

日中機械翻訳のためのとりたて詞の意味解析と課題

卜朝暉 池田尚志

岐阜大学工学研究科

1. はじめに

「も、さえ、でも…」などのとりたて詞は日本語の機能語の特殊な一族であり、意味上も構文上も複雑な特徴を持っている。言語上でそれに関する論述には文献[1][3][6][8]など数多くあるが、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、取立て表現に起因する誤訳が多く見られる。とりたて詞の意味解析が十分なされていないのが要因の一つだと思われる。本論文は、とりたて詞により取立てられる部分と述語部の統語的、意味的な属性と、その他のパターン特徴などにより、とりたて詞の意味解析を行った。さらに中国語側からも制限条件を作り、中訳語を特定するための翻訳規則を作成した。この翻訳規則を手作業で評価し、良好な結果を得た。またとりたて詞の意味解析を巡る問題についても分析を試みた。

本論では、とりたて詞の「も」、「さえ」、「でも」の3つを検討の対象とした。

2. とりたて詞の意味上の特徴

とりたて詞とは「文中のさまざまな要素(自者)をとりたて、これに対する他の要素(他者)との関係を示す」作用を持つ助詞である[3]。「他者」は常に暗示されていて、文の表に現れないことが多く、文脈や、社会知識などによって、推測される。両者の関係は場合によって、「他者への肯定、全面的な肯定(否定)、強調、制限、譲歩、意外性、例示」などさまざまである。

- ① 先生も今日のパーティに出席した。
- ② 戦争中16歳の子供も(でも)徴兵される。
- ③ ご飯さえ食べられれば満足だ。
- ④ 誰も私の話を聞いてくれない。
- ⑤ ここから駅まで30分もあれば足りる。

①は、「先生」以外の人もパーティに出席したことを暗示している。②は、16歳の子供が徴兵されることに対する「意外」の意味を表す。③は、満足の十分な条件は「ご飯を食べること」で、ほかはいらないという「最低限の条件」を示す。

とりたて詞のこの意味上の特徴は、「が、を」のような格助詞の「格関係の意味」とは異質のものであり、名詞や動詞などが表現する「明示的な物

事的な意味」とも異なる。暗示の意味を表すのが主役だと言えるであろう。

一つのとりたて詞は一般にいくつかの意味用法を持っている。例えば「も」は、「他者への肯定」①、「意外性」の意味②以外、「全面的な否定」④、「数量に対する評価」⑤)などの様々な意味用法を持っている。

とりたて詞のもう一つの特徴は、話し手がどのように状況を認識するか、人類の社会常識ではどのように理解されるのかによって、意味が異なってくることである。いわゆるとりたて詞の pragmatic な性質である。例②の「も」の「意外性」の意味を理解するには、「16歳の子供を兵士にするのは常理には合わない」という社会常識が必要である。

3. 取立て表現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、たいていとりたて詞と同じ意味を持つ表現があるが、直接的に対応する品詞は無く、副詞、介詞、連詞など様々な品詞を用いた表現に翻訳することになる。日本語で一つのとりたて詞が種々の意味を含むのと異なり、中国語では、意味が違えば、異なる訳語に対応する場合もあるし、また同様な意味も異なる訳語で表す場合もある。逆に日本語側で同じの意味を異なるとりたて詞で表現されるものがあるが、中国語では一つの訳語で対応できる場合もある。また、訳語が訳文の構造などにより変化する場合もある。訳語の訳文での位置もそれに修飾される成分により変化するが、日本語側の位置とは必ずしも一致していない。

従って、中訳語を特定するための意味解析は、中国語側の事情も考える必要がある。

4. 中国語への翻訳のための意味解析

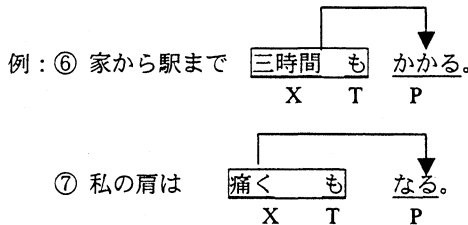
4.1 とりたて詞の意味分類

3つのとりたて詞の意味分類を中国語との対応関係を考慮に入れて、また文献[1]、[3]、[4]、[5]などに参考して表1に纏める。

4.2 とりたて詞の意味解析法

とりたて詞の意味について、寺村氏は「とりたて詞の意味は、係部分と、結び部分のそれぞれの文法的、意味的特徴と相関している」と指摘している[1]。我々もとりたて詞を含む大量の例文に対する観察により、たいていの意味用法はこの二つの要

素で区別できることを確認した。以下ではとりたて詞を T、係り部分(取り立てられる部分。単語、句、文など)を X、結び部分(述部)を P、また X の中訳語を X' とする。



取り立てられるもの(X)は⑥のように通常名詞あるいは名詞+格助詞であるが、用言を取立てる場合もある。その場合には、⑦のように「形容詞連用形+T+なる」や、「走ってもいけない」のように「動詞て形+T+いる」などの形をとる。この場合 X は用言の連用形或はて形であり、P は「なる/いる/する/ある」などの補助動詞となる。

しかし中訳語を特定するには、X、P の属性のみでは不十分で、中国語側の状況なども考えなければならない。我々は表2に示す5つの項目により、とりたて詞を中国語に翻訳するための意味解析を行った。表3に、これらの属性を用いた「さえ」と「も」の意味に関する解析規則を示す(表3の意味欄の数字は表1の意味分類の数字と対応する)。

属性についての記述は、中国語への訳し分けが可能な程度に分解されればよい。属性には意味素、品詞、活用形、肯(否)定式、ムート、成分及び文の構造特徴などが含まれる。

表2. 意味の判定条件の類別

1. X の属性
2. P の属性
3. 日本語側のパターン特徴
4. X' の属性
5. 中国語文の関係の構文特徴

5. 評価

『朝日新聞』の「天声人語」の約6万文を本研究室の解析システム「IBUKI」で解析した結果、「も」、「さえ」、「でも」を持つ文はそれぞれ7897、168、1039文が存在した。その中から、同形の他品詞と、慣用句を構成するとりたて詞を除外した後、各100文を抽出し、4節の意味解析規則を手作業で評価した。なお評価に使用した標本は、検討で使用したものと異なり、オープンテストである。またある市販翻訳ソフトの翻訳との比較も行った(表4)。評価は語順と分離し、訳語のみにより判断した。我々の結果は市販翻訳と比較して、正訳率は各々20%、77%、50%まざっており、本手法の有効性が確認できた。

表1 とりたて詞「も」、「さえ」、「でも」の意味分類

	意味分類	対応の中訳語	例文
さえ	1.条件複文の中で用いられ、X は主節の最低限の条件を示す。	①只要…就… ②只要…	①寝てさえいれば/寝さえすれば、風邪は治る。 ②あなたさえよければ、皆いいと言えるのか。
	2.X を極端の場合として際立てて、他場合も当然だという意外の意味。	①連/甚至…都/ 也… ②甚至…	①自分の名前さえ書けない。 ②太郎が授業中に寝てさえいた。
も	1.二つ以上の類似物を提示(列挙)する。	①和…都… ②既…也…	①私は金も暇もない。 ②彼女は親切でもあり、裕福でもある。
	2.数量に対する強調(評価) I 数量は予想より多い。 II 数量は予想より少ない。 III 数量は予想通りだ。	①競然有 ②競然 ③連…也/都 ④φ ⑤好	①今日五人も遅刻した。 ②家から駅まで三時間もかかる。 ③家から駅まで三分もかからない。 ④千円があれば足りる。 ⑤その日、何人も遅刻した。
	3.全面的な①肯定或は②否定の意味。	①都 ②也/都	①誰もが信じられる。 ②誰も信じられない。
	4.極端な例を挙げ、他を類推する。	連…也/都	その国で、十六歳の子供も徴兵される。
	5.他者に対する肯定の意味	也	私も東京へ行く。
	6.和らげの働き	φ/也	①速くも1年たった。 ②19世紀も初めの頃。
でも	1.X をP に適応するものとして提案し、それに限らないの意味。	一点儿/什么的/ φ(固定訳なし)	①お茶でも飲もうか。 ②風邪を引かせでもしたら大変だ。
	2.X の仮定の事柄を提示し、結果はそれと関わらない譲歩の意味。	即使…也…	①先生でもこの問題を解けない。 ②あいつは人を随れてでも自分の出世を考える。
	3.全面的に①肯定或は②否定(「どんな+名詞」の形だけ)の意味。	①都 ②也/都	①どこでも行ける。 ②どんな人でもできない。

表3 とりたて詞「さえ」と「も」の意味解析規則

意味分類	意味判定の条件とその属性					中訳語
	日本語パターン	X(X1, X2)の属性	P(P1)の属性	X'(X1, X2)の属性	中国語文の構文特徴	
さえ	1 X さえ P1...P2		仮定形		B 主節が反問文 或は「是…」文	A: 只要…就… B: 只要…
	2 X さえ P		仮定形以外	A: 主語/賓語/状語 B: 謂語		A: 連/甚至…都/ 也… B: 甚至…
も	1 ①X1もX2もP ②X1でもありX2でもある ③X1もX2もする	①名詞句の主語/目的語/状語 ②名詞句/形容動詞語幹 ③サ変動詞語幹		A: 主語/賓語/状語 B: 謂語		A: …和…都… B: 既…又…
	2 I XもP II XもP III XもP1…P2	普通数量詞/疑問数量詞	肯定	A: 普通数量詞の主語 B: 普通数量詞の賓語 C: 疑問数量詞		A: 竟然有 B: 竟然 C: 好
		普通数量詞/疑問数量詞/最小数量詞/最小数量副詞	否定			連…也/都…
		普通数量詞/疑問数量詞/最小数量詞/数量副詞	仮定形			φ
	3 ①XもP ②XもP	疑問詞/疑問詞+格助詞	肯定(「…しかない」も含む)			都
		疑問詞/疑問詞+格助詞	否定			也/都
	4 Xもしない	動詞の連用形	否定			連…也/都…
4/5/6 (※問題点2)					「也」をとる	

表4 評価 (A: 我々の訳 B: 市販ソフトの訳)

取立詞	例文数	正訳数		正訳率	
		A	B	A	B
も	100	98	78	98%	78%
さえ	100	99	20	99%	20%
でも	100	95	45	95%	45%

6. とりたて詞の意味分析の問題点

評価の問題点と関連して、言語と機械翻訳の角度からとりたて詞の意味分析の困難さを考察する。

(1) とりたて詞と同形の他品詞との区別がはっきりとしない場合がある。

「でも」には、次のような同形の他品詞がある。

- (a) 接続助詞の「でも」
- (b) 場所/道具格助詞「で」+とりたて詞「も」
- (c) 助動詞「だ」の連用形「で」+とりたて詞「も」

- ⑧ 日曜日でも出勤しなければならない。
- ⑨ そんなことは子供でも知っている。
- ⑩ レーガン総統の構想については、米国内でも強い批判がある。
- ⑪ 宴会の席でも、いい考えが浮かぶと急に立ち上がって消えてしまうことがあった。
- ⑫ 彼はダンサーであると同時に俳優でもある。
- ⑧と⑨は(a)と関わる問題である。どちらが接続詞

かどれがとりたて詞か言語上でも究明できていないが、機械翻訳では、意味も中訳語も同様であるため、区別して処理する必要がない。

⑩と⑪は(b)と関わる問題である。表現の形のみから見ると、2例とも『「意外性」のとりたて詞の「でも』と、『場所格「で」+「も』の解釈があり得る。人間なら、社会知識から⑩を「意外性」のとりたて詞の意味と知れるが、機械には困難である。⑩は人間にも二つの解釈があり得る。

⑫の例は(c)の場合である。『助動詞「で」+とりたて詞「も』の用法は、常に「ある/ない」が後続しているため、「意外性」のとりたて詞「でも」との区別が明白である。ただし「ある/ない」が後続しても、従属節の中に含まれていれば、「提案」の意味のとりたて詞の「でも」にもなり得るため、また曖昧性が残っている。

5節の評価では、(b)の場合の用法ととりたて詞の「でも」との区別は扱っていない。区別ができたことを仮定し、表1のとりたて詞に属する三つの意味用法の評価のみを行った。表5に100文のとりたて詞を抽出する際に行った「でも」に関する品詞別の比率を示した。

表5 とりたて詞「でも」と他品詞「でも」の統計

総文数	T	(b)	両方	慣用句	(c)
141	100	30	16	10	1

※表5で、「両方」は(a)と⑩例のような曖昧性が保留で

きる場合。(b)(c)は、各々『場所/道具格助詞「で」+とりたて詞「も』』と、『助動詞「だ」の連用形「で」+とりたて詞「も』』の意味用法しか取れない場合。

表5の統計によれば、他品詞の用法も含めた場合「でも」の正訳率は約79%に下がる。

(2) とりたて詞の pragmatic な性質のため、文脈や状況、社会知識などに頼らなければ意味を特定できない場合がある。

⑬ その国では、7歳の子供も入学できる。

⑭ その国では、4歳の子供も入学できる。

人間は社会知識によって、⑬の「意外性」を感じ、⑭の「他者への肯定」の意味と区別して理解できる。コンピューターには、構造も前後起語の属性も同じである⑬と⑭を区別するのは困難である。

4節の規則では、頻度が遥かに高い「他者への肯定」の意味という解釈をとるように扱っている。5節の評価で誤った例は、ともにこの区分に関するものである。

ただし、X=動詞の場合には、Pが否定の形なら(Xもしない)「意外性」の用法、Pが肯定の形なら(Xもする)、「他者への肯定」の意味であると推定できる。また、中国語でこの意味の訳語の「也」は、ある場合には「意外性」の意味も体現できるため、そのような場合には訳語の曖昧性が保留できる。

また「も」の4番と6番の意味の区別も困難であり、我々の規則ではやはり頻度が大きい4番の解釈をとるようにしている。ただし6番の中訳語の一部は4番と同じの「也」になっていることもあって、評価では、誤った例は出現しなかった。

「でも」の場合、1と2の意味を誤ったのが1件あった。

(3) 省略された表現の場合、判定条件が適用できない場合がある。

⑮ 回答は「美味しいサンドイッチでも」だったが、それを上回る接待ぶりだった

4節の評価で、誤訳した「でも」の3例は⑮の例と同じく述語が省略されたあるいは述部に省略があるものであった。

(4) その他。

「提案」の意味の「でも」は、不定代詞「什么的」、不定数量詞の「一点見」及び「φ」などに訳しうる。しかしながら、どのような条件で訳し分ければよいか特定しにくい。

⑯ こんな暑いときに、アイスクリームでも食べられればいいなあ。

現在訳：这样热的时候，能吃上冰淇淋就好了。

完全訳：这样热的时候，能吃上冰淇淋什么的就好了。

⑰ 仕方がない、弟とでも行くとするか。
(没办法，还是和弟弟去吧。)

⑱ まるで野中の一つ家に住んでもいるような隣近所に少しの遠慮もない。
(好象住在荒野似的一点也不用顾虑周围的人。)

現在の処理ではすべて「φ」に訳しており、⑯の完全訳のようなニュアンスは表せない。しかし、⑰や⑱の例では、「φ」に訳すのは適当である。なぜならば、⑱は提案されている「アイスクリーム」と関係の他要素は思いつくものであり、⑰はその他要素は実には存在しないためである。

⑱のように、「でも」を含む節の述語が「よくな、みたい」などの比況助動詞がついている場合は、「φ」に訳すのも正確である。

(5) とりたて詞は慣用句を構成する場合が多いため、慣用句の対訳辞書が必要である。

7. 終わり

とりたて詞の意味解析について、言語学では従来その取り立てられる部分と述語部の統語的、意味的な特徴という2要素から議論されてきた。本研究では、日中機械翻訳の角度から両言語を分析し、5要素まで拡大して、中訳語を特定する翻訳規則を提案した。

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。また、とりたて詞の意味解析を巡る課題についても検討した。

意味用法が複雑で、文法の用法と語彙的な意味を両方兼ねて持つとりたて詞の機械翻訳に、一つの手がかりを示したと考える。

しかし今回の規則は、他品詞との区別は扱っていないし、評価の例文も十分ではない。今後例文を更に増やし、他品詞の意味用法との区別も考慮して、規則を調整し再評価することが課題である。

参考文献：

- [1] 『日本語のシンタクスと意味』 寺村秀夫黒潮出版 1991年
- [2] 「日中機械翻訳における取立て表現の曖昧さ解消」
ト朝暉、謝軍、池田尚志 情報処理学会 63回大会論文集
- [3] 『時・否定と取り立て』 沼田善子ほか 岩波書店 2000年
- [4] 『日本文法大辞典』 松村明編 明治書院 昭和43年
- [5] 『標準日語語法』 顧明耀主編 中国高等教育出版社 1997
- [6] 『副助詞「も」と副詞「也」など』 大河内康憲 『日本語と中国語の対象研究』 第2号 1997.3
- [7] 『現代漢語八百詞』 呂淑湘 主編 商務印書館 1980年
- [8] 「も」、「でも」、「さえ」の含意について 山中美恵子 『日語と中国語の対照研究本語と中国語の対照』 第14号