

日中機械翻訳システム jaw/Chinese における助詞「で」の翻訳処理

張 津一 楊 鵬 松本 忠博

岐阜大学 大学院 工学研究科

1. はじめに

「で」は格助詞、接続助詞、形容動詞や助動詞「だ」の連用形の一部、接続詞と種類が多く、さらに格助詞「で」は場所、手段、材料、期限、原因・理由など多様な意味を持つ。そのため、日中機械翻訳において誤訳を引き起こしやすい。横山ら[2]は機械処理のために助詞

「で」を分析しているが、とくに日中翻訳を対象としてはおらず、実装には至っていない。

本研究では助詞「で」の中国語との対応について分析し、対応の曖昧性を解消するための翻訳規則を考案して、開発中の日中機械翻訳シ

テム jaw/Chinese[1]上に実装した。また、その評価のため、語学学習書や日中辞書などから抽出した例文を用いて、Web上の翻訳サービスと比較しながら、翻訳実験を行った。

2. 「で」の中国語との対応関係

助詞「で」に対応する中国語の表現は「で」に前接する名詞と文の述語によって変化する。

本研究では、語学学習書等から抽出した助詞

「で」を含む多数の例文を分析し、文献[2]の分類を参考にして、中国語との対応を表1のように分類した。

表1 助詞「で」の中国語との対応関係

格助詞 (名詞+で +動詞)	「で」の役割		例文と訳語	日本語パターン	中国語のパターン
	因果		雨で遅刻した/台風で休みました 「因为下雨而迟到了/因为台风而休息了」	事柄を表わす名詞+で……	因为……/由于……
	主体		委員会で決定する 「由委员会决定」	名詞で表される組織・人が動作主+で…… /自分で	由……/由……
	場所		公園で遊ぶ 「在公园玩」	地点・建物+で……	在……
	時間系	期間	三日間で行う 「三天之内进行」	長さを持つ時間帯+で……	……之内……
		時点	三日付で閉店する 「三日闭店」	瞬間的な時刻+で……	NULL
	人数		四人で行く 「四人一起去」	～人+で……、みんなで、グループでなど	……一起……/NULL
	金額		千円で買う 「花了一千元买」	～円+で……、無料で、無償で、有料で、有償で、いくらでなど	花了…… /NULL/NULL/NULL・ /NULL
	材料 要素		木で作る 「用树做」	変形や加工を受ける道具+で……	用……

	形容系	状態	はだかで行く 「裸着去」	状態を表わす名詞+で……	……着
		形容動詞	本気で怒る 「当真生气」	形容動詞+で……	NULL
	目的・方針		この方針で行う 「用这个方针进行」	ある目的や意志を持った名詞+で……	用……
	環境・条件		常温で実験する 「在常温试验」	科学論文・測定など表現+で……	在……
	単位		ミリ単位で計算する 「以厘米单位计算」	単位系+で……	以……
	判断		独断で実行する 「通过独断实行」	自分または他人の判断+で……	通过……
	割り算・微積分		3で割る 「用三除」	数学などの場合・数値+で……	用……
	放送媒体		テレビで放映する 「以电视机放映」	マスコミ的な名詞+で……	以……
	交通手段		バスで学校に行きます 「乘公共汽车去学校」	交通手段を表す名詞+で……	乘……
	範囲		世界で一番長い川 「世界上最长河流」	事柄が成立する範囲+で…… 世界で/片道で/往復で	在……/世界上…… /是……/是……/是……

3. jaw/Chinese と「で」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。「入力日本語文の木構造(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。

3.1 機能語翻訳テーブルに基づく翻訳

「で」の翻訳規則は主に 2 種類の機能語翻訳テーブルに記述する。一つは日中両言語で対応が 1 対 1 になる場合に用いるテーブル (Table1)。もう一つは対応が 1 対多になる場合に用いるテーブル (Table2) である。訳し分

けの条件は表現構造中の要素の値を使用して訳語に対し条件を設定するものである。

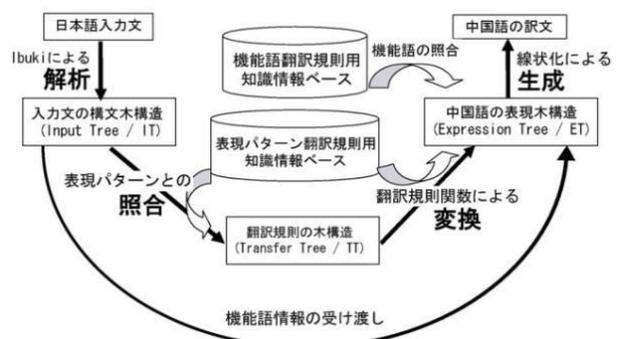


図 1 jaw/Chinese の翻訳の流れ

機能語	専用テーブル名	訳語 1	...
が	Table1	吗	...
で	Table2
...

ID	条件変数名	関係	条件値	訳語	...
01	category	==	主体	由	...
02	category	==	場所	在	...
...

図 2 機能語テーブルの例

名詞に助詞「で」が付加されているとき、「で」の機能語翻訳テーブルを利用して、名詞の意味属性 (NTT コミュニケーション科学研究所日本語語彙大系を参考にした) によって、「で」訳語を決定する (表 2)。

表 2 名詞の条件値と「で」の訳語対 (一部分)

名詞の条件値	「で」の訳語
地名/ 場所/ 場	在
因果	因为
主体	由
時間(期間/時点)	之内/NULL
金額	花了
単位	通过
状態	着
...	...

機能語翻訳テーブルで、「で」はほぼ翻訳できるが、例外もある。

例 彼の名前は故郷で有名だ。

他的名字在家乡有名。

この場合、「で」の後の用言と体言を考えなければ正しく翻訳できない。そのため、字面パターンの規則を利用して翻訳を行う。

3.2 字面パターン規則に基づく翻訳

Jaw/Chinese のパターン変換規則では、キーワードに係る文節の自立語の意味属性のほか、単語の字面を指定することもできる。この字面条件を使った日本語表現を字面パターンと呼ぶ。日本語表現パターンは Base パターンと Addition パターンの二つに分類される。表 3 に「で」の翻訳規則の一部を示す。

表 3 「で」の翻訳規則(一部分)

ID	日本語パターン	「で」の訳語
1	「材料」+で+「名詞」+作る	用
2	「災難」+で+休み	因为
3	「乗り物」+で+行く	乘
4	「場所」+で+有名	在
...

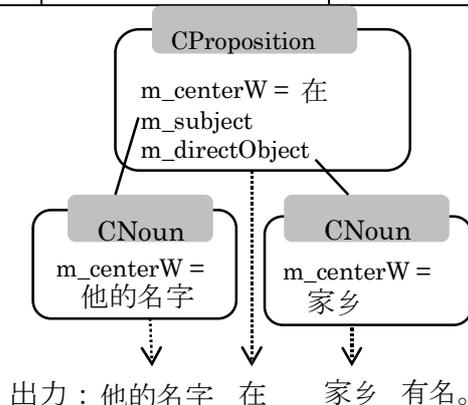


図 3 「彼の名前は故郷で有名だ」の翻訳結果

Jaw/Chinese ではその翻訳規則によって、目的言語の表現構造木(ET)を生成する。次に、線状化関数を利用して、ET の各オブジェクトを目的語の語順に並べて、訳文を生成する(図 3)。

4. 翻訳実験と評価

日中辞書や東京外国語大学言語モジュール等から抽出した「で」を含む100例文を対象に翻訳規則を実装し、その翻訳結果を評価した。

評価は、「で」を含む部分だけを対象とした。訳語と語順を考慮した実験結果を表4に示す。Web上の翻訳サービス（エキサイト翻訳を使用）と比較して、多くの文を正しく翻訳することが出来た。翻訳結果の例を表5に示す。

表4 翻訳実験の結果

正訳数 (100 文中)	
jaw/Chinese	Web 翻訳サービス
90 文	55 文

表5 翻訳結果の例

I D	例文	翻訳結果 (上段:jaw 下段: Web 翻訳サービス)	評 価
1	彼らは魚を生 で食べる。	他们生吃鱼	○
		他们直接吃鱼	×
2	牧場はここか ら車で10分 です。	牧场从这里乘车是 10分钟	○
		牧场是从这里乘车 10分。	×
3	ローマではロ ーマ人のごと くせよ。	在罗马罗马人的一个 样子。	×
		在罗马好象罗马人 一样做。	×
4	1分違いで汽 車に乗り遅れ た。	1分钟(对/给)车没 赶上了之差。	×
		1分由于差异为火 车赶不上。	×
5	2と3で5で す。	用2和3是5。	×
		是以2和35。	×

問題点の分析

- （表5の3）In Rome, do as the romans do
の日本語訳である。中国語に直訳しても格言
の意味が伝わらないため意識が必要である。

ここで「入郷随俗」と翻訳すれば適切である。

- （表5の4）「で」の訳語は語順により正しく
翻訳できなかった。「で」の訳語は“NULL”
（訳語なし）が適切である。
- （表5の5）計算結果を表す「で」の訳語は
正しく翻訳できなかった。

5. おわりに

本稿では、助詞「で」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。日中辞書などから抽出した例文を用いて、Web翻訳サービスと比較しながら、翻訳実験を行った。「で」に関する現状の問題点を解決するため、さらに多くの例文を用いて意味役割と翻訳規則を整理し、評価を行うことが必要である。

参考文献

- 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.
- 横山晶一, 菅野孝文, 機械処理のための助詞「で」の分析, 言語処理学会第2回年次大会発表論文集, 393-396, 1996.