ITS-2 : an interactive personal translation system

Eric Wehrli and Mira Ramluckun
Dept. of Linguistics – LATL
University of Geneva
1211 Geneva 4

email: wehrli@uni2a.unige.ch

ITS-2 is an interactive sentence translation system under development at the LATL lab of the University of Geneva. In its current status, ITS-2 translates French sentences into English or English into French over a still restricted vocabulary of approximately 3'000 bilingual entries. The main objectives of this project are (i) to show some of the advantages of interactive approaches to NL translation, (ii) to demonstrate the merits of generative grammar as a syntactic model for MT, (iii) to show the feasibility of personal translation on small personal computers (under MS-Windows).

ITS-2 is based on the familiar transfer architecture, with its three main components, parser, transfer and generation. The parser — which is the IPS parser described in Wehrli (1992)— associates with an input sentence a set of syntactic structures corresponding to GB S-structures, i.e. surface structures enriched with traces of moved elements and other empty categories. The role of the transfer component is to map source structures onto target structures. Transfer, which occurs at the D-structure level, is to a large extent a matter of lexical correspondence. For each lexical head of a SL structure, the lexical transfer component consults the bilingual lexicon to retrieve the most appropriate TL item, which is then projected according to the X-bar specifications of the TL. Applied recursively over the whole SL D-structure, this process determines an equivalent TL D-structure. From these structures, the generation component derives well-formed S-structures, which are finally converted into the target sentence by morphological process.

The current demonstration version of ITS-2 runs under MS-Windows. Integrated within a small editing environment, it translates in real time a wide selection of sentences (French to English, or English to French) over a large range of grammatical constructions including simple and complex declaratives, interrogatives, relatives, passives, cliticization, some cases of coordination, etc.

The system is interactive in the sense that it can request on-line information from the user. Typically, interaction takes the form of clarification dialogues or selection windows. Interaction can occur at several levels of the translation process. First, at the lexicographic level, if an input sentence contains unknown words or typos. In such cases, the user is asked to correct or modify the sentence. At the syntactic level, interaction occurs when the parser faces difficult cases of ambiguity, such as, for instance, when the resolution of an ambiguity depends on contextual or extra-linguistic knowledge, as the case of some prepositional phrase attachments or coordinate structures.

By far, the most frequent cases of interaction occur during lexical transfer, due to the fact that lexical correspondences are generally of the many-to-many variety, even at the abstract level of lexemes. It is also at this level that our decision (not yet implemented) to try to restrict dialogues to the source language is the most challenging. While some cases of polysemy can be disambiguated relatively easily for instance on the basis of SL gender distinction, as in (1), other cases such as the (much simplified) ones in (2)-(3) are obviously much harder to handle, unless additional information is included in the bilingual dictionary.

(1)a. Jean regarde les voiles.
'Jean is looking at the sails/veils'
(1)b. voiles:
     -masculin (le voile)
     -féminin (la voile)

(2)a. Jean n'aime pas les avocats.
'Jean doesn't like lawyers/avocadoes'
(2)b. avocats:
     -homme de loi
     -fruit

Another common case of interaction that occurs during transfer concerns the interpretation of pronouns, or rather the determination of their antecedent. In an sentence such as (3), the possessive son could refer either to Jean, to Marie or (less likely) to some other person, depending on contexts.

(3) Jean dit à Marie que son livre se vend bien.
'Jean told Marie that his/her book is selling well'

In such a case, a dialogue box specifying all possible (SL) antecedents is presented to the user, who can select the most appropriate one(s).

In future work we intend to restrict the clarification dialogue to SL, to make ITS-2 available to monolingual users.

REFERENCES

Wehrli, E. 1992. "The IPS system". COLING92.