

日中機械翻訳における構文上の対応のずれに関する考察 —受動態と能動態のずれ、品詞のずれを中心に—

ト朝暉† 浅井良信‡ 王軼謳‡ 穆貴彬‡ 池田尚志‡

† 岐阜大学工学部/学振外国人特別研究員/中国広西大学外国語学部 ‡ 岐阜大学工学部

Email: bu@ikd.info.gifu-u.ac.jp

概要：翻訳を行う際、原言語と目的言語の対応に構文上でずれがある表現が存在している。このようなずれ表現の翻訳は機械翻訳の難点である。本研究は機械翻訳の角度から、日中両言語の構文上の対応にずれがある表現について分析し、その翻訳方法を検討することを目的とする。今回はその第一歩として、「ずれの表現」の中で受動態/能動態のずれ、及び品詞のずれに関わる表現を中心に考察し、日中機械翻訳のための翻訳規則をまとめた。また小規模であるが、760 受動文に対して、まとめた受動態/能動態のずれがある分類に属する文の出現率を調査し、本規則による正訳率を手作業で評価した。出現率は約 28%、翻訳精度は約 76%の結果を得た。本研究の必要性と有効性を一定の程度で確認した。

Analysis of the Syntactic Disconnection between Japanese and Chinese in MT

Zhaohui BU†, Yoshinobu ASAI‡, Yio WANG ,‡Guibin MU‡, Takashi IKEDA‡

† Faculty of Engineering, Gifu University/ the JSPS Postdoctoral Fellow/ Guangxi University, China {Email:bu@ikd.info.gifu-u.ac.jp} ‡ Faculty of Engineering, Gifu University.

Abstract: Some expressions in one language cannot be translated with equal syntactic structure into another language (we call it syntactic disconnection here). This syntactic disconnection causes a difficult problem in MT. In this paper, from the viewpoint of Japanese-Chinese MT, we analyzed the disconnection of the passive voice and the active voice, and the disconnection of the parts of speech between Japanese and Chinese, and proposed the machine translation rules for translating these two types of disconnected expressions. Our manual evaluation showed that about 28% of 760 Japanese passive voice sentences contain the syntactic disconnection in Japanese-Chinese translation (cannot be translated into Chinese passive voice correspondently), and the translation accuracy of our rules on the passive voice sentences was about 76%.

1. はじめに

原言語と目的言語で構文構造上、同型の対応をとることが困難な場合がある。本稿ではこのような直接に同型の構文ジャンルで対応できない表現を「構文上にずれがある表現」と呼ぶ。例えば、受動態を主動態に、連体修飾構造を連用修飾に、

名詞を動詞に翻訳するなど。これらのずれがある表現は人間の翻訳のみではなく、機械翻訳にとっても難題であり、特別な翻訳処理を加える必要がある部分である。人間による言語翻訳では、増訳、減訳、転(変)訳などの手法を用いて対応することがある。言語学の側では、このようなずれのある

表現の翻訳に関する論述があるが[1][2][3]、機械翻訳の角度からの論述は未だ少ない。また言語上の分析は内容が豊富であるが、機械翻訳にとって考慮しなくてもよい内容も混在しているため、機械翻訳の観点からの整理と洗練が必要になる。我々はずれがある表現の個別的な事例について開発中のシステム jaw/Chinese での個別な対応処理を述べたことがあるが[5]、このようなずれ表現の全体に関する分析には至らなかった。本研究はこれらのずれ表現について日中機械翻訳の特徴を考慮して体系的に分析を行い、その翻訳方法を検討することを目的とする。今回はその第一歩として、「ずれの表現」の中で受動態/能動態、及び品詞のずれに関わる表現を中心に考察し、日中機械翻訳のための翻訳規則をまとめた。また小規模ではあるが、まとめた受動態/能動態のずれがある文の出現率と本規則の手作業による正訳率の評価を行った。このようなずれのある表現に関する分析は日中機械翻訳の問題発見と翻訳レベルのアップに大きな意義があると考えられる。

2. 対応にずれがある表現の定義と分類

2.1 ずれがある表現の定義

本稿で議論する「構文上の対応にずれがある表現」は日中両言語では通常それぞれ対応する構文上のジャンルがあるが、場合によって通常のジャンルで対応できず、異なるジャンルの表現で翻訳するものである。例えばある条件の下で、連体修飾構造を連用修飾に、名詞を動詞に翻訳するなど。また本稿の議論の対象は機械翻訳であり、構造を変換してもしなくても翻訳できる場合は、容易な方、すなわち変換しない翻訳を採用する。また文学翻訳の「更なる自然な文、更なる美しい文に翻訳する」というような目標は求めない。

以下ではこのような方針で議論を進めていく。

2.2 ずれがある表現の大別分類

文献[1][2][3]などで述べている人間翻訳の日中

翻訳手法(増訳、転訳、減訳など)を参考にし、機械翻訳の必要も考慮して、我々は日中両言語の構文上にずれがある表現を下記のように大別する。

[A] 相(態)のずれ

(1) 川の両側にはバラの花が植えられている。

→ 河的两岸种植着玫瑰花。

(日：受動態⇔中：能動態)

[B] 品詞のずれ

(2) 大勢の人が困んでいるから、何か起こったに決まっている。

→ 因为有很多人围住，一定是发生什么事了。

(日：動詞⇔中：副詞)

[C] 成分のずれ

(3) 田中は果物栽培の知識を豊富に持っている。

→ 田中有丰富的种植水果的知识。

(日：連用修飾語⇔中：連体修飾語)

[D] 単語の数のずれ

[D-1] 単語の数の増加

(4) 頭からの血液はすでに固まっている。

→ 从头流出来的血液已经凝固了。

(動詞「流れる」の増加)

[D-2] 単語の数の減少

(5) 私がまず手本をみせるから、ちゃんと見て。

→ 我先做示范, 你们好好看着。

(「から」の訳語の減少)

3 節からはこのうちの[A][B]について考察する。

3. 能動態と受動態のずれに関する分類と分析

2.2[A]で分類した態(相)のずれとは受動態/能動態、可能態、使役態などに関するずれのことである。ここでは日中における受動態/能動態に関するずれについて分析する。

日本語の受動態は基本的に「N1はN2にVられる/れる」のように表現されるが、中国語では動詞の変形がなく、「N1被N2V」という表現が対応する。すなわち「られる/れる」は中国語では基本的に受動態のマークの介詞(前置詞)「被」で対

応できる。(口語では「譲、叫、挨、給」などを使う場合もある)。しかし、意味上で「受け身」であり、日本語では受動態であっても、中国語では能動態として翻訳する場合がある。中国語の受身形の使用は日本語ほど頻繁ではない。逆に日本語は能動態で、中国語では受動態に翻訳される場合もある。頻度的には著者の観察では前者の方が多い。また受動のマークが「被」ではなく、必ず「遭受」、「由」のような前置詞で表現しなければならない場合もある。文献 [1][2][3][4]などを参考にし、日中の能動態と受動態におけるそれを以下のように分類し、分類ごとに機械翻訳の規則としてまとめた(以下のパターンでは「られる」のみ表示するが「れる」のことも代表する)。

3.1 日本語の受動態を中国語の能動態に翻訳する場合(即ち「られる」=「Φ」)

(1) Nは/がVられる

① 自然言語処理学会は来月に開かれる。

→ 自然语言处理学会下个月召开。

分析:(a) Nは動作を発することができない抽象的な物事;(b) Vは他動詞、抽象的な動作を表す;(c) 動作の執行者は表現されない場合が多い。

翻訳規則:→NV (N=抽象, V=抽象動作の他動詞)

(2) N1には/にN2がVられている

① 彼の涙にはある種の期待が込められている。

→ 他的眼泪包含着某种期待。

② 川の両側にはバラの花が植えられている。

→ 河的两岸种植着玫瑰花。

分析:(a)意味的にはある状態で存在する、あるいはあるものに対する処置の結果などを表す;(b)N1=場所を表す語;(c) V=他動詞。また上記の2文は存在動詞を使わないが、意味上は存在を表す。存在は客観的なこと、日本語では客観的なことは受動態で表すのが好みであり、中国語ではその客観性は動詞自身に含ませる。

翻訳規則:→(在)N1VN2 (N1=場所,V=他動詞)

(3) (a) (Nは/が)PとVられる

①中国は人民元を更に値下げするだろうと伝えられている。

→ 据说,中国可能进一步贬值人民币。

(b) (Nは/が)PものとVられる

② 会談は成果が挙げられないものと見られる。

→ 看来会谈不会取得成果。

分析:Vは伝達、思考、予測などの意味を表す動詞(「言う、思う、見る、推測する」など)。この場合、日本語では主観的な言い方を避けるために動詞の受動態を使って表現することが多い。中国語ではその客観性は動詞で表現できるので日本語のように受動態を使った婉曲的な表現を用いる必要はないようである。

翻訳規則:→(据)VP (V=伝達、思考動詞)

(P=小文あるいは小文の述語部)

(4) N1に/でVられた/られているN2...

① 壁に掛けられている絵はとても美しい。

→ 挂在墙壁上的画很美。

② この雑誌に発表された作品を読んでください。

→ 请读发表在这本杂志上的作品。

分析:(a)N1=範囲や場所を表す語(b) V=他動詞

翻訳規則:→V在N1的N2,(在N1V的N2)

(N1=場所,範囲,V=他動詞)

(5) NはVられた/ているものだ

① この桜の木は去年植えられたものだ。

→ 这棵樱花树是去年种的。

分析:この場合の受動文は発生した事実や状態などを表し、存在の意味を含んでいる。

翻訳規則:→N是V的

(6) (a) N1は(Adv)Vられてある

① 家具はきちんと並べられてある。

→ 家具摆的整整齐齐。

(b)N1はN2にVられてある

② 離婚届はテーブルの上に置かれてあった。

→ 离婚申请书放在桌子上。

分析:意味上では着目されるのは動作発生後の状

態の存続であり、目で見える状態である。
N1は場所を表す語。

翻訳規則：→NV(在N2)(N2=場所 V=他動詞)

(7) 間接受動文の場合

間接受動文はある動作(V)がある有情物(N3)に間接に「被害」や「迷惑」をもたらす意味を示す。直接受動文には必ずしもそのような含意はない。構文上では、直接受動文は能動文の目的語を主語に変えて作られるため、動詞は他動詞からしかできない。それに対して、間接受動文は能動文になかった新たな名詞句(N3)(もとの文の動作のメンバーではないもの)を新たな主語として追加して作れるため、自動詞からも他動詞からも作れる[6][7]。以下の(a)と(b)は間接受動文である。(c)は構文上では間接受動文と相似しているが、意味上は直接受動に近い。間接か直接かその中間に置くか種々の論説がある。中国語との対応から見ると、(c)タイプは中国語や英語などでは受動文として翻訳でき、(a)と(b)の間接受動文は中国語と英語などには直接対応しにくい。(a)タイプの動詞の数は少なく予想され、動詞ごとにパターンを登録することで解決できるが、(b)タイプの動詞の数は推測し難く、また訳語も特定しにくく、課題である。

(a) N1が(は)N2にVIられる(自動詞間接受動)

- ① 花子がおかあさんに死なれた。
→ 花子死了妈妈。
- ② おかあさんが子供に泣かれた(困っている)。
→ 孩子哭(得)妈妈很难办。
- ③ 私は雨に降られた。
→ 我被雨淋了。(普通の受動文に訳す)

翻訳規則：(a)→N1V(了)N2 あるいは N2V, [N1…]

(N1=人/動物 V=「降る, 吹く」など以外の自動詞)

(b) N3はN1にN2をVTられる(他動詞間接受動)

- ④ 私は太郎に下手な歌を歌われた。
→ 我被迫听了太郎的糟糕的歌唱。(太郎硬要我听了他的糟糕的歌唱。)

このような間接受動文は中国語では確定的な翻訳がなく、前後の文脈によって「被迫、让、Φ」など意識的に対応するため、課題である。

(c) N3はN1にN2をVTられる(持ち主の受け身)

- ⑤ 彼はすりに財布を盗まれた。
→ 他被小偷偷了钱包(他的钱包被小偷偷了。)
- このタイプは通常受動文と同じく「被」構文で対応できるが、形式上は(b)と同様であるため、その区別は問題である。

3.2 日本語の受動態を中国語の特殊な受動態に翻訳する場合

(a) 「N1はN2に見舞われる」を「N1遭受N2」に翻訳する。

- ① 今年陽子江の南部は大きな水害に見舞われた。
→ 今年, 长江南部遭受了大水灾。

(b) 「N1はN2に付き添われる/見守られる」を「N1由N2V」に翻訳する。

- ② 彼女は両親に付き添われ病院へ検査に行った。
→ 她由双亲陪同, 去医院做检查了。
- ③ 大統領の専用は戦闘機三機に守られて空港に着陸した。
→ 总统的专用机由三架专机守护着降落在机场。
分析: 動詞は「見舞う」、「付き添う」と「守る」など動詞の受動態は中国語では「被」ではなく、動詞「遭受」のみか、「由」の前置詞で表す。動詞は小数に限ると推測する。翻訳法としては、見舞う/付き添う/守るをキーワードとするパターン規則で解決できる。

3.3 日本語の受動態を中国語の使役形に翻訳する場合

N1はN2に言われてN3をV

例: 兄に言われてこれをやった。

→ 哥哥让我做这个事。

分析: この文を中国語の受動文に翻訳すると、とても不自然であり、使役文に翻訳する。

翻訳規則： →N2 让 N1VN3

3.4 日本語の能動態を中国語の受動態に翻訳する場合

日本語の能動態を中国語では受動態に翻訳する場合は少ないが、日本語の動詞が能動態で受動の意味を表す場合、中国語では受動態に翻訳することがある。即ち、N1 は N2 に V→N1 被 N2V

① 顔は炎に赤く染まっていた。

→ 脸被火光染红了。

② 彼女はこの映画に感動した。

→ 她被这个电影感动了。

分析：このような場合、N2 は無生物で、V は自動詞である。他に「石につまずく、雪に埋もれる」などがある。このような日本語動詞は受動態になれず、動詞自身が受動の意味を含んでいる。個々の動詞のパターン規則の記述で解決できる。

規則例： 「つまずく」

<N1>は<N2>につまずく (N1=人/動物 N2=具体物)

→ N1 被 N2 绊倒

以上の分析に基づき、「られる」に対応する中国語表現の訳しわけ規則を表 1 にまとめた。

4. 品詞のずれの表現に関する分析

(1) 名詞を動詞に翻訳する(日本語のコプラ文に関連するずれ)

(1-1) 動作性を表す名詞を動詞に翻訳する

① 私は来月アメリカへ行く計画だ。

→ 我计划下个月去美国。

分析：①のようなコプラ文(「うなぎ文」)は中国語では通用しない。「計画」に対応する名詞訳語もあるが、この場合は「…と計画する」のように、動詞として翻訳すべきである。パターン翻訳では「だ」をキーワードとし、「だ」の名詞部を字面条件とするパターン規則で解決できる。

例：<N1>は<V><計画>だ (N1=人/動物)

→ N1 计划<V>

名詞の「計画」は中国語の述語動詞として訳し、「だ」に対応する訳はない(「だ」=「Φ」)。「アメリカへ行く」の部分は連体修飾語から動詞フレーズの目的語に変換する。

もし、この文の「計画」の位置に現れる名詞がすべてサ変名詞であるという推測が成り立つならば、次の翻訳規則が可能である。

N1 は VN2 だ→N1+N2´+V

(N2=サ変名詞, N2´=N2 に対応する中国語動詞)

(1-2) 思想, 気持ちを表す名詞を動詞に翻訳する。

① そのとき苦しい思いだった。

→ 那个时候觉得很苦。

分析：「苦しい思いだった」は「苦しく感じた」に翻訳する。次のパターン翻訳規則で解決できる。

① (Nは)<adj><思い>だ

→ N 觉得<adj>

このような形容詞/名詞にはどのようなものがあるかについてはさらに検討する必要がある。

(2) 形容詞(「近い、強い」)を動詞に翻訳する

① 大会は終わりに近い。→ 大会接近结束。

② 北極熊は寒さに強い。→ 北极熊耐寒。

分析：①の形容詞の「近い」は中国語では動詞「接近(接近する)」に翻訳、②の「強い」も中国語では動詞「耐(耐える)」に翻訳する。翻訳規則としては形容詞ごとに字面のパターンとして登録すればよい。

① <N>は<終わり>に近い N=体言

→ N1 接近结束

このような形容詞がどれくらいあるかについての検討は今後の課題。

(3) 動詞を副詞に翻訳する

① 大勢の人が困んでいるから、何か起こったに決まっている。

→ 因为有很多人围住，一定是发生什么事了。

② 太郎はいつも一つ事を繰り返して話す。

→ 太郎经常重复说话。

③ 何十回の登りを経て、とうとう山頂に辿り着くことに成功した。

→ 经过几十回的攀登，终于成功地到了山顶。

④ 私は一貫して彼を支持する。

→ 我一贯支持他。

⑤ 悪天候の中太郎は苦勞して車を運転した。

→ 在恶劣的气候中，太郎辛苦地开着车。

分析：① は動詞「決まる」を副詞の「一定(必ず)」に翻訳する。翻訳規則としては、

〈用言〉に決まっている → 一定〈用言〉

というパターンで翻訳できる。

② は動詞の接続形「繰り返して」を中国語の副詞の「重複」に翻訳する。

〈N〉を繰り返して〈V2〉→ 重複 V2N

というパターン規則で翻訳できる。このような動詞がどれぐらいあるかは今後の課題。

③-⑤はサ変動詞の問題でもある。③は文末で現れ、④-⑤は「して」接続形として現れる。翻訳規則としてはそれぞれのサ変動詞をキーワードとするパターンの構成で翻訳できると思われる。ただし、数が相当あると予測され、また同じサ変動詞でも「運動して汗をかく」のような場合は、副詞ではなく動詞に翻訳したりするケースもあるため、サ変動詞の属性に関する分類から訳しわけの条件に関する考察が必要である。

(4) 動詞を形容詞に翻訳する

① 正常では空気中の酸素と二酸化炭素の配分率は決まっている。

→ 正常情况下，空气中的氧和二氧化碳的分配比率是一定的。

分析：「決まる」は日本語では動詞であるが、中国語では「一定的」という形容詞に翻訳し、コブラ動詞「是」を用いる必要がある。このような動詞は限られていると思われるので、動詞をキーワードとするパターンの構成で解決できると思う。ただし、このような動詞がどれぐらいあるかについての分析は課題である。

5. 評価

(1) 2、3、4節で列挙したずれがある日本語例文をある市販日中翻訳ソフトで翻訳したところ、32文に関して、わずか約31%の精度(ずれ部分の評価)を得たのみであり、本研究の必要性が伺えた。(2) 日英対訳コーパス[9]の中からランダムに抽出した1万文を対象とし、我々の研究室で開発しているIbuki解析システムの文節解析結果に基づき、さらに人手のチェックを経て、760文の受動文を抽出した。次に、表1の(1)-(10)で示しているずれがある受動文210文を抽出した。さらに抽出したずれ受動文の760受動文における出現率と、表(1)で提案した翻訳規則に基づく翻訳精度を手作業で評価した。全体として、ずれ受動文の出現率は約28%、翻訳精度は約76%の結果を得られ、本研究の必要性和規則の有効性を確認できた。またそれぞれのずれタイプの出現率と、翻訳規則の正訳率も評価し、詳細を表2に示す。ただし、翻訳精度は受動助動詞「られる」に対応する訳語は正確かどうかのみを注目して行ったものである。

問題考察：

以下は評価実験で現れた主な問題点である。

[1] 現在のずれに関する分類ではカバーできない場合があった。

(a) 表1のタイプ(3)のずれに関する問題

(1) これが必要な条件と考えられている。

(2) これが必要な条件だと考えられている。

(3) 彼はスーパーマンと言われる

例(1)の「られる」も「Φ」に翻訳すべきだが、この例の「必要な条件」は名詞節であるため、表1の規則(3)の条件「(Nは/が)PとV」というパターン条件に照合できないため、正解は得られない。「(N1は/が)N2とVられる」というパターンを増やして対応すると、例(3)も照合できるが、この場合の「られる」は「被」に訳すべきであるため、例(3)は誤訳になる。上記の3例とも正訳を得る

ために規則を更に詳細化する必要がある。

表1 受動文の翻訳規則

タイプ 番号	翻訳条件		「られる」に対 応する中訳	注
	日本語パターン構成	構成要素の条件		
(1)	N は/が V られる	a N=抽象 b V=他動詞, 抽象性 の動作を表す(※)	Φ	抽象名詞は文献[8]を参 考するが、抽象性の動詞 の属性の特定は困難
(2)	N1 には/に N2 が V られる	a N1=場所(<地名/組 織名/場所/場>) b V=他動詞	Φ	N1 の属性としては、 抽象的な場所の限定 は困難
(3)	(a) (N は/が)P と V られる (b) (N は/が)P ものと V られる	V=「思う,考える,見 る,見なす,推測する, 言う,伝える,見込 む,予想する,見積も る,話す,語る,推定 する」。	Φ	
(4)	N1 に/で V られた/られている N2	a. N1=範囲/場所 b. 連体修飾語の V=他 動詞	Φ	N1 の属性の限定は困 難な場合がある。
(5)	N は V られた/ているものだ		Φ	
(6)	(a) N1 は (Adv)V られてある (b) N1 は N2 に V られてある	V=他動詞 N2=場所	Φ	
(7)	(a) N1 が (は)N2 に VI られる (b) N13 は N1 に N2 を VT られ る	N1=人/動物 (a) V=「降る,吹く」 以外の自動詞(死ぬ, 泣く,倒れる,来る, なる) N2=省略不可 (b) V=他動詞	Φ	(a)には動詞ごとにパ ターン規則で翻訳で きるが、(b)には現在 の条件では普通の受 動文との区別は困難
(8)	N1 は N2 に言われて N3 を V	V=他動詞	让	
(9)	N1 は N2 に見舞われる		遭受	
(10)	「N1 は N2 に付き添われる/守 られる」		由	
(11)	default		被	

表2 ずれのある受動文の出現率と翻訳の正解率(表1の(1)-(10)のタイプ)

対象文数	抽出した受動文	ずれがある受動文およびその出現率	ずれのある受動文の正訳率
10000	760	総文数とその出現率 210 210/760≒28%	総正訳率 159/210≒76%
		タイプ別の受動文とその出現率	タイプ別の正訳文と正訳率
		(1) 111 (2) 19 (3) 34 (a) 16 (b) 9 (4) 19 (5) 7 (6) 0 (7) 19 (a) 8 (b) 11 (8) 0 (9) 1 (10) 0	111/760≒15 19/760≒3% 35/760≒5% 19/760≒3% 7/760≒1% 0 19/760≒3% 0 1/760≒0.1% 0
			タイプ別の正訳文と正訳率 (1) 75/111≒68% (2) 17/19≒89% (3) 34/34=100% (4) 17/19≒89% (5) 7/8≒88% (6) 評価文なし (7) (a) 8/8=100% (b) (未評価) (8) 評価文なし (9) 1/1=100% (10) 評価文なし

(b)使役受動文の問題

「V せられる」のような使役受動の中訳語は「被」、「Φ」、「让」、「被迫」などに翻訳し、対応にずれがあるがその訳し分けに関する考察

は今後の課題である。

(1) 君はこんなひどい目に逢わせられたとは思わなかった。

→我没想到你Φ遭受了这么大的不幸。

(2) 父親は百円払わせられて不機嫌な顔をした。

→ 父亲被迫付了一百元, 显得很不高興。

[2] 翻訳規則の条件の属性の付与に関する問題

(a) 表1のタイプ(1)のずれの属性問題

分析: タイプ(1)の動詞(V)の属性は「目で見えない抽象的な動作を表すもの」というように想定している(「(会議を)開く、把握する、思う」など)。

しかし、抽象性の動作を表す動詞の中身はどこまで包括するかすべて明確とは言えない。名詞の属性は基本的に文献[8]を参考にしたが、用言の属性に関して、[8]では<抽象的関係>という付与があるが、本論で想定している「抽象性の動詞」と異なる部分が多い。そのため今回はタイプ(1)の「V=抽象性動詞」という判断は筆者の判断のみであり、次の例(1)-(3)のようなNの属性あるいはVの属性が曖昧な場合は今回の評価では除外した。従って、今回評価したタイプ(1)の出現率はおおよそそのものとどまった。また、正確率は約68%で低い。NとVの属性を再分類し、より詳細な翻訳条件が必要であると考える。

(1) 何も入力されないと、スリープモードが動く。
(2) 生産開始の判断は、この調査の結果に基づいてなされる。

(3) 各個人を識別するユーザ IDが付与される。
分析: 例(1)と(3)には、N「何」と「ユーザ ID」、例(2)はV「する」の属性は判断しにくい。

(b) 同じ表(1)の(2), (4)タイプの「場所」と「範囲」の属性の機械での付与は困難な場合もある。

(1) この都市で開催されたファッションショウはすばらしかった。
(2) 本仕様書で規定された品質の資材を使用しなければならない。

分析: この2例とも表(1)タイプ(4)の条件と照合でき、手作業評価では一応正解とした。しかし、機械では訳例(1)の「都市」に対して「場所/範囲」の属性の付与は可能であるが、「仕様書」の属性を「場所/範囲」と付与しにくい([8]では「都市」

=<場所>、「仕様書」=<人工物>/<文書類>)。

6. まとめ

本研究は日中機械翻訳のために、両言語の構文上にずれがある表現の翻訳について考察することを目的とした。今回はその第一歩として受動態と能動態のずれ、品詞のずれを中心に分析し、その翻訳の規則をまとめた。また受動態と能動態のずれについて手作業によってそのずれ表現の出現率と翻訳規則の精度の評価を行った。わずかな評価サイズであるが、760受動文に対して約28%の出現率と約76%の正訳を得て、本研究の重要性と有効性を一定の程度で確認した。

今後はずれ表現の分析と評価をさらに進め、また規則のシステムへの実装を進めていきたい。

謝辞: 本研究は主に日本学術振興会外国人特別研究員事業の支援により実施したものである。

参考文献:

- [1] 刘绩生编著, 实用日汉对译技巧, 湖南人民出版社, 1997.
- [2] 苏琦编著, 汉日翻译教程, 商务印书馆, 1998.
- [3] 今富正巳, 新訂中国語⇔日本語翻訳の要領, 光生館, 1973.
- [4] 陶振孝編著, 动词, 外语教学与研究出版社, 1997.
- [5] 浅井良信, 宇野修一, 卜朝暉, 池田尚志等, 機械翻訳システム jaw と多言語への翻訳実験, 言語処理学会第11回年次大会発表論文集 pp. 237-240.
- [6] 金水敏, 日本語の受動文および関連する現象, <http://www.let.osaka-u.ac.jp/~kinsui/zyugyou/history/2002/passive.pdf>
- [7] 高橋太郎, 他著, 日本語の文法, 2005年4月.
- [8] 池原悟他, NTT 語彙体系, 岩波書店, 1997年.
- [9] 村上仁一, 池原悟他, 日本語英語の文対応データベースの作成, 第7回 LACE 研究会, pp1-10, 2002, 12, 1.