

# The prerequisites of MT for an industrial user

## A corpus-based approach

Sylvie REGNIER-PROST  
Eva DAUPHIN  
AEROSPATIALE - Louis-Blériot Research Centre  
Suresnes • France

Since the beginning of the 1980's, the Information Department of the Common Research Centre of AEROSPATIALE has been involved in the experimentation and evaluation of Natural Language Processing tools, ranging from terminology extraction and management tools to translation memories and Machine Translation systems. Through its experience in comparative testing and diagnostic evaluation of the main existing machine translation systems, it has developed two user-oriented evaluation methods.

The first evaluation method used at AEROSPATIALE was essentially based on error identification and classification, representative of the traditional approach to Machine Translation, i.e. the assessment of general purpose MT systems in a black box situation. The results of this first type of evaluation highlighted a number of limits. In such assessment environment, the user cannot evaluate the characteristics of the heart of the system and its components. Neither the long term evolution capabilities and limitations of the system nor its linguistic reliability and viability can be evaluated by the user in a black box situation.

The second evaluation method AEROSPATIALE has developed represents a new, user-oriented approach to Machine Translation, where users are directly involved in the specifications of the system and in its tailoring to industrial requirements and can work in close cooperation with developers.

The main evaluation resources used in this method are corpora and application-oriented test suites. The major evaluation criteria are the availability of re-usable and standardised lexical, terminological and grammatical resources, the availability of assistance tools for the translators as well as the capacities of the system to be integrated in the industrial environment and in the translator's workbench.

The characterisation, by the user, of the corpora to be translated is an essential part of the whole evaluation process. Monolingual and bilingual corpora studies aim at describing three kinds of text-specific criteria, i.e. the logical structure of the corpus (e.g. SGML tags), its syntactical characteristics and lexical and terminological specificities. The use and/or definition of various types of controls on the source texts may be an additional evaluation criterion.

Designing application-specific test suites is another important constituent of this evaluation method. Test items built by users have to be both representative of the most frequent linguistic and extra-linguistic phenomena identified in corpora and of the problematic phenomena which are not properly processed by any existing NLP system. Such test suites can thus be used for both adequacy and progress evaluation and can be re-used to test other types of NLP systems.

In addition to these mostly linguistic evaluation criteria, it is necessary to assess the performance of peripheral assistance tools which guarantee the maintainability of the system in the user's specific context. Such tools may include translation alignment and memories; terminology databases; post-editing and revision assistance tools; computer-assisted generation of bilingual standardised repetitive sequences; evaluation and diagnostic assistance tools and maintenance and upgrading tools.

The last but not least evaluation criterion is the availability of re-usable and standardised resources and the possibility for the user to have a user-friendly access to them and also to upgrade and maintain them.

Since the design of such standardised resources is a highly time and cost consuming process which results are crucial for the actual evaluation and use of MT systems and NLP applications, in general, AEROSPATIALE has decided to participate to their development in the framework of European projects.