

日中機械翻訳における「NはNです」文型の分析

陳麗 山本いずみ

名古屋工業大学大学院情報工学専攻

1. はじめに

「NはNです」という文型は、日本語学習者にとって最も基本的な文型である。日本語の「NはNです」文型は名詞文であるが、対応する中国語の訳は名詞文だけでなく、動詞文の場合もある。本稿では、日本語の「NはNです」文型を中国語に翻訳する時の問題点、特に現有の機械翻訳ソフトにおける「NはNです」文型翻訳の問題について考察し、名詞の意味範疇に着目しながら翻訳規則を提案する。

2. 日中両言語の名詞文の対応の多様性

日本語では、名詞は他の品詞より、同じ範疇に属する語数が圧倒的に多い。人名や地名などの固有名詞、一般的なものの名前、年月日や時間、場所など、日常生活でよく使うことばは名詞である。つまり、名詞だけで多くの情報を交換することができると言える。

一方、中国語名詞文には「是」を用いる文と用いない文がある。「是」字を用いる文を「是字句」と呼ぶ。文中に「是」をはさみ、名詞を主語とするともに、述語ともする文型である。日本語の名詞文と中国語の対応関係について、具体的に示すと次のようになる。

(1) 中国語も名詞文で対応するもの

① 「是字句」で対応するもの

中国語の「是字句」は日常生活などで多用される表現形式である。日本語の「N1はN2です」文型（以下、前項NをN1、後項NをN2とする）は一般的に中国語の「是字句」で翻訳される。例えば、以下の通りである。

例 1-1 これは本です。

1-2 这是书。

② 名詞文「NN」で対応するもの

この表現の述語は、時間・月日・曜日・季節・年齢・値段などを表す名詞である。この場合、中国語は「是」を省略した名詞文「NN」で対応する。

例 2-1 今日は6月20日です。

2-2 今天6月20日。

(2) 「在字句」で対応するもの

日本語の名詞文が場所や存在を表現している時、中国

語では「在字句」で対応する。中国語の「在字句」はあるものはある場所に存在することを表現する文であり、常に場所名詞を伴う。

例 3-1 日本語科は図書館の隣です。

3-2 日语系在图书馆的旁边。

(3) 動詞文で対応するもの

日本語には「する」をつけることでサ変動詞となる、強い動作性を持つ名詞がある。こうした名詞を動作性名詞と呼ぶ[1]。N2に動作性名詞を用いる名詞文は、中国語では動詞を用いて表現する。

例 4-1 試合は中止です。

4-2 比赛中止。

また、N2に動詞連用形が名詞化した名詞を持つ文も、中国語では動詞文になる[1]。

例 5-1 あの本は売切れです。

5-2 那本书卖完了。

さらに、日本語で天候名詞を用いる名詞文も中国語では動詞文で表現される。

例 6-1 今日は雨です。

6-2 今天下雨。

以上のように、日本語の「N1はN2です」文型は中国語との対応が複雑である。それにもかかわらず、現在公開されている機械翻訳システムには、これらのことが十分反映されていないため、名詞文を誤訳することになると考えられる。

3. 日中機械翻訳における誤訳

3-1. 誤訳率

現在公開されている機械翻訳システムにおける誤訳の例について考察する。ここでは、「みんなの日本語初級I・II」の各課に挙げられた基本文型を資料とし、現在よく使われている“エキサイト翻訳” “livedoor 翻訳” “Infoseek マルチ翻訳” “Yahoo! 翻訳” “翻訳: @nifty” の5つの翻訳サイトを用いて翻訳を行った。そこで出現した誤訳について具体的に分析する。

「みんなの日本語初級I」の総文型数は79である。このうち、「N1はN2です」文型（ここでは形式名詞等

も含め、なるべく大きな枠で洗い出した)は17あった。これらを各翻訳サイトで訳した結果、最も誤訳率が低かったのは“エキサイト翻訳”であり、18%であった。「みんなの日本語初級Ⅱ」の総文型数は73であり、「N1はN2です」文型数は8であった。「Ⅰ」に比べ「Ⅱ」では文型が多少複雑となる。“エキサイト翻訳”を用いて翻訳した結果、「N1はN2です」文型の誤訳率は50%になった。これらの結果を表1「誤訳率」で示す。

表1 誤訳率

	全体			「NはNです」		
	○	△	×	○	△	×
初級Ⅰ	80%	14%	6%	82%	12%	6%
初級Ⅱ	62%	26%	12%	50%	40%	10%
合計	72%	19%	9%	72%	20%	8%

※注：○：正しい文
△：意味が分かるが、文法的に正しくない文
×：意味も文法も正しくない

3-2. 誤訳の具体例

現在公開されている翻訳サイトを用いて行った翻訳の結果を具体的に示す。表2「全正訳文」は5つの翻訳サイト全てで正しく訳された文である。表3「正訳・誤訳文」では“エキサイト翻訳”以外の4つのサイトで誤訳となった文を示し、表4「全誤訳文」では、“エキサイト翻訳”を含む全てのサイトで誤訳になった文を挙げた。

まず、表3について分析する。N01, 2, 3, 6では外国人の人名(サントス, ミラー)の訳が正しくなされなかった。外国人名が外来語の一般名詞あるいは意味不明の言葉と認識され、正しく訳されないことは多々起こり得ることである。N04では、「わたしの」の訳語として

表2 全正訳文

N0	例文	対応文型	訳文
1	これは辞書です。	N1 是 N2	这是词典。
2	これはコンピュータの本です。	N1 是 N2	这是计算机的书。
3	それはわたしの傘です。	N1 是 N2	那是我的伞。
4	ここは食堂です。	N1 是 N2	这里是食堂。
5	今4時5分です。	N1 是 N2	现在是4点5分。
6	富士山は高い山です。	N1 是 N2	富士山是高的山。
7	今日は僕の誕生日だ。	N1 是 N2	今天是我的生日。
8	あしたから旅行なんです。	N1 是 N2	从明天开始是旅行。

「東西」が余分に加わっている。N05では「很好看」と訳すべきところを「的」を抜かして「很好看」と訳し、逆にN06では「在那里」の訳語として「在那里」を訳している。N07でも「始まる」の訳語として「开始」が加わっている。N08では「明日は寒くなる」の語順が間違っており、N09では「隣の部屋に」の訳語と語順が違っている。

次に、表4の例文の誤訳原因について分析する。N01は、外国人の名「マイク・ミラー」を訳し間違えている。N02は場所名詞「あそこ」に対して存在を示す「在」を使うべきところを「是」を用いたことによる誤りである。N03は「下雨(雨が降った)」と訳さなければならないところをやはり「是」を用いて訳している。N04, N06は日本語と中国語の時制の捉え方の違いによって起こった誤りである。N05は「という」の訳語と位置を誤っている。N07では「はず」の訳語「应该」の位置を誤っており、N08では「大学を卒業」の訳語「大学毕业了」の語順を誤っている。

以上、現有の機械翻訳システムにおける「N1はN2です」文型翻訳の問題点を見てきた。単語自体の誤訳、語順の誤り、対応する文型の違いから、最も基本的な名詞文「N1はN2です」が、現有の機械翻訳システムではきちんと翻訳されていなかった。以下では、これらを解決するための翻訳規則を提案する。

4. 日本語名詞文の翻訳規則

日本語の「N1はN2です」文型を中国語に翻訳する場合、対応する中国語の文型は「是字句」、名詞文「NN」、「在字句」、「動詞文」である。ここでは、「N1はN2です」のN1, N2の意味に注目して分類し、対応する中国語を示して翻訳規則を提案する。なお、複雑な「うなぎ文」については考察の対象としない。

「N1はN2です」文型を中国語に翻訳する時の最も一般的な文型として、「是字句」が考えられる。以下ではこれに当てはまらない場合について述べる。

(1) 名詞文「NN」で訳される場合

①N1が相対的な時を表し、N2が絶対的な時(年月日・季節・時間・曜日など)を表す

例 7-1 今日(相対)は土曜日(絶対)です。

7-2 今日(相対)は(絶対)星期六。

②N1に関係なく、N2が年齢・値段を表す

表3 正訳・誤訳文

NO	例文	対応文型	訳文	
1	サントスさんは学生じゃありません。	N1 不是 N2	○	Santos 先生不是学生。
			△	<u>圣者们</u> 先生不是学生。
2	ミラーさんは会社員ですか。	N1 是 N2 吗	○	米勒先生是公司职员吗？
			△	<u>镜子</u> 是公司职员吗？
3	サントスさんも会社員です。	N1 也是 N2	○	Santos 先生也是公司职员。
			△	<u>圣者们</u> 先生也是公司职员。
4	この傘はわたしのです。	N1 是 N2	○	这个伞是我的。
			△	这把伞是我的 <u>东西</u> 。
5	桜はきれいな花です。	N1 是 N2	○	樱花是漂亮的花。
			△	樱花是 <u>很好看</u> 花。
6	あそこにいる人はミラーさんです。	N1 是 N2	○	在那里的人是米勒先生。
			△	在那里 <u>有</u> 的人是 <u>镜子</u> 。
7	会議は今から始まるところです。	N1 正要 V	○	会议正要开始于现在。
			△	会议想要现在开始 <u>开始</u> 。
8	天気予報によると、 明日は寒くなるそうです。	据说 N1V	○	据说根据天气预报，明天变得冷。
			△	据气象预报说， <u>变得明日好象冷</u> 。
9	隣の部屋に誰がいるようです。	好像 V	○	好象谁在隔壁的房间。
			×	<u>象有邻居的屋子谁</u> 。

表4 全誤訳文

NO	例文	対応文型	訳文	
1	わたしはマイク・ミラーです。	N1 是 N2	○	我是迈克·米勒。
			△	我是 <u>麦克风</u> · <u>镜子</u> 。
2	電話はあそこです。	N1 在 N2	○	电话在那里。
			△	电话是 <u>那里</u> 。
3	きのうは雨でした。	N1V	○	昨天下雨了。
			△	昨天是 <u>雨</u> 。
4	これはミラーさんが作ったケーキです。	N1 是 N2	○	这是米勒先生制作的蛋糕。
			△	这是米勒先生制作 <u>了</u> 的蛋糕。
5	立ち入り禁止は入るなという意味です。	N1 是 V 的意思	○	进入禁止是别进入的意义。
			×	是 <u>所说的</u> 进入禁止别进入的意义。
6	わたしが日本へ来たのは去年の3月です。	N1 是 N2	○	我是去年的3月来的日本。
			△	<u>我来了日本是去年的3月</u> 。
7	書類は速達で出しましたから、 あした着くはずです。	N1 应该 V	○	文件以快信出了，明天应该到达。
			×	文件 <u>应该</u> 以快信出了，明天到达。
8	彼は3月に大学を卒業したばかりです。	N1 刚刚 V	○	他3月刚刚大学毕业了。
			△	他3月刚刚 <u>毕业了</u> 大学。

例 8-1 彼は13歳です。

8-2 他13岁。

(2) 「在N1的N2」で訳される場合

※場所名詞の種類を以下のように分類する[3].

a 地名(日本, 北京など)

b 場所を表す意味を含んだ語(家, 学校など)

c 位置を表す語(方位詞, 指示代名詞など)

①N1が場所名詞a b c, N2が有情物

例 9-1 テーブルの下は猫です。

9-2 在桌子下面的是猫。

(3) 「在字句」で訳される場合

①N1が場所名詞a b, N2が場所名詞c

例 10-1 事務所は東側です。

10-2 事务所在东面。

②N1が場所名詞ではなく, N2が場所名詞a b c

例 11-1 電話はあそこです。

11-2 电话在那里。

(4) 動詞文で訳される場合

①N1が時詞, N2が天候を表す名詞

例 12-1 今日は雨です。

12-2 今天下雨。

②N1に関係なく, N2がサ変名詞・動詞連用形名詞

例 13-1 今日は休業です。

13-2 今天休業。

例 14-1 電車は15分の遅れです。

14-2 电车晚点15分钟。

5. 誤り分析と新たな問題点

提案した規則を評価するために、インターネットから拾い上げた300文を用いて手作業で誤り分析を行った。その結果、場所名詞に関し、前述の提案規則では対応できない例が現れた。

提案した規則では「是字句」で対応する一般的な文型とした「N1が場所名詞b, N2が場所名詞a」において、誤った翻訳がなされる場合があった。以下に正訳と誤訳を示す。

例 15-1 中国の首都は北京です。

15-2 中国的首都都是北京。

この例15は正しく翻訳されたが、次の文では誤訳が起こった。

例 16-1 私の学校は東京です。

16-2 (誤) 我的学校是东京。

16-3 (正) 我的学校在东京。

例16は、「学校」は「東京」のある場所(一部分)に存在することを表しているので、「是字句」ではなく、「在字句」で訳されなければならない。

また、形式名詞の翻訳でも問題が見られた。

例 17-1 旅行に行くつもりです。(意志)

17-2 (誤) 是去的想法旅行

17-3 (正) 打算去旅行。

例 18-1 火を消したつもりです。(思いこみ)

18-2 (誤) 是关闭了火的想法

18-3 (正) 以为熄灭了火。

例17の意志を表す「つもり」の対応語は「打算」であり、例18の思いこみを表す「つもり」の対応語は「以为」である。同じ「つもり」でも使用される文により意味が異なる。こうした意味の異なりを正しく訳すためには、前後にどのような語が来るか、どのような語と一緒に使われることが多いかについて分析しなければならない。「みんなの日本語初級Ⅱ」の誤訳と考えると、形式名詞についてはさらなる検討が必要である。

6. おわりに

本稿では日本語の「NはNです」文型の中国語への対応の多様性について述べ、現有の機械翻訳システムにおける問題点について考察した。これを基として、日中機械翻訳のための翻訳規則を提案し、検証を試みた。

今後は、さらに例文数を増やして検証することで、提案した規則を見直し、より精選した翻訳規則を提案する予定である。

【参考文献】

- [1]張 建華「日中翻訳における表現構造に関する考察」.
<http://www2.kokken.go.jp/jallic/group5/98.8p77.pdf>
- [2]鈴木義昭「初級文型の日中対照—「は」と「が」をめぐって—」(1986)『講座日本語教育』第22分冊.
pp54-67
- [3]丸尾 誠「中国語の場所詞について:モノ・トコロという観点から」(2004)『言語文化論集』Vol. 25,
No. 2. pp151-166
- [4]王軼諷, 池田尚志「日中機械翻訳における存在表現の翻訳について」(2007)『自然言語処理』
Vol. 14, No5. pp64-105