



Issue No. 34 (vol. 12, no. 1)

Winter 2003

ISSN 0965-5476

Published three times a year

MT News International

Newsletter of the International Association for Machine Translation

In this issue . . .

Special Conference Report	1
Announcements/Products . .	3
Conference Reports	4
Features	6
Conference Announcements	10

Of special note ...

- **Special MT Summit IX Report** (starts page 1)
- **In Memoriam: Antonio Zampolli** (page 2)
- **Conference Reports** (starts page 4)
- **Professor Tsujii's Inaugural Address the IAMT** (page 7)
- **AMTA 2004 Announcement** (page 10)

Special Conference Report

Machine Translation Meets New Orleans

MT Summit IX

October 2003

By Jaro Lajovic

What comes to your mind when you hear "New Orleans?" The cradle of jazz? Deep South? Romantic plantations? Be that as it may, it is probably not machine translation. However, some 200 persons will, from now on, remember the Queen of the Mississippi also for MT: from September 23-27, 2003, the city hosted the ninth MT Summit.

Welcome to New Orleans!

The conference was held at the elegant Fairmont hotel, originally dating from the 1890s, but luxuriously renovated and situated just a few minutes walk from the famous French Quarter. So in leisure hours it was possible to take advantage of the vicinity and plunge into the district in which the spirit of the times gone by can still be felt. One could opt for a pleasant day-

light walk through alleys of the Quarter—with their antiquity galleries, shops and restaurants—to Jackson Square with its attractive park, cross Decatur street to have a coffee in Café du Monde, finally making the last few steps to the Mississippi. On the other hand, with dusk falling one could wander along the same streets to visit quite a different city, throbbing with music and bubbling with people along Bourbon street, and visitors heading for a dinner, a drink—or a fortune teller, who were abounding in front of the cathedral behind the Jackson Square.

One can not but compare cities where the last three MT Summits took place: there was Singapore with its modern life beat, there was Santiago de Compostella with its magnificent scenery and tradition, and there was now New Orleans with colorful remnants of its almost three hundred years of history.

Continued on page 2 ►



MT News International

Issue No. 34 (vol. 12, no. 1)
Winter 2003

EDITOR-IN-CHIEF:

Laurie Gerber
Tel: +1 (619) 200-8344
Fax: +1 (619) 226-6472
E-mail: mtnei@eamt.org

MANAGING EDITOR

David Clements
E-mail: dclemen1@san.rr.com

CONSULTING EDITOR:

John Hutchins
E-mail: info@eamt.org

REGIONAL EDITOR, AAMT

Hitoshi Isahara
Fax: +81-774-98-6961
E-mail: isahara@crl.go.jp

REGIONAL EDITOR, EAMT

Jörg Schütz
Fax: +49 (681) 389-5140
E-mail: joerg@iai.uni-sb.de

Copyright 2003 by IAMT. Permission is hereby granted to reproduce articles herein, provided that they appear in full, and are accompanied by the following notice:

"Copyright 2003 International Association for Machine Translation (IAMT). Reprinted, with permission, from the IAMT newsletter, Machine Translation News International, issue #34, November, 2003."

Electronic copies available upon request: mtnei@eamt.org

MT Summit IX

...continued from previous page

MT in all settings, so to say. And that is not a bad denotation of the Summit itself.

MT in All Settings

The diverse main three-day program was preceded by a day of tutorials (such as Introduction to statistical MT) and followed by a day of workshops (including the workshop on teaching translation technologies, which is becoming somehow traditional).

The main conference program was every day introduced by an invited lecture, dedicated to: soaring translation demands and the role of MT (Donald Barabé, Canadian government translation bureau), to speech translation for hand-held devices (Akitoshi Okumura, NEC Corporation) and multilingual document processing at Xerox Research Centre Europe (Pierre Isabelle, XRCE). Afterwards, the program continued in 3 parallel sessions, comprising 72 talks and presentations grouped under several headings. Nine papers presented were dealing with MT evaluation, and (in the rapid deployment section) the same number focused on statistical MT, followed by talks regarding lexical issues with six contributions. The most numerous, but of course most heterogeneous was the system presentations group. Other discussed topics included EBMT, user studies, speech translation, translation memory and others. The conference itself was rounded out by interesting exhibits offered by a dozen companies.

Even if the facts above do not tell the whole story, they certainly make obvious that the organizers—program chair Elliott Macklovitch, local arrangements chair Florence Reeder and all others—faced a challenging task. Let me therefore without hesitation note that they performed up to expectations and did an excellent job.

Let this short summary be followed by a less dry view. As New Orleans spoke French in her youth, let's start this visit in her mother tongue:

Continued on page 13 ►

In Memoriam Antonio Zampolli

By Maghi King

Computational linguistics in general and machine translation in particular owe much to Antonio Zampolli. He created in Pisa one of the first European research groups, and by organising a series of summer schools there in the early 1970's played a major role in establishing and spreading a relatively new field of intellectual endeavour. The schools inspired enthusiasm and conviction amongst the many participants, who came from all over Europe to learn from the great and the good whom Zampolli invited to introduce them to a discipline still in its rambunctious and exciting childhood. An early COLING also took place in Pisa in 1973, and helped to confirm growing academic recognition of the field. Anyone who took part in these events cannot fail to have vibrant memories: there was so much excitement, so much discussion, so much the sense of being involved in something important.

A lot of this came directly from Antonio Zampolli: he was himself a passionate believer in the field, dedicated to creating a political and academic environment in which it could flourish, and full of excitement about what could be achieved. He never spared himself in pursuit of his vision. I remember being aghast at the amount of time he spent in aeroplanes, in meetings, on the telephone – pouring out energy in a seemingly endless stream. His enthusiasm could not help but rub off on those associated with him; only Antonio could have convinced me to write a draft of part of a proposal during a plane trip, wheedle my way into a University not my own to type it up and then spend a rainy day in Edinburgh searching for a fax machine (still rare at the time) to send him the result.

He had worked on machine trans-

Continued on page 9 ►

Announcements and Products for the MT Community

It's Time to Renew Your Membership in IAMT for 2004

As the end of the year approaches, it's time to renew your membership in the International Association for Machine Translation. How do you get to be a member? By joining one of the regional associations: AAMT (Asian Pacific Association for Machine Translation), AMTA (Association for Machine Translation in the Americas) and EAMT (European Association for Machine Translation).

Please see the association website or contact your regional administrator for information:

AAMT website: aamt.info

AAMT administrator: Kayoko Takada, takada@aamt.info

AMTA website: www.amtaweb.org

AMTA administrator: Priscilla Rasmussen, focalpoint@amtaweb.org

EAMT Website : www.eamt.org

EAMT administrator: Gisella Anspach, secretariat@eamt.org



AFNLP Website Serves the NLP Community

The AFNLP has a very useful website, with calls for papers and announcements of events that will interest people in the NLP community. Some past conferences have made their papers available online, and there are links to these in the "archives" section. This section also has the

handy feature of allowing search and retrieval of papers from these conferences. Conferences with papers so indexed include NLPRS 2001 and PACLING 2001.

See the AFNLP website: www.afnlp.org.



FreeTranslation Now Offers Human Services

[from press release]

FreeTranslation, a website that has offered free translation by SDL's Transcend translation engine is being upgraded to a full translation resource portal. The website, at www.freetranslation.com, will offer professional translation services from SDL International, as well as access to the instant translation service from within Microsoft Office.

FreeTranslation receives more than 1.5 million visitors per week and translates 25,000 words every minute. With the website enhancement, when the user receives an instant translation, they will see how much it would cost to have the same text professionally translated by SDL International. Also available is an "editing" service, a fast, inexpensive way to improve machine-translated text. Users continue to have the option of free, instant translation. However, when a text requires the highest standard of translation quality they can now submit it for fast-turnaround, professional translation.

In addition, the FreeTranslation portal now plays host to a subscription service that allows you to obtain translations directly from your desktop. SDLDesktopTranslator is a downloadable toolbar, providing instant translation from within Microsoft Word 2000/XP or

Microsoft Outlook 2000/XP via an online connection. This subscription service costs less than \$40 a year. The toolbar also provides a button in Office from which to initiate professional translations.

Webmasters can also add FreeTranslation's services to their own website using one of the custom designed applets, or you can personalize your browser by adding buttons to your toolbar, allowing you to translate web pages and text on the Internet with one easy click.

FreeTranslation.com is owned by SDL International, a leading multilingual services and software solutions company.

Press Contact: Fiona Camden;
Email: freetranslation@sdl.com

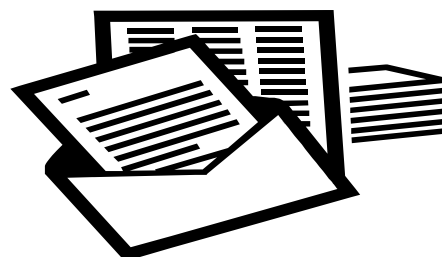


Got News? Be a Contributor to MTNI

Got an idea for MTNI? Have an urge to see your name in print? We need editorials, letters, news and features with relevance to the MT community.

If you've got a news item or a story that you'd like to see published in MTNI, just contact Laurie Gerber, mtni@eamt.org, David Clements, dcllemen1@san.rr.com, or one of the regional editors.

We want to hear from you!



Conference Reports

AAMT 2003

June 2003

Tokyo, Japan

The Asia-Pacific Association for Machine Translation (AAMT) Annual Conference, which included the 2002 AAMT activity report and lectures, was held at the Chuo University Surugadai Memorial Hall in Tokyo, Japan on June 20, 2003.

Reports

Following the opening remarks from the President of AAMT, Professor Jun'ichi Tsujii (University of Tokyo), three reports were presented: (1) "Machine Translation (MT) Evaluation using TOEIC* (Test of English for International Communication)," by the Chair of the Technical Research Committee, Dr. Hitoshi Isahara (Communications Research Lab.); (2) "Market Research on Vendors and the Questionnaire Survey on End Users," by the Chair of the Market Research Committee, Professor Yoshiyuki Sakamoto (Tsukuba Women's College); and (3) "Evaluation of Machine Translation Systems on the Internet," by the Chair of the Network Translation Workshop,



Professor Makoto Nagao, President of Kyoto University

Professor Shoichi Yokoyama (Yamagata University).

Lectures

(1) "Machine Translation; the Past, the Present, and the Future," by Makoto Nagao.

President Makoto Nagao of Kyoto University, a leading researcher in MT, looked back upon the history of Japanese MT research in this lecture and talked about a wide range of technical problems.

(2) "Knowledge Innovation by Machine Translation," by Nobutoshi Hatanaka.

Mr. Nobutoshi Hatanaka of Canon Inc. reported on the MT work they are tackling in their company. Their dictionary maintenance and Internet interface helped it expand from several users at the time of its introduction to 350. Its use increased in their business planning and customer service sections, which enabled them to meet the requirements for prompt replies. He mentioned that it also improved productivity by allowing the saved time and cost



**Professor Jun'ichi Tsujii,
AAMT President**

to be devoted to creative work.

(3) "Intercultural Collaboration Experiment among Japan, China, Korea, and Malaysia (ICE2002)," by Toru Ishida

Professor Toru Ishida at Kyoto University reported on a joint experiment of open-source software development conducted by a multinational team in their

Continued on page 14 ►

The Localization Community Considers MT

Localization World

October 2003

Seattle, Washington, USA

The localization community has recently been thinking much more seriously about machine translation. Recent presentations by major corporations at LISA (Localization Industry Standards Association) forums suggest that many have initiatives underway to explore operational use of machine translation.

At the recent Localization World conference held October 14-16 in Seattle, Washington—organized by the Localization Institute, GALA (the Globalization and Localization Association) and Multilingual Computing magazine—there was a full day pre-conference workshop, led by Jaap van

der Meer, joined by Laurie Gerber. Later in the conference, Laurie gave a presentation on recent trends in machine translation research, as well as the commercial world. Jaap van der Meer also spoke on a panel about the return on investment (ROI) from increasing automation in the translation process. Mr. van der Meer's company, Cross Language, has invested considerable effort in researching this subject, and coming up with realistic, quantified terms for understanding the benefits of introducing machine translation, and the cost of leaving strategic materials

Continued on page 14 ►

CICLing-2003

February 2003
Mexico City, Mexico

This is the first part of a two-part report on CICLing-2003. It will be continued in the next issue of MTNI. — ed.

By Alexander Gelbukh

There was good news and bad news about CICLing-2003 (www.CICLing.org), the 4th Conference on Intelligent Text Processing and Computational Linguistics held in February in Mexico City. The good news was the significantly increased number of submissions and, as a consequence, participants—even though this year, only the papers that were unanimously accepted by the reviewers were included in the program. The bad news was that it was a rather tight squeeze for all comers to get in the hall and, above all, in a 46-seater bus for far all-day excursions, see below. The Proceedings volume published by Springer-Verlag in the series Lecture Notes in Computer Science (www.Springer.de/comp/lncs), N 2588, was a record-sized 648 pages.

Keynote Speakers

Perhaps one of the reasons for such a vivid interest to the conference was a good selection of keynote speakers. Each keynote speaker in one day presented a formal talk published in the Proceedings, and in another day organized an informal discussion.

In his keynote talk, Aravid Joshi—the first scientist awarded, in 2002, the Lifetime Achievement Award by the ACL—argued for the utility of complex primitives to simplify grammar development. To put it simple for those who did not understand much of the lexicalized tree-adjoining math, his motto was: “Complicate locally to simplify globally!”

Following the appeal, Ted Pedersen in his keynote talk complicated things enough by presenting a generalization of his adapted Lesk algorithm for word sense disambiguation. He compared

various measures of semantic relatedness between words and showed that the combination of the adapted Lesk algorithm with the Jiang-Conrath relatedness measure gives 38% accuracy on a sample of Senseval-2 data with WordNet senses.

But Eric Brill of Microsoft seems not to like complications. His keynote talk caused perhaps the most controversial opinions, starting from the very title: “Processing Natural Language without Natural Language Processing.” On numerous examples from different applications Eric showed that extremely simple methods combined with a really *huge* amount of raw language data can surprisingly outperform sophisticated knowledge-rich approaches or state-of-the-art machine learning algorithms trained on traditional (that is, tiny) text corpora. Indeed—he said,—thousands of people have devoted decades to, and wrote tons of papers on, perfecting sophisticated tagging algorithms; if we had instead stuck with the very first tagger developed 30 years ago and all these 30 years one grad student in the world were busy full time marking up a corpus to train it, then we would have now a better tagger than what we call today the state of the art! Thus—no more eggheads sharpening their wits on formalisms and learning algorithms! What we need to advance is data, data, more data!

And Adam Kilgarriff, *our last keynote* speaker, knows where the data is. You’ve guessed right: in the Web. The Web is so great, so interesting, so exciting, you can do so many interesting things with the Web—Adam couldn’t help singing when he began with his talk! (And he actually sang!) However, said Adam, the access to all this treasure is difficult for the linguists since the major Internet search engines do not present the results in the way good for linguistic statistical research. A solution for the computational linguistics community would be to build our own search engine, with its own crawling agent, which would present the information we need the way we need, including lemmas, parts of speech, syntactic constructions, etc.

This suggestion caused controversial opinions among the public. Stephen

Nightingale noted that such marks would be assigned according to a specific theory or formalism, which might render the whole thing unusable for all researchers not working in frame of this theory. Alexander Gelbukh mentioned that for some class of statistical research, first of all collocations, much smaller special corpora can be constructed through existing search engines, as described in his last year’s CICLing paper. In any case, Adam’s idea is something worth the closest attention of the CL (and MT) community.

Presentations of Interest

In total there were 67 presentations (including posters) by 150 authors from 23 countries, so that we cannot give here even a short mention of each of them. The following ones seem to be of most interest for the MT community. Grzegorz Kondrak presented an algorithm that, given a bilingual wordlist of cognates, catches the regular correspondences that in 90% of cases allow predicting the words in one language given the words in another one: how to say *sprint* in Japanese? Stephen Nightingale discussed accumulating translation resources from parsed parallel corpora. Donia Scott explained how to generate text with a specific style, an ability with most NLG—and consequently MT—systems lack. The style choice affects nearly all levels of language, from text planning, discourse structure and layout to syntax and lexical choice. Bengt Nordström presented a system for multilingual syntax editing in the Grammatical Framework, a grammar formalism built upon Logical Framework introduced by Martin-Löf, which can be used for multilingual authoring of documents in restricted languages.

Dictionary and lexicon was discussed in a number of presentations. Patrick Cassidy and Tom O’Hara discussed Factotum, a Roget-derived ontology with the idea similar to that of WordNet but a significantly different vocabulary. Stan Szpakowicz showed how WordNet can be integrated with Longman’s dictionary. Marie-Laure Reinberger, Caroline Varaschin-Gasperin, and

Continued on page 15 ►

Special Feature: Speaking of MT

A Chat with Dr. Heyan Huang, Founder and CEO of HuaJian Technology

Huajian Technology is a Chinese MT Developer that has been quite successful in China, but is not that widely known outside. Dr. (Ms.) Heyan Huang, the company's founder, kindly provided us with some company background information, and agreed to an interview to let the rest of the world know more about her company. — ed.

Background

Huajian was incorporated in 1997, and 10 years before its incorporation, its founders had begun to do MT research. Until now, the company, with a registered capital of 50 million RMB, has more than 11 subsidiaries across the country. The business of Huajian covers e-government, e-commerce, digital library, system integration, embedded application software and specific industry-oriented application system etc.

Business

At the beginning, Huajian focused on computer language information processing and machine translation business. However, today Huajian's business has expanded quickly to fields of e-government, e-commerce, digital library, system integration, embedded application software and specific industry-oriented application system etc. Through close cooperation with big names such as China Telecom, China Unicom and PCCW, Huajian developed different information processing platforms accommodating partners' various demands. In tandem with developing commercial applications, Huajian submitted total e-government solutions as well, to different departments of the central and local governments, satisfying their needs to streamline working process and honor particular missions ranging from environ-

mental protection to highway management etc.

Technology and Awards

Through years of development, Huajian has become a national leading total solution provider in computer language information processing and relevant application fields. For this objective, Huajian attaches great importance to technological innovation that can be evidenced by several honors Huajian has been awarded from 1993, for example, the First Prize in National Scientific & Technological Progress (the highest prize of scientific and technological invention in China), the First and Second Prize in CAS Scientific and Technological Progress. Moreover, Huajian has got 8 inventive patents and 43 computer software copyrights, which are Huajian's core intellectual property. Accordingly, in 2002, Huajian passed ISO9001 International Quality Administration System Authentication, and obtained "First-grade Qualification Certificate of Computer Information System Integration" as well as "Certificate for Computer Information System Integration Qualification Involving State Secrets."

Human Resources

Huajian attributes its success to the continuous innovation in technology, which is undoubtedly accomplished by its energetic and creative staff. Young as the staff members are, over 90% of them hold a bachelor's degree, and 30% master's or doctor's degree. These people are devoted to Huajian, and vice versa, Huajian regards them as invaluable treasures.

Huajian's Specialty—Computer & Language Information Engineering Research Center.

One of Huajian's distinct features is its perfect integration in "production, education and research." This is com-

pletely reflected by the establishment of Computer & Language Information Engineering Research Center. As a pilot research institute under CAS, it is completely sponsored and controlled by Huajian Group. With a high-level scientific research team led by doctoral tutors and computer experts, the research center is entitled to grant Master's and Doctor's Degree, and able to educate 15-20 master or PhD candidates per year. Since its establishment, it has been undertaking key research projects of the state and international cooperation, which has provided Huajian with advanced scientific and technological achievements as well as back-up talents. The research center's specific influence on Huajian will definitely accelerate Huajian's development.

Interview

MTNI: You mentioned that your Computer & Language Information Engineering Research Center can grant Masters and PhD degrees. Is it then considered an accredited university in China? Or is there some other organizational status that can grant degrees there?

HYH: The Computer & Language Information Engineering Research Center is an academic research arm affiliated with the China Academy of Science. This center has a close relationship with several famous Universities in China, such as the University of Sciences & Technology of China and Tsinghua University and the Polytechnic University of Nanjing. Therefore, some Master's and PhD applicants can be instructed by me and other professors under this center. Upon completing their research work and papers here, they will be granted Master's and PhD degrees in the name of the relevant universities.

Besides accredited universities, the state-owned academic institutions in China, such as CAS, Chinese Academy of Social Science, People's Bank of China (the central bank) etc, can grant Master's and PhD degrees.

MTNI: What is your background, and how did you decide to start an MT company?

HYH: I graduated and obtained a PhD from the Institute of Computing Technology, CAS in 1989, since then I have engaged in the research and development on machine translation and information process. When we were aware that our MT system was developed so well and the relevant market was growing bigger, we decided to incorporate this company to expand MT business. Since Huajian's incorporation, I have been working here and in charge of 7 projects covering national research programs, government-sponsored projects. Now, I am in the capacity of board's chairperson and president of Huajian, and at the same time, I still work as the professor of University of Science and Technology of China as well. Generally, I instruct about seven to eight Master's and PhD students every year.

MTNI: Some people are concerned about selling software in China because of the lack of protection against illegal copying of software. Is this changing? And do you feel this is an issue for Hua Jian?

HYH: Several years ago, there was really a problem on copyright of software. However, the situation is changing fast. More and more people understand the importance of copyright protection, and the parliament and government issued laws and rules to protect it. Therefore, I feel that copyright protection is not a very serious problem for Huajian anymore.

MTNI: You offer quite a number of language pairs involving Chinese. (English-Chinese, Chinese-English, German-Chinese, Russian-Chinese, Japanese-Chinese, Chinese-Japanese, Chinese-Spanish, and Chinese-French. See "Data Driven MT," MTNI #33). What percentage of your market is in China? Have you localized any products for foreign markets?

HYH: In the mainland of China, Huajian's MT products consist mostly of network-translation platforms, professional translation tools, bilingual Internet browsers, bilingual translation tools for PC users and embedded translation software solutions. They have occupied markets such as telecom, publishing

Continued on page 14 ►

Professor Jun'ichi Tsujii's Address to the IAMT

The following speech was given by Professor Tsujii upon assuming the role of IAMT President during MT Summit IX in 2003. —ed.

Diversity, Cooperation and Understanding

I am very honored to succeed Eduard Hovy as president of IAMT, who has been dynamic and devoted to the development of IAMT. IAMT has had a series of distinguished and devoted presidents, and I realize that it is very hard for me to match their performance.

Our association has several unique characteristics as an association. Because of the nature of our field, machine translation, the association is truly international. We have three daughter associations, AMTA, EAMT and AAMT, which cover different regions in the world. Since the regions they cover are different from each other in terms of their linguistic, cultural and social backgrounds, they have had different focuses and aspirations and subsequently have different ways of managing themselves. Due to the rapid development of information technology and subsequent surge of demand for translation among different languages, the daughter associations will have to cover regions that they have not covered yet. Expansion of regions to cover may bring further diversity of needs and demands of members. IAMT, as well as the daughter associations, will have to cope with this inevitable challenge.

Our association also has diverse membership. Our members include computer scientists and engineers who are interested in recent technological development, professional translators who are the major users of MT systems, managers of multinational companies who are concerned with productivity and quality of translation, and

MT vendors who are interested in, among others, future markets for their products. Such heterogeneity of membership is the unique strength that distinguishes our association from other associations of academics or professional translators. By putting people with different backgrounds together, we have been successful in promoting synergy among these constituencies and keeping our field intellectually relevant to society.

Though MT is the center of our interest, we have witnessed that recent development of NLP and its related fields have accelerated, and with this some historically separate fields, including MT, are being integrated into the whole of human language technology. Almost all NLP applications have become multilingual in nature and MT has come to constitute the core of all NLP applications, such as multilingual IR, multilingual document management, multilingual automatic summarization, etc. At the same, we cannot ignore technological developments in other fields like speech recognition and synthesis, ubiquitous computing environments, etc. Our association has to be open to and aware of developments in these wider technological fields and their potential impacts on the field of MT.

Co-operation and understanding will be the key for success of our association as such. We have to understand the needs and demands of members with different backgrounds who live in different linguistic and social realities.

The challenges we will face are many, but I hope I will do my job with help of you all.

Professor Tsujii teaches in the Department of Computer Science at the University of Tokyo. He can be reached at: tsujii@is.s.u-tokyo.ac.jp.



Feature Article: From the Garage to the Attic

An Insider's View of Entrepreneurial MT

By David Clements

The Lernout & Hauspie saga first brought welcome attention, and then unwelcome attention, to the machine translation and language technology world. This installment of David Clements's story continues a first person account of the early years of MicroTac and Globalink that began in MTNI 32. David Clements, a veteran MT developer, is also the AMTA managing editor of MTNI. —ed.

Part 3: Hitting the Road

It was late 1990, and a new version of Language Assistants was about to appear. Version 4 of the Language Assistants was the first to attempt "real" translation. Several things led to this transition into MT. First, MicroTac hired a new software engineer, Jim Hicke, another friend of Tac's from the Young Adult Group days. Tac had also been a frequent visitor to the foreign language forums on CompuServe (FLEFO), and started testing small translations by posting them on FLEFO to see what users' reactions would be. Feedback from the translation crowd was mixed; some were very enthusiastic and encouraging, others made it plainly clear that "engineers should stay out of translation." Tac, however, remained both resilient and optimistic. There was probably nothing that annoyed a set of pessimistic detractors, such as the human translation community, more than a bright, optimistic, articulate spokesman for MT.

A Lively Upstart

There was a broad sense in which MicroTac was an irreverent, lively upstart. Not only was it poised to open the general public's eyes for the first time to user-friendly consumer MT, but also it brashly proclaimed a new era of accessible translation in the MT commu-

nity itself. This had some far-reaching effects on both the MT and HT industries.

Jim Hicke brought a new level of professionalism to the young company's software development "team." That "team" consisted only of Tac and Jim, but represented more than a doubling of resources. Tac began to be able to look at the "big picture" of MT, and Jim took over most of the inner mechanics of the programs. The software was no longer programmed in Pascal, but C++. Jim instituted a kind of rewrite rule system, which in later incarnations came to be known as Barcelona. Most importantly for us as an office, the addition of a second programmer meant we had to look at how to share work. We had no network then, and used floppy disks (irregularly) as backups. After a few incidents in which source code was mangled or lost, the company instituted a source control program, and assembled a network.

The telephone was "ringing off the hook" with new orders. Most of our sales work at that time was direct ordering. Customers would read advertisements and call the offices to place orders, usually small or medium. Spanish Assistant was very popular in Texas, California and Florida, among schools, churches and government agencies. French Assistant caught on in Europe, Australia and Canada. I used to exchange almost weekly letters with a human translator in Arizona, who just loved Spanish Assistant, but always wanted to help us improve the product in ways that would help him, specifically. For example, he was working on plumbing and roofing documents, and wanted to enrich the lexicon with terminology of construction contracting. Then there was Monash University in Australia, which always bought French Assistant, loved the product, and offered to help us fix the "flaws." One of my favorite letters (in the days before e-mail blossomed, hard-copy letters were the surest form of customer support after the telephone), was from a

French user who sent reams of lexical mistakes in French Assistant to me, demanding with Gallic righteous linguistic indignation: "Where did you learn to speak French?" The software only translated from English to the "foreign" language, but many customers found all the features useful, particularly the "verb-conjugator," dictionary and accent-entry utilities. The included Random House dictionaries were bilingual, even though the translator was not. There was a built-in text editor, where the users could type their documents in English, translate them and save output files in line-by-line or side-by-side format. Included also were the grammar-help files, but as time went on, they received less emphasis as the product moved from a language-learning program to a full-featured translation program.

We had so many orders coming in by phone, that all of us assumed telephone duty when necessary. As the only linguist on the staff (and thus not in as *important* a position as the engineers), my job was split between answering the general telephone, doing written and phone tech support, and taking sales orders. Sometimes the volume was overwhelming, and none of us on the first line of phone "defense" was able to pick up a ringing receiver. So, occasionally, Tac or Jim would take a call, though this came to a screeching halt one day when one of them, after answering three calls, slammed the receiver down and shouted, "Let me pass you to someone who gives a damn!"

Engineering vs. Linguistics

This occasion marked the beginning of one of the fatal flaws of the MicroTac organization: what I later came to call "the two-class system." From 1991 until the final days of the office as part of Bowne Global Solutions, engineers were considered a "class above" everyone else in the office, especially the linguists. Certainly this isn't unusual in the NLP industry, and there are many factors involved, including the much higher job-demand for engineers than for linguists. Nevertheless, since MicroTac was iconoclastic in so many ways, it was disappointing over time to observe the "class system" developing, with repercussions that would finish by

wasting time and talent.

MicroTac was very much an engineering-driven company. Though Tac was fascinated by language, he remained an engineer at heart. This, combined with an overweening entrepreneurial spirit, meant that he often preferred a “blunt-force trauma” approach to language issues, rather than subtle but time-consuming linguistic research and development. The success of this approach meant that MicroTac soon achieved sales in the \$3 million annual range. The downside was that this attitude also carried on long after the founders left the company, and imbued its corporate culture for the duration of the MT group. Since everything had to be developed on the fast track, there just wasn’t time to think through a lot of issues, or correct mistakes.

Since I was only working part-time in the early 1990s, Tac decided he needed more linguistic resources. Besides me, the only other “linguist” on staff then was our office manager, who, probably not coincidentally, was also a longtime friend of Tac. This office manager not only handled all phone duties, but also worked on Spanish and German. Since Spanish was our biggest seller, Tac looked for help in that area first, and hired his sister. This sister had a degree in linguistics and was fluent in Spanish, and seemed a good match for the company. She did provide some theoretical underpinnings for the direction of the Language Assistant lexicons. Try as he might, however, Tac was unable to convince her to stay on with the company, as her interests lay in field linguistics, not NLP.

Direction for Linguistics

Before she left MicroTac, Tac’s sister was involved in hiring our first “director of linguistics,” John O’Hara. John was fluent in German and Spanish, which was a plus for the young company. The direction of the hiring program then headed towards finding more native speakers of German and Italian. Somehow engineers often believed that finding linguistically apt native speakers (with or without formal linguistics education) was “easy” compared to finding other capable engineers. With this attitude, the company advertised in local papers and some

academic publications, and had very unfavorable results. Tac paid all the moving expenses for one “hard-core German linguist” to come work for us, only to fire the unfortunate fellow two months later. Another German consultant we hired was a blue-haired “school marm” type who fondly recalled to me one afternoon, as we pored over the German lexicon, “In Germany in the 40s we never had to lock our doors, there was no crime.” We were particularly hard pressed to find Italian linguists. One whole set of Italian Assistants was botched because the consultant we had hired insisted that Italian used the “personal a” like Spanish. Also, a lot of the Italian references we looked at were inconsistent regarding placement of diacritical marks, so the verb conjugator feature was badly implemented. The company hired one non-native Italian speaker as a lexicographer, only to fire her after a few months, when one of the engineers complained her work “messed up” his programming. (This was after we had already hired two more engineers).

Looking for a Lexicon

All this lexicography work was necessary because there were issues with the Random House dictionaries we had licensed. Not only were Random House’s dictionaries small, but also there were errors that we were unable (or Tac was unwilling) to fix. The biggest problem was one of control. As long as the dictionaries belonged to Random House and not MicroTac, fees would have to be paid to Random House for every Language Assistant sold, and yet MicroTac would be unable to add to or correct the lexical data itself. The solution Tac and the company’s VP, Garet Juntunen, came up with was to drop the Random House dictionaries and develop our own. The only problem with this was, as always, there was “no time” to do the job right. So, the company went in search of the fastest and cheapest solution: pre-developed, ready-to-license dictionaries. They found these in Utah, with a company called Transoft.

A deal was hastily cut with Tran-

Antonio Zampolli

...continued from page 2

lation to some extent during the early days of the Pisa group, but his major involvement came with the birth of the Eurotra project. From the very start he was keen to follow everything that was going on. He used his splendid mind to dissect suggestions and ideas, and his negotiating skills to bring about agreement between opposing intellectual parties. Perhaps even more important than any of this, he was always there, always at the end of the telephone to give advice when asked or simply to listen when one needed to think something through.

Thus, what I shall remember most is not so much his intellectual talents, great though they undoubtedly were. His warmth of personality and his kindness will linger long, as will his sometimes outrageous sense of fun. Who could forget the mock bull fights at the Granada LREC, or the time he hid the bus during a Eurotra excursion, causing major panic amongst the excursion organisers and justly punishing them for their pomposity? To each of us our own memories: all will agree that time spent with Antonio was never boring, whether in a meeting, in a working group, in a conference session—or in a dinner party.

Antonio Zampolli was a visionary whose like we may not see again. He always seemed to know what was needed next and was prepared to devote himself ferociously to achieving it. He was a European par excellence, who thought instinctively on a large scale. He played a key role in defining, stimulating and supporting research in the European Union. But above all he was a man, rich in feeling and unafraid to show it. We are the poorer for his passing.

Editor’s note: Antonio Zampolli was the founder of ELRA, www.elra.fr.



Continued on page 15 ►

Conferences and Events

AMTA-2004

**September 28-October 2, 2004
Washington DC, USA
Preliminary Announcement**

AMTA 2004 will return the MT community to the historic site of the 1954 GAT (Georgetown Automated Translation) project, one of the first operational tests of machine translation. In fact, the ideas developed during the GAT project formed the foundation for a number of well-known MT systems that are still in operation today. We'll hear from and about pioneers in the industry and about the progress in 50 years.

One of the founding goals of

AMTA was to bring together researchers, developers and users (a group that includes

sponsors, evaluators, and observers of MT). AMTA 2004 seeks to maintain the high quality technical content that has become a hallmark of AMTA events, while significantly increasing participation by users and prospective users, of translation technology. This event will include presentations and panels including current users, as well as reaching out to the upcoming user community at translation and business schools. We would particularly like to encourage dialog among the three groups. This will give users and future users an opportunity to learn about what is in the pipeline in research and development, as well as giving researchers and developers a chance to hear the interests, concerns and desires

of users.

Commercial Exhibit

The conference will include a 2-day exhibition of commercial machine translation products and services. The exhibition provides a wonderful opportunity for commercial developers to meet with prospective customers and show off their latest products, while giving shoppers a chance to compare offerings side by side.

Research and Deployment Showcase

The conference will include a "science fair" of current university research systems, as well as deployments of existing products. This showcase will emphasize US Government use of translation technology, giving participants from various government agencies

and offices the chance to see how others are making use of translation technology.

Related Events

As usual, this conference will include a day of tutorials aimed at

making the conference presentations more accessible and giving participants a chance to boost their knowledge outside their specialty. Special interest groups will organize full day hands-on workshops. Users and prospective users of MT can gain insight on operational use of MT from a half-day excursion to PAHO, the Pan American Health Organization, which has been using MT to produce first-draft translations in a production translation operation for 25 years. We may also fit in a visit to the recently opened and hugely popular Spy Museum in downtown Washington DC.

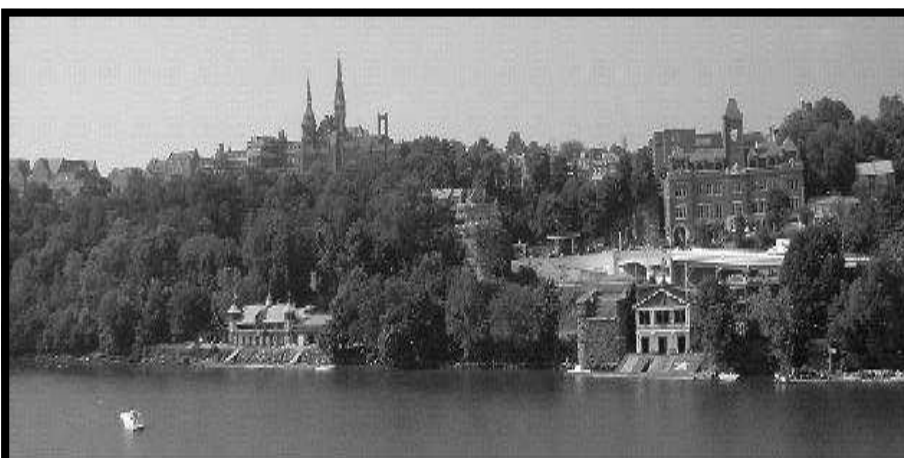
For updates, please see the conference website:

www.amtaweb.org/AMTA2004. □

Mark Your Calendar for these Upcoming Events!

EAMT 2004 Workshop
April 26-27
Malta
See: www.eamt.org

MT Summit X
2005
Phuket, Thailand
Organized by AAMT
See: www.aamt.info



September 28-October 2, 2004

Georgetown University, Levy Conference Center, Washington DC

CICLing-2004

February 15-21, 2004

Seoul, Korea

The Fifth International Conference on Intelligent Text Processing and Computational Linguistics will be held in Seoul, Korea, from February 15-21, 2004. It will be hosted by ITRI of the Chung-Ang University.

Keynote speakers include Martin Kay of Stanford, Philip Resnik of the University of Maryland, Ricardo Baeza-Yates of the University of Chile and Nick Campbell of ATR, Spoken Language Research Labs, Japan. The topics of the talks will be posted on the conference Web site as they are made available.

In general, we are interested in whatever helps, will help eventually, or might help *computers* meaningfully process *language* data. These include, but are not limited to, some of the following:

Computational linguistics research:

- Computational linguistic theories and formalisms
- Representation of linguistic knowledge
- Morphology
- Syntax
- Semantics

Intelligent text processing and applications:

- Document classification and search
- Information retrieval
- Information extraction
- Text mining

The conference is intended for the exchange of opinions between scientists working in different areas of the growing field of computational linguistics and intelligent text processing. Our idea is to get a general view of the state of art in computational linguistics and its applications.

For more information, including registration fees and deadlines, please see the CICLing-2004 Web site: www.cic.ipn.mx/~gelbukh/cicling/2004/. □

Asian Symposium on Natural Language Processing to Overcome Language Barriers

Hainan Island, China

March 25-26, 2004

The Internet has made it easy to access to vast amounts of information and to connect with people all over the world. However, there are numerous barriers to effective information access and efficient communication. One of these is language differences. This symposium aims to bring together NLP researchers in Asia and to discuss how language barriers can be overcome by applying Natural Language Processing.

The symposium will be held immediately following the 1st International Joint Conference on Natural Language Processing (IJC-NLP-04). The official language will be English.

The symposium is organized by The Institute of Electronics, Information and Communication Engineers (IEICE); The Technical Group on Natural Language Understanding and Models of Communication (NLC); and The Technical Group on Thought and Language (TL). The symposium is supported by the Communications Research Laboratory (CRL)

Symposium co-chairs are Hitoshi Isahara, Communications Research Laboratory; and Yoshihiko Nitta, Nihon University. The program committee chair is Tatsuya Izuha, Toshiba

Corporation.

Call for Papers

Papers are solicited on topics with a focus on NLP aimed at overcoming of language barriers. In addition to completed work, work in progress and ongoing projects are welcomed. Topics of interest include, but are not limited to: Cross-lingual information access and communication; Cross-media information access; Information access and communication for the handicapped and the aged; Basic technology for making language barrier-free.

For more information, see the symposium website: www.rcl.cityu.edu.hk/ijcnlp04/satellite.htm or contact program chair Tatsuya Izuha, tatsuya.izuha@toshiba.co.jp. □

Symposium Important Dates

Submission deadline	December 2, 2003
Notification to authors	December 23, 2003
Final papers due	February 23, 2004
IJC/NLP-04	March 22-24, 2004
Symposium	March 25-26, 2004

IJC/NLP-04

Sanya, Hainan Island, China

March 22-25, 2004

The 1st International Joint Conference of Natural Language Processing is being organized by the Asia Federation of NLP associations (AFNLP). Sponsoring organizations include: The Chinese Information Processing Society of China, The Association for Natural Language Processing of Japan, and The Association for Computational Linguistics

The conference program will include four thematic sessions designed to gather people with the same special interest, and let them meet each other at a specific time-space to discuss and exchange ideas. The thematic sessions are: Natural Language Learning using Both Labeled and Unlabeled Data; Natural Language Technology in the Text Processing User Interface; Mobile Information Retrieval; and Text mining in Biomedicine.

See: www.cipsc.org.cn/IJC/NLP-04/. □

International Workshop on Spoken Language Translation

September 30 - October 1, 2004
Kyoto, Japan

Call for Papers

The importance of spoken-language translation technology is increasing with our increased chances for cross-language communication in face-to-face and telephone conversation, especially in the domain of tourism.

Novel technologies have been proposed to tackle the problems in spoken-language translation research. A number of institutes are developing huge bilingual or multilingual spoken-language corpora. MT technologies based on machine learning, such as statistical MT and example-based MT, are being applied to the translation of spoken language by using these huge corpora. Some of the characteristics of spoken language seem suitable for the application of machine-learning-based MT in comparison with written language. However, there is still no concrete standard methodology for comparing the translation quality of spoken-language translation systems.

Prominent Research Activities

One of the prominent research activities in spoken language translation is the work being conducted by the Consortium for Speech Translation Advanced Research (C-STAR III), which is an international partnership of research laboratories engaged in automatic translation of spoken language. Current members include ATR (Japan), CAS (China), CLIPS (France), CMU (USA), ETRI (Korea), ITC-irst (Italy), and UKA (Germany). One of C-STAR's ongoing projects is the joint development of a speech corpus that handles a common task in multiple languages. The creation of such a corpus will not only enable translation among multiple languages but will also facilitate exchange and discussion of research results among member labs. As a first result of this activity, a Japa-

nese-English speech corpus comprising tourism-related sentences, originally compiled by ATR, has been translated into the native tongues of C-STAR members.

Evaluation Campaign

In this workshop, an "evaluation campaign" of spoken-language translation technologies will be held by using the multilingual speech corpus containing the tourism-related sentences developed by ATR and C-STAR members. Two types of submissions are invited: 1) participants in the evaluation campaign of spoken-language translation technologies, and 2) technical papers on related issues. Evaluation will focus on Chinese->English and Japanese->English translation of the Basic Travel Expression Corpus (BTEC). Both subjective and automatic evaluation methods (BLEU, NIST, WER, etc) will be included.

Paper Topics

The workshop also invites technical papers related to spoken language translation. Possible topics for the session include, but are not limited to:

- MT Evaluation Measures
- MT Algorithms
- Word / Phrase Alignment
- Multilingual Lexicon / Translation Rule Extraction
- Multilingual Parsing

Workshop Important Dates

Evaluation specifications Jan. 15, 2004
Application submission Feb. 15, 2004
Notification of acceptance Mar. 1, 2004
Papers/posters notification Jan. 23, 2004
Camera-ready submission Sept. 17, 2004
Workshop Sept. 30-Oct. 1, 2004

4th Workshop on Asian Language Resources

Hainan Island, China
March 25, 2004

This is a satellite workshop to the IJCNLP conference and immediately follows it. Language resources play a very essential role in corpus-based, stochastic, and learning approaches to natural language processing research. The 4th workshop on Asian Language Resources will be held with the following purposes:

- To investigate and discuss the problems related to the construction, dissemination and NLP research based on Asian Language Resources
- To establish the collaborative effort on Asian language resource construction, management, keeping the resources accessible, distribution, and sharing
- To launch the Asian Language Resources roadmap
- To publicize the status of Asian language resources to researchers in other regions

Following the former three preceding workshops, we invite papers on all topics related to language resources, in particular Asian language resources and their development including, but not limited to:

- Text corpora
- Lexicons
- Grammars

For more information, see the workshop webpage: www.rcl.cityu.edu.hk/ijcnlp04/workshops1.htm.

Workshop Important Dates

Submission deadline Dec. 12, 2003
Notification of acceptance Jan. 10, 2004
Camera-ready copy Jan 24, 2004
Workshop March 25, 2004

MT Summit IX

...continued from page 2

Sacré BLEU!

It is common knowledge that MT evaluation is a tough and multifaceted issue. The recently (2001) proposed BLEU metric (and its variant NIST) was thus greeted as a major step in the field. However, in the meantime the advantages and disadvantages of the approach are being realized and—not surprisingly—a vivid and interesting discussion accompanied the presentations dealing with this topic. BLEU was the buzz-word, for better or worse: wherefrom the title. The take-home message: the n-gram metric does have a useful role for the MT development; it should, however, be used with utmost caution as the basis for formal evaluation. The specific developers' view of evaluation was complemented by Dr. Margaret King's lecture on FEMTI (Framework for MT evaluation in ISLE), emphasizing the variety of evaluation perspectives, including the users'. In practical terms the latter was also discussed in a presentation on scalability of MT systems as one of the most important aspects of assessment in practice.

Statistical MT

The second thematic central point—not only by the number of contribu-



A Meeting of the Minds at MT Summit IX

tions—was the statistical MT. The once alternative approach recently attracted much attention, to a great extent also due to the circumstances after September 11. To put it briefly, the approach seems to have come of age. This claim is not confirmed only by the release of the first commercial SMT system by Language Weaver, but also by the demonstrated ability of statistical techniques in providing the possibility for rapid development of

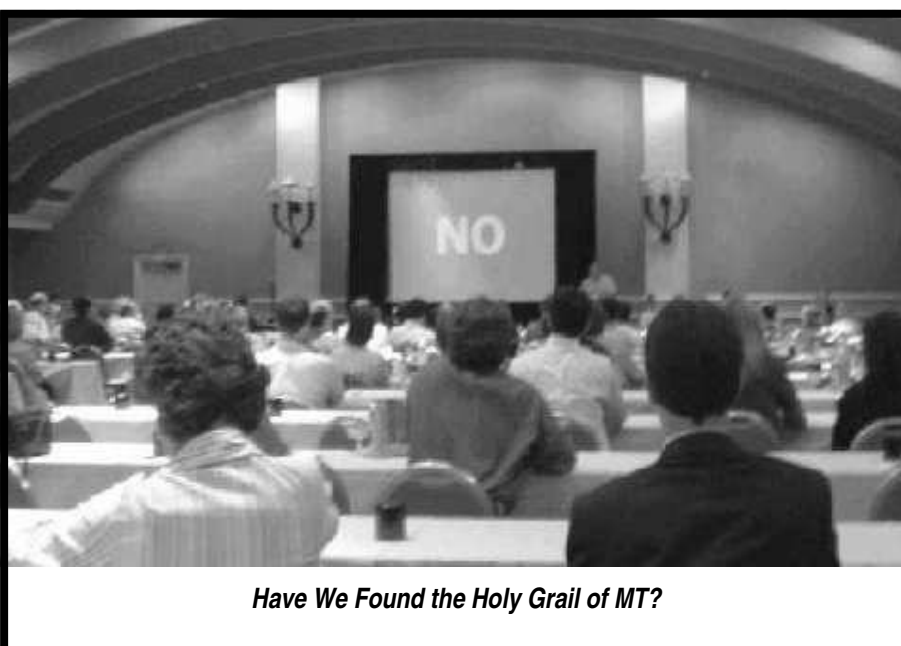
MT in response to unexpected requirements, as well as by projects like TransType2, embedding SMT within an interactive translation environment.

User-Oriented Focus

And finally there was a third, more personally colored and heterogeneous group of very stimulating papers, like the one on the rapid postediting service development at Swiss-based language services provider CLS, presentation of MITRE's Foreign Language Resource Center or Dr. John Hutchins' most interesting analysis of whether MT systems improved over the last decades (to satisfy your curiosity: the preliminary and tentative answer seems to be 'no'). Maybe the unifying theme of the majority of presentations in this third group would be their user-oriented focus. And mentioning users it should be added that they seemed somewhat underrepresented group among the Summit participants—this being most probably the only reproach to the conference.

Have We Found the Holy Grail?

This was the provocative title of the closing panel discussion led by Elliott Macklovitch and joining as panelists



Have We Found the Holy Grail of MT?

Continued on page 15 ►

Localization World

...continued from page 4

untranslated.

The upcoming LISA conference to be held in Washington DC December 8-12 likewise includes a full-day MT seminar, in this case led by MT industry expert Mary Flanagan.

To explore the online localization community, see:

www.crosslang.com
www.gala-global.org
www.lisa.org
www.localizationinstitute.com
www.multilingual.com

□

AAMT 2003

...continued from page 4

first languages (Japanese, Chinese, Korean, and Malaysian). With a built-in MT bulletin-board and Internet browsing system, the students attempted to conduct their daily communication and knowledge sharing by sending information in their first languages. During the “self-initiated repair” (refinement by adjustment) process, they modified and adapted the documents to be translated satisfactorily. Then, during the “collaborative repair” (refinement by cooperation) process, when the message on the board was unclear, they asked the writer and confirmed their intention. Through this collaborative process, the usefulness of MT was confirmed.

In this conference, it was recognized that MT is used not only by professional translators, but also by general end users who conduct overseas correspondence and multilingual communication. It is now a reality that MT has entered into more practical use at long last. It will be extremely interesting to follow the further development of MT.

*TOEIC is a general trademark of the Education Testing Service (EST).

□

Heyan Huang

...continued from page 6

industry, digital library, governmental departments, etc.

Huajian constructed four large-scale network translation platforms successively for China Telecom, China Unicom, China Jitong Telecom and Luneng Group. In addition, Huajian developed multi-lingual information process platforms for several governmental departments. Furthermore, Huajian won the bid for construction of a multi-lingual translation platform for the Organizing Committee of the 2008 Olympic Games.

With regard to embedded translation software solutions, Huajian produced the first pocket translation device with Hong Kong Sense Group, and realized total revenue of 1.27 billion HKD. Now Huajian is developing a translation chip for mobile phone under the sponsorship of the Ministry of Information.

Concerning MT products such as professional translation tools, bilingual Internet browsers and bilingual translation tools for PC users, Huajian either established its own distribution channels to reach end clients or cooperated closely with several giant IT companies in China to bundle MT products with millions of computers sold in Mainland China. For the purpose of expanding overseas markets, Huajian set up a joint venture with the most reputable Telecom and IT company in Hong Kong, PCCW, to sell Huajian’s MT products in Hong Kong and other markets.

MTNI: You mentioned revenue of 1.27 Billion HKD from your pocket translator.

In what year was that, or over what time period? Is the pocket translator something like a phrase book? Or an open-domain machine translation system?

HYH: The revenue had been materialized over three years in the last decade. The pocket translator has an embedded open-domain machine translation system that can translate almost any input sentence automatically without restriction.

MTNI: Concerning the “translation chip” for mobile phones, does this

mean that there would be an additional microprocessor inside mobile phones that would just do translation? Would this be text or speech translation? Or maybe SMS?

HYH: An additional microprocessor will be installed inside mobile phones. The input information like speech, conventional text and handwriting text will be processed by the translation microprocessor. Short messages will be processed and translated automatically by another system that will be operated jointly by Huajian and relevant telecom operators.

MTNI: By the way, when I visit your website, there is a button that says “English,” but nothing happens when I click it, so I don’t know if you have an English website or not. Is there an address I can use to go directly to the English website?

HYH: I feel very sorry about the inconvenience of browsing our English website. Our website is under construction for an updated version and the new English version is not completed. In several weeks, you can browse Huajian’s English website.

MTNI: You mentioned that the English website is under construction. Have any of your products been localized for English or other non-Chinese markets? For example, do you have any products where the user interface and documentation are available in other languages?

HYH: All our machine translation products featured translation between Chinese and other languages. The end users can easily select the interface of the Chinese version or the interface of the English and other languages. As I mentioned before, our products have been sold to markets like the USA and Canada.

See: www.hjtek.com

□

MT Summit IX

...continued from page 13

Hermann Ney, Dekai Wu, Steve Richardson and Ed Hovy. The question is of course rhetorical—and the panel's answer was the expected “no.” But obviously a new chapter has begun, opening new perspectives and possibilities, while still making it recommendable that lessons learned in the previous one not be forgotten.

Summit Proceedings Available

Last but not least: proceedings of the MT Summit IX will be available shortly on the AMTA's website, www.amtaweb.org, while an unofficial gallery of photos taken by Dr. Adriane Rinsche can be found at www.langtech.co.uk/eng/summit.asp.

Jaro Lajovic is a physician and medical translator based in Slovenia:
jaro.lajovic@mf.uni-lj.si.



CICLing-2003

...continued from page 5

Héctor Jiménez-Salazar presented different ontology-extraction tools that learn lexical relationships from text corpora and build semantic clusters of words. Stefan Bordag suggested an approach to build dictionaries that better suite the needs of language processing tools; he uses, again, the relationships found in the very texts. Accordingly, corpus linguistics issues were tackled in several presentations. For example, Ted Pedersen showed his n-gram statistics package designed to be easily used by other researchers incorporated into their programs.

To be continued in MTNI #35....

Alexander Gelbukh can be contacted at gelbukh@cic.ipn.mx, www.gelbukh.com.



Garage to the Attic

...continued from page 9

soft, without a thorough linguistic “due diligence” having been performed. The result was that we had our first real crisis, as we needed Transoft's bilingual dictionaries in order to produce the next-generation, fully bi-directional Language Assistants. Jim and a couple of the IT staff members had to fly to Utah to help Transoft prepare the lexical data for delivery. Then as the situation grew more grave, Tac flew out. Every day, in our common room, the rest of the staff would get updates on a blackboard from Jim. Since Jim was a pilot, he used aeronautical metaphor: “Spanish: on time. French: on radar. German: late. Italian: not on radar.”

Lexical Lemons

After weeks of grinding away in Utah, Tac and Jim brought the incomplete lexicons to San Diego. Our job was now to fix them and finish them. Unfortunately, the data were widely inconsistent. Transoft's terminology came largely from human-translator terminology data, i.e., from translation jobs that Transoft had done. Thus, it wasn't terribly surprising that their best-performing database was the Spanish one, since that is typically the most commonly translated language in the USA. Even though Transoft was owned by an Italian businessman, the weakest lexicon was Italian. Even as I used to plow through the data, I would find mistakes that would make me laugh, such as the Italian translation for “ambivalent” as “omosessuale” (homosexual).

Besides the mistakes, we also needed to define what features we would place on lexical entries. Given the many hundreds of thousands of words we had just “purchased” from Transoft, this was also a daunting, time consuming and resource-intensive task. Actually, defining this feature set was another of our weak spots, as we never developed a consistent set of features

to be applied across languages, and never worked out a hierarchical order for the features we did use. Finally, since the purchased data was specific to translation jobs, we had issues with incompleteness. Some common words would be missing from one language's lexicon or another, and some translations for common words were skewed towards whatever industry had paid Transoft to do translation jobs.

Growth and Prosperity

Despite the linguistic weaknesses, MicroTac continued to grow and prosper. We greatly rejoiced when the major software wholesale distributors (Ingram, Merisel and Kenfil) began to carry the products. Now, MicroTac would be a force to be reckoned with in the retail marketplace, which—unlike now—was vibrant and full of promise.

Major stores like CompUSA, Egghead and Computer City carried the MicroTac name to the public. For the better part of the 90s, MicroTac's products, with a street price of around \$69, were the best-selling MT products the world had ever seen. They beat the much pricier—or non-retail—products of Systran, Globalink and LOGOS handily. In fact, years later a Globalink executive confided to the staff that one of the reasons they looked at merging with MicroTac was that, “We looked at the computer store across the street from our office and it was filled with MicroTac's products, not Globalink's.”

Shaking up MT

MicroTac's success was starting to make people notice. In fact, at one time, MicroTac seriously looked at purchasing or licensing Systran's lexicons, though Systran was much older, well respected and seemingly stable than MicroTac. Just knowing this made me realize how much our company of than 10 people was shaking up the small MT world.



MT News International

Subscription Order Form for non-members of IAMT Associations

Subscription to **MTNI** is a benefit of membership in any of the three regional IAMT Associations. Non-members may also subscribe. This form should be sent to the appropriate region together with a remittance in the currency specified. The fee covers a one-year airmail subscription (individual or institutional) for three issues, starting in the spring of the current year.

I/we wish to receive a one-year subscription to MT News International:

Name: _____

Organization: _____

Address: _____

City: _____ State: _____ Postal code: _____

E-mail: _____ Fax: _____

For individuals or institutions located in the Asia-Pacific region, please return this form with payment of ¥4,000 (bank draft or international money order) to:

Association for Machine Translation in the Americas
3 Landmark Center
East Stroudsburg, PA 18301

For individuals or institutions located in Europe, the Middle East, or Africa, please return this form with payment of Sw.fr. 70.00 (check or money order) to:

Association for Machine Translation in the Americas (AMTA)
3 Landmark Center
East Stroudsburg, PA 18301

For individuals or institutions located in North, South, or Central America, please return this form with payment of US\$ 75.00 (check, money order, or credit card) to:

Association for Machine Translation in the Americas (AMTA)
3 Landmark Center
East Stroudsburg, PA 18301

Publications Order Form

Please return this form with payment or credit card information, to:
International Association for Machine Translation (IAMT c/o AMTA)
3 Landmark Center
East Stroudsburg, PA 18301

Please send the items marked at the right to:

Name: _____

Organization: _____

Address: _____

City: _____ State: _____ Postal code: _____

E-mail: _____ Fax: _____

Price (in U.S. dollars)¹

Title	Member ²	Non-member
<input type="checkbox"/> Compendium of Translation Software (on-line version)	FREE	\$20.00
<input type="checkbox"/> Proceedings of MT Summit VI	\$40.00	\$60.00
<input type="checkbox"/> Proceedings of AMTA-96	\$40.00	\$60.00
<input type="checkbox"/> Proceedings of AMTA-94	\$40.00	\$60.00
<input type="checkbox"/> Proceedings of Workshop on MT Evaluation (1992)	\$55.00	\$55.00

¹ Prices include shipping and handling.

² Member of AAMT, AMTA, EAMT.

Method of Payment

Check or M.O. Visa MasterCard American Express

Card number _____ Exp. Date: ____/____

The proceedings of AMTA-98 and AMTA-2000 appeared as #1529 and #1934 in the Springer series Lecture Notes in Artificial Intelligence. To order, contact the publisher at www.springer.de.

Asia-Pacific Association for Machine Translation

CORPORATE / INDIVIDUAL APPLICATION FORM

Please fill out the appropriate form in both your native language (NAT) and English (ENG) and send it to the address below.

CORPORATE

Corporate name: _____ Seal: _____
NAT: _____
ENG: _____
Capital: _____ Date established : _____ # employees: _____
Name of company president:: _____ Seal: _____
NAT: _____
ENG: _____
Name of person responsible for this application: Seal: _____
NAT: _____
ENG: _____
His/her office/department: _____
NAT: _____
ENG: _____
Mailing address: _____
NAT: _____
ENG: _____
Tel: _____ Ext.: _____ Fax: _____
E-mail: _____
Business category:
 Government agency Manufacturer
 Service industry Translation business
 Other (please specify): _____
All corporate application forms should include a company prospectus.

INDIVIDUAL

Name: _____ Seal: _____
NAT: _____
ENG: _____
Mailing address: _____
NAT: _____
ENG: _____
Tel: _____ Fax: _____
E-mail: _____
Occupation: _____
Company name: _____
Company address: _____
 Translator Other: _____
 University/Institution/Researcher
Specialty: _____

Individual Membership Fees

Initiation fee for individual members	¥ 1,000
Annual dues for individual members	¥ 5,000
Total payment	¥ 6,000

The foregoing amount will be paid by ___/___/___, or within one month from the date this form is mailed.

Corporate Membership Fees

Initiation fee for corporate members
(1 unit = ¥10,000; minimum 1 unit) ___ units = ¥ _____
Annual dues for corporate members
(1 unit = ¥50,000) ___ units = ¥ _____
Please check appropriate box below:
 MT system developer (minimum 10 units)
 Other, capital over ¥10 million (minimum 2 units)
 Other, capital up to ¥10 million (minimum 1 unit)
Total payment (initiation fee + annual dues) ¥ _____
The foregoing amount will be paid by ___/___/___, or within one month from the date this form is mailed.

Method of Payment

Wire transfer to:
Mizuho Bank
Ochanomizu Branch
Tokyo, Japan
A/C No. 1737479
For: Asia-Pacific Association for
Machine Translation (AAMT)
 International postal money order in
Japanese yen to AAMT at ►
Applicants are responsible for all bank
charges. Please retain your copy of bank
draft or money order as proof of payment.

Today's date: ___/___/___

Please send this form to:

Asia-Pacific Association for Machine Translation (AAMT)
c/o Japan Electronics and Information Technology Industries Association (JEITA)
Mitsui Kaijo Bekkan Building, 3F
3-11, Kanda-Surugadai, Chiyoda-ku
Tokyo 101-0062, Japan
Fax: +81 (03) 3518-6472

Association for Machine Translation in the Americas
MEMBERSHIP APPLICATION / RENEWAL FORM

Type of member and membership fee per calendar year:

- Individual US\$ 60
- Institutional (nonprofit) US\$ 200
- Representative: _____
- Corporate US\$ 400
- Representative: _____

Please return this form, together with your payment or credit card information, to:

Association for Machine Translation
in the Americas
3 Landmark Center
East Stroudsburg, PA 18301

Last name(s): _____ First name(s): _____ Title: _____

Address: _____

Home tel.: _____ Work tel.: _____ Fax: _____

E-mail: _____ Website: _____

Affiliation: _____

Professional associations: _____

Area of specialization:

MT User MT Developer MT Researcher Translator Manager Other _____

Method of Payment

- Check enclosed
- Credit card

Type of credit card: Visa MasterCard American Express

Card number _____ Exp. Date: ___/___

European Association for Machine Translation

APPLICATION FOR MEMBERSHIP

Please return this form, together with your payment or credit card information, to:

EAMT Secretariat, c/o TIM / ISSCO
 Université de Genève
 École de Traduction et d'Interprétation
 40, blvd du Pont-d'Arve
 CH-1211 Geneva 4, Switzerland

Type of member and membership fee per calendar year:

- Individual SFr 50
- Non-profit-making institution SFr 175
- Representative: _____
- Profit-making institution SFr 350
- Representative: _____

Last name(s): _____ First name(s): _____ Title: _____

Address: _____

Home tel.: _____ Work tel.: _____ Fax: _____

E-mail: _____ Website: _____

Institution / organization: _____

Area of specialization:

- MT User MT Developer MT Researcher Translator Manager Other _____

Method of Payment

- Cheque payable to EAMT, enclosed
- Banker's draft (copy enclosed) to account no. 351.091.40L
 Union Bank of Switzerland
 Bahnhofstrasse 45
 CH-8021 Zürich, Switzerland

Please note: All bank charges must be borne by the applicant

- Credit card (please provide information at right →)

Type of credit card: Visa Eurocard

Card number _____

Expiration date: ___/___