STYLISTIC CONSIDERATIONS OF
MACHINE-ASSISTED AND HUMAN TRANSLATION

by J. Richard Ruffino

Today's machine-assisted translation (MAT) systems have reached a high level of sophistication and are having a powerful impact on the world of translation as an art and as a business. It is of course fitting that in this computer age the magical machines composed of bits, bytes and "buts" have infiltrated the world of language translation. Although MAT has been with us for a number of years, it is with the advent of the age of computer mania that more systems have come on the scene and personal computers will soon be making help in language translation a commonplace phenomenon. Please note that I insist upon the use of the word "assisted" and will only drop it from the definition of a system when its output no longer requires post-editing.

The most fascinating aspect of this technology lies not in what the computer can do as a translation tool but in what it cannot do! There is no great mystery involved in the programs and dictionaries that in combination can mimic the structure of a sentence in a foreign language. The computer can place subject, verb and object in correct order by managing to check the correct tables at the correct place through the correct bytes, which are most obligingly turned on or off by recognition of the combination of letters that form a word in English or another language. What it can't do, however, is create a sentence from the source language (SL) that truly has the style and feel of a human translation. A simple directive sentence or phrase may or may not elude the computer stylistically or syntactically, depending on the smallest factor. The more idiomatic and complex sentence, on the other hand, will almost always baffle the computer technically and always stylistically.

Recently I used ALPS and Systran MAT systems to translate a series of sentences which I composed for the experiment this paper describes. Those familiar with MAT have certainly heard of ALPS (Automated Language Processing Systems), a language translation

system based in Provo, Utah. Systran, now based in *La Jolla, California,* is the "grandaddy" of MAT systems, with a history going back to the 1960's and considered a descendant of the Georgetown system of the 1980s

Besides translating these sentences using Systran and ALPS, I also asked a colleague of mine, a talented and experienced translator and a native speaker of Spanish, to translate the same sentences as she would normally do. She was not exposed to the computer outputs.

The ALPS system is interactive, which means that the computer asks the user a series of questions about ambiguous areas in analyzing a given sentence. These questions might include choices about homograph resolution, problems of word enumeration, choice of preposition translation, and so forth. So as not to be influenced in my selection of any options presented by the ALPS system, I did not read the results of the manual translation until I had completed the ALPS interactions. In addition, I did not utilize the word-processing feature of ALPS to alter or correct any of the translations, as this would be tantamount to post-editing and the results would no longer be valid for comparison.

We expect any system that attempts to translate to commit blatant errors in analyzing the source text and target-language (TL) rearrangement. But what about those situations where the translation is not blatantly wrong, but is still unacceptable from an esthetic point of view? (I hope there is no objection to my use of the word "aesthetic" in this context, but I truly feel that aesthetics are a part of language and translation.) Here's a simple sentence hat may help demonstrate my point:

MODEL SENTENCE 1

Not only products, but service will be included.

TRANSLATIONS:

ALPS: No solamente productos, pero el servi-
         cio se incluirán.

SYSTRAN: No solamente los productos, pero el
         servicio serán incluídos.

HUMAN: Se incluiran no sólo productos sino
         también servicio.

Neither the ALPS nor the Systran translations contain errors of any great proportion, but the translations are awkward. The mistaken translation of "but" as pero instead of sino in both cases is not an error which leaves the sentences unintelligible, In both the ALPS and the Systran versions, however, the failure to rearrange the sentence as the human translator did and to use the extra word combination of sino también as in the human translation give the translation the feeling to which I am referring. In addition, both words, solo and solamente, are valid translations for the word "only." Computer dictionaries do not generally allow for variant equivalents of common words, unless the choice is added to the ALPS interactive selections or special parameters are identified and a unique entry is made to the Systran dictionary. This is not to say that the word solamente does not fit, but <u>solo</u> is more in tune with the sentence as it has been translated by the human translator

The style in the sentence produced by the human translator is a "total package," which conveys a true feeling of naturalness. As a result, the human translation is the most "aesthetically pleasing" because it is not artificial.

At this point, it is necessary to add that there is no ideal human translation of any sentence or phrase, nor is such an idea suggested as the premise of this paper. Rather, the human translations provide a model here and are considered acceptable and stylistically natural translations of the English sentences used in this study. As well, the English sentences have not been contrived for use in this comparison. They are commonly used constructions in technical writing. Only the terminology has been "generalized" to avoid identification with any specific industry.

MODEL SENTENCE 2

A repair is defined to be any corrective action.

TRANSLATIONS:

ALPS: Una reparacion se define a ser cual-
         quiera accion correctiva.

SYSTRAN: Se define una reparacion para ser cual-
         quier accion correctiva.

HUMAN: Cualquier accion correctiva se consi-
         dera una reparacion.

"TO BE + DEFINED TO BE" is a very common writing convention among technical writers. Unfortunately, its correct translation eludes the computer. Like so many stylistic conventions in English, this phrase is often used simply because it seems to add an air of "sophistication" to a technical text. Basically, this phrase can be translated by the verb "to be" alone, but the human translator correspondingly picked up on the slight affectation of the SL text and used <u>se considera</u> as a stylistically valid equivalent translation. Both Systran and ALPS present us with literal translations of this phrase, which with minor post-editing give an intelligible yet not stylistically acceptable translation. One of the problems with the average technical writer's input is just such stylistic conventions. MAT has its greatest application for the translation of technical texts, but we cannot count on the absence of "stylistic twists," as evidenced in this sentence.

MODEL SENTENCE 3

Many different procedures generate problems in the carriage.

TRANSLATIONS:

ALPS: Muchos procedimientos diferentes gene-
         ran problemas en el carro.

SYSTRAN: Muchos procedimientos diferentes gene-
         ran los problemas en el carro.

HUMAN: Una **serie de** diferentes procedimientos
         crean problemas en el carro.

In this case, the literal translations produced by ALPS and Systran are valid and acceptable. It is most likely that a post-editor of a technical text containing this sentence would not take the time to rewrite it in the way that the human translator rendered it, although the editor would probably make minor changes, perhaps with articles or adjective placement. But it is my contention that the machine output, although acceptable, lacks "naturalness." For practical purposes of analyzing style, a sentence can be assessed in terms of its degree of intelligibility, acceptability and naturalness. Let's look at a brief example in English which illustrates what I mean by these terms.

"I'm happy that John finally got his Ph.D.

No, he didn't get his Ph.D.

John is a

1. medicine doctor." (INTELLIGIBLE)
2. doctor of medicine." (ACCEPTABLE)
3. medical doctor." (NATURAL)

As we see in this example, what is intelligible may not be acceptable, and what is acceptable may not be the most natural way of expressing a thought. Style in translation seems to be most exhibited by the most natural way of expression. A translation should not be dictated literally by the structure of the SL. Instead, it must be molded from the thought of the SL, reflecting the most natural way of expressing a particular thought in the TL. Machine-assisted translations fail to do this because of the absence of the thought process. In model sentence 3, the Systran output is intelligible and acceptable, but a question arises as to its naturalness. The human translation has taken a simple thought and converted it in the most natural way to another language. The translation is not dictated by the SL but based on the SL. Let's look at another example which further illustrates this point.

MODEL SENTENCE 4

Describing methods of delivering signals is shown here.

TRANSLATIONS:

ALPS: Describiendo metodos de distribuir
         señales se muestra aqui.

SYSTRAN: Se muestra la descripción de los méto-
         dos de distribuir las señales aqui.

HUMAN: Aqui aparece una descripción de
         métodos para emisión de señales.

The ALPS translation gives us an intelligible but not altogether acceptable rendering of the sentence. The Systran translation, with its sentence rearrangement and use of the nominal form instead of a gerundive, presents an acceptable translation. But the human translation does several things which contribute to its assessment as natural. First, the adverb "here" (aqui) is placed at the beginning of the sentence. Second, the translator eliminated the awkward but common use of the passive voice. More subtly, however, the translator made a contextual judgment and

translated "delivering" by <u>emisión</u> rather than by <u>distribuir</u> as in the two MAT versions. All of these considerations show how the human translator interacts with the original sentence to create a thought reflecting that original rather than simply mimicking or calquing the SL lexicon and syntax.

MODEL SENTENCE 5

The larger the office, the higher the level of interaction and the remoter the possibility of achieving high standards.

TRANSLATIONS:

ALPS: El mas grande la oficina, el más
         alta el nivel de la interacción y el
         mas remota la posibilidad de lograr
         estándares altos.

SYSTRAN: El mas grandes la oficina, el más
         alto el nivel de la interacción y el
         mas remoto la posibilidad de lograr
         los altos estándares.

HUMAN: Cuanto más grande es la oficina,
         mayor es el nivel de interacción y
         más remota
         la posibilidad de lograr altos están-
         dares.

Sentence length and complexity of construction do not necessarily stump the computer to the extent of nullifying the sentence's intelligibility. Usually, as we see here in both MAT versions, the first casualty of a more complex sentence is agreement, or exactly "what goes with what?" Neither raw machine translation is acceptable because of this factor, but even if the problem were corrected, both renderings would be barely acceptable due to being strictly patterned after the SL text. The English constructions presented here require subtle changes in the Spanish in order to capture the idiomatic features of the SL sentence. The more idiomatic the constructions are, the less adept the computer is at reproducing anything that can be called stylistically palatable. These last two model sentences will illustrate my point even more dramatically.

MODEL SENTENCES 6 & 7

6. Light may cause changes in the photograph.
7. Light may bring about changes in the photograph.

TRANSLATIONS:

ALPS: Luz puede causar cambios en la
        fotografia.

        Luz puede traer sobre cambios en la
        fotografia.

SYSTRAN: La luz puede hacer los cambios en la
         fotografia.

         La luz puede traer sobre cambios en la
         fotografia.

HUMAN

      La luz puede causar cambios en la foto-
      grafia.

      La luz puede causar cambios en la foto-
      grafia.

The obvious inability of the computer to deal with the verbal phrase "bring about" is striking in both machine translations. Specialists in the field would contend that the solution to the problem is simple. We need only add the phrase "bring about" as a verbal unit to our dictionaries, giving it the meaning of causar. Unfortunately, the answer is not that easy. Syntactically, an intervening object could totally change the required translation. For example, "I brought the book about which we spoke." But even if programming or dictionary entries could allow for this eventuality, stylistic considerations demand a different translation depending on context. "The experiment will bring about the required findings." Wouldn't resultar en be a more suitable translation in this case? What about the sentence, "Irrigation brings about good crops."? Definitely, verb producir is the more stylistically natural translation of the phrase "bring about" here. To carry the point even further, one Italian translator consulted said she would translate the second sentence as "Il risultato dell'irrigazione è una buona raccolta."

Computer-aided translation systems cannot capture style. They can produce acceptable translations, but rarely is their output the most natural way of expressing a given thought. This has been the major complaint of post-editors and translators — not with what the computer can do but rather with what it cannot do. The results of machine translation can be surprisingly accurate, but often they leave a stylistic gap that can only be filled by human interpretation and creativity.

The true benefits of MAT can be seen when its limitations are recognized. Repetitive, directive text of a technical nature lends itself well to the use of the computer as a translator's tool. When conveying clear technical content is the translation's goal, the machine can churn out line after line of text that can be useful to the translator who is willing to limit corrections to grammatical points, disregarding stylistic aspects. When the translator feels the need to rewrite for the sake of style, he or she should no longer be using the machine output.

Translation is an art. Not unlike computerized music or drafting, MAT has its limitations. Computer graphics systems can create impressive three-dimensional technical drawings, but no one would want to hang one in the living room! Likewise, MAT has some impressive results, but true style can only be a product of the human psyche.