英語の認識構造を介した日中間用例翻訳方式

4L－11

周 強　　宮崎 正弘

新潟大学大学院自然科学研究科

1　はじめに

基軸言語を介して多言語間での用例ベース翻訳 [1] を行なう場合、２つの目的言語に対応する別々の基軸言語表現の類似性を抽出することは、この方式の基本となるものである。しかし、基軸言語である異なった２つの言語表現に対して、単に表層的な情報、統語情報などから、類似性を判定するだけでは不十分で、言語表現の意味を反映させた認識構造を介して類似部分を判断する必要がある。そこで、本稿では、英語を基軸言語として、目的言語との対訳用例ベースから得られた２つの英語表現の認識構造を抽出し、これを手掛かり、類似性を抽出することによって、対応する２つの基軸言語表現の関係をとる方法について述べる。そして、日本語－英語間用例ベースおよび英語－中国語間用例ベースを利用するより、日本語－中国語間での用例ベース翻訳を実現する可能性を論じる。

2　英語表現の類似性のとり方

本稿では、三浦つとむの言語過程説 [2] と宮下眞二の英語文法にこの説を適用した考え方 [3] に基づいて、ＳＧＬＲパーザを用いて抽出された英語の対象認識構造 [4] を利用し、基軸言語である２つの英語表現の認識構造を構築する。そして、この２つの認識構造において、主体表現、客体表現および主体＋客体表現などの３つの観点から、単語、句、単文レベルで類似する部分を抽出することによって、２つの英語表現の対応関係をとる（図１）。

Example-Based Japanese-Chinese MT System through the Speaker's Recognition of English
Qiang Zhou
Masahiro Miyazaki
Niigata University

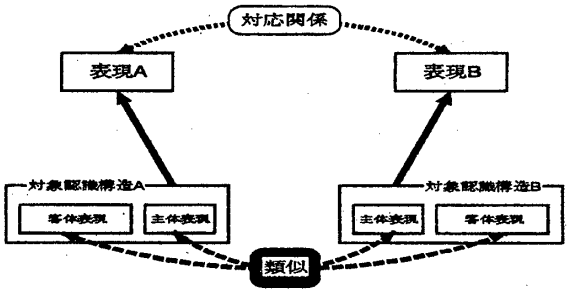

図１　認識構造から言語表現の類似性の抽出

2.1　主体表現からの類似性の判定

主体表現は、話者の感情、欲求、意志、判断などを直接表現したものである。これに対応する言語表現の認識構造を抽出すれば、異なった２つの認識構造における主体表現の類似性をとることによって、対応する言語表現の関係を得られる。例えば、

[E1] Could you join us ?

[E2] Would you join us ?

図２を示すように、認識構造を調べると、対応する２つの主体表現は、[E1] の方が [E2] より'丁寧'な言い方であるが、認識構造としては同じ構造'意欲'を持つ。よって、２つの認識構造を類似と判断し、対応する２つの言語表現 [E1] と [E2] との対応をとることができる。

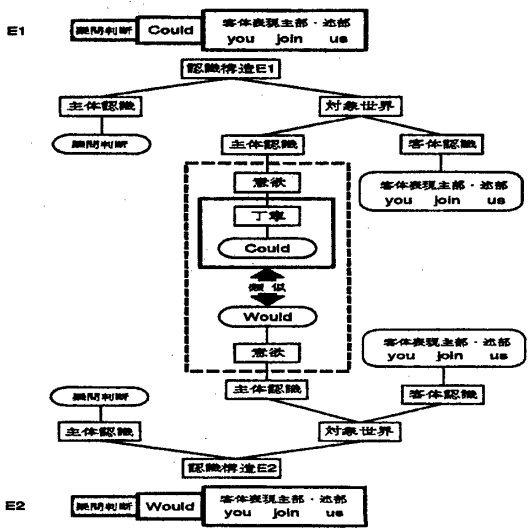

図２　主体表現から類似性の抽出

2.2 客体表現からの類似性の判定

客体表現は話者が対象を概念化して捉える表現である。認識構造の客体認識において、実体、属性およびその概念的な関係などの類似性をとれば、対応する2つの言語表現の類似性を判定することができる。例えば、

[E3] I wrote him a letter .

[E4] I wrote a letter to him .

図3のように、この2つ表現の認識構造の中で、客体表現の与格対象の部分は'実体'と'前置詞句'のように表層的、統語的には異なるが、意味的には'前置詞句'がこの'実体 (him)'への方向性を示している点が異なるのみで、2つの認識構造が類似であると判断し、対応する2つの表現 [E3] と [E4] との対応がとれる。

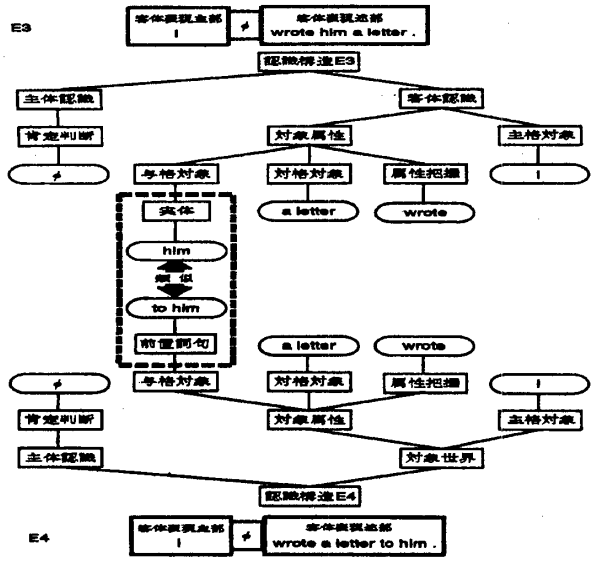

図3　客体表現から類似性の抽出

2.3 主体+客体表現からの類似性の判定

英語表現の認識構造において、ある主体表現Aが客体表現B+主体表現Bの組合せから構成されていることがしばしばある。このような表現の類似性を判定するには、認識構造上で対応する主体表現と主体+客体表現との類似性を判定することによって行なう。例えば、

[E5] I may have one .

[E6] It is possible that I have one .

図4のような認識構造で、言語表現'E5'の'may'は主体表現であり、'I have one'の'可能性'と認識されるが、言語表現'E6'の'It 'と'that'とも認識的に、'I have one'という対象を示して'可能性(possible)'を意味するから, この2つ認識構造は類似であるとを判断し、対応する2つの言語表現'E5'と'E6'の類似性を判定することができる。

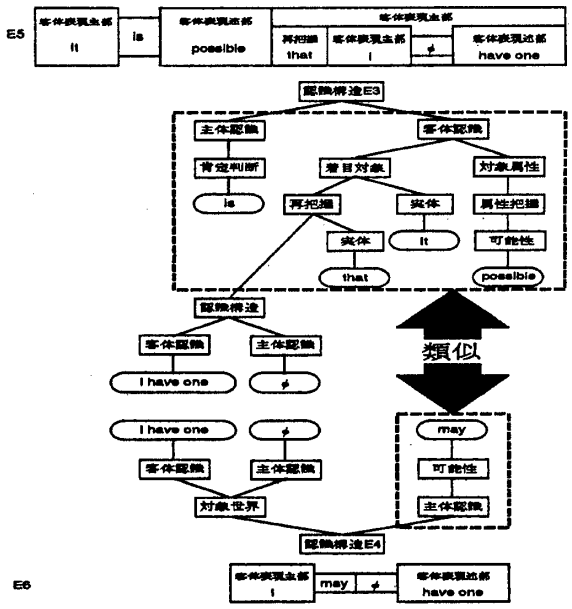

図4　主体表現+客体表現から類似性の抽出

3 おわりに

本稿では、2つの目的言語に対応して、基軸言語である2つの英語表現が存在する場合、2つの英語表現の認識構造から、主体表現と客体表現の類似性を抽出することによって、2つ英語表現の類似性を判定する。そして、この関係を利用し、用例ベースから2つの目的言語間の表現の対応をとることにより、日一中間での用例ベース翻訳が可能となることを論じた。今後、対象の認識過程に基づいて、英語の認識構造を深く検討する必要がある。

参考文献

[1] 周 強・宮崎 正弘：基軸言語を介した多言語間用例ベース翻訳方式，第 *52* 回情報処理学会全国大会 *N0.6B-1(1996)*

[2] 三浦 つとむ：日本語とはどういう言語か，講談社学術文庫（*1976*）

[3] 宮下 眞二：英語はどういう言語か，季節社（*1985*）

[4] 五百川 明・沼崎 浩明・宮崎 正弘：話者の対象認識構造を分析する英語文パーザの基本的枠組，第 *49* 回情報処理学会全国大会 *NO.2G-9(1994)*