Issue No. 22 (vol. 8, no. 1)
July 1999
ISSN 0965-5476
Published three times a year

MT News International

Newsletter of the International Association for Machine Translation

In this issue . . .

Editor's note . . .

MTNI is now in its eighth year. With this, our twenty-second issue, we start a parallel numbering system that will indicate the volume (or year) of publication. Members and other subscribers receive three issues a year. The next issue (#23, or vol. 8, no. 2) will be available in September at MT Summit VII in Singapore.

As always, we welcome letters to the editor, contributions in the form of news and features, and all offers of assistance, especially volunteers to head up the departments. —MV

Spotlight on the News

Compendium of Translation Software Now Available to Members

After greater delay than we would have wished, the long-promised Compendium of commercial machine translation systems and computer-aided translation support tools is now available free to IAMT members. Many members will have already received it, or will be receiving it very soon. This is still very much a *preliminary* edition; the intention is that later this year we will published a true 'first edition' to be put on sale to the general public (and at a very low nominal price to IAMT members).

In order to produce a listing which is as accurate and up-to-date as possible, members are being strongly urged to look in the Compendium for any systems and translation tools that they know about, to check whether they are listed and, if they are, to check for any errors or gaps in the information (however small). Any information about any systems should be sent to *WJHutchins@compuserve.com*. In the interests of being as reliable and authoritative as possible, the compiler would like members to provide full information regarding their sources, and, if possible, copies of vendors' publicity.

The Compendium is intended to provide a detailed listing of all systems of machine translation and computer-based translation support aids and systems that are currently available for purchase on the market. It should be noted that it does not, and will not, include any commercial systems still under development, nor any systems or translation tools of limited availability, such as those developed for particular

Continued on next page . . .

MT News International

Issue No. 22 (vol. 8, no. 1)
July 1999

. . . ***Compendium*** *continued*

clients by software or other IT companies.

In addition to information about products, the Compendium includes details of companies (mail and e-mail addresses, telephone, fax numbers, Websites) – which members are also asked to check – and two indexes, one for language pairs and the other for translation support tools.

Full information on what types of systems and tools the Compendium will cover and what kind of information it will include may be found in the Preface and in the section Arrangement of Entries.

IAMT members will have seen in the last issue of ***MTNI*** (#21) that AMTA has launched an initiative for the certification of MT products. Although the compiler has been involved in the discussions of the ad hoc committee working towards proposals to be put to the IAMT Council, it must be stressed that the categorisation of systems used in the Compendium should not in any way be regarded as 'official' IAMT categories.

There are a number of benefits for IAMT and its three regional associations that may well come as a consequence of publication of the first edition of the Compendium later this year. It ought to be the most complete and reliable source of information about products in this field, it should help to demonstrate to the public at large that the IAMT associations are the main sources of information and advice about machine translation and computer-based translation tools, and it may attract a wider range of new members to the associations. Because of this, the Compendium must be seen as a publication by the *IAMT membership as a whole*, and not just the work of one individual or group. It is particularly important that every member makes some contribution to its content, to ensure its accuracy and currency. Please help to make this a publication worthy of the IAMT and its associations. ❑

Systran, Infoseek Launch GO Network Translation Service

[Based on press release]

Infoseek Corporation, operator of GO Network, has introduced the GO Network Translator, an Internet translation service that lets users instantly translate entire Websites, e-mail, and any text in ten language combinations of French, German, Italian, Spanish, and Portuguese. The GO Network, launched in January 1999 by Infoseek in partnership with the Walt Disney Company, already has over 12 million registered users.

"While half of all Internet users have traditionally been English-speaking, that percentage is expected to decline over time. With the introduction of Translator, we are making the Internet a truly International medium by making it more accessible to more people than ever. By enhancing Translator with our own powerful front end and integrating it with GO Network, we're taking a great step toward achieving this," said Patrick Naughton, executive vice president of products for Infoseek.

Thanks to Infoseek's technology, the system's intelligent front end can preprocess and normalize a vast array of html types, and it can handle large, complex Web pages—entire Websites and even linked pages—with ease.

The GO Network Translator will soon have integrated links throughout the service in e-mail, chat, and search and directory pages, as well as across Infoseek's foreign language sites.

GO Network Translator can be accessed via the www.go.com home page or at http://translator.go.com. For further information, contact. Amanda Higgins, Infoseek Public Relations, tel. +1 (408) 543-6930, e-mail higgins@infoseek.com; Dave Allen, Infoseek Investor Relations, tel. +1 (408) 543-6664; e-mail: dallen@infoseek.com.

Visit SYSTRAN at www.systransoft.com. ❑

L&H iTranslator Service Now on Internet

[Based on press release]

Lernout & Hauspie announced that two European Internet service providers have signed on to offer L&H's iTranslator Internet translation service to their users. The agreements, awarded just one month after L&H announced the service which combines its machine and human translation expertise, point to growing market demand for fast, cost-effective translation services to facilitate multilingual communications. These agreements, will bring the service to many Internet and intranet users across Europe.

"We believe that use of the Internet and intranet to facilitate document translation delivery has unleashed a huge demand from multinational companies and users for translation services," said Florita Mendez, president of L&H's Consulting & Services division. "Our approach combines the Web with our unparalleled skills in human and machine translation to give customers a more easily accessible, reliable and affordable translation solution."

L&H iTranslator Service employs L&H's expertise in human translation with its machine translation software to provide fast, accurate translation of electronic documents over the Internet. L&H believes the administrative cost and turn-around time for the service is more affordable than competitive services because it leverages the Internet to help facilitate document and cost estimate delivery and receipt. The service, which is ideal for documents in a variety of sizes, is available to one-time as well as repeat customers.

The service offers human translation as well as MT (which provides a "gist" translation) of documents, e-mails, and Web pages. Human translation will be available in all major languages, including Western and Eastern European, South American and Asian languages. Specialized language domains for vertical markets, such as finance, computers, and medicine, help increase accuracy of machine translation.

Lernout & Hauspie: www.lhs.com. ❑

NL-Translex to Develop MT for Dutch

Demonstrating a commitment to machine translation, the Dutch Language Union recently announced its NL-Translex project and launched a call for tenders for the development of MT components that will handle unrestricted text and translate Dutch from and into English, French, and German. The work will be done within an already existing MT system. The components to be developed are intended for use by the translation services of official bodies of the European Union Member States and by the translation services of the European Union.

The development of these components is to take place within the framework of an MLIS project, in which the Dutch Language Union cooperates with several official bodies in the Netherlands and in Belgium.

Further information: Nederlandse Taalunie,, tel.: +31/70/346 95 48; fax: +31/70/365 98 18; e-mail: secr@ntu.nl; Website: /www.taalunie.org/_/actueel/translex.html. ❑

Accent LanguageWare Offers Free Internet MT

[Based on press release]

Accent Software International has debuted machine translation as a free Internet service available at their Website, LanguageWare.net. The machine translation is powered by Lernout & Hauspie's MT engine.

"Free machine translation will drive more business people to our Website and generate more interest in our automated, Internet-based translation service, which represents faster and more efficient methods of language translation," explained Todd Oseth, Accent's president and CEO.

"For rapid understanding of an e-mail or a paragraph in a document," he continued, "a customer can use LanguageWare's machine translation. For publication-quality translation, LanguageWare's rapid, high-quality translation is ideal and both are accessible over the Internet."

Translation of English to and from French, Italian, German, and Spanish is available today at no cost. Future enhancements, such as additional language pairs and the ability to accept longer documents, will be available in the future for a small subscription fee.

"LanguageWare.net offers users a direct translation portal through the Internet, enabling access from any location and from Macintosh, Unix, or PC operating systems. This portal expands our overall market opportunity and gives customers immediate access to all our translation services," said Accent's vice president of marketing, Bob Antoniazzi.

About Accent Software International

Accent Software International, founded in 1988, provides software products and services that help internationalize software and media in any language over the Internet. Accent's current products and services are enabling the Internet with Language Information Technologies that facilitate global communication.

Further information: www.accentsoft.com/ or www.languageware.net. ❑

Products

Transparent Language Announces Desktop Translator for Documents, E-mail, Web Pages

Transparent Language, Inc., a provider of language learning and translation software, has introduced Desktop Translator™, a comprehensive professional translation solution for Spanish, French, German, English, Italian, and Portuguese home office and small business users.

Desktop Translator translates a wide variety of documents, including plain text, e-mail messages, Web pages, documents in Microsoft Word and Corel WordPerfect, and more. Interface options—including the Desktop Translator toolbar, NotePad, and System Agent—provide users with the opportunity to select their personal preference for accessing translations. Users can also perform translations directly in their Web browser and in Microsoft Word or Corel WordPerfect applications. Desktop Translator translates Web pages and Microsoft Word and Corel WordPerfect documents with full format preservation.

Desktop Translator is based on the advanced TranscendRT™ natural language translation engine, the fastest and one of the most accurate computer translation technologies available, thanks to its superior analysis of the text. Moreover, Desktop Translator's interactive Dynamic Dictionary and links to 100,000-word supplemental dictionaries in each language allow users to quickly tailor translation results to best reflect the intended meaning of the source document.

Another feature is Desktop Translator's Dictionary Wizard, which allows users to add key terminology such as company and product names into a "do not translate" dictionary to improve the accuracy of their translations.

Desktop Translator is further enhanced with Dragon NaturallySpeaking speech recognition software from Dragon Systems, Inc., with full support for dictation, command-and-control, and text-to-speech functionality. Users now have the unique opportunity to dictate text into their PC, instruct Desktop Translator to translate the dictated text into Spanish, French, German, or Italian, and then listen to a high-quality audio synthesis of the translation. They can also choose to hear the English source text.

"Speech and translation are natural companions, and we are very pleased to have Dragon NaturallySpeaking as part of this exciting new program," said Roger Matus, vice president of North American Marketing at Dragon Systems. "The combination of Transparent Language's world-class translation technology with Dragon's award-winning speech programs will provide users with a useful application for multilingual communication and correspondence."

Transparent Language's TranscendRT engine powers the company's EasyTranslator™2, Transcend®, the Enterprise Translation Server™, and the TranscendRT SDK packages.

Availability and Pricing

With an expected street price of US$ 139.95, Desktop Translator will be available August 1999 directly from Transparent Language and through leading computer software retailers such as CompUSA, Best Buy, MicroCenter, Fry's, and others.

Further information: Transparent Language, tel.: +1 (800) 752-1767; fax: +1 (603) 465-2779; Website: www.transparent.com.. ❑

LanguageForce Launches Internet Product, UTStarGate™

LanguageForce Inc., a leader in Internet technology solutions designed to enhance global e-commerce trade and communications, has announced the launch of a new LanguageForce tool called UTStarGate™. This new product incorporates the latest innovations in IBM ViaVoice, Microsoft Explorer, and Universal Translator Technology to enable Internet users to instantly communicate across language barriers with anyone, anywhere in the world regardless of language. UTStarGate will be available via retailers, Internet service providers, numerous download sites, portals, and corporate intranets.

UTStarGate is a complete Internet, global communication solution that employs LanguageForce's proprietary Universal Translator Technology system.

The company claims it will provide fast, accurate translation of electronic documents, voice and typed messages, voice chat, and Web pages in ten different languages. Twenty additional languages will soon be available. Full keyboard and font support is available for all the languages, including Japanese, Chinese, Korean, and Arabic, on any PC-based operating system, including non-Asian operating systems.

UTStarGate is enabled for IBM ViaVoice Dictation, Voice Command, and Text-to-Speech; all applications are "voice-aware," allowing users to dictate thoughts and notes, hear spoken text in English or another language, and use voice-activated commands. It supports both IBM ViaVoice and Microsoft Speech.

Availability and Pricing

Pricing for LanguageForce's UTStarGate will vary according to how many languages users elect to use. As a launch promotion, it is available free at *www.UTStarGate.com*. UTStarGate will be made available in localized foreign versions that support Japanese, Chinese, German, Spanish, Portuguese and French by late 1999. Other LanguageForce products include Universal Translator™, Universal Translator Deluxe™, Universal OCR™, and the Instant Language 2000™ Learning Series (French, German, Russian, and Spanish).

Further information: LanguageForce, Inc., 1601 East Lincoln Avenue, Orange, CA 92865; tel: +1 (714) 279-9080; fax: +1 (714) 279-9368; Website: www.LanguageForce.com. ❑

From the Archives

Warren Weaver Memorandum: 50th Anniversary of Machine Translation

John Hutchins
Consulting Editor

In July of this year the MT community celebrates the 50th anniversary of one of its most significant milestones: the memorandum circulated by Warren Weaver, director of the Natural Sciences Division of the Rockefeller Foundation., to some 30 acquaintances on the possibility of using the recently invented digital computers to translate documents between one natural human language and another.

The Weaver memorandum was probably the single most influential publication in the early days of machine translation, since it formulated goals and methods before most people had any idea of what computers might be capable of, and since it was the direct stimulus for the beginnings of research first in the United States and then later, indirectly, throughout the world (Hutchins 1997).

Weaver had first mentioned the possibility of using the computer to translate in March 1947 in a letter to the cyberneticist Norbert Wiener and in a conversation with Andrew Booth, a British x-ray crystallographer, who was visiting various locations in the United States where computers were being built. In the following two years, he had been urged by his colleagues at the Rockefeller Foundation to elaborate on his ideas. The result was a memorandum, entitled simply "Translation," which he wrote in July 1949 at Carlsbad, New Mexico (Weaver 1949).

In his position at the Rockefeller Foundation, Warren Weaver was responsible for instigating and approving grants for major projects in molecular engineering and genetics, in agriculture (particularly for developing new strains of wheat and rice in Central and South America and Southeast Asia), and in medical research. He was a mathematician and had a special interest in probability and statistics. During World War II, Weaver had been seconded from the Foundation to head the Applied Mathematics Panel at the U.S. Office of Scientific Research and Development, where he directed the work of several hundred mathematicians on operations research of all kinds. Because of this background, he was fully familiar with the development of electronic calculating machines and well aware of the successful application of mathematical and statistical techniques in the deciphering of enemy messages.

The impact of Weaver's memorandum is attributable not only to his widely recognized expertise in mathematics and computing, but also, and perhaps even more, to the influence he enjoyed with major policy-makers in U.S. government agencies.

The memorandum began with a brief account of what had been done already to apply computers to the task of translation. Firstly, there had been some experiments with punched cards by Richard H. Richens and Andrew D. Booth in England, which had produced crude word-for-word translations of scientific abstracts (later published as Richens and Booth 1955). Secondly, there had been newspaper reports of a computer in Los Angeles which was intended to be used for simple experiments in translation. (Although Weaver does not say so, the computer was based at the Institute for Numerical Analysis at the University of California Los Angeles, a branch of the U.S. National Bureau of Standards, and the research was directed by Harry Huskey, who had previously worked on computers at Princeton University and the National Physical Laboratory in England.) These were, of course, just the beginnings, and Weaver was quick to point out the grave limitations of any simplistic word-for-word approach. He put forward four proposals. The first was that the problem of multiple meanings might be tackled by the examination of immediate context:

> If one examines the words in a book, one at a time through an opaque mask with a hole in it one word wide, then it is obviously impossible to determine, one at a time, the meaning of words. "Fast" may mean "rapid"; or it may mean "motionless"; and there is no way of telling which.
>
> But, if one lengthens the slit in the opaque mask, until one can see not only the central word in question but also say N words on either side, then, if N is large enough one can unambiguously decide the meaning. . .

The problem was, of course, to determine how much context would be required, and Weaver expected this to vary from one situation to another. However, Weaver thought that "relatively few nouns, verbs and adjectives" were actually ambiguous, so that the problem was not large. How wrong he was!

His second proposal started from the assumption that there are logical elements in language. He drew attention to a theorem proved by McCulloch and Pitts (1943)—developed in fact in the context of research on the mathematical modeling of the neural structure of the human brain—that "a robot (or computer) constructed with regenerative loops of a certain formal character is capable of deducing any legitimate conclusion from a finite set of premises." The mathematical possibility of computing logical proofs suggested to Weaver that "insofar as written language is an expression of logical character," the problem of translation is formally solvable. (A further influence in this regard may have been Rudolf Carnap's *Logical Syntax of Language* (1937), but Weaver does not refer to it, so this must remain speculation.)

The third proposal concerned the possible applicability of cryptographic methods. Weaver had been impressed at the success of cryptography based on, as he put it, "frequencies of letters, letter combinations, intervals between letters and letter combinations, letter patterns, etc. *which are to some significant degree*

independent of the language used" (emphasis Weaver's own). He illustrated with a wartime experience of deciphering a Turkish text. It had been given to a mathematician who, without knowing what the original language was, had succeeded in "recreating" the Turkish source text.

Weaver's ideas on cryptography were linked to *information theory*, which had recently been advanced by Claude Shannon. In fact, Weaver was writing a book about information theory with Shannon at the time (Shannon and Weaver 1949). The theory is concerned with the basic statistical properties of communication, including the effects of noise in telecommunication channels and of relative frequencies of signals. In particular, it embraced "the whole field of cryptography." Shannon was himself the author of one of the most influential texts on cryptography, originally written in 1945 but not published until 1949 (Shannon 1949), and it is quite likely that Weaver had seen it before publication. Weaver admitted that the validity of the cryptographic approach was difficult to assess, but he was obviously attracted:

It is very tempting to say that a book written in Chinese is simply a book written in English which was coded into the "Chinese code." If we have useful methods for solving almost any cryptographic problem, may it not be that with proper interpretation we already have useful methods for translation?

As it happened, researchers in machine translation were to recognize very soon the fallacy of Weaver's argument. The mistake lay in a confusion between the activities of decipherment and translation, which arise whenever the same person does both, as indeed is often the case in cryptanalysis.

For his fourth proposal, Weaver became more utopian. It was based on the belief that, just as there may be logical features common to all languages, there may also be linguistic universals. Earlier in his memorandum he commented on a paper by a sinologist, Erwin Reifler, who had remarked that "the Chinese words for 'to shoot' and 'to dismiss' show a remarkable phonological and graphic agreement." Weaver's comment was: "This all seems very strange until one thinks of the two meanings of 'to fire' in English. Is this only happenstance? How widespread are such correlations?" Obviously, Weaver thought that such universals may be very common.

At the end of the memorandum, Weaver asserted his belief in the existence and applicability of language universals with what is one of the best known metaphors in the literature of machine translation:

> Think, by analogy, of individuals living in a series of tall closed towers, all erected over a common foundation. When they try to communicate with one another, they shout back and forth, each from his own closed tower. It is difficult to make the sound penetrate even the nearest towers, and communication proceeds very poorly indeed. But, when an individual goes down his tower, he finds himself in a great open basement, common to all the towers. Here he establishes easy and useful communication with the persons who have also descended from their towers.
>
> Thus it may be true that the way to translate from Chinese to Arabic, or from Russian to Portuguese, is not to attempt the direct route, shouting from tower to tower. Perhaps the way is to descend, from each language, down to the common base of human communication—the real but as yet undiscovered universal language—and then re-emerge by whatever particular route is convenient.

Weaver realized, of course, that this approach involved a "tremendous amount of work in the logical structures of languages before one would be ready for any mechanization." However, he believed that some steps towards it had been made, particularly in the proposed Basic English of Ogden and Richards, which was then at the height of its popularity (Ogden 1930).

Response to the memorandum was mixed. Some rejected the very idea of mechanizing the complexity of translation, in much the same terms as many professional translators reject machine translation today. Others, however, were less negative.

One of the first to pick up on Weaver's proposals was Erwin Reifler, the sinologist. Over the next few months, Reifler proposed possible uses for crude word-for-word "translations," introduced the notions of *pre-editing* and *post-editing*, and suggested the use of regularized languages (Reifler 1950).

Another who rose to the challenge was Abraham Kaplan at the Rand Corporation, who followed up Weaver's suggested statistical approach to resolving problems of multiple meaning (Kaplan 1950). In addition, at the University of California Los Angeles there were some early studies on possible methods of automating syntactic analysis (Oswald and Fletcher 1951).

In the long term, however, perhaps the most significant outcome of the Weaver memorandum was the decision in 1951 at the Massachusetts Institute of Technology to appoint the logician Yehoshua Bar-Hillel to a research position. Bar-Hillel wrote the first report on the state of the art (Bar-Hillel 1951) and convened the first conference on machine translation in June 1952.

References

Bar-Hillel, Y. (1951): 'The present state of research on mechanical translation.' *American Documentation* 2 (4), 229-237.

Carnap, R. (1937): *The logical syntax of language*. London: Kegan Paul. [Translation of: *Die logische Syntax der Sprache*, Vienna: Springer, 1934.]

Hutchins, J. (1997): 'From first conception to first demonstration: the nascent years of machine translation, 1947-1954. A chronology.' *Machine Translation* 12, 195-252.

Kaplan,. A. (1950): 'An experimental study of ambiguity and context.' Santa Monica: The RAND Corporation. Repr. in: *Mechanical Translation* 2 (2), 1955, 39-46.

McCulloch, W.S. and Pitts, W. (1943): 'A logical calculus of the ideas immanent in nervous activity.' *Bulletin of Mathematical Biophysics* 5, 115-133.

Ogden, C.K. (1930): *Basic English: a general introduction with rules and grammar*. London: Kegan Paul, Trench, Trubner.

Oswald, V.A. and Fletcher, S.L. (1951): 'Proposals for the mechanical resolution of German syntax patterns.' *Modern Language Forum* 36 (3/4), 1-24.

Reifler, E. (1950): 'Studies in mechanical translation, no.1' [Seattle: University of Washington.]

Richens, R.H. and Booth, A.D. (1955): 'Some methods of mechanized translation.' In: Locke, W.N. and Booth, A.D. (eds.) *Machine translation: fourteen essays* (Cambridge, Mass.: Technology Press of the Massachusetts Institute of Technology), pp. 24-46.

Continued on page 15 . . .

Conferences

EAMT-99: MT in Eastern Europe

Accession to the EU, Other Multilingual Challenges Faced by the Community's Eastern Neighbors

Colin Brace
Language Industry Monitor

The delightful Czech city of Prague was the site of this year's EAMT Workshop, and our hosts, the famed Institute of Formal Applied Linguistics at Charles University, long a bastion of MT research. The 1999 Workshop, held April 21-22, 1999, was the first EAMT event to take place in Central Europe, and as such provided a welcome opportunity to learn more about MT in the former Eastern Bloc countries. The region boasts rich linguistic diversity and a long tradition of linguistic research, but until only very recently it has had comparatively less experience with the implementation of working MT systems.

The two-day programme offered a diverse mix of presentations from representatives of commercial software developers, government, and research, with a particular focus on challenges posed by forthcoming accession to the EU by a number of Central European countries, notably the Czech Republic, Hungary, Poland, Slovenia, and Estonia. Underrepresented at this year's event was the user perspective, customarily an important dimension of the EAMT workshops. The deficiency was undoubtedly due at least in part to the relatively underdeveloped market for translation technology in this region.

At the opening session on Thursday, April 22, Local Organizer Eva Hajicová welcomed participants, who came from at least 17 countries, to Prague, reminding us of the Institute's long involvement in MT.

In the first presentation, Pavlína Obrová, of the Czech Coordinative Centre for Translation of EU Materials, detailed her organization's efforts to bring Czech national law in line with EU law in the run-up to membership in the EU. A team of 35 translators is currently translating the 80,000 pages of the so-called *acquis communautaire* (the EU laws and regulations) with the help of various translation packages.

Dimitri Theologitis, Head of the Computer Translation Aids Unit of the European Commission Translation Service (SdT) in Luxembourg, gave a lively description of Commission policy with regard to the ever-increasing number of languages to be dealt with in his organization, possibly the largest users of MT in Europe.

Gábor Prózesky of MorphoLogic (Budapest) described how translators in Hungary have been using his company's dictionary server software for translation of the *acquis communautaire* into Hungarian. MorphoLogic has developed a number of innovative linguistic software packages for the Hungarian market. Elena Paskalcvea of the Bulgarian Academia of Sciences in Sofia also tackled the subject of translating the EU's *acquis communautaire*. Her organization has been working on aligning and extracting translation equivalents from EU documents in preparation for Bulgaria's application for EU membership.

Johannes Rizke, Lernout & Hauspie's GMS group, presented his company's strategy for developing tools for new languages. Rizke provided a fascinating glimpse into the challenge faced by L&H in providing a unified interface to the variety of translation packages it has acquired in the past few years. Among other things, his group is working on Serbian, Croatian, and Albanian, making it not too difficult to imagine in what sector its (potential) customer is active.

Svetlana Sokolova of PROject MT (St. Petersburg), who claims 50,000 users worldwide of its MT system Stylus, described the company's diversification into additional European languages, including English, French, and German language pairs, marketed by the Paris-based company Softissimo. Theó Hoffenberg of Softissimo then went into further detail about these new packages, marketed under the name Reverso. French-German (bi-directional) was launched in June 1998, and French-English (bi-directional) was released in April 1999.

Jan Hajic and colleagues at the Institute of Formal and Applied Linguistics in Prague have developed a translation system for two very closely related languages, Czech and Slovak. Hajic told the audience that what is essentially a word-for-word approach works well in this case, and the group has found at least one customer for its system within the software localization business. Further down the road, when Slovakia readies itself for EU accession, their system could also be deployed for translation of the EU legislative texts into Czech.

Poul Andersen of DGXIII at the Commission inventorized the surprisingly large number of translation support packages for Slavic languages.

During the course of the two days, presentations of a more theoretical orientation where offered by Igor Boguslavsky (Russian Academy of Sciences), Petr Sgall (Institute of Formal and Applied Linguistics, Charles University), and others.

For the year 2000, the EAMT is maintaining its focus on our Eastern neighbors, with a workshop planned in Ljubljana, Slovenia, in May. The exact dates and programme details will be posted on the EAMT Website, *www.lim.nl/eamt*, towards the end of the year. ❐

Commentary

"New Languages" Are Not "Virgin Languages"

EAMT-99 Workshop from the Eastern Point of View

Vladislav Kubson
Charles University, Prague

The decision to hold EAMT-99 in Prague reflected the Workshop's main topic, namely, the accession of Central and East European Countries into the European Union from the point of view of translation problems. The prospect of the increase of the current 11 official languages to potentially 22 after the accession of all candidate countries will not only mean that the number of language pairs will grow from currently 110 to potentially 462, but it will probably also increase the demand for the whole range of translation support tools, from databases of terms to full-fledged machine translation systems.

Designed with the intention to map the "possibilities, policies, and practicalities" of translation, the program ranged from talks of officials both from the European Union and from the Czech government through presentations of scientific results, demonstrations of commercial systems and presentations of industrial companies active in the field of MT.

The workshop had some very positive aspects. The composition of participants, who represented a mixture of people from industry, governmental institutions, and research, indicated that MT is still one of the fields of computational linguistics which has the potential to bridge the industry-research gap. There was an even more important aspect to be noticed: the industrial participants represented not only the developers of tools and systems, but also potential users, who came with the intention to find out what the developers and researchers can provide. The fact that at least some users feel the need to participate in such a workshop and that they actively seek help with an enormous task of translation and localization is quite important. On the one hand, it means that there is a good market potential for MT tools and systems, and, on the other hand, it also means that commercially available systems do not provide the help customers expect to get. That could have a positive influence on application-oriented MT research in the near future.

The positive aspects of the workshop prevailed from the general point of view. However, from the standpoint of researchers investigating the so-called "new" languages, not all impressions from the workshop were entirely positive. During the past decade we had time enough to get accustomed to the European version of market economy. Most of us working in the MT field understand that there is a long journey from scientific research to commercial application, that the development of an M(A)T system is very costly, and that no commercial company is going to invest into such a venture if the market for the product is not big enough. All that is quite clear—to us, at least. But it seems that some companies go by to different rules. This must be so, for otherwise Lernout & Hauspie would not invest in Albanian prior to investing, for example, in Czech, Slovak, or Bulgarian. Nor would they undertake all the development of systems for "new" languages (those above the threshold of minimum market size) from scratch, rather than trying to adapt linguistic resources that already exist. As Elena Paskaleva from the Bulgarian Academy of Sciences has pointed out: "New languages are not virgin languages." In several candidate countries there are numerous linguistic resources already available which might be useful for any kind of translation software. They represent many man-years of work and would provide a good basis for a successful commercial application. The example of the two successful companies specialized in NLP (Morphologic, Hungary) and MT (ProMT, Russia) shows that both the developers and resources available in these countries are comparable to those from "old" countries and that the arrogance of some companies is really not wise. The results presented during the workshop had shown that ProMT is a serious competitor to any company well-established on the M(A)T market, while Morphologic represents an example of a company founded by former researchers who are commercially exploiting the results of their previous linguistic research.

There is one more point worth mentioning. The languages of "new" countries belong to language types different from the current official languages of the European Union and pose a slightly different type of challenge for automatic translation. The attitude of Western companies "Our experts can handle everything and our methods are already good enough" (expressed on more than one occasion) may in the end give rise to the establishment of more Morphologics and ProMT's in the "new" countries. That would finally be a very positive outcome of the "Let's cooperate, but . . ." attitude.

Vladislav Kubon is a research at the Institute of Formal and Applied Linguistics at Charles University Prague. His main interest is syntactic parsing of free-word order languages. He has been active in MT since 1987. E-mail: vk@ufal.ms.mff.cuni.cz. ❑

MACHINE TRANSLATION SUMMIT VII
"MT in the Great Translation Era"

Singapore, September 13-17, 1999

It's Time to Register!

The seventh Machine Translation Summit, organized by the Asia-Pacific Association for Machine Translation (AAMT), will be held at Kent Ridge Digital Labs on the campus of the National University of Singapore on September 13-17. These days will be devoted to the following activities:

Mon, September 13	Tutorials		Reception
Tues, September 14	Main conference	Exhibition	
Wed, September 15	Main conference	Exhibition	Banquet
Thurs, September 16	Main conference	Exhibition	
Fri, September 17	Workshop		

In this issue of ***MT News International*** we give you an update of the final program and provide you with a registration form on page 11, which you can fax or mail to the MT Summit Secretariat. The Secretariat will also make your hotel reservations, so be sure to fill out the requested information on the back of the form (page 12).

An exact timetable, the list of accepted papers, and complete information about the program, including the tutorials and the workshop, "Machine Translation for Cross-Language Information Retrieval," is available on the conference Website: www.jeida.or.jp/aamt/mts99.html. Information about Singapore, the conference venue, accommodations may be obtained from the Website of the MT Summit Secretariat: www.krdl.org.sg/mts99/.

We hope to see you there!

Hozumi Tanaka, President, AAMT
Jun'ichi Tsujii, Chair, Program Committee
Low Hwee Boon, Chair, Local Organizing Committee

History of MT Summits

I	Hakone, Japan	September 17-19, 1987	
II	Munich, Germany	August 16-18, 1989	
III	Washington DC, USA	July 1-4, 1991	
IV	Kobe, Japan	July 19-22, 1993	AAMT
V	Luxembourg	July 10-13, 1995	EAMT
VI	San Diego CA, USA	Oct 29-Nov 1, 1997	AMTA
VII	Singapore	September 13-17, 1999	AAMT

Invited Speakers

PL1: Opening Session

Welcoming Remarks

Tsujii Jun'ichi *(Chair, Program Committee)*, University of Tokyo, Japan, and UMIST, UK

Hwee Boon Low *(Chair, Local Organizing Committee)*, Kent Ridge Digital Research Labs, Singapore

Area Policies

(TBC) Representative of the Singapore Government

(TBC) Representative of the Japanese Government Ministry of International Trade and Industry

Keynote Speech

Hozumi Tanaka *(President, International Association for Machine Translation)*, Tokyo Institute of Technology, Japan

Invited Talks

I1: Martin Kay *(MT research community)* Xerox PARC and Stanford University, USA

I2: Toru Nishigaki *(user/evaluation/social impact perspective)*, University of Tokyo, Japan

I3: Jo Lernout *(MT developers and vendors)*, Lernout & Hauspie Speech Products nv, Belgium

Special Talk (Banquet)

Makoto Nagao *(founder and first president of IAMT)*, Kyoto University, Japan

→

Special Sessions

S1: *"Controlled Language"*

Hiroyuki Kaji, *Coordinator*, Hitachi, Ltd., Japan

Uus Knops, LANT nv, Belgium

Teruko Mitamura, Carnegie Mellon University, USA

S2: *"Localization"*

Virginia Cha, *Coordinator*, Star+ Globe Technologies, Singapore

Michael Anobile, Localisation Industry Standard Association, Switzerland

Jeong-Yong Kim, International Transformational Information Corp., Korea

Nikolai Puntikov, STAR SPB Ltd., Russia

S3: *"Survey Reports on MT"*

Asia

Hemant Darbari, Center for Development of Advanced Computing, India

Aiping Fu, Chinese Academy of Social Sciences, China

Yoshiyuki Sakamoto, Tokyo Kasei Gakuin, Japan

Se-Young Park, Electronics and Telecommunications Research Institute, Korea

Europe

Rose Lockwood, Equipe Consortium Ltd., UK

North America

Elliott Macklovitch, University of Montreal, Canada

S4: *"MT and Speech"*

Key-Sun Choi, *Coordinator*, Korean Advanced Institute of Science and Technology, and Korea Terminology Research Center for Language and Knowledge Engineering, Korea

Hitoshi Iida, ATR Interpreting Telecommunications Research Laboratories and SONY Computer Science Laboratories Inc., Japan

Gianni Lazzari, Instituto per la Ricerca Scientifica e Tecnologia, Italy

Alex Waibel, Carnegie Mellon University, USA

S5: *"MT from the Research Perspective"*

Dominique Estival, *Coordinator*, Syrinx Speech Systems, Australia

Yusoff Zaharin, Universiti Sains, Malaysia

Christian Boitet, Joseph Fourier University, France

Anette Frank, France

S6: *"Views from Users"*

Hirosato Nomura, *Coordinator*, Kyushu Institute of Technology, Japan

Marjorie León, Pan American Health Organization, USA

(TBC) **Alan Sockett** , UK

S7: *"Multilingual Information Access"*

Vilas Wuwongse, *Coordinator*, Asian Institute of Technology, Thailand

Sung H. Myaeng, Chungnam National University, Korea

Yoshihiko Hayashi, NTT Cyberspace Laboratories, Japan

S8: *"MT for the Next Century"*

Tsujii Jun'ichi, *Coordinator*, University of Tokyo, Japan, and University of Manchester Institute of Science and Technology, UK

Eduard Hovy, University of Southern California Information Sciences Institute, USA

Shin-ichiro Kamei, NEC Corporation, Japan

Keh-Yih Su, Behavior Design Corporation, Taiwan

Yorick Wilks, Sheffield University, UK

Alan Barrett, Lotus Development, Ireland

Panels

PN1: *"International Cooperation from Funding Agencies' Point of View"*

Antonio Sanfilippo, *Moderator*, European Commission, Luxembourg

Roberto Cencioni, European Commission, Luxembourg

Ronald Larsen, Defense Advanced Research Projects Agency, USA

Gary Strong, National Science Foundation, USA

(TBC) Japan: Representative from Ministry of International Trade and Industry

(TBA) China:

(TBA) Korea:

PN2: *"MT Evaluation"*

Margaret King, *Moderator*, University of Geneva, Switzerland

Eduard Hovy, USC Information Sciences Institute, USA

Benjamin K. Tsou, City University of Hong Kong, Hong Kong

John White, Litton PRC, USA

Yusoff Zaharin, Universiti Sains Malaysia, Malaysia

MT SUMMIT VII REGISTRATION FORM

For hotel accommodations, see back of this page.

Send to:
MT SUMMIT SECRETARIAT
Attn: Ms Vicky Toh
Fax: (65) 779 6958 Tel: (65) 874 2003
E-mail: vicky@krdl.org.sg
Address: Kent Ridge Digital Labs
21 Heng Mui Keng Terrace
Singapore 119 613
(Mark *MT SUMMIT* on the envelope)

Last name: _____ First name: _____

Address: _____

_____ Country: _____

Affiliation (for badge): _____

Tel.: _____ Fax: _____

E-mail: _____ Website: _____

Registration

Please mark your choices by indicating the dollar amount on the blank line at the right. All rates are quoted in Singapore

Tutorials (Monday, September 13)

Fee includes coffee breaks and lunch. Participants may attend one tutorial in the morning and one in the afternoon. Please mark your choices and fill in the fee on the line at the right.

***Morning** (9:30 a.m. -- 12:30 p.m.)*

"A High-Level Overview of Machine Translation" — Eduard Hovy ☐

"English/Chinese MT: Past, Present, Future" — Zhendong Dong ☐

***Afternoon** (2:00 – 5:00 p.m.)*

"MT and the Localization/Translation Industry" — Jane Ee ☐

"MT and the Internet" —Mun Kew Leong ☐

S$150 _____

Conference (September 14-16)

General (with Proceedings) S$380 _____

Student (without Proceedings) S$200 _____

Conference fees include lunches and welcoming reception.

Banquet (Wednesday, September 15)

One person $ 50 _____

Extra person(s): ____ x S$ 50 = _____

Workshop (Friday, September 17)

"Machine Translation for Cross-Language Information Retrieval" $ 50 _____

Fee includes proceedings and coffee breaks.

Total Amount

S$ _____

A 3% goods and services tax (GST) will be added to the above fees.

Method of Payment

☐ By Visa/MasterCard (please circle one)

Card # _____ Exp. date ___/___

Name on card _____

Signature _____

☐ By check

Bank draft or money order in Singapore dollars made payable to KENT RIDGE DIGITAL LABS and mailed to the Secretariat.

Cancellation Policy

In the event of cancellation, there will a processing fee of $50. No refunds will be made after September 6. Substitutes may be accepted upon prior notice.

→

Hotel Accommodations

Both hotels offer: free return transfer from hotel to conference venue • complimentary hourly shuttle service to main shopping areas • swimming pool • coffee/tea making facilities in rooms • free cable TV • hair-dressing salon.

Please make reservations through the MT Summit Secretariat. Mark your choice in the space provided. Guests are expected to make payment to the hotel upon their departure.

Garden Hotel
14 Balmoral Road
Singapore 259800
Tel: (65) 235 3344; fax: (65) 235 9730

Single/double without breakfast	S$ 85 nett ☐
Single with American breakfast	S$ 95 nett ☐
Double with American breakfast	S$105 nett ☐

Arrival date: September _____ , 1999
Departure date: September _____ , 1999

Copthorne Orchid (Conference hotel)
214 Dunearn Road
Singapore 299 526
Tel: (65) 250 3322; fax: (65) 250 9292

Single without breakfast	S$110.00 ☐
Twin without breakfast	S$120.00 ☐
Single with American breakfast	S$120.00 ☐
Twin with American breakfast	S$130.00 ☐

Arrival date: September _____ , 1999
Departure date: September _____ , 1999

The Copthorne Orchid rates are subject to 10% service charge, 1% cess, and 3% government sales tax.

Conference Organization

Organizing Committee

Honorary Chair
Makoto Nagao
Kyoto University, Japan

Chair
Hozumi Tanaka
Tokyo Institute of Technology, Japan

Vice Chairs
Hwee Boon Low
Kent Ridge Digital Labs, Singapore

Jun-ichi Tsujii
University of Tokyo, Japan, and University of Manchester Institute of Science and Technology, UK

Executive Committee

Chair
Hirokazu Amemiya
Toshiba Corporation, Japan

Vice Chair
Hideki Morimoto
Fujitsu Ltd., Japan

Local Organizing Committee

Chair
Hwee Boon Low
Kent Ridge Digital Labs, Singapore

Secretary
Victorine Chen-Toh
Kent Ridge Digital Labs, Singapore

Program Committee

Chair
Jun-ichi Tsujii
University of Tokyo, Japan, and University of Manchester Institute of Science and Technology, UK

Vice Chair
Loong Cheong Tong
Kent Ridge Digital Labs, Singapore

Members
Virginia Cha, Singapore
Jason S. Chang, Taiwan
Zhao Xiong Chen, China
Key-Sun Choi, Korea

Program Committee (*cont.*)

Robert Dale, Australia
Zhendong Dong, China
Dominique Estival, Australia
Changning Huang, China
Hitoshi Iida, Japan
Etsuo Ito, Japan
Hiroyuki Kaji, Japan
Shin-ichiro Kamei, Japan
Asanee Kawtrakul, Thailand
Kunio Matsui, Japan
Se Young Park, Korea
Wanchai Rivepiboon, Thailand
Trihono Sastrohartono, Indonesia
Youlwon Seong, Korea
Satoshi Shirai, Japan
Keh-Yih Su, Taiwan
Chew Lim Tan, Singapore
Benjamin K. Tsou, Hong Kong
Dekai Wu, Hong Kong
Yusoff Zaharin, Malaysia

Updates

One Web, One Language: UNL

Jörg Schütz
Institute for Applied Information Sciences Saarbrücken

The Universal Networking Language Project (UNL) is an international enterprise under the auspices of the United Nations University (UNU) in Tokyo. The project was set up in 1996 at the Institute of Advanced Studies (IAS) of the UNU under the lead of Prof. Tarcisio Della Senta, director of the IAS, and Dr. Hiroshi Uchida, UNL project director. The mission of the project is to provide the methods and tools for overcoming the language barrier on the World Wide Web in a systematic way.

The UNL has been defined as a digital metalanguage for describing, summarizing, refining, storing and disseminating information in a machine-independent and human-language-neutral form. Other comparable systems exist for annotating sentence meaning, leading to a serious problem: the existing systems vary greatly in detail and conception, depending in many cases on the human language they were developed for, thus generating fundamental incompatibilities between them and inadequacies in dealing with different languages. A language-neutral metalanguage can circumvent this problem, permitting the coding, storage, dissemination, and retrieval of information independently of the original language in which it was expressed.

UNL can be seen as a kind of mark-up language which represents not the formatting but the core information of a text. As such it can be embedded in the eXtensible Mark-up Language (XML), and just as HTML/XML annotations already can be realized differently in the context of different applications, machines, displays, etc., so UNL expressions can also have different realizations in different human languages.

The principle of the UNL presentational system is its concept orientation, which is anchored in three basic mechanisms:

- Labeled links (binary relations);
- Universal words (UWs); and
- Attributes.

The UNL represents information and meaning sentence by sentence for each sentence in a given text. Sentence information is represented as a list of interrelated (semantic) labeled links, each between two of the concepts present in the sentence. Concepts are represented as character strings, called Uws, which are used to index the UNL knowledge base (KB). UWs can be annotated with attributes that provide further information about how the concept is being used in the specific environment of the sentence. The UNL KB is dynamic in the sense that it evolves as information is added. The semantic links that build structures out of concepts are signaled in human language texts by different grammatical means such as word order, agreement, suffixes, etc., in different languages. The links can also be interrelated in complex ways to represent very complex relations between concepts or groups of concepts—for example, coordinated structures. Thus, representations across sentence boundaries are possible, but they are not yet on the UNL development agenda.

It is important to understand that the UNL does not provide a single way of representing a given meaning. Rather, it provides tools (UNL editors and UNL-to-human-language generators) and a framework for exploring different alternatives for conceptual representations that are adequate for a wide variety of languages. During the development effort, sublanguages, or "dialects," of the UNL will surely arise. However, for worldwide acceptance as the digital language of the Web, the best of them will become de facto standards for the development community. Thus, the way is prepared for Web standardisation.

At this point in time it would be foolish to say that it is possible for the UNL to represent the full meaning of any word, sentence, or text in any language. Subtleties of intention and interpretation make the full meaning, whatever concept we might have of it, too variable and subjective for any systematic treatment. The UNL avoids these pitfalls of trying to represent the full meaning of sentences or texts, targeting instead the core or consensual meaning that is most often attributed to them. In this sense, much of the subtlety of poetry, metaphor, figurative language, innuendo, and other complex indirect communicative behaviour is beyond the current scope and goals of the UNL. Instead, the UNL targets direct communicative behaviour and literal meaning as a tangible, concrete basis for much or most of human communication in practical everyday settings. Therefore, the current application portfolio of the UNL comprises:

- Information dissemination, in particular the communication of scientific and technical information, product descriptions, and operating, repair, and service guidelines;
- Cross-language information retrieval and information filtering;
- Electronic services and electronic commerce;
- Personal e-mail and chat (dialogue settings);
- Personalized presentation of content.

Beyond this, applications with speech input and speech output are also imaginable within the above areas.

Recently the project got some very positive international press. For example, in Germany we had over 20 presentations and interviews in different media. Although the project was primarily intended to support the six official languages of the United Nations (Arabic, Chinese, English, French, Russian, and Spanish), in the meantime early prototypes of UNL technology modules already exist for Arabic, Chinese, English, Japanese, Russian, French, Hindi, Thai, Brazilian Portuguese, Italian, Spanish, Indonesian, Swahili, Mongolian, Latvian, and German in differing degrees of complexity in terms of the coverage of human language constructions.

Approximately 120 researchers and developers are collaborating worldwide in this important project.

Further information: www.unl.ias.unu.edu (Japanese coordinator), or www.iai.uni-sb.de/UNL (German collaborator). ❑

Users' Forum

Editor's note: The contribution below from Edith Westfall is based on her extemporaneous talk at MT Summit VI in San Diego. ***MT News International*** *has arranged to transcribe and summarize her presentation so that readers may benefit from her many cogent suggestions.—MV*

What Do Users Need From MT Systems?

Edith Westfall

Trados Corporation

There are four different categories of machine translation users that I would like to speak about here. *Casual users* make use of MT for general information. *Business users* are government agencies, large corporations, and small businesses that use machine translation to garner information and make business decisions. *Language professionals*, including translation bureaus, use MT to increase productivity. And finally, *multilingual developers* use MT by incorporating it in the products they offer in order to increase value added.

These four groups of users share at least three major characteristics in common: they are computer-savvy, they need to get information quickly in order to make decisions, and they have access to documents in electronic form. They are all faced with the challenge of the information explosion, and they need to have information that is not in their native language. Also, they must be sensitive to the needs of other countries in the course of conducting their business, and they may even be required by law to provide documentation and product information in a country's language.

End users want improvements in MT products in the areas of tools and standards. The most important issue for users is not the quality of output but the ease of use. They want to know how they can enhance the use of machine translation ino their business process.

Pre-Processing

In terms of pre-processing, end users often ask about different products and when they will be available. One important thing that users need is a translatability index: Is there a way to run a document through a tool to find out if the MT output is going to be useful? They may want analysis of a document based on the number of words not found, or they may want to know if the sentence structure is of a type that their MT engine can handle effectively. One way to determine translatability would be to run a grammar checker to see if the sentences are logically constructed.

Users also ask for filters in order to screen their electronic documents and be sure that all the text will export before putting them into text files. Users want to have spell-checkers built into their MT systems. They would like to be able to incorporate word lists that they have already developed. At the same time, they need a way of handling words that are not found in the MT dictionary, and they want to know when a word is part of a multi-word phrase. Finally, users need to be able to add words easily to their electronic dictionaries, glossaries, and word lists.

The Translation Phase

Users want to see improved quality in "raw" machine translations, and they would like to have optional ways of working on the MT system output, both interactively and in batch mode. Some users need to do batch processing. Users want the ability to process—that is, run translations—when they are inside other programs.

Users also want more language pairs, particularly non-English pairs and European language pairs. New programs should include instructions on how to use new features such as toolbars and templates. When users customize their dictionary, their output goes up, so they need a dictionary interface. There needs to be an efficient way to incorporate new word lists in their electronic dictionaries. There also needs to be a feedback loop, or channel, so that users can tell developers what terms they would like to see included in future updates or releases of the MT dictionaries.

Post-Processing

For the post-editing process, users are asking for the presentation of alternate translations that would be visible on the screen as they are working on the translation. Also useful are macros that allow the user to transpose or replace words, sentences, or sentence fragments. Any operation that makes the postediting process easier and faster is more cost-effective. Grammar editing and display of translation options are useful. In some instances it is better to deal with one sentence at a time, whereas in other cases it might be better to see whole paragraphs, or it may or may not be desirable to display the source text on the screen.

Operational Factors

Users want handbooks and guide books that give *clear* instructions on how to use machine translation and integrate it. Programs need to have manuals that give full information on how to use different features. Users need to know when it is best to use MT, and how MT fits into the rest of their business routine.

Users want their machine translation programs to be integrated into products they are already using. They want to have easy access to machine translation from their e-mail, from their browser, or from speech recognition softward. Train-

Continued . . .

. . . **Westfall** *from previous page*

ing and documentation are big issues for users.

Standards

It is essential that standards be developed for evaluating products. Such standards will help buyers to know more about the products, enabling them to make well-informed decisions and purchase products that fit their needs. There are more than 450 different MT products on the market now, and a great variety of different features within those programs. Users want to be able to compare products. They want to know what MT can and cannot do for them. They want more guidance on how to choose products and less hype from vendors and developers.

In terms of format, users want standardized tools that will enable them to move dictionaries from one product to another. MT users want to be able to call up different products on the same machine. Standards are time-consuming to develop, but they are really beneficial for all concerned—customers and developers alike.

Conclusion

In conclusion, the industry is very healthy, but it is clear that the users of machine translation want more. I would especially like to recommend that IAMT take the lead in developing standards in a timely fashion.

Edith Westfall has extensive experience with the translation process from both a business and a workflow perspective. She may be reached at the offices of Trados Corporation, in Alexandria, VA; tel: +1 (703) 683-6900; e-mail: edith@trados.com. ❐

. . . **Hutchins** *from page 6*

Shannon, C. E. (1949): 'Communication theory of secrecy systems.' *Bell System Technical Journal* 28 (4), 656-715.

Shannon, C. E. and Weaver, W. (1949): *The mathematical theory of communication*. Urbana: University of Illinois Press.

Weaver, W. (1949): 'Translation'. Repr. in: Locke, W.N. and Booth, A.D. (eds.) *Machine translation of languages: fourteen essays* (Cambridge, Mass.: Technology Press of the Massachusetts Institute of Technology, 1955), pp. 15-23. ❐

People

Mary Flanagan Founds New Business: Maverick Translation Solutions

Throughout her career, Mary Flanagan has been a trailblazer in the field of on-line MT applications. In December 1998, she embarked on a new path as founder of Maverick Translation Solutions (MTS) in Framingham, Massachusetts

Maverick develops custom software for integrating MT within corporate e-mail systems, Websites, search engines, real-time communication, and customer service applications. MTS also provides consulting and engineering services for MT evaluation, dictionary development, translation tools, and on-line service development. The Maverick staff has three linguists in addition to Dr. Flanagan: Philip Jensen, Sophie Toole, and Douglas Chinnock, all former members of the CompuServe Advanced Technologies Group. The company's clients include Transparent Language, Boss Broadcasting, and Lernout & Hauspie.

Mary Flanagan

Prior to founding Maverick, Dr. Flanagan served as Director of Online Translation Applications at Linguistix, a Bowne subsidiary, where she directed development of an Outlook-integrated e-mail translation service.

From 1992-1998, Dr. Flanagan led the pioneering research team at CompuServe that developed the world's first applications of online translation technology. The services she and her team developed included fully automatic translation of forum message boards, document translation, and chat translation. The CompuServe services introduced more than 2 million users to machine translation while translating upwards of 40 million words a year.

At MT Summit VI, Dr. Flanagan's team demonstrated a ground-breaking spoken chat prototype service for English and French using Dragon Dictate and Transcend RT. The prototype delivered spoken translations of chat room contents in near real time.

Dr. Flanagan has a patent pending for her role in the invention of that program, and already holds two patents for her invention of programs related to the online submission process for MT and for automated pre-editing of text for MT. A fourth patent is pending for a technique for protecting HTML code from MT.

As a computational linguist for Globalink, Dr. Flanagan helped develop the English-French GTS product. At INSO (formerly Houghton-Mifflin), she developed grammar and spelling correction rules for the company's GCS and ICS software.

Dr. Flanagan holds a Ph.D. in Computational Linguistics from Georgetown University. The author of numerous publications on on-line translation applications and MT evaluation, she is also a former Vice President of the Association for Machine Translation in the Americas.

For further information, contact Mary Flanagan, tel: +1 (508) 877-3430; fax: +1 (508) 877-1681; e-mail: mt4all@compuserve.com. Maverick Translations Solutions, 61 Nicholas Road, Suite B3, Framingham, MA 01701. ❐

In Memoriam

Gary G. Erickson (1940-1998)

Gary G. Erickson, President and CEO of GARJAK Research Inc., passed away on November 23, 1998, at the age of 58.

GARJAK was recognized as a key player in the machine translation of English/Chinese/Korean. Dr. Erickson was instrumental in developing a commercial machine-aided translation system, BI-LING Writer, for professional translators working in language combinations between English and Chinese/Korean/Japanese. With BI-LING Writer, translators are able to double or even triple their speed. He also successfully adapted BI-LING Writer for communication requirements for the Army and the Navy.

Dr. Erickson earned his Bachelor's degree in physics in 1963 from the University of Washington in Seattle, his Master's in engineering physics in 1971, and his Ph.D. in applied mechanics in 1976 from the University of California in San Diego.

Erickson was well known in the defense industry for the development of sophisticated microcomputer simulations of air base survivability and operability. In a separate business area, he directed the growth of artificial intelligence techniques applied to the computer translation of English and Far Eastern languages.

Dr. Erickson signed up GARJAK as one of the first corporate members of AMTA and remained active in the association over the years. Machine translation had always been his professional passion. He believed that MT will be indispensable in the new century. As computer technology explodes, new possibilities will open up for the integration of computer language processing into everyday communication between people.

Submitted by Dr. Mei-Du Li Erickson of GARJAK Research, Inc. GARJAK remains in business. Tel.: +1 (858) 625-9799; fax; +1 (858) 625-3818. ❑

Calendar

1999

August 5-8: Terminology Summer School, Krems, Austria. *Information: www.termnet.at/frm_tss99.htm.*

August 9-11: Training of Terminology Consultants, Krems, Austria. *Information: www.termnet.at/frm_tsc99.htm.*

August 9–20: ESSLLI-99: European Summer Schools in Logic, Language and Information, Utrecht, Netherlands. *Information: www.coli.uni-sb.de/esslli/.*

August 22–26: TMI-99: 8th International Conference on Theoretical and Methodological Issues in Machine Translation, University College, Chester, UK. Program chair: Francis Bond. *Information: www.ccl.umist.ac.uk/events/tmi99/.* **August 22**: tutorials; **August 26**: Workshop "Problems and Potential of English-to-German MT Systems," convened by Claudia Gdaniec. *Information: cgdaniec@us.ibm.com.*

August 22–27: HCI-99: 8th International Conference on Human-Computer Interaction, Munich, Germany. *Information: HCI99@iao.fhg.de.*

August 23-27: TKE-99: 5th International Conference on Terminology and Knowledge Engineering, University of Innsbruck, Innsbruck, Austria. Organized by Association for Terminology and Knowledge Transfer (GTW), International Information Centre for Terminology (Infoterm), and International Network for Terminology (TermNet). *Information: http://gtw-org.uibk.ac.at; tel.: +43 (512) 507-4261; fax: +43 (512) 507-2966; e-mail: peter.Sandrini@uibk.ac.at.*

September 13–17: MT Summit VII, Kent Ridge Digital Labs, National University of Singapore, Singapore. General Chair: Hozumi Tanaka; Local Organizing Committee Chair: Hwee Boon Low; Secretary: Victorine Chen-Toh; Program Committee Chair: Junichi Tsujii; Vice Chair: Loong Cheong Tong. *Information: www.krdl.org.sg Contact: Vicky, phone: +65 874-2005. fax: +65 776-8109.*

[See further details on pages 9-12 of this issue]

November 10-11: Translating & the Computer 21, London, UK. Organized by Aslib, the Association for Information Management, and supported by the Institute of , the British Computer Society , the International Association for Machine Translation, and the European Association for Machine Translation. Program chairs: Daniel Grasmick, Ruslan Mitkov, and Chris Pyne. *Information: www.aslib.co.uk/conferences/tc20form.html;tel.: +44 ((0)20) 7903 0000; fax +44 ((0)20) 7903 0011; e-mail: nicole.adamides@aslib.co.uk.*

2000

April 29-30: CLAW-2000: 3rd International Workshop on Controlled Language Applications, Westin Hotel, Seattle, Washington, USA. *Information: www.up.univ-mrs.fr/~veronis/claw2000.*

November 20–22: "MT-2000": International Machine Translation Conference, Exeter, UK. Organized by Exeter University and Natural Language Translation Specialist Group of the British Computer Society. *Information: David Wigg, Chairman, BCS-NLTSG, wiggjd@sbu,ac.uk.*

2001

September: MT Summit VIII, Santiago de Compostela, Spain. ❑