

209 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

桃内佳雄 大友雄介 越前谷博

北学園大工

1 はじめに

アイヌ語^[1]と日本語は構文的に類似した点が多く、アイヌ語文から日本語文への機械翻訳においては、各単語の逐語翻訳処理によっても、ある程度の意味的に理解可能な日本語訳を得ることが可能になると思われる。しかし、構文的、さらには、語彙的、意味的な相違点も存在し、より自然な日本語訳を得るためには、さらに進んだ処理が必要となるであろう。本報告では、逐語翻訳処理を出発点として、自然な日本語訳を得るためには、さらに、どのような処理が必要になるだろうかという問題について基礎的な考察を行う。

漸進的な自然言語解析^[2]は、例えば、文の解析において考えると、単語が入力されると、単語そのものの解析も含めて、その単語が入力されたところまでの部分的な解析を行い、その結果を文脈として蓄積、利用しながら段階的に解析を進めるという手法である。解析過程途中で、曖昧な部分や、情報不足で構造などを決定できない部分があると、それを保留し、処理の進行とともに新しい情報を得ながら、その解決を図りつつ解析を進める。本報告では、逐語翻訳処理の自然な発展形として、漸進的な解析手法を適用した翻訳処理を位置付け、その適用の可能性についても検討する。

2 考察の基礎としての資料

考察を進めるにあたっての基礎的な資料として、次の著作におけるアイヌ語-日本語対訳を参考にした。

(B1) 中川裕, 中本ムツ子: エクスプレス・アイヌ語, 白水社, 1997.

(B2) 中本ムツ子, 片山龍峯: アイヌの知恵 ウバシクマ [1], 片山言語文化研究所, 1999.

(B3) 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.

B1とB2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、それぞれをEXP, UPAと名付け、それらへの参照を基礎として考察を進めた。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。アイヌ語は、話し言葉であり、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、本考察では、アルファベット表記を採用することとした。従って、

本報告では、B1, B2に依拠して、アルファベット表記のアイヌ語の翻訳について考察することとなる。

exp0100601: usey e=ku rusuy ya?

exp0100602: お湯 あなた=飲む たい か?

exp0100603: 名詞 人接=他動 助動詞 終助?

exp0100604: お湯を飲みたいか?

3 アイヌ語-日本語逐語翻訳とその問題点

アイヌ語-日本語逐語翻訳はアイヌ語に対応する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。逐語翻訳に関する基本的な規則は次のようである。

- ・辞書中の対応する日本語を割り当てる。
- (辞書のアイヌ語見出しの単位となる語は、ほぼ日本語の語に対応しているものとする。)
- ・人称接辞は、人称のみに対応する逐語訳をあて、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。
- ・動詞の逐語訳には、終止形をあてる。

辞書として、次の3冊を参考にしている。

・中川裕: アイヌ語千歳方言辞典, 草風館, 1995.

・田村すず子: アイヌ語沙流方言辞典, 草風館, 1996.

・萱野茂: 萱野茂のアイヌ語辞典, 三省堂, 1996.

逐語翻訳の具体的な例と基本的な問題点について、2章で示した例を用いて考えてみよう。

usey	e=ku	rusuy	ya?
↓	↓	↓	↓
お湯	あなた=飲む	たい	か?
この逐語訳は、自然な日本語とは言えず、引き続き、次のような処理が必要であると考えられる。			
お湯	あなた=飲む	たい	か?
<格助付加>	<格助付加>		
↓	↓		
お湯を	あなたが=飲む	たい	か?
	<人接ゼロ>	<語形変化>	
↓	↓	↓	
お湯を	φ	飲み	たい か?
(φ: ゼロ化(省略))			

逐語訳を自然な日本語へ変換するために、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となるという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、動詞の意味(語彙的知識)と文脈に依存し

て進められる。アイヌ語のより多くの文の逐語訳を自然な日本語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 逐語訳からの基本的な変換処理

3章で考察した変換処理も含めて、逐語訳を自然な日本語へ変換するための基本的な処理として、EXPとUPAを参照した具体例の検討を通して、現在のところ次のような処理を抽出している。

- (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

これらの変換処理に加えて、逐語訳そのものの生成に伴う処理として、多義語の処理が重要な課題である。

(6) 多義語処理

アイヌ語にも、多義語がたくさん存在する。多義語による曖昧さの解消は文脈に依存して行われることになるであろう。多義語のいくつかの例を次に示す。

・品詞カテゴリーが異なる。

kor: 持つ (他動詞), て (接続助詞)

e: 食べる (他動詞), はい (間投詞), …

・品詞カテゴリーが同じで、多義である。

kor: (他動詞) ①持つ, 所有する, ②つける, かぶる, ③(子供を)生む, …

・品詞カテゴリーは同類であるが、下位区分カテゴリーが異なる。

wa: て (接助), よ (終助), から (格助)

5 漸進的な翻訳処理の可能性

漸進的な翻訳処理は、1章での漸進的な自然言語解析の考え方を翻訳処理に適用するものである。アイヌ語の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蓄積、利用しながら、翻訳処理を進めるというモデルとして構成される。基本的な処理の流れを、3章での例でみてみよう。まず、日本語の名詞と動詞の基本的な辞書情報を次のように設定する。

・名詞の基本的な辞書情報

[お湯: 名詞; [liq]], [あなた: 名詞; {hum}]

・動詞の基本的な辞書情報

[飲む: 他動詞; [N {hum} が+N {liq} を+V]]

< 結合係パターン >

動詞[ku]の逐語訳「飲む」の辞書情報として、上のような結合係パターン⁽³⁾を格納し、これを利用して、格助詞の付加処理、節構造の構成を行う。

・まず、最左の名詞 usey を読み、次に、人称接辞 e を読み、逐語訳を行い、日本語訳を得る。

```
usey          e
  ↓           ↓
お湯 $1      あなた $2
[liq]        [hum]
```

\$1, \$2 は、格助詞付加の保留を表わす。

・次に、動詞 ku を読み、逐語訳を行い、結合係パターンを参照して、前方文脈に格要素を探索し、\$1 に対して「を」、\$2 に対して「が」と決定する。

```
usey    e    =    ku
  ↓      ↓      ↓
お湯 $1 あなた $2 飲む
[liq] [hum] [N{hum}が+N{liq}を+V]
[お湯を あなたが 飲む $3]
```

ここまで、左から右への漸進的な処理によって、名詞の逐語訳の結果を保持しつつ、動詞の逐語訳が得られたところで、結合係パターンへの格要素の割り当てと格助詞の付加を行い、次のような構造が構成された。

[お湯を あなたが 飲む \$3]

「飲む \$3」の \$3 は、まだ動詞「飲む」の語形 (語尾) 変化が決定していないことを表している。

・次に、人称接辞「e」の逐語訳に対応する「が格」[あなたが] をゼロ化する。

[お湯を 飲む \$3]

・次に、rusuy を読み、助動詞として逐語訳「たい」を得て、[飲む \$3] と [たい] の接続処理を行い、[飲み+たい \$4] と語形変化が処理されることになる。

ここまでで、次のような構造が構成される。

[お湯を 飲み たい \$4]

・次に、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] の接続処理を行い、最後に、「？」を読んで、最終的な構造が構成される。

[お湯を 飲み たい か ?]

○4章でまとめた基本的な処理の各々について、漸進的な翻訳処理過程の中に埋め込むことの可能性について検討を進めている。また、具体的な処理アルゴリズム/プログラムとしての実現も今後の課題である。

謝辞

本研究の一部は、北海学園大学ハイテク・リサーチ・センターの研究費による援助を受けて行われました。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学工学部・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析モデルについて, bit, Vol.26, No.8, 共立出版, 1994.
- [3] 石綿敏雄・荻野孝野: 結合係から見た日本文法, 「文法と意味」, 2章, 朝倉書店, 1983.