
Computers in translation

A practical appraisal

Edited by
John Newton



London and New York

First published 1992
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
a division of Routledge, Chapman and Hall, Inc.
29 West 35th Street, New York, NY 10001

© 1992 John Newton

Typeset in Baskerville by
Columns Design and Production Services Ltd, Reading
Printed in Great Britain by Biddles Ltd, Guildford and King's Lynn

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

A catalogue record for this book is available from the British Library.

ISBN 0-415-05432-X

Library of Congress Cataloging-in-Publication Data

Computers in translation : a practical appraisal / edited by John Newton.

Includes bibliographical references and index.

ISBN 0-415-05432-X

1. Machine translating. I. Newton, John. 1944-
P308.C67 1992

418'.0270285—dc20

91-41614
CIP

Machine translation in a high-volume translation environment

Muriel Vasconcellos and Dale A. Bostad

The real test of machine translation is whether or not it is effective in large-scale operations. These may be specialized applications, as in the case of Canada's METEO (see Chapter 3), or they may involve the translation of a broad variety of text types. In the latter case, the purpose of the translation will dictate the characteristics of the installation, particularly the human post-editing component. The purpose can run the gamut from publication for dissemination to 'information only'. For the product that has a public and undergoes scrutiny, a translator post-editor must be enlisted to eliminate all problems from the output, interpret connotations appropriately and make certain that the reader will understand precisely what the author intended to say. On the other hand, MT, by automatically generating a product where nothing existed before, has created a new mode of work in which a less than ideal translation can be provided for the consumer who merely needs to assimilate information and is not concerned with disseminating it.

The first type of application is exemplified by the Spanish to American (English) MT system (SPANAM) and English to Spanish MT system (ENGSPAN) at the Pan American Health Organization (PAHO) in Washington, DC, and the second by the SYSTRAN operation at the US Air Force's Foreign Technology Division in Dayton, Ohio.

TRANSLATION FOR CLOSE SCRUTINY: PAHO

MT has been enlisted in the service of general-purpose practical translation at PAHO¹ since January 1980. The first language combination to be implemented was Spanish-English, using SPANAM, the Organization's batch MT system developed in-house

(Vasconcellos and León 1988). ENGSPAN, a second, more sophisticated in-house product with a strong English parser,² followed suit with English–Spanish in 1985 (León and Schwartz 1986).³ Later, SPANAM was entirely rebuilt using the conceptual design of ENGSPAN. MT is now the principal mode of translation for these two language combinations and is an integral part of the Organization's translation and terminology service.⁴

PAHO undertook to develop MT with two broad purposes in mind: to meet its internal translation needs more efficiently, and to disseminate health information in the Latin American and Caribbean countries. Very few of the translations done at PAHO are for information only. There are occasional applications of this kind but the majority must be of a quality that will stand close scrutiny. Some are for publication although few serve as the basis for important decisions, covering diverse subjects ranging from the commitment of resources to primary care for sick children.

SPANAM and ENGSPAN, each with a robust parser and large deeply coded dictionaries (63,000 and 55,000 terms respectively), are capable of handling a broad range of subjects and dealing with free syntax in a variety of genres. Among the most important subject areas are medicine, public health, sanitary engineering, agriculture, computer science, management and law, and there are a number of others as well. Cross-cutting this spectrum is an even broader range of styles and discourse genres: journal articles and abstracts, textbooks, manuals (both for human health and for software), proposals for funding, reports of missions, contracts and agreements, minutes of meetings, business letters, diplomatic exchanges, certificates, product specifications, supply lists, captions for displays and even promotional materials and film scripts. This type of variety actually presents more of a challenge than differences in subject matter but SPANAM and ENGSPAN tackle them all. For most applications the texts are post-edited by professional translators working on-screen. The input texts are never pre-edited or customized in any way. The only preprocessing is a quick review by the clerical staff to ensure that the word-processing document conforms to one of the standard formats used in the Organization. Thus, from any perspective, it is safe to say that MT at PAHO has been both designed and implemented as a general-purpose product. Examples of SPANAM and ENGSPAN output are shown in Figures 5.1 and 5.2.

Figure 5.1 Spanish-English unedited machine translation at PAHO.

ILRAD	OK	ILRAD
LABORATORIO INTERNACIONAL DE INVESTIGACIONES SOBRE ENFERMEADES ANIMALES	OK	INTERNATIONAL LABORATORY FOR RESEARCH ON ANIMAL DISEASES
El Laboratorio Internacional de Investigaciones sobre Enfermedades Animales (ILRAD) se fundó en 1974 con el objeto de ayudar al desarrollo de controles eficaces de dos importantes enfermedades que afectan a la ganadería: la tripanosomiasis y la teileriosis. En conjunto, estas dos enfermedades afectan a la producción ganadera de extensas zonas en unos 50 países en desarrollo de Africa, América Central y del Sur, el Oriente Medio, el subcontinente indio y Asia. Las pérdidas totales que causan, de recursos humanos y económicos, son incalculables, no sólo en materia de leche y carne, sino también de cuero, lana, fertilizantes, tracción animal y otros subproductos animales, y en posibles recursos de capital. Cientos de millones de personas, entre ellos algunos de los más pobres del mundo, resultan gravemente afectados. El ganado rumiante transforma estos elementos a partir de vegetación que el hombre no puede comer, a menudo en terrenos que no puede utilizar para cultivos. En otras zonas, es conveniente la producción integrada de ganado y cereales.	OK TU OK OK PP 39 TU NO 05 PP 15 TU GT PP 12	The International Laboratory for Research on Animal Diseases (ILRAD) was founded in 1974 with the purpose of helping the development of effective controls of two important diseases that affect livestock raising: trypanosomiasis and theileriosis. Together, these two diseases affect the livestock production of extensive areas in some 50 developing countries of Africa, Central and South America, the Middle East, the Indian subcontinent and Asia. The total losses that cause, of human and economic resources, are untold, not only in terms of milk and meat, but also of hide, wool, fertilizers, animal power and other animal by-products, and in possible capital resources. Hundreds of millions of persons, among them some of the poorest in the world, result critically affected. Ruminant livestock transforms these elements from vegetation that man cannot eat, often in lands that it cannot utilize for "crops. In other areas, it is desirable the integrated production of earned and grains.
LA ESTRATEGIA DEL ILRAD	OK	THE STRATEGY OF ILRAD
Las dos enfermedades mencionadas son causadas por parásitos que son transmitidos por insectos.	OK	The two diseases mentioned are caused by parasites that are transmitted by insects.

La mosca tsetse transmite los tripanosomas y la teileriosis OK
es transmitida por las garrapatas. En ambos casos las OK
relaciones entre parásitos, huéspedes y vectores son
complejas y sutiles, y por tanto la intervención es
difícil. Además, en ambos casos, otros animales salvajes y PP 20
domésticos sirven también como huéspedes de los parásitos,
creando así reservas de infección prácticamente
inaccesibles a las medidas de control.

Aun en el caso de que el control de los vectores fuera PP 35
técnicamente posible, ello requeriría la aplicación de
medidas sistemáticas, oportunas y muy costosas, que están
evidentemente más allá de las capacidades del pequeño
agricultor del mundo en desarrollo. Por consiguiente, las OK
investigaciones del laboratorio hacen hincapié en la
identificación y elaboración de medios de control de
enfermedades basados en las reacciones inmunológicas de los
animales huéspedes.

Además de mantener un programa completo e innovador de PP 35 TU
investigaciones originales, el ILRAD sirve también como
centro de intercambio para integrar los diversos esfuerzos
de otras instituciones nacionales e internacionales que
laboran en los mismos problemas desde diferentes
perspectivas.

The tsetse fly transmits the trypanosomes and theileriosis
is transmitted by the ticks. In both cases the relations
between parasites, hosts and vectors are complex and
subtle, and accordingly the intervention is difficult. In
addition, in both cases, other wild animals and domestic
serve as well as hosts of the parasites, creating thus
reservoirs of infection practically inaccessible to the
measures of control.

Even if the control of the vectors were technically
possible, this would require the application of systematic,
timely and very expensive measures, that are evidently
beyond the capacities of the small farmer of the developing
world. Consequently, the research of the laboratory
emphasizes the identification and preparation of means of
disease control based on the +immunological+ reactions of
the host animals.

In addition to maintaining a complete and innovative
program of original research, ILRAD serves as well as
clearinghouse in order to integrate the various efforts of
other national and international institutions that work in
the same problems from different perspectives.

Figure 5.2 English-Spanish unedited machine translation at PAHO.

S0015	NT	S0015
EPIDEMIOLOGIC STUDIES OF CERVICAL CANCER IN LATIN AMERICA	OK	ESTUDIOS EPIDEMIOLOGICOS DEL CANCER DEL CUELLO UTERINO EN AMERICA LATINA
During the last five years, two large epidemiologic studies were conducted in Latin America to determine risk factors for cervical cancer in this high-risk population and to compare the prevalence and effects of different risk factors in this region and in Spain, where the incidence of invasive cervical cancer is very low.	OK TU TU	Durante los últimos cinco años, dos grandes estudios epidemiológicos se realizaron en América Latina para determinar los factores de riesgo de cáncer del cuello uterino en esta población de alto riesgo y para comparar la prevalencia y los efectos de diferentes factores de riesgo en esta región y en España, donde la incidencia del cáncer invasivo del cuello uterino es muy baja.
The Latin American Cervical Cancer Study was conducted in Bogotá, Colombia, Mexico City, Mexico, Costa Rica and Panama, by the National Cancer Institute of the United States (NCI) in collaboration with the "Gorgas Memorial Laboratory and research institutes in the participating countries, between January 1986 and July 1987. The study examined cases of invasive cervical cancer in patients younger than 70 years of age at the major cancer referral hospitals in the areas of the study. For each case, two controls were selected: in Bogotá and Mexico City, two hospital controls were selected, matched according to age (in 5-year groups), excluding patients with diagnoses related to the exposures of interest.	SD PP 32 PP 24 TU OK OK TU TU	El Estudio de Cáncer del Cuello Uterino Latinoamericano se realizó en Bogotá, Colombia, la Ciudad de México, México, Costa Rica y Panamá, por el Instituto Nacional del Cáncer de los Estados Unidos (INC) en colaboración con los Gorgas Memorial Laboratory y los institutos de investigaciones en los países participantes, entre enero de 1986 y julio de 1987. El estudio examinó los casos de cáncer invasivo del cuello uterino en las pacientes más jóvenes que 70 años de edad en los hospitales de referencia de cánceres principales en las áreas del estudio. Para cada caso, se seleccionaron dos controles: en Bogotá y la Ciudad de México, se seleccionaron dos controles del hospital, armonizado según la edad (en grupos de 5 años), excluyendo a las pacientes con diagnósticos relacionados con las exposiciones del interés.

In Costa Rica and Panama, one hospital and one community control were selected from census listings of the same areas of residence of the cases.

Overall, 99% of 766 eligible cases and 96% of 1532 controls agreed to participate and were interviewed on demographic, sexual, reproductive, medical, dietary and contraceptive history. A cervical swab was utilized to obtain cells from the cervical lesions of the cases and the cervical os of the controls to determine the presence of human papilloma virus (HPV). In addition, a blood sample was drawn for determination of micronutrients and antibodies against specific sexually transmitted agents. A physical examination was performed on participating males (78% of eligible case husbands and 71% of eligible control husbands) and a swab of the coronal sulcus and urethral canal were obtained to test for HPV.

OK

OK

OK

OK TU

A PP 30 TU

En Costa Rica y Panamá, un hospital y un control de las comunidades se seleccionaron de los listados del censo de las mismas áreas de la residencia de los casos.

En total, 99% de 766 casos aptos y un 96% de 1532 controles acordaron participar y se entrevistaron en la historia demográfica, sexual, reproductiva, médica, alimenticia y anticonceptiva. Un hisopo cervical se utilizó para obtener células de las lesiones cervicales de los casos y el os cervical de los controles para determinar la presencia del +virus de papiloma humano+ (VPH). Además, una muestra sanguínea se extrajo para la determinación de los micronutrientes y los anticuerpos contra agentes de transmisión sexual específicos. Un examen físico se realizó en los varones participantes (78% de esposos del caso apto y 71% de esposos de control apto) y un hisopo del surco coronal y el canal uretral se obtuvieron para examinar para detectar el VPH.

The PAHO environment is distinguished by the fact that production and system development are carried on side by side as part of the same operation. As those who are working on a particular translation are in the best position to suggest appropriate target glosses, post-editors are encouraged to mark their copies of machine output with suggestions for the dictionaries (new technical terms being subject to review by the terminologist) and for system improvement in general. The computational linguists, for their part, conduct their research on production text and are constantly monitoring the output to ensure that the two systems are performing up to standard. Problems that are easy to solve are dealt with from one day to the next. In this way, the post-editors have the satisfaction of seeing their feedback incorporated. (More difficult problems take their place on a 'wish list,' in order of their priority). Some of the post-editors have learned how to update the dictionaries themselves, and whether they do or not they still get to see that they can have a personal hand in improving the quality of the output.

Over 8 million words delivered

Getting under way

As of 1990, SPANAM and ENGSPAN had generated some 34,800 pages (8.7 million words) in the service of production translation, with current machine output averaging around 6,000 pages (1.5 million words) a year. In 1989, MT was officially recognized as the Organization's primary mode of translation, as by that time it was supporting at least 60 per cent of all regular production in the two language directions. But the road to success had been bumpy at times. The use of MT at PAHO went through several phases, and it is of interest to take a look at this experience as it unfolded.

When SPANAM made its *début*, MT was operated as a separate unit in PAHO, offering direct competition to the existing human translation (HT) service. In view of the broad range of texts to be translated, its managers realized from the beginning that efficient and effective post-editing would be crucial to successful acceptance of MT. With this in mind, a certified translator was hired on a full-time basis to review and correct the output. This person also updated the dictionaries based on problems that came up in the course of day-to-day production. The first major application was

the Organization's biennial program budget, for which it was possible to demonstrate savings of 61 per cent compared with human translation, as well as a reduction in staff days of 45 per cent (Vasconcellos 1984).

For the first five years, post-editing was done at no cost to the requesting office. The choice of whether to use MT was left to the end consumer. Many of the receiving units were pleased with the service, and the good word spread. Cost was undoubtedly a factor in the early popularity of MT, since the service was being offered for free, whereas HT sometimes had to be farmed out and the resulting cost charged to the client. Also, turnaround was faster. In addition, MT was welcomed with enthusiasm because the delivered product was machine readable, which was not true of HT at the time. Also, MT with light post-editing was occasionally used for information purposes only, and this service addressed a previously unmet need – yet another factor that added to its successful reputation. Since no one was forced to use MT, and since the service was free, there were virtually no complaints.

By 1984, MT production from Spanish into English was reaching 100,000 words (400 pages) a month, and ENGSPAN was being tried on an experimental application from English into Spanish. At the end of that year, in the wake of a series of personnel changes (including departure of the MT post-editor), the Organization's upper management decided to merge the MT and HT services and delay the recruitment of new translators until a study could be done on the overall allocation of human resources for translation. In the interim, MT continued to be an option but post-editing had to be done by contractors with the cost charged back to the requesting offices. The contractors, most of them freelance professional translators, were paid by the job at slightly more than half the then prevailing rate for HT, and they came to the PAHO offices to use the Organization's word-processing equipment. As a result of these changes, there was in fact a slight drop in MT production levels but the monthly average never fell below 67,000 words (268 pages).

In the fall of 1987, still without any conclusive data in hand and with several translator vacancies to be filled, PAHO's management decided to conduct an experiment that would yield concrete data for decisions about the appropriate mix of the four possible resources – namely, in-house MT, in-house HT, contract MT and contract HT.

The experiment

A formal eleven-month experiment was undertaken from October 1987 through August 1988 to establish whether machine translation in the PAHO environment was cost-effective, fast in turnaround and capable of supporting a higher level of year-round average production than human translation. The experiment also attempted to establish whether MT output post-edited by a professional translator was as serviceable as human translation.

Two temporary translators, one for each target language, were recruited to serve as post-editors for the duration of the experiment. MT ceased to be optional. All incoming work was screened to determine whether or not it could be submitted to the computer. A directive was issued instructing the requesting offices to submit their texts for translation on diskette. In this way new MT applications could be captured and use of the technology was maximized.

Interim statistics for the first eight and a half months of the experiment showed that MT, in-house and contract combined, was used to process 79 percent of the jobs into English (839,635 words or 3,359 pages, corresponding to 262 jobs) and 60 percent of those into Spanish (427,310 words or 1,709 pages, from 120 jobs).⁵ MT was used less for Spanish because a large proportion of the input texts were from published sources outside the Organization and therefore not machine-readable. Optical character recognition (OCR) was used whenever possible to scan hard copy but it often turned out that cleaning up the scanner's mistakes was too time-consuming to be worth the effort.⁶ In total, 74 percent of the jobs into English were machine readable and, of the rest, 5 percent could be scanned using the OCR but 21 percent could not. Into Spanish, on the other hand, only 54 percent were machine readable and, of the rest, 7 percent were scannable and 39 percent were not.

Data for the full twelve months showed a significantly greater volume translated by MT during the experiment than in previous years. The texts represented a broad range of subject areas and discourse types, as evidenced by the fact that more than 40 different 'clients' had been served.

The requesting offices, which usually did not know which mode was used, were asked to provide feedback on the serviceability of

the product delivered to them. Responses were received for 30 percent of the jobs processed. Both modes drew praise and criticism but there was a slight preference for MT: 85.1 percent indicated satisfaction with jobs processed using MT, compared with 78.1 percent for HT. Interestingly, the lowest percentage of satisfaction (60 percent of total response) was for Spanish HT. Sometimes the client incorrectly assumed that the job had been done using MT when in fact it was HT. For example, the following comments were made about *human* translations: 'machine translation is not good enough. I practically had to rewrite it,' and 'the quality of the translation was more "rigid" than at other times; it seemed more like a machine translation without having a "human hand" go over it.'

The English in-house post-editor and some of the contractors were able to achieve daily outputs of 7,000 words and higher in the short term. In the long term, average daily in-house productivity for English, calculated according to standard methods used in the international organizations (i.e. net of leave, weekends and holidays) and *including periods without incoming translations* during which the translators worked on the dictionaries, was 2,546 words. This result exceeded the United Nations target of 2,000 words a day by 27 percent and far surpassed the real average of 1,200 words a day informally reported by many in-house services in the international organizations. As MT does not involve transcription and certain other types of general services support that were used for HT during the period, it was considered that the goal of 30 percent savings over in-house HT had not only been achieved but in fact surpassed. For contract translation, the cost differential between HT and MT was self-evident, as MT post-editors were being paid 55 percent of the HT rate.

As a result of the experiment, a new policy was announced in February 1989 under which MT became the primary mode of translation in PAHO. The policy entailed a restructuring of the translation service so that as much incoming work as possible could be channeled in the direction of MT.

Current situation

The use of MT has not been stabilized in PAHO. The new technology continues to do the lion's share of the work. The decision to use MT, which rests entirely with the terminology and

translation service, is based on the following characteristics of the input text:

- 1 machine readability (or optical 'scan-ability');
- 2 complexity of format; and
- 3 linguistic characteristics (e.g. grammar, discourse genre, need for between-the-lines interpretation, etc.).

These factors intersect with time-frame considerations, the availability of post-editors and the special strengths of the individual translators.

In-house translator positions have been established with post-editing built into the primary duty in the job description. At the same time, contractors are still employed for much of the work which, as always, comes in spurts. Compensation for contractors has been brought in line with the three-tiered scale of the US Department of State (general, semi-technical and technical). The pay for MT post-editors is the same as for State Department reviewers. A recent development is that most of the MT post-editing is now being done off-site by contractors working on their home computers.

The translator and the MT output

The experience of the last three years, since the mechanism for client feedback was formalized, has brought home increased awareness that it is the translator's skill, not the mode of work, that determines the quality of the final translation product.

At the same time, not all the PAHO translators have chosen to adapt to post-editing; some of the ones that we rely on have preferred not to try it, and of those who have tried it, not all have been equally successful. In total, about 30 post-editors have worked with SPANAM or ENGSPAN and, of these, six into Spanish and eight into English are now on the roster of contractors regularly used. An effort is made to give these 'regulars' an HT assignment every third or fourth time. There is a core group of translators, all of them veterans, who consistently prefer the MT mode.

Post-editing seems to be a special skill, somewhat related to traditional editing. It involves learning how to preserve as much of the machine's output as possible and 'zapping' the text at strategic points rather than redoing it from scratch. The post-editor quickly develops a set of context-dependent techniques for dealing with the

patterns produced by the machine (Vasconcellos 1986, 1987a, 1987b; McElhanev and Vasconcellos 1988; Santangelo 1988).

There is a fine line to walk between allowing the output to stand and meeting the high standard of quality that needs to be applied to most of the translation jobs done at PAHO. In order to come up to this standard, the post-editor must be certain of all technical terminology, capture every nuance, provide accurate interpretations for coherence (Vasconcellos 1989a), ensure idiomatic equivalence, maintain a uniformly appropriate register, build in cohesiveness where it is lacking, provide adequate 'staging' (Grimes 1975: 323), and preserve the information structure (Halliday 1967; Vasconcellos 1986).

Sometimes, in addition to using approved terminology, the translation will call for special phraseology, as, for example, with resolutions of the organization's governing bodies. Here MT can present a problem: the post-editor cannot settle for an equally clear but unofficial version produced with the aid of the machine. In such cases it can be seen that in-house translators and 'regulars' are in a better position to make decisions about the MT output. To help the contractors in particular, the documentalist, who is very familiar with the workings of the Organization, reviews each incoming job and locates all the background materials the translator should have at hand.

When it is known that a translation will have only limited use – a first draft, for example, or an internal trip report – MT offers the possibility of delivering a less-than-ideal translation at a lower level of effort – i.e. by doing less post-editing. PAHO is currently attempting to implement an intermediate level of quality which would be acceptable in such situations. The text is syntactically and propositionally correct, but nuances may be sacrificed. So far, it has been difficult for the translators to 'lower their standards', but occasionally the press of time has been sufficient motivation for them to cut the right corners. In-house translators are more appropriate for such an approach, first because compensation is not at issue, but mainly because they are in a better position to sense the minimum level that will be acceptable. It is planned to experiment further with translations of intermediate quality.

Unfortunately, it does not necessarily follow that when a translation is of minor importance a raw or semi-polished MT product can be used. It often happens that when the input text has been prepared in a rush, maybe dictated but not read, or

transcribed from an oral presentation, it contains incomplete fragments and grammatical lapses which throw the machine – and the post-editor – into a tailspin. Not only must the output be carefully post-edited, but the work of the post-editor is greatly slowed down, to the point where sometimes the effort has to be compensated at the full HT rate.

By far the most difficult challenge for the translators has been to know when to leave PAHO-specific terminology untouched. Terms that may look ‘funny’ to the uninitiated – such as *fertility regulation*, where *birth control* is the more common expression, or *harvesting* and its Spanish counterpart *cosecha* in reference to the culling of monkey populations – are apt to get changed by post-editors. Hypercorrections of this kind are the main source of complaints from clients. Again, the advantage of in-house translators or ‘regulars’ who know the technical terminology is obvious. To help deal with the problem, the entries in the SPANAM and ENGSPAN dictionaries can be coded so that a word or phrase will be flagged in the output, telling future post-editors that the term has been researched and is reliable. This resource has been implemented to a certain extent, but at the same time care needs to be taken not to clutter up the output with so many flags that the text becomes unreadable. And, of course, there is never a guarantee that a flagged term will not require a different translation in another context. Some of the other ways in which the post-editor can have some control over the output are described in the next section.

The translator and the dictionary

PAHO post-editors can gain a certain amount of power over the MT output by learning to manipulate the SPANAM and ENGSPAN dictionaries, which offer a wide variety of approaches for triggering expressions that are appropriate to the given context. As a first step, they are expected to mark up their hard side-by-side copy as they move along with each translation. The following kinds of information can be recorded: glosses for words that were not in the dictionary at all (which have averaged less than 0.5 percent since 1984), different or additional glosses for words that are in the dictionary, special translations for multiword phrases, reliability flags, and notes about syntactic errors the machine has made.

Perhaps the most immediately useful resource for the translator

is the phrase-building capability. Whether or not the translator learns to enter these units in the dictionary personally, it is still rewarding to see them come up in future translations. If a particular term happens to have required research, the translator's efforts are captured once and for all, avoiding the duplication that is so common in translation services.

SPANAM and ENGSPAN can provide special translations in a wide variety of situations. Some examples are: different subject areas (English *core* for Spanish *núcleo* in an atomic energy text vs *nucleus* in an ordinary text); singular and plural forms of the same source word (*toxicity* for *toxicidad* and *toxic effects* for *toxicidades*); fixed phrases of two to five words in length (*in general* for *por lo general*) or *Pan American Health Organization*; long names of institutions and titles of publications; multiword phrases that can be parsed in more than one way if necessary; discontinuous phrasal verbs; glosses based on the syntactic or semantic features of associated words (the English source word *raise* with an inanimate object gives Spanish *levantar*; with an animate object, *criar*; with a human object, *educar*; and collocated with the word *flag*, *izar*). The post-editor can request these translations or learn to actually enter them in the dictionary.

The post-editor also brings recurring structural problems to the attention of the computational linguists.

Future applications

Thanks to ongoing dictionary work and system improvement, ENGSPAN now produces raw output of sufficiently reliable quality that consideration is being given to the translation of data bases and other applications in which users can access MT directly. Of particular interest are data bases that are available on compact disk read only memory (CD-ROM). Several proposals have been made and some of these may materialize into active projects.

The Consultative Group on International Agricultural Research (CGIAR) has been collaborating with the PAHO MT project since 1986 and provided support for the installation of ENGSPAN at the International Rice Research Institute in the Philippines and the International Center for Tropical Agriculture (CIAT) in Colombia. CGIAR is helping to form a donor consortium that will provide PAHO with funds to adapt ENGSPAN to a micro-computer and develop parallel systems from English into French

and Portuguese, as well as establish an MT center within the CGIAR network that will support users and build specialized MT dictionaries in agriculture.

The availability of PAHO's MT systems on microcomputer will enable the organization to outpace this technology in its member countries, and this should contribute to the exchange and dissemination of types of needed health information which have not been translated in the past because of the cost and the limited availability of translators.

TRANSLATION FOR INFORMATION: FOREIGN TECHNOLOGY DIVISION, US AIR FORCE

Machine translation has undergone a long evolutionary development at the Air Force's Foreign Technology Division (FTD), extending over more than twenty years. While for many years the SYSTRAN Russian-English system was the sole MT operation, with time additional language pairs were acquired: French-English, German-English, Japanese-English and, recently, Spanish-English. A large-scale machine translation environment has been created, with some fifteen people inputting data, six or seven people editing the data, and, since 1975, a group of four to seven people constantly working on improving the dictionaries and the parsing algorithms. At all times, development of the various systems has gone hand in hand with use of the systems to translate texts on a daily basis.

The SYSTRAN machine translation systems are used to translate large amounts of information – roughly 50,000 to 60,000 pages of source text per year. The translations are used by scientific and technical analysts who need to keep abreast of foreign developments in a wide range of technical fields and to prevent technological surprise that could threaten the United States.

FTD got involved early with machine translation and has consistently been a supporter of MT because it was enamored with the prospects of what MT could do in the early years, and because it was believed that machine translation was the only way to translate massive amounts of material and get the translations back to the requesters while the information was still useful. The early intuition that machine translation could provide this service has proven correct, and FTD has not veered from its solid commitment to MT for over twenty years.

However, the way of producing machine-translation products

has evolved over time at FTD, and it is of interest to briefly discuss how FTD got started in the MT business and where it is now.

In the 1960s and the first years of the 1970s the Russian-English system was used to produce a very rough hard-copy translation. Post-editors edited the translation extensively – almost rewriting it – and then the edited version was sent to a ‘recomposition’ pool where it was retyped. Alternatively, some translators used the rough machine translation as a ‘pony’ when they dictated the translation on tape. The translator picked and selected those parts of the machine translation that were valuable – certain technical terms and phrases – and worked directly from the Russian text. The dictated translation was then transcribed. Machine translation was really only used as an aid to the translator. The end translation was a quite accurate finished product.

Because this method was slow and productivity was low, in 1974 it was discarded and a very lightly edited hard-copy translation was instituted. It was decided that this was the only way to cope with an immense backlog that had built up. The degree of post-editing was at the discretion of the editor, but speed was the name of the game. Changes were written in ink on large fanfold computer printout. There were no stylistic changes. The end product could only be called fairly crude and the format primitive – a low-quality utilitarian product whose one saving grace was that it got the information quickly to the requester and allowed an immense backlog to be eliminated. This was FTD’s first effort at producing what we call ‘partially edited machine translation.’

In 1976 a dedicated input/editing minicomputer system using a DEC PDP 11/05 and sixteen workstations was designed and installed. This allowed the first on-line computer revision of machine translation, and in 1977 the first version of EDITSYS was written. EDITSYS is a software module called at the end of the Russian-English analysis that identifies certain potential problem areas in the output and brings these conditions to the attention of the post-editor. When a given condition is met, the program generates a full-page-width string of characters (a ‘flag’) immediately in front of the condition. The post-editor must react to the ‘highlighted’ condition and verify the machine translation version or make a revision. The flag is removed during a later step in the post-processing. Such conditions as ‘not-found’ words, rearrangement, acronyms, problem words, mistyped words that produce spurious match-ups, etc., are highlighted. Post-editing is thus

determined by a software program that tells the editors the minimum number of conditions that have to be reviewed, although the editors can – and some of them do – post-edit beyond the flagged conditions. But sentences and blocks of text basically go through unscrutinized by an editor. This type of partially edited text has been our mainstay for the last fifteen years. The translation is then printed on high-quality paper on a laser printer and has a much better appearance than the previous product. This product has been well received by FTD analysts.

As stated above, some editors do edit to a certain extent beyond the flags. There are also other exceptions to the procedure just described. Some short translations for dissemination require careful sentence-by-sentence editing. Moreover, medical and biology texts, at present, are closely edited for technical accuracy. Approximately 5 per cent of MT translations receive this type of post-editing. Finally, some fifty to 100 translations per year, ranging from one paragraph to two pages, are removed from the ‘first-in-first-out’ queue, translated by MT, and then edited on hard copy by translators who do not normally do post-editing. The idea is to get the jobs done quickly and out the door; otherwise they tend to stagnate behind larger jobs. Requesters expect fast turnaround on very short translations; they understand that a 300-page book, for example, takes longer.

In 1987, FTD developed a new MT application which we call ‘interactive machine translation.’ This system gives all users individual access to MT at their own terminals. It is now available to users via hookup with approximately 1,600 PCs within FTD. This is raw machine translation without the mediation of translators. The system is designed so that a user can rapidly determine the significance of the material he wants translated and weed out extraneous information. It is best used for rapid translation of titles of books, tables of contents, captions of tables and graphs, and individual sentences and paragraphs. However, it can also be used effectively to translate complete short articles and to get back a rapid translation instead of going through the sometimes time-consuming operation of routing translations through the formal bureaucracy. One very effective use of the system is for ‘gisting’ a large book – that is, determining the significant parts of the book and then routing this material through the normal translation procedures. For example, if a user has a 350-page book, the system might be used to determine that only Chapters 3, 7 and

12-15 are really pertinent to the research in question. Obviously, the use of such a tool makes for tremendous cost savings by eliminating the translation of irrelevant material.

FTD has conducted three extensive surveys of machine translation over the last ten years to analyze the effectiveness and use of MT and gain insights into how to improve the product. The two most important insights coming out of the surveys are:

- 1 speed of translation is the most important consideration for FTD analysis; and
- 2 the existing product, partially edited MT, is deemed satisfactory in meeting most users' translation requirements.

There has been some feedback spelling out deficiencies in certain technical disciplines, and there was a small percentage of users who found MT unsatisfactory as a translation product but overall the acceptance rate has been very high.

The most recent survey, taken in December 1990, covered users of both partially edited MT and interactive MT. Sixty surveys were sent out; thirty-three were returned, for a return rate of 55 percent. In addition, within this number, six users were personally interviewed. According to the survey, 73 per cent of the respondents felt that the current post-edited MT product was meeting or exceeding their requirements for technical accuracy. An even higher proportion - 82 percent - stated that the post-edited product was meeting or exceeding their requirements for readability! There was high praise for the interactive system. However, because the analysts themselves have to input the data, it was not deemed effective for longer documents. More analysts would be willing to use raw MT if the data could be input for them. They would forego the current post-editing if the translation directorate would input the documents and send the raw MT product electronically to them directly after translation. Very recently the utility of raw MT was emphasized by the fact that 600 pages of French had to be translated in two weeks and, due to the time constraint and the lack of post-editors from French, the only translation that could be produced was raw MT. The requester accepted the raw MT in order to get quick access to the data and meet his deadline.

Recently an assessment of Russian partially edited MT was made by an independent group of scientists conducting Air Force research in a subfield of physics. For one particular portion of a

book they had access to both a human translation (done for another government agency) and a machine translation. Their final report contained the following unsolicited comments on the quality of the two English translations:

While the [human] translation read somewhat more smoothly, it seemed to use inappropriate or erroneous terminology more often than the [machine] translation did. Consequently, we relied primarily on the [machine] translation, using the [human] translation mainly for reference.

FTD is now embarked on an ongoing project to improve the efficiency of its MT operation. The areas being addressed include: incorporation of software for individual users to modify dictionaries; OCRs to scan Russian, French and German; expansion into other language pairs; and continued refinement of specialist dictionaries.

NOTES

- 1 Regional Office of the World Health Organization for the Americas. WHO is a specialized agency in the United Nations system.
- 2 Developed with partial assistance from the US Agency for International Development (AID) under Grant DEP-5443-G-SS-3048-00. ENGSPAN is installed at AID and runs there on an IBM 3081 (OS/VMS).
- 3 SPANAM and ENGSPAN are written in PL/1 and run on PAHO's IBM mainframe computer, currently an IBM 4381 (DOS/VSE/SP), which is used for many other purposes.
- 4 The Organization's working languages are Spanish and English. The English-Spanish and Spanish-English combinations account for 90 per cent of the translation workload. Portuguese and French, which together make up the other 10 per cent, are also official languages of the Organization but are handled by a separate service.
- 5 For a detailed review of the data from the experiment, see Vasconcellos 1989b.
- 6 The equipment on hand was already old at the time of the experiment. Current OCR technology would undoubtedly do much better.

BIBLIOGRAPHY

- Grimes, J.E. (1975) *The Thread of Discourse*, The Hague: Mouton.
Halliday, M.A.K. (1967) 'Notes on transitivity and theme in English', part 2, *Journal of Linguistics* 3: 199-244.

- Hartmann, R.R.K. and Stork, F.C. (1976) *Dictionary of Language and Linguistics*, New York: Wiley.
- Hutchins, W.J. (1988) 'Future perspectives in translation technologies', in M. Vasconcellos (ed.) *Technology as Translation Strategy* (American Translators Association Scholarly Monograph Series, vol. II), Binghamton, New York: State University of New York.
- León, M. and Schwartz, L.A. (1986) 'Integrated development of English-Spanish machine translation: from pilot to full operational capability: technical report of Grant DPE-5543-G-SS-3048-00 from the US Agency for International Development', coordinated by M. Vasconcellos, Washington, DC: Pan American Health Organization.
- McElhaney, T. and Vasconcellos, M. (1988) 'The translator and the postediting experience', in M. Vasconcellos (ed.) *Technology as Translation Strategy* (American Translators Association Scholarly Monograph Series, vol. II), Binghamton, New York: State University of New York.
- Santangelo, S. (1988) 'Making an MT system work: perspective of a translator', in M. Vasconcellos (ed.) *Technology as Translation Strategy* (American Translators Association Scholarly Monograph Series, vol. II), Binghamton, New York: State University of New York.
- Vasconcellos, M. (1984) 'Machine translation at the Pan American Health Organization: a review of highlights and insights', *Newsletter*, British Computer Society Natural Language Translation Specialist Group.
- (1985) 'Theme and focus: cross-language comparison vis translations from extended discourse', unpublished PhD dissertation, Georgetown University, Washington, DC.
- (1986) 'Functional considerations in the postediting of MT output: dealing with V(S)O versus SVO', *Computers and Translation* 1, 1: 21-38.
- (1987a) 'Postediting on-screen: machine translation from Spanish into English', in C. Picken (ed.) *A Profession on the Move: Proceedings of Translating and the Computer 8*, London: Aslib.
- (1987b) 'A comparison of MT postediting and traditional revision', in K. Kummer (ed.) *Proceedings of the 28th Annual Conference of the American Translators Association*, Medford, New Jersey: Learned Information.
- (1989a) 'Cohesion and coherence in the presentation of machine translation products', in *Georgetown University Round Table on Languages and Linguistics 1989*, Washington, DC: Georgetown University Press.
- (1989b) 'Long-term data for an MT policy', *Literary and Linguistic Computing* 4, 3: 203-13.
- Vasconcellos, M. and León, M. (1988) 'SPANAM and ENSPAN: machine translation at the Pan American Health Organization', in J. Slocum (ed.) *Machine Translation Systems*, Cambridge: Cambridge University Press.