

## 日中機械翻訳システム jaw/Chinese における取り立て詞「も」の翻訳処理

陳 劭毓 方 丹 松本 忠博  
 岐阜大学 大学院 工学研究科

### 1 はじめに

取り立て詞は、その意味上と構文上の多様さのために、更には、中国語との対応関係の複雑さのために、日中機械翻訳において、曖昧さを引き起こしやすい。現在の日中市販翻訳ソフトでは、取り立て表現に起因する誤訳が多く見られる[1]。

本稿では開発中の日中機械翻訳システム jaw/Chinese[2]における取り立て詞「も」の翻訳処理について述べる。

### 2 取り立て詞「も」の中国語との対応関係

孤立語とされる中国語には、意味上から見ると、

取り立て詞と同じ意味を持つ表現が大抵存在するが、直接的に対応する品詞は無く、副詞などを用いた表現に翻訳することになる。日本語側では一つの取り立て詞が複数の意味を持つ場合があるが、中国語では、意味が違えば異なる訳語で表す場合もある。また、日本語側で同じ意味を異なる取り立て詞で表現できる場合があるが、中国語ではそれらが一つの訳語に対応する場合もある。本研究では、語学テキスト等から抽出した取り立て詞「も」を含む多数の例文を分析し、中国語との対応を表1のように分類した。

表1 取り立て詞「も」と中国語との対応

「も」の役割	日本語パターン	中国語のパターン
事物の類比	名詞+も	也…
	名詞+に/で/まで/+も	
	時間副詞+も	
	形容詞「く」+も+ない	
	形容動詞「で」+も+ない	
同類の事物の並列	名詞+も+名詞+も	和…都…
	時間副詞+も+時間副詞+も	
	形容詞/形容動詞+も+ない	
	名詞+に/で/まで/+も+名詞+に/で/まで/+も	
	名詞+で+も+名詞+で+も+ない	既不…也不…
数量詞への修飾	普通数量詞+も (肯定文)	竟然…
	普通数量詞+も (否定文)	连都…
	疑問数量詞+も (肯定文)	很多…
	疑問数量詞+も (否定文)	几…

	普通数量詞+も+動詞+ば	只要…
疑問詞への修飾	疑問詞+ (格助詞) +も (肯定文)	都…
	疑問詞+ (格助詞) +も (否定文)	
事物の類推	名詞+ (格助詞) +も (否定文)	连…都…
最大の限度	形容詞+も	最…

### 3 jaw/Chinese と「も」の翻訳処理

jaw/Chinese はパターン変換型の日中機械翻訳システムである。翻訳の流れを図 1 に示す。「入力文の木構造(Input Tree)の作成」「翻訳規則の木構造(Transfer Tree)への変換」「目的言語の表現構造木(Expression Tree, ET)の生成」「ETの各オブジェクトの線状化」の4つの処理で目的言語への翻訳を行う。

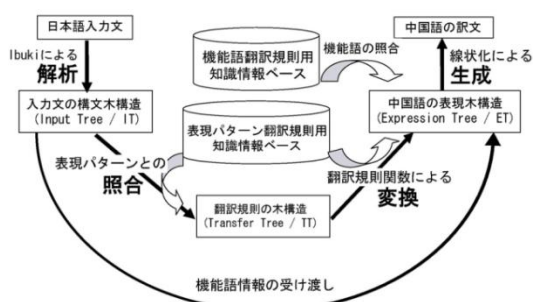


図 1 jaw/Chinese の翻訳の流れ

#### 3.1 数量詞+「も」の翻訳

次の文に対する翻訳は以下のように行われる。

例 子供はご飯を 1 回も食べていない。

##### Step1

入力文を解析し、下記の情報を得る。

- ・取り立て詞 (T) = も

- ・取り立てられる部分 (X) = 1 回(数量詞)
- ・述語 (P) = 食べていない (否定式)
- ・中国語の訳語 : (X') = 一次 (p') = 不吃
- ・日本語文の構造「主語+目的語+X も P」

##### Step2

X=数量詞 P=否定式なので、表 1 より「も」が“连…都”に翻訳される。

##### Step3

訳語“连”と“都”を jaw/Chinese システムの語順処理規則と照合し、目的語「ご飯」の前に“连”，数量詞「1 回」の後ろに“都”を置く。jaw の線状化関数により、文全体の語順を調整し、訳文“孩子连饭一次都不吃”を得る。

#### 3.2 「でも」の翻訳

日本語の「でも」には以下のような場合があり、それぞれ翻訳も変わる。

- 1) 場所/道具格助詞「で」+取り立て詞「も」
- 2) 取り立て詞「でも」
- 3) 逆接の接続助詞「でも」

1) の場合は前述の通りである。ただ、格助詞「で」とその前の名詞にの処理によって、得られた訳語と一緒に語順を調整し、訳文を得る。

2) の場合の翻訳手順を図 2 に示す。

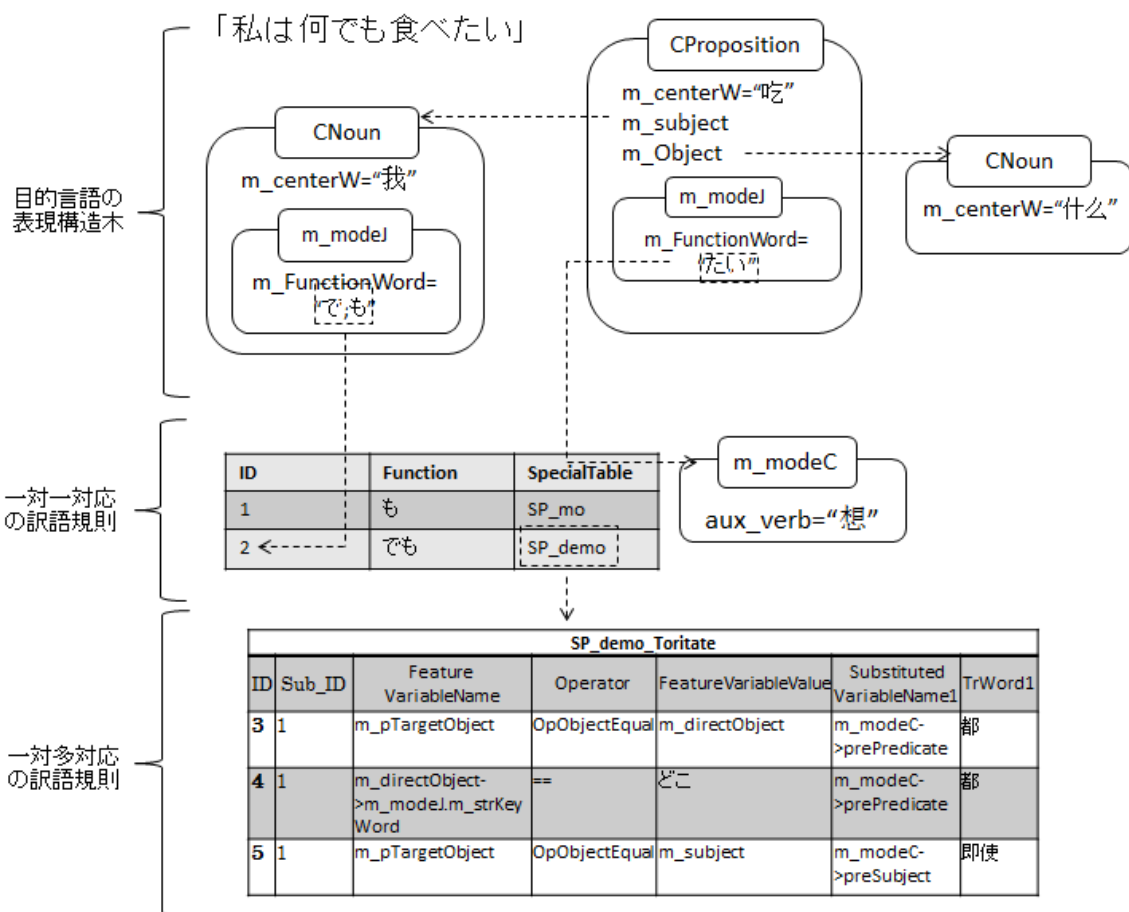


図2 取り立て詞「でも」についての翻訳手順

#### 4 翻訳実験と評価

登録済みの「も」の翻訳手順について、現代日本語書き言葉均衡コーパス

(BCCWJ) から抽出した「も」を含む 200 例文を対象に翻訳実験を行い、その結果を評価した。現在システムに登録されている翻訳規則では、「も」以外の表現が原因で翻訳結果が得られない場合があるため、翻訳実験は手作業で行った。

評価は、「も」を含む部分だけを対象とし、訳語とその語順も考慮した。実験結果を表 3 に示す。ウェブ上の翻訳サービス<sup>1</sup>

と比較しても、我々の方法は十分な有効性が期待できると考える。

表 3 翻訳実験の結果

取り立て詞	正訳数	
	A	B
も	93	77
でも	87	63

(A : jaw/Chinese B : livedoor 翻訳)

その翻訳結果の例を表 4 に示す。

<sup>1</sup> livedoor 翻訳 (高電社)  
<http://livedoor-translate.naver.jp/>

表 4 翻訳結果の例

ID	例文	翻訳結果 (上段:jaw 下段:市販ソフト)	評価
1	最近はデジカメで何枚も撮影できた	最近用数码相机拍了 <b>很多</b> 张	○
		最近用数码相机 <b>多少</b> 张也摄影完成了	×
2	良くも悪くも頭が固い人が多いようだ	脑袋顽固的人 <b>好的和坏的都多</b>	○
		<b>竟敢</b> 坏也头硬的人好像多	×
3	3回も同じ文章を入力した	<b>竟然</b> 3次都输入了同样的文章	○
		输入了一样的文章3次	×
4	家族が誰も起きてこない	家族谁都 <b>没</b> 起床	○
		家族谁都 <b>不</b> 起来	○
5	家でも学校でも戸は開けっ放し	家和学校门都 <b>敞</b> 开着	○
		<b>在家也</b> 在学校也门大敞大开	×

### 問題点の分析

1 取り立て詞「でも」は、中国語で「即使」に訳されるが、話者の語気により、「什么的」および「Φ」などに訳せるが、その条件の設定は難しい。

(1)日本の学校でも退学は悪いイメージがあるよね。

(2)とりあえず、学校でも行こう。

文1の訳語は“即使”が適当であるが、文2の訳語は“还是”または“Φ”が適切である。

2 取り立て詞が二つ以上ある場合は、正しく翻訳されない場合がある。

(1)夏はショートオールをパジャマにしてもいいと思う。

(2)切手にしても金券にしても、それが出品者に届くまでに時間がかかるね。

文1の訳語は“即使...也” jaw で正しく翻訳されたが、二つ以上ある場合は、文2が「切手と金券とも」と同じように“也好...也好”に翻訳されるべきである。

### 5 おわりに

本稿では、取り立て表現「も」の中国語との対応を分析し、翻訳規則を jaw/Chinese 上に実装した。コーパスから抽出した例文を用いて、市販ソフトと比較しながら、翻訳実験を行った。「も」に関する現状の問題点を解決し、取り立て表現を扱う辞書を拡充するため、さらに多くの調査を元に翻訳規則を整備して、評価を行う必要がある。

### 参考文献

- [1] 卜朝暉, 日中機械翻訳に関する研究—取り立て表現、否定表現を中心に—岐阜大学博士論文, 2004年4月.
- [2] 池田尚志, 日本語からアジア諸言語への機械翻訳システムの構築奮闘記, 日本語学, 28(12), 62-71, 2009.