

## 日中機械翻訳における存在文および関連する問題について

王軼謳 ト朝暉 宇野修一 浅井良信 池田尚志  
岐阜大学工学部

**あらまし：** 存在文はそれぞれの言語において特殊な位置を占めていると思われるにもかかわらず、それを対象とする研究は少なく、日中機械翻訳の角度からの研究は殆ど見あたらない。現在の日中市販翻訳ソフトでは、存在文に関する誤りが多く見られる。本論文では、日中両言語の存在文における異同について考察し、日中機械翻訳のために、日本語文の構文特徴、対応名詞の属性、中国語文の構文構造などを利用して存在動詞の翻訳規則をまとめ、存在文の翻訳方法について提案した。これらの翻訳規則を我々の研究室で開発している日中機械翻訳システム Jaw/Chinese に組み込んで、翻訳実験を行った。更に手作業による翻訳実験も加えて、この規則を検証し、良好な評価を得た。

### A solution for the problem of Existential Sentence in Japanese-Chinese Machine Translation

Yiou Wang, Zhaohui BU, Syuichi Uno, Yoshinobu Asai, Takashi Ikeda  
Faculty of Engineering, Gifu University

**Abstract:** Existential sentence as the basic sentence pattern is very important for each language. However in Japanese-Chinese machine translation, the translation of the existential expression is still one of the difficult problems. In this paper, we propose a method for handling the existential verbs based on the constraint of Japanese syntactic and semantic features, Chinese syntactic features, the attributes of the related nouns and so on. We implement the translation rules in Jaw/Chinese and evaluate our rules manually over 500 existential sentences and get an accuracy of about 93%, which is rather high compared to the currently commercially available translation software.

### 1 はじめに

存在文はたいていの言語で特殊な文型をとり、言語により異なりが現れてくる。日本と中国の言語界において存在文について論述があるが[1][2]、日中機械翻訳の角度からの研究は殆ど見あたらない。現在の日中市販翻訳ソフトでは、存在文に関する誤りが多く見られる。日中機械翻訳において、日本語の存在文を中国語に翻訳する際、中国語存在動詞の訳し分けと存在動詞とその同形詞の混淆は二つの大きな問題だと思われる。本論文では、日本語文の構文特徴、中国語文の構文特徴、対応名詞の意味属性などを利用して存在動詞の翻訳規則をまとめ、存在文の翻訳方法について提案する。

### 2 「ある」と「いる」

存在文は日本語の基本文型の一つであり、存在主（いわゆる「主語」）と存在場所に関する情報を必須のものとする。まず、存在主の意味的性質により、二種に大別される[3]。これは日本語の特徴らしい。存在主が非情物（無生物や現象など）の場合は、存在の意を「ある」で、非存在の意を「ない」で表す。一方、存在主が有情物（生物や人間など）の場合は、存在の意を「いる」で、非存在の意を「いない」で表す。中国語では、非情物と有情物にかかわらず、「存在」そのものを「有、在」などの存在動詞により表現する。

存在の概念をわれわれの言語を通じて明らかにしたいならば、「存在」という名詞にではなく、

まず「ある」と「いる」という動詞に着目するのが当然と思われる。飯田隆[1]は「「存在する」の用法を検討し、それと「ある」や「いる」の用法とを比較することは、十分意義のあることである」と指摘している。

本論文では「ある」、「いる」、「ない」を対象として、中日両言語における存在動詞の翻訳規則を中心に述べる。

## 2.1 非情物存在主文「ある」

「ある」は中国語に翻訳すると、意味上の「所在」か「所有」に焦点に置くことによって、基本的に「有」か「在」で表現し、ほかに場合によっては「在于、发生、进行」などの別の動詞で対応することもある。例えば、

(1)机の上に本がある。

桌子上有书。

There is a book on the desk.

これらを次のように言い換えると、ナニカよりドコカが焦点となって、存在というよりも、ものありか、所在を示す文になる。中国の訳語も変わる。

(2)本は机の上にある。

书在桌子上。

The book is on the desk.-

「ある」の意味分類を中国語の訳語との対応関係を考慮に入れて、また文献[1]、[2]、[3]、[5]などを参考にして表1にまとめた。

以降の表中のN1、N2、N3は名詞を、Xは単語、句、文などを、Pは結び部分(述部)を表す。

表1 「ある」の意味分類

分類	基本文型	日本語の意味	対応する中訳語	例文
1	N1はN2にある。	所在地の説明	在、位于	郵便局は角を曲がった所にあります。
2	N1はN2にある。	帰属する	属于, 在	責任は彼にある
3	N1はN2にある。	よる、関係する	在于,	西洋の美は均整にあり、東洋の美は不均整にある。
4	N1にN2がある	存在物の説明	有	ここから目黒へ行く間にとっても静かな自然教育園があります。
5	N1にはN2がある	持っている	有, 具有	恵美は自分に音楽の才能があるとうぬぼれている。
6	N1でN2がある	行われる	进行, 开展 举行	イースターは3月または4月にあるキリスト教徒たちのお祭りです。
7	N1でN2がある	起こる	发生	男はその事件があつてから全く生まれ変わった。
8	N1がN2ある	高さ、重さ、距離、面積など数量の表示(判定詞)	Φ、有、是	あの人は背の高さが182センチ、体重は68キロある。
9	……とある	書記活動の結果存続	写有	右へ行け、とある。
10	……したことがある	経験	曾经…	先生はアメリカへ行ったことがある。
11	……することがある	時、場合	有时	群衆の中にいても寂しく感じることもある。
12	N1はN2に/である	存在状態、時期の説明	处于	彼には危険な状態にある。

「ある」の意味は、存在主と存在場所などの対応する名詞のそれぞれの文法的、意味的特徴と相関している。我々は広辞苑と日中辞書を調べ、日英対訳コーパス[6]から抽出した「ある」を含む存在文のうちの400例文に対する観察と日本語彙大系の「ある」のパターンなどを参考にして、たいていの意味用法は四つの要素で区別できる

ことを確認した。

以上の分析に基づき、日本語文の構文特徴(文型、助詞、テンスなど)と存在主と存在場所の意味属性、構文的な特徴を用いて、中国語の存在動詞の組み合わせの制限を総合して分析し、判定条件(翻訳規則)を機械で処理可能な形に整理した(表2)。

表2 「ある」の翻訳規則

意味分類	判定の条件とその属性						中訳語
	日本語パターン	名詞 N1 の属性	名詞 N2 の属性	名詞 N3 の属性	Xの構造と成分		
1	(1)	N1 は N2 にある	具体/固有名詞	場所/場/地名/組織名			N1 位于/在 N2
2	(2)	N1 は N2 にある	抽象	具体			N1 在 N2
3	(3)	N1 は N2 にある	抽象	関連/性質/状態/存在			N1 在于 N2
4	(4)	N1 に/は N2 がある	場/具体/固有名詞	具体/固有名詞			N1 有 N2
	(5)	XN1 もある。	N1=人/人々/者			動詞連体形	有 X 的 N1
5	(6)	N1 に/には N2 が/の /null ある。	体言	抽象			N1 有 N2
	(7)	N1 は N2 に/に対して N3 がある	主体	体言	抽象		N1 对 N2 有 N3
	(8)	N1 に/は N2 が/の /null ある	人	人間(親族関係)/人間(交際関係)			N1 有 N2
	(9)	N1 に/は X もがある	体言			用言連体形	N1 有 X 的一面/
	(10)	N1 に/は X くらいがある	体言			用言連体形/「体言+の」	N1 有点儿 X
6	(11)	N1 で N2 がある	場所/場	式、行事			在 N1 举行 N2
	(12)	N1 で N2 がある	場所/場	相談			在 N1 进行 N2
7	(13)	N1 で N2 がある	場所/場	来事/天災			在 N1 发生 N2
8	(14)	N1 に/は N2 の N3 がある	具体物/人	数量	度量衡(重さ/重量/面積/広さなど)		N1 N3 是 N2
	(15)	N1 は N2 が N3 ある	具体物/人	度量衡	数量		N1 N2 是 N3
	(16)	N1 が N2 にある。	限度	数量			N1 是 N2
9	(17)	N1 に/は X とある	体言			動詞連体形/体言	N1 上写有 X
	(18)	X とあって、P (文)				動詞連体形/体言	因为说是 X, 所以 P
10	(19)	N1 に/は X ことがある	主体			動詞連体形過去形	N1 曾经 X
11	(20)	X ことがある				動詞連体形現在式	有时 X
	(21)	X ことがよくある				動詞連体形現在式	经常 X
	(22)	N1 に/は N2 だけのことはある。	体言	体言			N1 不愧为 N2
	(23)	X だけのことはある。				動詞連体形過去形	没有白/值得 X
12	(24)	N1 は N2 にある	体言	期間、最中、立場			N1 处于 N2
	(25)	N1 にあって、X	体言			用言	N1 的时候 X
	(26)	X 運命にある				動詞連体形	命中注定 X

以上の 26 個の日本語パターンのうち、意味分類 5 の(6)と(7)の中には、状態名詞存在主文(「熱、金、暇」など、状態を表す名詞が存在主になるもの)、意思素質名詞存在主文(「やる気、勇気、才能」など、意思や素質に関する名詞が存在主になるもの)と出現物名詞存在主文(「効果、疑問、責任、」など、作用や行為の結果出現したことが存在主になるもの)が多い。この中には語彙の組み合わせの制約条件として、個々の語彙そのものを制約条件とする(字面照合)のが適切なパターンがある(現在のところ 31 個)。例えば

- (1) (N1:主体/精神/状態)には落ち着きがある  
→ N1 沉着
- (2) (N1:人/抽象物/性質)には表裏がある  
→ N1 表里不一

これらを含めると「ある」の翻訳パターンとして全体で 57 個を設定した。

## 2.2 有情物存在主文「いる」

「いる」は中国語に翻訳すると、意味と文法上の規定により、「在、有、处于…」などの可能性がある。日本語では「いる」とひらがな表記されると同音語彙でいくつの意味があり(居る、要る)、中国語に翻訳する際意味を混同する場合が多い。

「いる」の意味分類を中国語の訳語との対応関係を考慮に入れて、また文献[1]、[3]、[4]などを参考にして表 3 にまとめた。「ある」の場合と同様に、「いる」の翻訳規則を表 4 のようにまとめた。

表 3 「いる」の意味分類

分類	基本文型	日本語の意味	対応する中訳語	例文
1	N1がN2にいる。	存在する	有, 在	会社にいる間は愛想よくするようにしている。
2	N1がN2にいる。	住む, 居住する	居住, 在, 逗留	子どもたちは長い間海外にいましたので日本語がうまく話せません。
3	N1でいる	ある状態にとどまる	始終停留(在…)保持	兄はまだ独身でいる。

## 2.3 存在否定文

「今日は、風がある。:今日は、風がない。」「あそこに、猫がいる。:あそこに、猫がいない。」のように、述語が「ある」の否定「ない」と「いる」の否定「いない」等で構成される文を存在否定文という。丁寧体では、「ありません/いません」の形が用いられる。

「いる」の否定「いない」の存在否定文の翻訳処理は通常の否定の処理(「いる」の翻訳プラス否定の翻訳)でできる。しかし、「ない」は単純に「ある」の否定としては翻訳できない場合が多い。「ない」そのものは意義をもつ単語であり、「ある」の意味分類の外の意味も持っている場合がある。例えば、「類がない」、「またとない」は「とてもすばらしい」の意味であって、「今さら泣くことはない」の「ことはない」は「不必要なこと、あってはならないことをいう」の意味であ

る。これらは「ある」の否定として翻訳することはできない表現である。それで、我々は「ない」を存在詞の一種として、翻訳規則を整理した。

まず、「ある」の場合と同様に、状態名詞存在主文、意思素質名詞存在主文と出現物名詞存在主文に対して、語彙の組み合わせの制約条件にして、個々の語彙そのものを制約条件とする(字面照合)のが適切なパターンがある。「ある」と「ない」非対称型の字面照合パターンの整理を行って、現在のところ 21 個のパターンにまとめた。例えば:

- (1) N1は表裏がない → (N1 单纯)
- (2) N1は身長がない → (N1 不高)
- (3) N1に無理がない → (N1 自然)

次は、非存在を表す形式しかない成句と熟語を整理し、現在のところ 11 の個パターンにまとめた。例えば

- (1) しかたない → (没办法) 以上述べた場合の外の「ない」は、「ある」と  
 (2) 申し訳ない → (对不起) 対応できると考えて、存在文「ある」の翻訳プラ  
 それら以外の「ない」の用法と慣用文型の翻訳 ス否定の翻訳という方法で翻訳する。

規則を表5のようにまとめた。

表4 「いる」の翻訳規則

意味 分類	番号	判定の条件とその属性				中訳語
		日本語パターン	名詞 N1 の属性	名詞 N2 の属性	修飾文 X の構造と成分	
1	(1)	N1 は N2 にいる	人/動物	地名/場所/場		N1 在 N2
	(2)	N1 は N2 にもいる	人/動物	地名/場所/場		
	(3)	N1 はいる	動物			有 N1
	(4)	N1 はいる	人(個有名詞)			N1 在
	(5)	N1 はいる	人(個有名詞以外)			有 N1
	(6)	X N1 は/もいる	人		用言連体形	有 X 的 N1
	(7)	N1 に/は N2 が/の/null いる	人	人間(親族関係)/人間(交際関係)		N1 有 N2
	(8)	N1 に/には N2 がいる。	地名/場所/場	人/動物		N1 有 N2
	(9)	X N1 はだれもいない	人		用言連体形	没有 X 的 N1
	(10)	N1 の中には X (N2=もの/の) がいる	人/動物		用言連体形	N1 中有 X 的 N2
	(11)	N1 は N2 と一緒にいる	人	人		N1 和 N2 在一起
	(12)	N1 にいる	地名/場所/場			在 N1
	(13)	N1 にいる N2	地名/場所/場	人		(在) N1 的 N2
	(14)	上には上がいる				人外有人, 天外有天
3	(15)	N1 が N2 でいる。	人	衣服		N1 穿着 N2
	(16)	N1 が N2 でいる。	人	気分		N1 有 N2

表5 「ない」の翻訳規則

番号	判定の条件とその属性					中訳語
	日本語パターン	名詞 N1 の属性	名詞 N2 の属性	名詞 N3 の属性	修飾文 X の構造と成分	
(1)	X ことがない				動詞連体形 現在式	不必/不会 X
(2)	N1 がない/なくなる	人				N1 去世
(3)	X までもない				動詞連体形	用不着/无须 X
(4)	X わけが/はない				用言連体形	不可能 X
(5)	X よりしかない				動詞連体形	只有 X
(6)	N1 ほど XN2 がない N1 ほど XN2 ほかにない	体言	体言		用言連体形	没有比 N1 更 X 的 N2
(7)	X より N1 がない	N1 = 「しかた、ほか、ほかにしかた」			動詞連体形	只好 X/除了 X 没有办法

## 2.4 日英と日中存在表現の翻訳パターンの比較

日本語語彙大系[4]は日英翻訳のための辞書である。日本語語彙大系の日英翻訳パターンと我々の日中翻訳パターンを比較して、日英翻訳パターン数と日中翻訳パターン数の対比を表6に示し、「ある」の三つの種類のパターンの比較を表7に示す。

表6 日英翻訳パターンと日中翻訳パターンの対比

	日英翻訳パターン数	日中翻訳パターン数
「ある」	219	57
「いる」	0	16
「ない」	168	39

表7 「ある」の三つのパターンの日英と日中の比較

日本語パターン	日英パターン数 (全部字面照合)	日中翻訳	
		パターン	字面照合パターン数
①N1はN2がある	117	N1(体言)に/にはN2(抽象)がある	12
②N1はN2に/に対してN3がある	36	N1(体言)はN2(体言)に/に対してN3(抽象)がある	8
③N1はN2のN3がある	39	0	0

表6を見ると、日英「ある」と「ない」のパターンは日中翻訳のより多いことがわかる。我々の日中翻訳の場合と日英翻訳の場合の翻訳パターンの対応する関係は以下に示す：

- ① 一つの日中翻訳パターンに多数の日英翻訳のパターンが対応する。例えば表7の①と②。
- ② 日英翻訳のパターン(41個)に対応する日中翻訳のパターンは不必要である。

例えば、表7の③「N1はN2のN3がある」の形の日英パターンが39個あるが、日中の翻訳パターンは「N1はN3がある」と「N2のN3」の二つのパターンで翻訳できるので、「N1はN2のN3がある」の日中パターンは不必要である。

③ あるパターンは日中翻訳パターンにあるが、日英翻訳のパターン中にない。この場合には日英翻訳パターンにおける慣用文型の整理が完全ではないのが原因である。例えば、

日本語：N1はN2に恩義がある  
 英語：N1 be indebted to/for N2  
 中国語：N1欠N2的恩情

関係①と②の主な原因は次の2点にある：

- ① 英語では中国語と比べると存在主に対して前置詞(介詞)を細かく区別して使い分ける。
  - (1) N1はN2に効果がある  
 N1 be effective in N2  
 N1对N2 有效果

- (2) N1はN2に同情心がある  
 N1 be sympathetic to N2  
 N1对N2 有同情心
- (3) N1はN2に責任がある  
 N1 be responsible for N2  
 N1对N2 有责任
- (4) N1はN2に偏見がある  
 N1 have prejudice against N2  
 N1对N2 有偏見

以上の例は日中の翻訳規則では、すべて「N1はN2に/に対してN3がある→N1对N2有N3」で表現できる。

② 英語では形容詞で物の性質、素質と状況を説明するが、日本語では存在文(状態名詞存在主文、意味素質名詞存在主文、出現物名詞存在主文)で表現することが多い。そのために比日英の翻訳では個々の形容詞を対応させるためにパターン数が増える。中国語では、日本語とほぼ同様に存在文で表現することが多いので、パターン数は少なくすむ。

- 例(1) N1は暇がある。  
 N1 have time                      N1有时间
- (2) N1は勇気がある  
 N1 be courageous                  N1有勇气
- (3) N1は意義がある  
 N1 be significant                  N1有意义

- (4) N1は教養がある  
N1 be educated N1有**教**养
- (5) N1は価値がある  
N1 be valuable N1有**价**值

### 3 翻訳実験

日英対訳コーパス[6]の約15万文から初めの2万文を我々の研究室の解析システムIBUKIで解析した結果、「ある」、「いる」、「ない」を持つ文はそれぞれ989、379、832文が存在した。更に重複している文と存在表現ではない文を除外し、残りの917、305、787文を存在文の例文とした。このうち、「ある」の400文、「いる」の100文と「ない」の100文を翻訳するための調査対象として分析し、観察した。2節の翻訳規則を我々の研究室

で開発している日中機械翻訳システムJaw/Chineseに組み込んで、調査対象文中の「ある」の72文と「いる」の18文をJaw/Chineseで全文翻訳し(中国語の翻訳文生成まで)、良好な翻訳結果を得た。

調査対象としなかった文の中から、「ある」273文、「いる」177文、「ない」92文を抽出し、オープンテストの評価対象として、評価文の存在表現を2節の翻訳規則を用いて手作業で翻訳実験した。個人判断(語順、存在動詞の選択と助詞の翻訳など)で評価を行った。同時にある市販機械翻訳ソフトを用いて翻訳し、その結果も評価した(表8に示す)。

表8 例文翻訳結果統計と比較

	2万文中の存在文	評価文	正訳文		正訳率	
			A	B	A	B
「ある」	917	273	261	167	95%	61%
「いる」	305	177	165	57	93%	31%
「ない」	787	92	84	35	91%	38%

(A: 我々の訳 B: 市販ソフトの訳)

### 4 関連するいくつかの問題

#### 4.1 数量表現と分量表現の問題について

日本語の存在文では存在動詞と分量表現の副詞あるいは数量表現の名詞を組み合わせる用いることがよくある。存在文を正しく翻訳するためには、分量表現と数量表現の正しい翻訳方法が必要である。

分量表現は副詞(「いっぱい」、「たくさん」、「おおぜい」、「少し」など)を中心とする表現で、人の数や物の数量の多寡を表す表現である[2]。

分量表現は日本語では述語にかかる副詞としてあるが、中国語では形容詞を用いたり、副詞を用いても補語になったり、状語になったりと異なった構造の文として表現される場合が多い。

例:①母はぼくらが汚す衣類を毎日たくさん洗濯します。

→母亲每天洗**很多的**我们弄脏的**衣服**。

(中国語では形容詞、目的語を修飾する)

②彼は大酒飲みでたくさん**飲む**。

→他是个大酒鬼，**喝得很多**。

(中国語では副詞であり、動詞を修飾する補語表現となる)

③ベルトをゆるめて、たくさん**食べて**ください。

→**解开腰带，请多吃**

(中国語では副詞であり、動詞を修飾する状語表現となる)

④**猿**がたくさん木に登っている。

→**很多的**猴子在爬树。

(中国語では形容詞、主語を修飾する)

このように分量表現の翻訳処理は複雑であるが、動詞の自他性と日本語文の構文特徴(文型、テンス、アスペクトなど)を総合して分析し、翻訳規則を今検討中である。

数量表現について、日本語では「数詞+助数詞」の形を取るものを「数量詞」と呼び、数量詞とそ

れに対応する名詞が含まれた句や文を「数量表現」と呼ぶこととする。数量詞は名詞や格助詞との関係から、次の四つのタイプに分けられる。[5]

- ① 数量数詞Q+連体助詞ノ+名詞N+格助詞C+述語  
3人の学生がいる (QノNC型)
- ② 名詞N+連体助詞ノ+数量数詞Q+格助詞C+述語  
学生の3人がいる (NノQC型)
- ③ 名詞N+格助詞C+数量数詞Q+述語  
学生が3人いる。(NCQ型)
- ④ 名詞N+数量数詞Q+格助詞C+述語  
学生3人がいる。(NQC型)

QノNC型、NノQC型とNQC型では、数量詞はいずれの構文においてもある名詞句を修飾限定しているが、NCQ型の数量詞は、述語を修飾しており、数量詞の副詞用法である。しかし中国語では、①-④の表現はすべて存在主である名詞を修飾する表現になる(訳語は「有3个学生」)。また、「彼は身長が1.8メートルある」は「他身高1.8米」となり、数量表現そのものが述語部となる。このように、数量表現の翻訳は複雑であり、数量表現の翻訳処理ルールまだ整理中である。

#### 4.2 「ている」と「である」

日本語の存続体はある種の動作が終了しその結果が存続する状態にあることを表し、他動詞の連用形+「である」(助詞「を」は「が」に変更する)、あるいは自動詞の連用形+「ている」により構成される。例えば

- (1) 壁に絵を掛ける。  
→把画挂在墙上。
- (2) 壁に絵が掛かっている。  
→墙上挂着画。
- (3) 壁に絵が掛けてある。  
→墙上挂着画。

上例の文(2)と(3)はカカッタ状態が一時的に存在することを示す。そのような「ている」と「である」文をいまかりに「一時的存在文」と呼ぶ[5]。二つの文の中国の訳語はいずれも「墙上挂着画」であるが、現在の日中市販翻訳ソフトを

調べると、そのような文に関する誤り(語順の問題や動詞の誤訳など)が多く見られる。中国語の文法によって例文(1)は「把字句」という文型であるが、例文(2)と(3)は「存在句」という文型であり、異なる構造の文として表現される。この翻訳処理方法については現在検討中である。

#### 5 おわり

日中両言語の存在文における異同について考察し、日中機械翻訳のために、日本語文の構文特徴、対応名詞の属性、構文構造などを利用して存在動詞の翻訳規則をまとめた。中国語の構文上の組み合わせの制限を総合して考察し、判定条件を機械で処理の可能な形で示した。更に例文を用いて手作業でこの規則を検証し、良好な評価を得た。これらの翻訳規則は我々の研究室で開発している日中機械翻訳システム Jaw/Chinese に組み込んで、翻訳実験を行っている。

今後例文を更に増やし、規則を調整し再評価していく予定である。

#### 参考文献

- [1] 飯田隆：存在と言語—存在文の意味論。URL, 2001年。
- [2] 西山佑司：日本語名詞句の意味論と語用論，ひつじ書房，2003。
- [3] 小池清治・小林賢次・細川英雄・山口佳也：日本語表現・文型記事，朝倉書店，2002年。
- [4] 池原悟他：日本語語彙大系-5 構文体系。岩波書店，1997年。
- [5] 金田一春彦・林 大・柴田武：日本語百科大事典。大修館書店，1988年。
- [6] 村上仁一：日英対訳データベースの状況。「言語、認識、表現」第7回年次研究会プログラム，2002年。
- [7] 劉月華・潘文娛・故韡：現代中国語文法総覧，くろしお出版，1996。
- [8] グループ・ジャマシイ：日本語文型辞典。くろしお書店，2001年。