

産業日本語シンポジウム
2010年2月24日

産業日本語を育てよう

国立国会図書館長
長尾真

機械翻訳の必要性

- 今日、グローバル世界で活動するにしたがって、言語の翻訳の仕事が急激に増加している。
- 人手による翻訳は時間がかかり、コストが高く、言語によっては翻訳者の数も限られる。
- 機械翻訳システムを導入し、機械翻訳の結果を人手で後修正することが行われてきた。

機械翻訳の質

- 機械翻訳の質は原文テキストの分野、書き方によって大いに異なる。
- 翻訳の質としては、「文章のあちこちに誤りが見られるが、ていねいに読めば内容が理解できる」というのが、最低の受入れ可能なものと考えられる。
- 機器の取扱い説明書のように、無機質な文章の場合は比較的翻訳率は良い。
(70～80%)

システム設計における条件

- どのような工学システムや機械も設計に際して明確にすべきことがある。
- その第1は使用に際しての規定、範囲、条件である。
- その第2はこの条件の下での作動結果の性能である。

機械翻訳システムにおける 条件設定の難しさ

- どのような文章を入力として受付けるか。
- どのような翻訳結果を合格とするか。
- これらは読む人の持つ知識にもよる。
- 現在の機械翻訳システムにおいては、特に入力文の構造の複雑さが最大の問題である。

条件設定で注意すべきこと

(辻井潤一氏の指摘)

- 背景知識を持たない人でも理解できる文
- 解釈が人によって変わらない文
- そのためにはすぐれた(専門用語)辞書が必要

制限日本語

- 日本語の文章表現に制限をつけて、できるだけ機械翻訳システムで解析が正しく行われ、翻訳がスムーズに行えるようにする。
- そのような制限言語のあり方については、1983年から84年にかけて3つの論文を情報処理学会の研究會に発表している(長尾ほか)。

制限言語のための 文章作成援助システム

- これは1984年に発表した制限言語のための一種のワープロシステムである。
- 1つの文の長さが一定の数を超えると警告を発する。
- 「AのBとC」という表現が作られたら「Aの(BとC)」か「(AのB)とC」かを聞いて()でくることを行う。
- 主語省略等のある時、警告を発する。
- 単語に幾つかの意味があった時、それらを示し、選択をさせる。

など

産業日本語

- 産業分野で使われる文章を制限日本語という考え方で書く。
- 制限の条件を列挙する。これらはできるだけ単純で明確でなければならない。
- 表現を制限することのほかに、システムが正しく構文解析できるように「Aの(BとC)」のように補助的記号を導入することも考える。

産業日本語ワープロシステムの開発

- 制限を守らない書き方について、on-line、real-timeに警告を出し、書き手に修正させる。
- 文を書いて行くにしたがって、並行して構文解析もできるだけ行い、構文的あいまいさが出る場合には、どの解析が正しいかを書き手に決めさせる機能をもたせる。

産業日本語のための辞書

- 多義の単語や句について充実した辞書を作り、そのような表現が使われた時に、どの意味で使ったかを書き手に決めさせる。
- 産業の各分野で使われる専門用語は徹底的に収集し、文の解析、翻訳が困難に陥らないようにする。

産業日本語、制限言語で 注意すべきこと

- 文章を書く人に大きな抵抗感を与えないような制限条件でなければならない。
- 産業日本語ワープロを作り、多くの人に使ってもらいながら、抵抗感が少ない制限の仕方などを追及する。
- 制限条件が翻訳の相手言語によって変えねばならないという状況がありうることも考えておく。