

TACnology Corner

A TERMINAL CASE!

The field of translation and computers, like many of today's dynamic and burgeoning areas of development, has begotten a terminology that is ambiguous, overlapping, and imprecise. At the 1986 conference in Cleveland the participants in the TAC wrap-up session asked that the TAC Committee help to bring some uniformity into the situation by proposing a set of standard terms for use by ATA.

The miniglossary that follows is the result of efforts undertaken over the past year. It has been prepared in accordance with the principles recommended by the International Standards Organization. The systematic approach to organization and definition of concepts is designed to eliminate overlap and ambiguity.

It will be noted that the catch-all computer-aided translation falls by the wayside because its prevailing usages violate the principle of "one term, one concept." Members are urged to begin to phase it out.

1. machine-aided translation (MAT). Translation generated by the human translator with the aid of word processing, as a minimum, and usually additional computer support, particularly lexical data bases (q.v.).

Synonym: machine-assisted translation, machine-assisted human translation (MAHT).

Deprecated: computer-aided/assisted translation (CAT). This term should be avoided because of its ambiguous application, since it has been used to mean, variously, machine-aided translation, machine translation, and both. It also has a proprietary association and may therefore be interpreted as favoring a particular software product.

1.1.1. word processing (WP). Computer software that facilitates the

preparation and editing of running text. It offers, as a minimum, the capability of easily manipulating segments of text by means of insertion, deletion, copying, and moving.

1.1.2. lexical data base (LDB). A computer file of terminology records, often organized systematically by subject area. Each record is organized into a set of data elements, which may include: the term, its definition, foreign language gloss(es), source(s), synonym(s), example(s) of context, subject classification code, reliability code, scope note(s), date of entry, identification number.

Synonym: term bank

2. machine translation (MT). Translation generated by a computer, with or without human intervention.

Synonym: automatic translation, mechanical translation.

Deprecated: machine-aided/assisted translation (reserved for the use identified above), computer-aided/assisted translation (CAT).

2.1. Types of human intervention

2.1.1 pre-editing. Human intervention in which the input text (q.v.) is specially adapted to the MT process. Pre-editing is often considered to include the use of customized language, in which vocabulary and/or syntactic structures are constrained at the time the original text is written.

2.1.2 interactive editing. Intervention by the translator at decision points which affect the outcome of the computer-generated translation (see 2.2 below).

2.1.3 postediting. Modification of computer-generated output, usually by a professional translator who has been specially trained in the appropriate techniques, with a view to improving the quality of the text. The term implies tht the translation is sufficiently polished to meet the purpose for which it was requested. For information-only translation, a large, sophisticated MT system may produce output that requires only a small amount of postediting or none at all,

and the process can be partially or fully automated.

2.1.4 foreign-word replacement. A level of intervention at which the translator provides glosses for not-found words but does not attempt to improve the quality of the rest of the text.

2.2 Forms of text

2.2.1 input text. The text to be translated, which must be in machine-readable form, i.e., the characters must have been converted into digitized representation and recorded on a medium (usually magnetic tape or disk) that can be interpreted by the computer. Text in the form of hard copy can be made machine-readable by keyboarding it manually into a word processor; it can also sometimes be converted automatically by means of optical character recognition (OCR).

2.2.2 raw (machine) output. The translation as it is generated by the computer, without any human intervention. It is most often displayed on a word-processing screen for postediting (q.v.) prior to being printed as hard copy.

2.3 Modes of operation

2.3.1 interactive. The translation process stops to await intervention by the human translator; sentences are interrupted at decision points that would affect the outcome of the translation.

2.3.2 batch. The text is submitted to and returned by the computer as a single job stream. There is no possibility for human intervention in the course of processing by the computer.

Comments on the definitions above are most welcome. They may be directed to the TAC Committee Chair.

Muriel Vasconcellos
Chair, Committee on Translation and Computers
1739-1/2 Corcoran Street, N.W.
Washington, DC 20009