

アイヌ語から日本語への漸進的な翻訳処理に関する基礎的考察

桃内佳雄[†]・大友雄介[†]・越前谷博[‡]

Fundamental Studies of Incremental Translation from the Ainu Language into Japanese

Fundamental Studies of
from the Ainu Land

Incremental Translation Language into Japanese

Yoshio MAMOUCHI¹, Yusuke OTOMO¹ and Hiroshi ECHIZEN-YA²

アイヌ語と日本語は構文的に類似した点が多く、アイヌ語から日本語への翻訳処理において、アイヌ語文を構成する各単語の逐語翻訳の結果得られる日本語の単語列で、そのおよその意味を理解することができると思われる。しかし、アイヌ語文と日本語文で語順が一致しない場合がある、また日本語の格助詞に対応するいくつかの要素がアイヌ語にはない、などの相違点も存在する。従って、逐語翻訳によって得られた日本語の単語列が必ずしも自然な日本語文となるとは限らず、より自然な日本語文を得るために、逐語訳からさらに進んだ処理が必要である。本報告では、アイヌ語文を構成する各単語の逐語翻訳を出発点として、より自然な日本語文を得るために、さらにどのような処理が必要になるか、また、それらの処理をどのような翻訳解析モデルとして構築したらよいかということについての基礎的な考察を試みる。

1. 特長の概要

アイヌ語と日本語は構文的に類似した点が多く、アイヌ語から日本語への翻訳処理の出発点として、アイヌ語文を構成する各単語の逐語翻訳の結果得られる日本語の単語列で、そのおよ

* 北海学院大学工学部電子情報工学科 材料系

Department of Electronics and Information Engineering, Faculty of Engineering, Hokkai-Gakuen University

† 北海学園大学大学院工学研究科電子情報工学専攻

Division of Electronics and Information Engineering, Graduate School of Engineering, Hokkai-Gakuen

University

その意味を理解することができると思われる。しかし、アイヌ語と日本語の文で単語の語順が一致しない場合がある。アイヌ語には日本語における格助詞に対応するいくつかの要素がない、また、日本語にはない名詞の所属形が存在する、などの相違点も多く存在する[1]。従って、逐語翻訳によって得られる日本語の単語列が必ずしも正しい日本語文となるとは限らず、より自然で、正しい日本語文を得るために、逐語訳からさらに進んだ処理が必要である。本報告では、アイヌ語文を構成する各単語の逐語翻訳を出発点として、より自然な日本語文を得るために、そこから先、どのような処理が必要になるか、また、それらの処理をどのような翻訳解析モデルとして構築したらよいかということについての基礎的な考察を試みる。

漸進的な自然言語解析は、例えば、文の解析において、単語が入力されるごとに、その単語が入力されたところまでの部分的な構造と意味の解析を行い、その結果を段階的に積み上げて行きながら、文全体の処理を終えたときに、その文全体の構造と意味の解析を終えているという解析手法である[2]。解析過程で、曖昧な部分や情報が不完全で決められない部分などがある場合、その状態をそのまま保持して、より後からの情報を得て、その解決を図りながら解析を進めて行く。このような解析手法を文の翻訳処理にも適用することができるであろう。つまり、単語が入力されるごとに、その単語が入力されたところまでの部分的な翻訳処理を行い、その結果を段階的に積み上げて行きながら、文全体を処理し終えたときに、その文の翻訳処理を終えているという翻訳手法として構成することができる。

従来の多くの機械翻訳システムが採用している構文トランスファー方式では、まず、解析処理で原言語の文の構文（意味）構造を抽出し、その後、変換（トランスファー：transfer）、生成という処理を経て目的言語の文が出力される[3]。解析が終わって変換、変換が終わって生成という処理単位ごとの段階的な処理過程となっている。漸進的な翻訳手法では、解析から生成までの処理が融合した形で翻訳処理が進められる。ほぼ線形的に発話される音声の翻訳システムにおいては、漸進的な翻訳手法が有効であると考えられ。今後、アイヌ語の音声からの翻訳システムの構築を考えるとき、本考察の結果の自然な拡張が可能となることが期待される。

2. アイヌ語-日本語対訳テキストベースの作成

本考察における基本的な資料として、次の著作におけるアイヌ語-日本語対訳を参考にした。

- (B 1) 中川裕、中本ムツ子：エクスプレス アイヌ語、白水社、1997。
 - (B 2) 中本ムツ子、片山竜峯：アイヌの知恵 ウバシクマ[1]、片山言語文化研究所、1999。
 - (B 3) 田村すず子：アイヌ語、言語学大辞典セレクション：日本列島の言語、三省堂、1997。
- B 1とB 2については、著作中のアイヌ語-日本語対訳を参考にして、アイヌ語-日本語対訳テキストベースを作成した。それぞれのテキストベースを、EXP、UPAと名づけた。それ

らのアイヌ語-日本語対訳テキストベースの基本的な構成要素は、付加コードを付与した「アイヌ語」・「逐語訳」・「品詞列」・「日本語」の4つの組である。品詞列はアイヌ語に対応する品詞列である。例を以下に示す。

例文1 exp 0100601 : usey¹ e=ku² るこか rusuy³ ya⁴? 見る 初子を 飲み物 か お湯 か
例文2 exp 0100602 : お湯 あなた = 飲む たい 飲 欲 か あなた お湯 飲み物 か
例文3 exp 0100603 : 名詞 + 人接 = 他動詞 + 助動詞 + 終助詞
例文4 exp 0100604 : お湯を飲みたいか ? お湯 飲み物 か

付加コードによる各構造の記述は、逐語訳の左側に括弧で示す。例文1の逐語訳は以下のようだ。
exp 逐語訳テキストベースの識別記号 (exp, upa) が最初に来る。その後は各構造の記述である。
01 本文段落章、節、単元、など 文や子文などを構成する構造の範囲を示す記号。日本語では「章」、「節」、「段落」、「句」、「文」、「子文」など。
006 パターンコード (文, 節) 番号 構成要素の番号を示す。例文1は「文」、「節」、「段落」、「句」、「文」、「子文」。
01 アイヌ語 ([01]) 路の名前を示す地名をカタカナ表記する。読み方の「アハクシ」とは異なる。
02 逐語訳 ([02]) 路の名前をカタカナ表記する。読み方の「アハクシ」とは異なる。
03 品詞列 ([03]) 例文1の「usey e=ku るこか rusuy ya」を品詞別に分類する。読み方の「アハクシ」とは異なる。
04 日本語 ([04]) 例文1の「お湯を飲みたいか？」を日本語で翻訳する。読み方の「アハクシ」とは異なる。

アイヌ語には書き言葉がなく、話し言葉に対応してアルファベット表記あるいはカタカナ表記が行われているが、本報告では、アルファベット表記を対象とする。従って、B1, B2, B3に依拠して、アルファベット表記のアイヌ語文の翻訳処理について考察することになる。

アイヌ語における品詞の種類は次のように設定した [1, 8, 9]。
名詞 (「普通名詞」、「固有名詞」、「代名詞」、「位置名詞」、「形式名詞」、「数詞」等)
動詞 (「完全動詞」、「自動詞」、「他動詞」、「複他動詞」)
連体詞：副詞；接続詞；助動詞；問投詞；接続助詞；接続形詞；接続形容詞；接続副詞；接続助詞

人称接辞 (「人接」)；接辞；
 アイヌ語では形容詞という独立の品詞がなく、日本語の形容詞で表現される様態はたいていの場合自動詞によって表現される [3]。[例：poro：自動詞：大きくなる(大きい)、大きくなる]。

3. アイヌ語-日本語逐語翻訳とその基本的な問題点

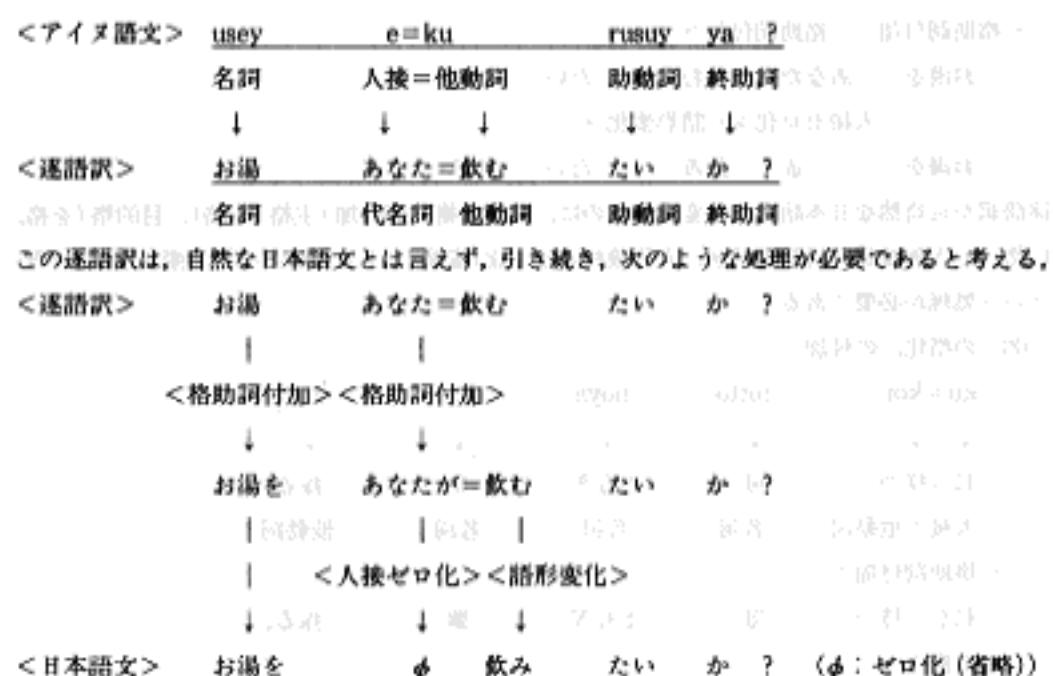
アイヌ語-日本語逐語翻訳は、基本的には、アイヌ語に対応する、対訳辞書中の日本語をその訳として割り当てる翻訳と考える。逐語翻訳の基本規則は次のようにまとめられる。

語、アイヌ語-日本語対訳辞書中の対応する日本語訳を割り当てる。アイヌ語-日本語対訳辞書の基本的な構成は次の辞書を参考にして作成する予定である。(D 1) 中川裕: アイヌ語千歳方言辞典(草風館, 1995)。(D 2) 田村すず子: アイヌ語沙流方言辞典(草風館, 1996)。(D 3) 蓼野茂: 蓼野茂のアイヌ語辞典(三省堂, 1996)。

- ・人称接辞には、人称の代名詞に対応する逐語訳をあてる。格助詞は付加しない。
- ・名詞の所属形の逐語訳には、「の」を付加する。
- ・名詞のうち位置名詞の逐語訳には、「の」を付加する。

3. 動詞、助動詞の逐語訳には、終止形をあてる。アイヌ語-日本語対訳辞書におけるアイヌ語-日本語対訳情報として、ここでは、簡単のために次のような情報を考える。この辞書情報の詳細化、特に多義の問題については後に検討する。

アイヌ語	品詞	日本語	品詞
usey	名詞	お湯	名詞
e	人称接辞	あなた	代名詞
ku	他動詞	飲む	他動詞
rusuy	助動詞	たい	助動詞
ya	終助詞	か	終助詞



逐語訳を自然な日本語文へ変換するために、名詞の格助詞付加、人称接辞の格助詞付加、人称接辞のゼロ化(省略)、そして動詞の語形変化という処理が必要であることが理解されるであろう。また、これらの処理は、主として、文法的な知識(動詞の意味)と文脈に依存して進められる。アイヌ語のさらに多くの文の逐語訳を自然な日本語文へ変換するためには、ここで述べた処理だけでは不十分である。さらに、どのような処理が必要となるであろうか。

4. 逐語訳から日本語文への基本的な変換処理

前章で考察した変換処理も含めて、本章では、逐語訳を自然な日本語文へ変換するために必要な基本的な処理について考察する。具体的な例に即して、漸進的な処理ということを念頭におきながらも、その枠組みにそれほど強く制約されない方向で検討を進める。

(1) 格助詞付加、人称接辞ゼロ化、語形変化

usey	e=ku	rusuy	ya?
↓	↓ ↓	↓	↓
お湯	あなた=飲む	たい	か?
名詞	人接=他動詞	助動詞	終助詞

このままでは、自然な日本語文とは言えず、次のような処理が必要になる。

お湯	あなた=飲む	たい	か?
----	--------	----	----

<格助詞付加><格助詞付加>

お湯を	あなたが=飲む	たい	か?
-----	---------	----	----

<人接ゼロ化><語形変化>

お湯を	あなたが=飲む	たい	か?
-----	---------	----	----

逐語訳から自然な日本語文へと変換するのに、名詞の格助詞付加(主格(が格)、目的格(を格、に格))、人称接辞の格助詞付加、人称接辞のゼロ化(省略)、そして動詞の語形変化(活用処理)という処理が必要である。

(2) の格化、の付加

ku=kor	otto	noya	ham	uk
↓ ↓	↓	↓	↓	↓
私=持つ	母	よもぎ	葉	採る。
人接=他動詞	名詞	名詞	名詞	他動詞
<格助詞付加>				
私が=持つ	母	よもぎ	葉	採る。

<の格化> もの=持つ 母の=母の=よもぎの=葉の=葉採る。<終助詞>

私の 母 よもぎ 薫葉 採る。私=衆人
<格助詞付加>

私の 母が よもぎ 薫葉 採る。

私の 母が よもぎ 薫葉 採る。<語形変化>

私の 母が よもぎの 薫葉を 採った。<格化>

この例では、ku=kor の連語訳+格助詞付加の結果である「私が=持つ」を「私の」とする処理である。また、<の付加>は、「よもぎ」と「葉」の名詞並びを「よもぎの葉」とする処理である。ku=kor がついに「私の」に変化するわけではなく、また、名詞並び「A B」がついに「A の B」に変化するわけではないところに問題がある。次のような例がある。

soy ta an eikuni ku=kor wa cise onnayke k=omare.
 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 私 外 に ある 薄葉 に 私=持つ て 家 中 で 私=入れる。
 「外」 名詞 格助詞 自動詞 名詞 人接=他動詞 接助詞 名詞 位置名詞 人接=複他動
 「ある」 外の物に存在する「薄葉」を「私」が持つて「家」の中に「私」を入れる。
 「wa (て)」が接続助詞であって、「の」に接続助詞「て」が連換することはない」という規則
 が働く。

kani isepo kina k=k=ere.
 ↓ ↓ ↓ ↓ ↓ ↓
 私 うさぎ 草を 私=食べさせる。
 代名詞 名詞 人接=複他動詞 食べる 食事 食べ
 私が うさぎに 草を 私=食べさせる。食べる
 (うさぎに 草を)

「食べさせる」の結合値パターンが [(hum) が, (anim) に, (food) を] であることが制約となる。代名詞の「私」と人称接辞の「私」の重複の処理も必要である。最後の語形変化は、「採る=>採った」のタ形への変化であるが、これは文脈(状況)に依存していて、文レベルの処理を超えているが、一つの例としてここに示した。

(3) 所属形連結

1. ku=tekehe 2. piro 3. hi 4. ta 5. hite 6. tsu
 手筋子 おひめ 姉妹 おひめ 手筋子 おひめ おひめ おひめ
 「私の手」の「手」は「手筋子」が「傷つく」「手筋子」に付いて「手筋子」の隣に

人接=名詞 (所属形) 自動詞 名詞 格助詞
 <所属形連結> <格助詞付加>
 私の手が る鮮 傷つく とき に,
 <所属形連結>は、人称接辞と名詞所属形あるいは名詞と名詞所属形を連結する処理である。人称接辞や名詞に対する格助詞「の」の付加とは考えない。名詞所属形は、所属形であることが辞書中に明記されているものとする。

(4) 人称接辞への格助詞付加、非ゼロ化
totto erumkina ~ sesekka wa ~ en=kotukka ~ akusu pirkano yenu.
 ハトトエヌトトキナセセッカワエヌコトウカアクスウピルカノユヌ。
 「母がオオバコをあぶって、やめ私=付ける」(toto=totokka) よく膚が出る。
 人接=名詞 (名詞) 他動詞 接助 人接=複他動詞 接続詞 副助詞 自動詞
 タトトカセセッカワエヌコトウカアクスウピルカノユヌ <格助詞付加>
 <非ゼロ化>
 タトトカセセッカワエヌコトウカアクスウピルカノユヌ。
 「母がオオバコをあぶって、やめ私につける」など、よく膚が出た。
 人称接辞の目的格形には、「格助詞“を”または“に”」を付加し、「ゼロ化(省略)」しないで表出する。この例で、格助詞を“に”とする処理は、動詞“付ける”の格要素の省略処理も含めて考えなければならない。“付ける”の結合箇所パターンが、(hum) が、(thing) を、(hum) に)であるとすると、“私=付ける”には二つの格要素が欠けている。

(5) 位置名詞 ("or" の逐語訳) ゼロ化

nisatta Nupurpet or un e=arpa chid ya? ~
 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 明日 登別 の所 ヘヤクタ あなた=行くか? イタ
 名詞 固有名詞 位置名詞 格助詞 人接=自動詞 終助詞 ?
 <位置名詞 ("or") ゼロ化> ある いのそ か?
 明日 登別 ヘ 行くか?
 [名詞=位置名詞 "or": の所=格助詞] という並びでは、位置名詞 "or" に対応する逐語訳をほとんどゼロ化することができるが、例外もある。また、位置名詞 "or" には、「の中」という対訳もある。
 makta huci ka sinrit or ta oka sekora =ye.
 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
 遠い昔の おばあさん も 先祖 の所 で 葬らす と 人=言う。
 連体詞 名詞 副助 名詞 位置名詞 格助 自動 副助 人接=自動詞
 この例で、“の所”をゼロ化して、“先祖で”とすると日本語としておかしい表現になる。ゼロ

化できない。つまり、日本語では、“先祖”という名詞が「場所を占める」という意味特徴を持っているないので、“の所”をゼロ化することができないということができる。

(6) 訓順寫作

ku=sinki	kusu	somo	k=arpa	わたしは疲れます。
↓	↓	↓	↓	わたしは疲れます。
わたし=疲れる	から	ない	わたし=行く	行きたい
人接=自動詞	接助	副詞	人接=自動詞	助動詞
<格助詞付加>				
わたしが=疲れる	から	ない	わたし=行く	たい。

疲れた から ない わたし=行く たい。
疲れた から ない わたしが=行く たい。
疲れた から ない 行き たい。

否定の副詞 “somo” が動詞の前におかれていて、日本語への翻訳では、語順の変化が必要となる。語形変化が “疲れる” なのか、“疲れた” なのかは文脈に依存する。アイヌ語では時制に関する動詞の語形変化あるいは語尾変化はないので、時制を過去と決定するためにはより広い意味での文脈情報が必要である。

以上の基本的な変換処理は、次のようにまとめられるであろう。

- [1] 付加処理：格助詞付加、の付加
 - [2] 削除処理：人称接辞ゼロ化、位置名詞ゼロ化
 - [3] 変形処理：語形変化、の格化
 - [4] 連結処理：所属形連結
 - [5] 並べ替え処理：語順変化

これらに、さらに多義語の処理が加わる。

 - [6] 多義語処理

- ・品詞カテゴリが異なる。日本語の複数形助詞「ます」は「終止」、複数形を表す「ます」であるが、英語では
 kor : 持つ (他動詞), て (接続助詞) 日本語の形容詞「美しい」は「形容」、物の特徴を示す「美しい」であるが、英語では
 e : 食べる (他動詞), はい (間投詞), 君 (かれ) (人称接辞) 日本語の形容詞「美しい」は「形容」、物の特徴を示す「美しい」であるが、英語では
 ku : 私 (人称接辞), 飲む (他動詞) 日本語の形容詞「美しい」は「形容」、物の特徴を示す「美しい」であるが、英語では
 - ・品詞カテゴリが同じで、多義である。
 kor : (他動詞) ①を持つ, を所有する, ②をつける／をかぶる, ③(子供)を生む
 - ・品詞カテゴリは同じであるが、下位区分カテゴリが異なる。形容 日本語の形容詞「美しい」は「形容」、物の特徴を示す「美しい」であるが、英語では
 wa : て (接続助詞), よ (終助詞), から (格助詞) 日本語の形容詞「美しい」は「形容」、物の特徴を示す「美しい」であるが、英語では

5. アイヌ語から日本語への漸進的な翻訳処理の可能性

アイヌ語の文（発話）を左から右へと読み（聞き）進みながら、日本語への漸進的な翻訳を進めてゆく構造の構成について、基本的な考察を進める。アイヌ語の話を読み、解析し、その結果である部分的な、あるいは曖昧な構造を保持しながら、左から右へと漸進的に変換を進めるというしくみで、どこまで翻訳が可能であろうか、その可能性を探る。

5. 1 減進的な翻訳処理の基本的な枠組み

例えば、前出の次のような例について考えてみよう。

＜アイヌ語文＞ usey-ku
rusuy-ya? お湯

＜逐語訳＞お湯 あなた = 飲む あなた = いいか？

名詞 依存式 人接 = 他動詞 いのやを 助動詞 終助詞 イルタマシテ

＜格助詞付加＞ <格助詞付加>

＜日本語文＞ お湯を ϕ 飲み たいですか？（ ϕ ：ゼロ化）

(a) アイヌ語—日本語翻訳辞書

【アイヌ語：口語品詞】を日本語訳す：品詞51】

[e]	人称接辞	: あなた [代名詞]	you
[ku]	他動詞	: 飲む [动] [他動詞]	drink
[rusuy]	助動詞 SB	[Imaiたい] [願う] [助動詞]	want
[va]	終助詞	: なま紫甘味 [終助詞]	sweet potato

(b) 目李體發音：由舌頭觸及齒齦或上顎，並捲曲舌頭，使發音部位與齒齦或上顎接觸。

[體見出上；品詞；連接情報；文法情報；意味情報] (3-1語段之空缺處) (本文)

「指標」(名詞) KT / K_{H} (No.)

「あなた」代名：KT／「」2人称：（hu m）】

[加工輔助：KS / 1 / 1]

[? : 論点 : KS / 既存の] 1000 項目

「 你れ！他動： 」¹²²〔五段：你主、你み、你む、你む、你ぬ、你ぬ〕

(Chum) to (Ia) &]

「 たい・動動・VV／た・（形容・たかろ・たかっ・たく・たい・たい・たけれ）」

· 38 ·

左側に連絡する情報の型を左連絡情報、右側に連絡する情報の型を右連絡情報としている。

¹⁰ 上の日本語訳書の構成では、必須的な連携情報のみを記述している。

<左連續模糊> / <右連續模糊> [啟用時] : 1

³ 1998 KT : 通用語彙叢書(3) : 了解生活用語辭典(上冊), 韓文版, 企劃編輯室。

KCS-1 适用颗粒物 小型：适用于实验室、车间、工场、仓库、办公室、家庭等场所的空气颗粒物浓度检测。

VX-1動詞未用形

· 10 ·

代名詞の人称・動詞の通用形（通用変化）などを記述する

• 300 pages

各欄については、意味統攝を記述している。

数据集：(H2O)、(液体)、(水在空气)、(Chem3D)、(Amm)

動詞については、動詞の結合類が名

能取：（burn） 焼（Ho） 烧（）

これまで、先の例について断続的な認証処理を試みてるよう、上記の二点が、本稿の核心となる。

② オオイ：最初の名詞「man」を読み、次に「人称接続「の(一)」を読み、材料検索を実行して

逐語翻訳を行う。次に、日本語接縫を用いて、意味構架を取り出し、翻訳の不完全部分構造

卷之三

usey **e** **u** **u** **u** **u** **y** **u**

お湯 あなたが お湯 格助詞人称代名詞
 [名詞, (liq)] [代名詞, (hum)] [名詞] [格助詞]
 【お湯 [名詞, (liq)] \$1 あなた [代名詞, (hum)] \$2 R】傾倒 [名詞]
 (\$1, \$2は、格助詞付加の保留を表わす) [格助詞]

② 次に、動詞「ku」を読み、逐語訳「飲む」を得る。「飲む」の結合箇バターンを参照して、前方文脈に格要素を探査し、\$1に対して「を」、\$2に対して「が」と決定する。

usey e = ku 丁寧 陽子 品川
 [名詞] ↓ 時人 [名詞] 吉澤 吉澤
 飲む [動詞, (hum) が, (liq) を] 飲む 飲む
 【お湯 [名詞, (liq)] \$1 あなた [代名詞, (hum)] \$2 I】 吉澤 吉澤
 オトム アタマ (jmu)

【お湯を ハ あなたが 飲む [他動詞] \$3】 (\$3は活用処理の保留を表す)

③ 次に、人称接辞「e」の逐語訳に対応する主格「あなたが」をゼロ化する。
 【お湯を ハ あなたが 飲む [他動詞] \$3】 倾倒 [名詞]
 オトム アタマ (jmu)

④ 「rusuy」を読み、助動詞として逐語訳「たい」を得て、【飲む\$3】と【たい】の接続処理〔左連接情報 VY により、左連接語は動詞連用形とする〕を行い、【飲み+たい\$4】と語形変化が処理されることになる。

usey e = ku rusuy 駒助人名
 [名詞] ↓ 時人 [名詞] 吉澤 吉澤
 たい [動詞]
 【動詞: VY / 名: 連接】 駒助 吉澤 吉澤
 (駒助 ← 動詞連用形) 駒助 吉澤 吉澤
 【お湯を 飲む\$3】 吉澤 吉澤 (吉澤) 吉澤 吉澤
 オトム アタマ (jmu)

【お湯を 飲み たい\$4】 吉澤 吉澤 (吉澤) 吉澤 吉澤
 ⑤ 「ya」を読み、終助詞としての逐語訳「か」を得て、「たい\$4」と終助詞「か」の接続処理〔左連接情報 KS により、左連接語は活用語終止形とする〕を行い、最後に「?」を読んで最終的な構造が構成される。

usey e = ku rusuy ya ? 駒助人名
 ↓

か おもてなしの心 30

[赞助商: KS / 11]

◎ 人情文字解題(1) 朝鮮文 金子正一著 **『活用語終止形』**

【 お湯を 飲み たいS4

【お撮りを 感みたいのか？】

得られた情報は、不完全なままに受け入れ、保持し、また、不完全な部分を、それまで保持されている情報や新しく得た情報の利用によって補充するという処理が構成された。ここで、最も基本的な情報は、動詞(用言)の結合価パターンである。基本的な処理をまとめてみると次のようになる。

- (1) 語を読んで、対訳辞書を引く。
(2) 語の品詞に応じて、辞書情報を利用しながら、構造を構成する。

名詞：格助詞の付加を保留する。 例文：おじいちゃんの本

動詞：結合形パターンに基づき格フレームを構成する。

参考文献: 機動組の付加処理について(第2回), あらすじ(機動組)

辞書動詞：結合個パターンに基づき格フレームを構成する。また日本語動詞の動詞化用法では、動詞の付加処理（接頭語「も」、「と」）、属性語の付加処理（接頭語「が」、「は」）、「人称接辞のゼロ化処理」（接頭語「を」、「に」）、「語尾変化（活用変化）」を保留する。

助動詞: 動詞への左連接・接続処理

終助詞：助動詞への左連接・接続処理

(3) 接続処理：保留されている語尾変化（活用形）の処理

5. 2 多義語の処理

前節の例について、漸進的な処理の中での多義語の処理について基本的な考察を行う。

<アイヌ語文> usey e=ku rusuy ya? [ウセイ エク ルスユヤ?]

上の例に含まれる語について、アイヌ語千歳方言辞典 [9] を引いてみると次のような品詞と日本語が対応している。

- ・usey [名] お湯 週
 例 e. 朝ごはんを食べる [動2] ～を食べる お湯を飲む 朝ごはんを食する お湯を飲む
 [間接] はいの音階の“はは”から、おはんを食べる お湯を飲む 朝ごはんを食する
 あるふの時 [人接] お前が、お前を、お前に、お前の頭や身をどうぞ、それと、あなた様
 おのまへども [接頭辞] ～の頭、～の先端 鮎魚の頭 “手前” “手前みどり” “手前” “手前
 おとする [接頭辞] ～について、～で以って、～<場所>で ～する おとする おとする
 おとする [接尾辞] ～させる オトセルを頭や身で頭にあわせ、おとする 破壊的

·ku [動2] ~を飲む

[名] 弓 ([概念形]: 【所属形: kuhu】 [9] では概念形、所属形の区別をしていいない。この区別は [8] による。本考察では区別して考える。)

[人接] 私が、私の

·rusuy [助動] ~したい

·ya [名] 網

る跡跡の海 [位名] 陸、浜辺、岸 [3] も、[4] お湯 [5] 入れ置けるを意味する。語義的には「始める」と「終る」 [終助] ~か? お湯 ne ya : 「~だの」「~も」は別々の意味で取ることや複数である。単純に上の多義をすべて組み合わせると、 $1 \times 6 \times 3 \times 1 \times 3 = 54$ 通りの組み合わせを考えられる。これを左から右への漸進的な処理の過程で一通りに絞って行くことができるかどうかが問題である。

usey e = ku rusuy ya ?

暖昧さ: (1) (6) (3) (1) (3) 1 種が5-4種の複合語

"usey" は一通りで述語訳は決定する。e は 6通りの暖昧さで、この時点では、[間接] と [接尾辞] の可能性は消える。e = "まで読み進んだところで、"e" は人称接辞であると決定できる。次の "ku" の暖昧さは 3通りであるが、述語訳 "弓" に対応する "ku" は、[8] に依拠すれば、所属形ではないので、その述語訳 "弓" は棄却される。また、人称接辞 "ku" に対する "私" も棄却される。<人称接辞> = <人称接辞> という連鎖はない。

usey e = ku rusuy ya ?

お湯 あなた 飲む したい 何か お湯 あなた 網 陸

弓 網

私 陸

綱の陸を く

従って、"ku" を読み、"rusuy" まで読み進んだところまでは、一通りの訳となる。

usey e = ku rusuy ya ?

お湯 あなた 飲み したい 何か お湯 あなた 網 陸

綱

弓 お湯 あなた

陸

弓 お湯 あなた

最後の "ya" に対する述語訳の暖昧さの解決は、"お湯をあなたが飲みたい" という節が連体修飾節として機能するかどうかにかかっている。動詞 "飲む" の結合価パターンの必須格は充足されている。このような状況で連体修飾節として機能する場合、次に来る名詞には制約がある。例えば、"こと" とか "気持ち"、"様子" などは連接可能であるが、"網"、"陸" などが次に来る可能性はきわめて低いと考えられる。従って、ここでは終助詞 "か" が優先されるとする。この問題は、隣り合った語の間の局所的な連接関係だけの手がかりでは解決することのできない

い問題である。連体修飾節というより大きな言語単位の構成と被連体語句との意味的関係についての解析が必要である。

6. おわりに

アイヌ語から日本語への漸進的な翻訳処理の可能性について基礎的な考察を行った。考察のもとになったアイヌ語のデータは書かれたアイヌ語の初步的な、そして部分的なデータである。データの範囲と量をさらに広げてゆく必要がある。特に、日本語にはない語順の変換パターンや連語の処理についての考察が一つの大きな課題である。アイヌ語のより多くの様々な言語現象の基本的な処理のしくみを明らかにしながら、それらを漸進的な翻訳処理の枠組みの中にどのように埋め込んでゆくかについても引き続き考察を進めて行きたい。また、アイヌ語-日本語対訳辞書と日本語辞書の構成も進めなければならない。両方とも既存の辞書を参考にして、また、構築するテキストベースを基礎にして、その電子化を進めて行きたいと考えている。

謝 辞

本研究の一環は、北海学園大学ハイテク・リサーチ・センター研究費による援助を受けて行われました。ここに記して謝意を表します。また、アイヌ語の文法についてご教示をいただいている本電子情報工学科 切替英雄先生に感謝いたします。

参 考 文 献

- [1] 田村すず子：アイヌ語、「言語学大辞典セレクション：日本列島の言語」、三省堂、1997.
- [2] 奥村学：漸進的な自然言語解釈モデルについて、bit, Vol. 26, No. 8, 共立出版、1994.
- [3] 佐藤理史：機械翻訳、「長尾真編、自然言語処理、第12章」、岩波書店、1996.
- [4] 有緒敏雄・荻野孝志：結合語から見た日本文法、「文法と意味Ⅰ：2章」、朝倉書店、1983.
- [5] 中川裕、中本ムツ子：「エクスプレス アイヌ語」、白水社、1997.
- [6] 中本ムツ子、片山竜三：「アイヌの知恵 ウバシクマ【1】」、片山言語文化研究所、1999.
- [7] 萩野茂：萩野茂のアイヌ語辞典、三省堂、1996.
- [8] 田村すず子：アイヌ語沙流方言辞典、草風館、1996.
- [9] 中川裕：アイヌ語千歳方言辞典、草風館、1995.