

Multilingual Generation and Japanese-European Machine Translation

Graham Wilcock

Sharp Corporation (Japan) and University of Manchester (England)

Abstract

In the design of effective machine translation systems the direction of translation is fundamental. “Import” translation from foreign languages into the user’s language should be contrasted with “export” translation from the user’s language into foreign languages. The structural transfer approach used in most current MT systems is very effective for import translation. But it is less effective for export translation, which is extremely dependent on the human post-editor.

An alternative approach to export translation ([Wilcock 91]) uses a source language functional interface to a target language systemic grammar for generation. By helping the source language user to make choices in target language systems, the extreme dependence on the post-editor is reduced. Further developments of this approach are being investigated, including incremental refinement of the systemic choices for “forward generation” of the target text. Other desirable developments would be confirmation of the choices by “back generation” into the user language, and checking for ambiguity in the target text by “back translation”.

In multilingual translation systems, structural transfer rules are multiplied, so an interlingual or pivot approach is more attractive. There are practical reasons for using English as a pivot in machine translation between Japanese and the languages of the European Community. An internal representation of the English text (similar to the Interface Structure in Eurotra) might be sufficient as a pivot for import translation based on structural transfer.

But for Japanese-European export translation, the dependence on post-editors is an especially difficult problem, due to the severe shortage of people with knowledge of both Japanese and the various EC languages. Therefore the alternative approach, based on a functional interface to a systemic grammar for generation, is particularly desirable to reduce the post-editing load.

In a multilingual application of this approach, all the EC target languages would share a common systemic grammar. Most of the functional systems would be shared, but some would be unique to specific languages, as described in [Bateman et al 91]. The realization rules need not be shared. The Japanese user would make functional choices in the target systems, using a Japanese interface, just as in [Wilcock 91].

English can be used in the role of a pivot language in this approach, during “forward generation” of the target texts. Only the English version would be displayed to the Japanese user during interactive incremental refinement. The systemic choices made for English would be re-used in the generation of the other EC languages, with relatively few further choices specific to the other languages.

Further research is required into the feasibility of this approach, initially with a small subset of EC languages.

References

- [Bateman et al 91] Bateman, J.A., Matthiessen, C.M.I.M., Nanri, K. and Zeng, L.: Multilingual text generation: an architecture based on functional typology. *International Conference on Current Issues in Computational Linguistics*, Penang, Malaysia (1991)
- [Wilcock 91] Wilcock, G.: An experimental Japanese interface to a systemic grammar for generation of English auxiliary verbs, 18th International Systemic Congress, Tokyo 1991.