

ミャンマー語の述語部の文法分析と日本語-ミャンマー語機械翻訳システム jaw/Myanmar について

福本真哉 ma ngin khaw cing 今井啓允 池田尚志

岐阜大学工学部

1 はじめに

我々はパターン変換型機械翻訳エンジン jaw の開発を行っており、現在日本語からアジアの諸言語への機械翻訳システムの構築を試みている。本報告では、日本語からミャンマー語への翻訳システム jaw/Myanmar について述べる。ミャンマー語は日本語と同じく膠着語であり、基本的な語順も「主語 + 目的語 + 述語」で日本語と同じである。また日本語と同様に述語部は述語と助詞・助動詞の類の機能語からなるが、その構成は複雑であり日本語との対応も単純ではない。本報告では、ミャンマー語の述語部の文法分析と日本語からの翻訳規則、その jaw/Myanmar 上での実現を中心に述べる。

2 jaw/Myanmar の概要

2.1 システムの概要

jaw では翻訳を命題部分の翻訳と機能語部分の翻訳の 2 段階に分けて行う。命題部分の翻訳規則は、日本語の係り受け木構造のパターンとそれに対応するミャンマー語の表現構造の対である。表現構造として VC++ のオブジェクトを利用している。用言後接機能語部分の翻訳は機能語部翻訳テーブルを介して行う (4 節)。システムの概要を図 1 に示す。

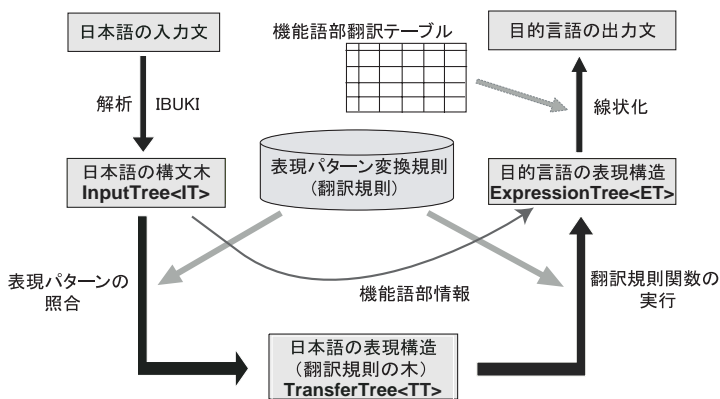


図 1: jaw/Myanmar の概要

2.2 命題部分の翻訳

入力された日本語文は、係り受け構造を表した構文木である InputTree(IT) に解析される。IT はパターン変換辞書

中の日本語パターンと照合され、対応する翻訳規則の木である TransferTree(TT) に変換される。翻訳規則は目的言語の表現構造を作り出すプログラム (dll) として格納されている。TT のそれぞれに対応する翻訳規則関数を根ノードから再帰的に実行すると、オブジェクト指向の枠組みを利用した目的言語の表現構造の木である ExpressionTree(ET) が生成される。最後に、ET の各オブジェクトのメソッドである線状化関数を根ノードから再帰的に実行することで、目的言語の翻訳文が生成される。

ミャンマー語の動詞系の表現構造のために動詞系クラスに持たせたメンバを表 1 に示す。

表 1: 動詞系クラス

クラス	メンバ	説明
CProposition	m_centerW	訳語の中心 (動詞部)
	m_centerW2	訳語の中心 (名詞部)
	m_subject	主格
	m_object	目的格 1
	m_object2	目的格 2
	m_adverb	副詞
	m_time	時間格
	m_pConnection	接続詞
	m_nounModifier	名詞系修飾
	m_adjective	形容詞

m_centerW2 は 3 節で述べる複合動詞の場合の名詞部への訳語のためのメンバである。動詞に対する修飾要素としてはこれらのメンバの他に time_start(時間の開始点)、time_end(時間の終了点)、place_start(場所の出発点)、place_end(場所の到達点)、method(手段) などがある。これらの修飾要素は CProposition の線状化関数によって図 2 に示される語順に並べられる。なおこの最後には文の完成を示す完成助辞を加える必要があるが、これについては次節で述べる。

CProposition =

```
topic+subject+accompany+time+time_start+time_end+
place+place_start+place_end+method+object+object2+
adverb+centerW2+centerW
```

図 2: ミャンマー語の基本語順

これらによって、助動詞類の付属しない簡単な例文 100 文のミャンマー語への翻訳実験を行った。その結果の一部を図 3 に示す。

① 仙台は日本の北部にある
 စန်ဒိုင် သည် ဂျပန်နိုင်ငံ ၏ မြောက်ပိုင်း တွင် ရှိ သည်
 仙台 は 日本 の 北部 に ある (完成助辞)

② 水は水素と酸素からなる
 ဓာတ် သည် အောက်စီဂျင် နှင့် ဟိုက်ဒရိုဂျင် မှ ဖြစ် သည်
 水 は 酸素 と 水素 から なる (完成助辞)

③ 熊が道を歩く
 ဝက်ဝံ သည် လမ်း ၏ လမ်း လျှောက် သည်
 熊 が 道 を 歩く (完成助辞)

④ 彼はとても泳ぎが上手だ
 သူ သည် ဓာတ်ပုံ အလွန် တော် သည်
 彼 は とても 泳ぎが 上手だ (完成助辞)

図 3: 翻訳実験例

3 ミャンマー語述語部の特徴

ミャンマー語の述語には多様な助辞が後接して述語部を構成し、様々な意味を表す。以下にミャンマー語述語部の主な特徴を示す。

○動詞

「行く」のような動態動詞、「良い」のような状態動詞の両方を動詞とみなす。日本語のような活用や、ヨーロッパ言語のような語尾屈折はない。動詞には一単語で意味を成す単一語と、「名詞+動詞」や「動詞+動詞」の組み合わせで一つの意味を成す複合動詞の二つの種類がある。

例

単一動詞：食べる、行く、等

複合動詞 (名詞+動詞)：

歩く (道+歩く)、うれしい (腹+心地よい)、等

複合動詞 (動詞+動詞)：

ついてくる (つき従う+来る)、等

法や時制は動詞に後接する助辞によって表される。

○完成助辞

文末に用いられ、文章を完成させるための働きを持ち、同時に時制や否定等の助辞と呼応してそれぞれの意味を表す。

- ・書き言葉と話し言葉では違いが大きい。
- ・時制を表現する完成助辞と表現しない完成助辞に分類できる。

表 2 に完成助辞の例を示す。

表 2: 完成助辞

完成助辞	意味	説明
သည်	非未来	肯定文の非未来で使用
မည်	未来	主格が自分の際の未来形で使用
လိမ့်မည်	未来	主格が自分以外の未来形で使用
လား	疑問	非未来形の疑問形で用いられる
မလား	疑問	未来形の疑問形で用いられる
ဘူး	否定	非未来形の否定形で用いられる。
စေ	命令	裁判所などが判決を下す際などに用いられる命令形
ပါစေ	許可	身分の高いものから低いものへの許可を与える場合
ပါရစေ	懇願	なんとかして願いを聞いてもらおうと頼むこと
ကြောင်း	伝聞	ニュースなどで用いられる。「〜と聞かされました」「〜との事です」といった意味

○疑問文

文末に疑問の助辞^{လား}をつけると日本語で「～か?」という文になる。ただし「何」「どこ」などの疑問詞が含まれる場合は^{လား}ではなく^{လိမ့်မည်}を使う。疑問詞の位置は肯定文と同じであって、英語のように先頭に移動するということはない。

○否定文

・動詞を^မ (否定助辞)と^{ဘူး} (否定の完成助辞)で挟み込んで表す。

မ စား ဘူး (食べない)
 食べる

・否定疑問の場合は否定の完成助辞を用いた後、疑問を表す助辞を文末につける。

မ စား ဘူး လား (食べませんか?)
 食べる (疑問助辞)

・疑問詞が文中に含まれる場合、否定の完成助辞は上記の^{ဘူး}とは異なり、未来の場合は^{မည်}、非未来の場合は^{လိမ့်မည်}を用いる。

・複合動詞の否定表現

「名詞+動詞」の場合は動詞の部分のみを否定する。

မိုးရွာ : 雨+降る (雨が降る)

မိုး မ ရွာ ဘူး (雨が降らない)

「動詞+動詞」の複合動詞を否定する場合は両方の動詞を否定する。

ထိုးတက် : 飛ぶ+上がる (飛び上がる)

မ ထိုး မ တက် ဘူး (飛び上がらない)

○使役

日本語の使役の表現に対してミャンマー語の表現は「^{ခိုင်း}」、「^{စေ}」、「^{ခွင့်ပေး}」、「非使役動詞を用いる」の 4 通りある。し

かし、現時点ではどのような場合にこれらを使い分けるかが整理できていないので、今回の実験では使役表現として最も使うことが多いと考えられる^{၆၆}を用いた。

また、日本語の「～てもらう」に対してミャンマー語で直接的に対応する表現はなく、「～させる」という使役の表現に翻訳される。

○受身

日本語の受身に対応するミャンマー語の表現には以下の3つの形がある。

1. ミャンマー語の受身の表現となる。
2. 同種の動詞で能動文で表現する。

例：家が建てられた 家が建った

အိမ် ဆောက် ခဲ့ သည် : 家が建った
家 建つ た

3. 全く違う動詞で能動文で表現する。

例：雨に降られる 雨に捕まる

မိုး မိ ခဲ့ သည် : 雨に捕まった
雨 捕まる た

ミャンマー語の受身表現は以下の形をとる。

動詞 + 名詞化接尾辞 + ^{၆၆} (目的格助詞)
+ 受身動詞 + 完成助辞

例えば「殴られる」は「殴りを受ける」というように表現される。受身動詞には次の二つがある。

自分の意思で自ら望んで「～される」場合 : ^{၆၆}
自分の意思と関係なく「～される」場合 : ^{၆၆}၇

受身に加えて過去や進行を表現する場合、これらの表現は^{၆၆}と^၇の二つの動詞の間に置かれる。

- 自ら進んで裁きを受けた(裁かれる)場合の表現
အပြစ်ပေးခြင်းကို ခံ သည်
裁く(V) (名詞化) を 受ける(受身) (完成助辞)
- 自分は望んでおらず殴られる場合の表現
ထိုခံရခြင်းကို ခံရ (ခံ+၇) သည်
殴る(V) (名詞化) を 受ける(受身) (完成助辞)
- 彼は叱られていた
သူ သည် ဆရာ ခြင်းကို ခံ နေ ခဲ့ ၇ သည်
彼 は 叱る(V) 名詞化 を 受身 ている た もらう (完成助辞)

○「～かもしれない」

他の助辞の多くが動詞の後ろに置かれるのと異なって次のように表現される。

動詞 + ^{၆၆} (「かもしれない」助辞)
+ 動詞 + လိမ့်မည် (未来系の完成助辞)

動詞は同じ語が繰り返して使われ、使役や過去や進行などの助辞がある場合は両方の動詞に付属する。また完成助辞は未来のものが用いられるが、主格が自分である場合とそうでない場合には区別がありそれぞれ「မည်」, 「လိမ့်မည်」が用いられる。

否定文「～ないかもしれない」あるいは「～したくないかもしれない」の場合は「かもしれない」助辞として「ချင်မှ」あるいは「မှ」が使われる。

- 彼女は行ったかもしれない
သူမ သည် သွား ခဲ့ လျှင် သွား ခဲ့ လိမ့်မည်
彼女は 行く た かもしれない 行く た 完成助辞
- 彼女は行かなかったかもしれない
သူမ သည် သွား ခဲ့ ချင်မှ သွား ခဲ့ လိမ့်မည်
彼女は 行く た かもしれない 行く た 完成助辞 (否定)
- 彼女は行きたくないかもしれない
သူမ သည် သွား ချင် မှ သွား ချင် မှ လိမ့်မည်
彼女は 行く たい た かもしれない 行く たい た 完成助辞 (否定)

4 用言後接機能語部の翻訳

4.1 述語部の助辞の語順

前節で述べたようなミャンマー語の多様な述語部の表現を生成するために、助辞の語の単位、種類を分類し、以下のようにその語順を整理した。

動詞(名詞部分) + 否定 + 動詞 + 受身(名詞化) + 受身
1 + 使役 + 許可 + 可能 + 判断1 + 希望 + 進行形 + 過去 + 受身2 + 判断5 + 判断6 + 判断2 + 判断3 + 丁寧 + 否定2 + 未来 + 疑問 + 判断4 + 問投助辞

複合動詞(名詞 + 動詞)の否定表現に対応するため、動詞は名詞部と動詞部に分けている

4.2 用言後接機能語部分の翻訳規則

日本語の用言後接機能語をひとまとまりとして考えると、その数は膨大でありその一つ一つに対して翻訳規則を書くことはできない。jaw では用言後接機能語の情報を文節構造解析により「使役・受身等」、「時制等」、「判断等」、「疑問等」の4つの要素に分割しており、要素ごとに機能語の翻訳規則を定め、各要素の翻訳規則を組み合わせることで機能語部の翻訳を実現している。

表 3: 用言後接機能語の翻訳規則テーブル

	受身 (名詞化)	希望	可能	受身	受身 2	使役	進行形	過去形
させる						ခိုင်း		
される	ခြင်း ကို			ခံ	ခံရ			
られる			နိုင်					
ていた							နေ	ခဲ့
てもらいたい		ချင်				ပေးစေ		
させられる	ခြင်း ကို			ခံ	ခံရ	ခိုင်း		

	否定	判断 5	判断 6	判断 3	未来	否定 2	間投助辞
なければならない	မ	လိုမ				ဘူး	
かもしれない			လျင်		လိမ့်မည်		
ね							နော်
らしい				သည့်ပုံရ			
てはならない	မ	ခံရ				ဘူး	
な	မ				နဲ့		

翻訳規則は機能語部の翻訳に必要な語句（一般には複数の語句の集合になる）をテーブルに記述している（表 3）。表 3 の属性は 4.1 で示した要素の列挙であり、生成関数によって 4.1 で示した語順に並べられる。

このような規則によって、3 節で述べたミャンマー語の多様な述語部表現への翻訳ができる見通しを得た。図 4 に用言後接機能語を含んだ日本語文の jaw/Myanmar による翻訳結果の例を示す。

⑤彼は私に長い手紙をよこした
သူ သည် ကျွန်ုပ် သို့ ရည်သော စ ကို ပေး ခဲ့ သည်
彼 は 私 に 長い 手紙 を よこす た 完成助辞

⑥熊は道を歩いていない
ဝက်ဝံ သည် လမ်း မှ လမ်း မ လျှောက် နေ ဘူး
熊 は 道 を 歩く ない 歩く ている 否定完成助辞

⑦彼は菓子を食べていたらしい
သူ သည် မှန် ကို စား နေ ခဲ့ သည့်ပုံရ သည်
彼 は 菓子 を 食べ ている た らしい 完成助辞

⑧あなたはご飯を食べなければならない
သင့် သည် ညးဖ ကို မ စား လိုမ ဘူး
あなた は ご飯 を 否定 食べる なければならない 否定完成助辞

図 4: 翻訳実験例

5 おわりに

日本語 - ミャンマー語機械翻訳システム jaw/Myanmar について述べた。ミャンマー語については会話用の書籍はあるが、書き言葉の文法規則については文献もほとんどなく、

ミャンマーからの留学生との議論を通じて徐々に作り上げている状況である。基本的な文章に始まり、使役等、時制等、判断等の機能語部を含んだ日本語例文の翻訳が出来た。今後はさらに多くの日本語例文の翻訳を試み、多彩なミャンマー語の言語構造の整理とともに日本語との対応、翻訳規則の開発、翻訳システムの改良をすすめていく予定である。

参考文献

- [1] 『日本語からアジア諸言語への機械翻訳の試み』
今井 啓允、他、情報処理学会 第 65 回全国大会、2003.3
- [2] 『パターン変換型機械翻訳システムに関する研究』
今井 啓允、岐阜大学工学部応用情報学科卒業論文、2002.3
- [3] 『パターン変換型機械翻訳システムにおける日本語パターン照合処理に関する研究』
河原 功幸、岐阜大学工学部応用情報学科卒業論文、2001.2
- [4] 『パターン変換型機械翻訳システムに関する研究-変換辞書に関する研究-』
大津 慎次郎、岐阜大学工学部応用情報学科卒業論文、2002.3
- [5] 『エクスプレスビルマ語』
加藤 昌彦著、白水社、1998.10