をそのまま使用した場合, 入力文には未登録語が多数出現する。今回の実験は, ウイグル語文全体の翻訳精度ではなく, 動詞句の翻訳精度を測定するためのものである。翻訳に必要な語を適宜登録し, 未登録語は無いものとして実験を行った。その結果, 実験に使用した辞書は, 約 16,000 語の単語を登録したものとなった。また, 訳語置換規則を 853 個登録して実験に使用したが, そのうちの 799 個がサ変動詞のための規則である。

実験では, 訳語置換表を用いた場合と用いなかった場合の翻訳精度の差を比較した。評価は翻訳システムの出力文に出現した動詞句のうち, まったく誤りの無いものを正解, それ以外を不正解とした。翻訳実験の正解率を表 6 に, また翻訳誤り箇所数の内訳を表 7 に示す。なお, 一つの動詞句の翻訳に失敗した場合でも, その原因が複数ある場合があり, 誤りの合計は翻訳に失敗した動詞句の数よりも多くなっている。

表 7 における誤りのうち, 終止形と連体形の区別, 派生語幹の不一致, および, サ変動詞の対応は, 6 章で述べた誤りである。終止形と連体形の区別が必要となる接尾辞は, 今回の実験では 176 箇所に出現したが, そのうち終止形であるものが 50 箇所あった。それらは, 訳語置換表を用いることにより, すべてを正しく翻訳できた。しかし, 正しくは連体形であるものを, 誤って終止形に置換してしまった箇所が 2 箇所あった。例えば, 入力文「氣ヲ配ラナ カッタ, アルイハ無神經ダツトコロ」の下線部「カッタ」は最後の「トコロ」を修飾するため連体形であるが, 読点の直前に出現するため, 訳語置換表の規則により終止形に置き換えられてしまった。そうした問題は構文的情報を利用しない本手法では正しく翻訳できない。

派生語幹の不一致 (7 箇所) やサ変動詞の対応 (60 箇所) の問題については, 訳語置換表を用

いることにより、すべてを正しく翻訳できた。

以上の結果、表 6 に示すように、訳語置換表を用いることにより、正しく翻訳できた動詞句が約 3 割増加した。このことから、訳語置換表の有用性が示せた。

以下では、翻訳失敗の原因について考察する。

まず、直訳不能に分類される誤りは、慣用句やそれに近い表現であり、単語ごとに翻訳しただけでは適切に翻訳できない。例えば、「環境ノ世紀ヲ迎エルニ当タッテ」という入力文は、ウイグル語に直訳しても元の文とは異なる意味になってしまう。正しく翻訳するためには「当タッテ」を「迎エル前ニ」といった意味的に等価な表現に変換する必要があると考えられる。

訳語の多義性とは、日本語の単語に対して複数のウイグル語の訳語が相当する場合に、正しい訳語が選択できないという誤りである。今回の実験では、一つの日本語単語に一つのウイグル語訳を対応させた辞書を用いたため、文脈によってはその訳が適切でない場合があった。例えば、日本語の「～シテイル」における補助動詞「i-」に相当するウイグル語には、動詞接尾辞 “-(i)wat-”, 補助動詞 “kél-”, “két-”, “tur-”, “oltur-”, “yür-” などがある。入力文には「生存ヲ 許サレテイル。(ruhsét kıliniwatidu)」 「宇宙全体ニ カカワッテイル(tağılip kélidu)」 「似テイル(ohxap ketidu)」 「自然界ヲ 動カシテイル 秩序(hérkétléndürip turidiğan)」といった句で補助動詞「i-」が出現したが、それぞれの下線部に対する自然なウイグル語訳は、括弧内に示すものであり、「i-」に対する訳語はそれぞれ異なっている。どの訳語を選択するかは「～シテイル」の意味に依存するため、現状では対応できていない。今回の実験では、動詞の多義性が原因で失敗した箇所は 10 箇所、接尾辞および補助動詞の多義性が原因で失敗した箇所は 39 箇所であった。

人称接尾辞の誤りは、適切な人称接尾辞を付加できなかった誤りである。ウイグル語の動詞句には、日本語に存在していない人称接尾辞が付加する場合がある。人称接尾辞は、動詞に対する動作主に依存して決まるが、現在のシステムでは動作主を特定する処理は行っていない。そのため、今回の実験では、人称接尾辞が出現した場合は、いずれも三人称単数を表す人称接尾辞を付加した。その結果、48 箇所出現した人称接尾辞のうちの 5 箇所において人称の選択を誤った。人称接尾辞を適切に補うためには、動作主の特定が必要になる。特にウイグル語文では、日本語文と同様に動作主が文中に明示されない場合が多く、その特定は今後の課題である。

また、現在の翻訳システムでは、ウイグル語整形システムの内部処理により、8 章で述べた音韻規則を処理している。今回の実験では、動詞句 282 個に対し、音韻処理を 335 箇所で行った。1 つの動詞句に対し、複数の音韻処理が適用される場合があるため、適用箇所の合計は動詞句の数よりも多くなる。なお、音韻処理を適用しても音韻変化する条件を満たしていない場合は動詞句の形は不変であるが、今回の実験では 175 箇所がそれに該当し、実質的に音韻変化した箇所は 160 箇所であった。また、音韻処理を施した 335 箇所のうち、7 箇所で間違いがあった。その原因は、母音 i と e が母音調和に関して中立であるためである。母音として i もしくは

eしか含まない動詞の場合, 後接する接尾辞が前母音を含んだものになるか, あるいは後母音を含んだものになるかは動詞ごとに異なる. そのため, 動詞に含まれる後母音の有無で接尾辞を決定している現在のウイグル語整形システムでは, 適切な接尾辞を選択できない.

10 おわりに

本論文では派生文法を用いた日本語-ウイグル語の動詞句の機械翻訳手法について述べた. 派生文法を用いてウイグル語動詞句も記述することにより, 動詞接尾辞間の対応が明確になり, 逐語翻訳でも高精度な動詞句の翻訳が可能になることを示した. また, 本論文では, 不自然な翻訳を避けるための訳語置換表を提案した. その結果, 動詞句の翻訳において, より自然な翻訳文を生成することが可能となった. また, 本手法に基づく翻訳システムを作成し, 実験によりその有効性を確かめた.

現在の翻訳システムでは, 一つの日本語の単語に対して複数のウイグル語の単語が相当する場合に, どの単語を選択するかを考慮していない. 特に日本語で受身・可能・尊敬の意味をもつ動詞接尾辞「-(r)are-」に対しては, すべて受身と仮定している. 今後は, そうした曖昧な翻訳語の選択方法について検討していく.

また, 今回の提案では訳語置換の条件として, ウイグル語の単語および品詞の情報に限定しているが, 意味情報などを扱えるように拡張することにより, 訳語の多義性の解消などにも応用が可能と考えられる.

今後は, 翻訳の対象を動詞句以外にも広げ, 本システムを使用した翻訳実験を進めると共に, 実用的な日本語-ウイグル語翻訳システムの実現を目指す.

また, 本手法が他の膠着語間の機械翻訳においても有効であることを示すため, 現在, 本手法に基づいたウイグル語-日本語翻訳システムを開発中である.

参考文献

- Bloch, B. (1975). ブロック日本語論考. 研究社. 林栄一 監訳.
飯沼英三 (1982). ウイグル語辞典. 穂高書店.
清瀬義三郎則府 (1989). 日本語文法新論-派生文法序説-. 桜楓社.
清瀬義三郎則府 (1991). 日本語学とアルタイ語学. 明治書院.
金政仁, 大駒誠一 (1996). “日韓機械翻訳における拡張翻訳テーブルを用いた韓国語の生成方法.” 情報処理学会論文誌, **37** (9), 1697-1707.
金政仁, 李鍾赫, 李根培 (1998). “日韓機械翻訳における様相テーブルに基づいた韓国語述部の生成処理.” 自然言語処理, **5** (2), 3-24.

- ムフタル・マフスット, 外山勝彦, 稲垣康善 (1994). “日本語-ウイグル語機械翻訳における助動詞のパラメータ化による処理.” 電子情報通信学会技術研究報告 NLC 94-13, 電子情報通信学会.
- 小川泰弘, 杉野花津江, ムフタルマフスット, 稲垣康善 (1998). “日本語-ウイグル語電子化辞書の自動生成.” 平成 10 年度電気関係学会東海支部連合大会講演論文集, p. 284.
- 小川泰弘, ムフタル・マフスット, 外山勝彦, 稲垣康善 (1999). “派生文法による日本語形態素解析.” 情報処理学会論文誌, **40** (3), 1080-1090.
- 李義東, 中嶋正之, 安居院猛 (1990). “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, **31** (6), 801-809.
- 李秀眩, 小沢慎治 (1992). “韓日機械翻訳のための音韻表現形式による用言の活用処理.” 情報処理学会論文誌, **33** (12), 1565-1577.
- 竹内和夫 (1991). 現代ウイグル語四週間. 大学書林.

略歴

- 小川 泰弘:** 1995 年名古屋大学工学部情報工学科卒業. 1997 年同大学院工学研究科情報工学専攻修士課程修了. 現在, 同博士課程在学中. 自然言語処理に関する研究に従事. 言語処理学会, 情報処理学会各会員.
- ムフタル マフスット:** 1983 年新疆大学数系卒業. 1996 年名古屋大学大学院工学研究科情報工学専攻博士課程満了. 同年, 三重大学助手. 現在, 名古屋大学計算理工学専攻稲垣研究室特別研究員. 自然言語処理に関する研究に従事. 人工知能学会, 情報処理学会各会員.
- 杉野 花津江:** 1961 年愛知学芸大学数学科卒業と同時に名古屋大学工学部に勤務. 現在, 名古屋大学大学院工学研究科助手. オートマトン・言語理論, 確率オートマトン, 自然言語処理に関する研究に従事. 情報処理学会, 電子情報通信学会各会員.
- 外山 勝彦:** 1984 年名古屋大学工学部電気学科卒業. 1989 年同大学院工学研究科情報工学専攻博士課程満了. 同大助手, 中京大学講師, 助教授を経て, 1997 年名古屋大学大学院工学研究科助教授. 工学博士. 論理に基づく知識表現と推論, 自然言語理解に関する研究に従事. 言語処理学会, 情報処理学会, 電子情報通信学会, 人工知能学会, 日本認知科学会各会員.
- 稲垣 康善:** 1962 年名古屋大学工学部電子工学科卒業. 1967 年同大学院博士課程修了. 同大助教授, 三重大学教授を経て, 1981 年より名古屋大学工学部・大学院工学研究科教授. 工学博士. この間, スイッチング回路理論, オートマトン・言語理論, 計算論, ソフトウェア基礎論, 並列処理論, 代数的仕様記述法, 人工知能基礎論, 自然言語処理などの研究に従事. 言語処理学会, 情

報処理学会, 電子情報通信学会, 人工知能学会, 電気学会, 日本ソフトウェア科学会, 日本 OR 学会, IEEE, ACM, EATCS 各会員.

(1999 年 12 月 7 日 受付)

(2000 年 2 月 15 日 再受付)

(2000 年 3 月 27 日 採録)