Chapter 8

Multilinguality

(Following section is taken from Chapter 8 "Multilinguality")
of the book: "Survey of the state of the art in human language technology")

8.2 Machine Translation: The Disappointing Past and Present

Martin Kay

Xerox Palo Alto Research Center, Palo Alto, California, USA

The field of machine translation has changed remarkably little since its earliest days in the fifties. The issues that divided researchers then remain the principal bones of contention today. The first of these concerns the distinction between the so-called interlingual and the transfer approach to the problem. The second concerns the relative importance of linguistic matters as opposed to common sense and general knowledge. The only major new lines of investigation that have emerged in recent years have involved the use of existing translations as a prime source of information for the production of new ones. One form that this takes is that of example-based machine translation (Furuse & Iida, 1992; Iida & Iida, 1991; Nagao, 1992; Sato, 1992) in which a system of otherwise fairly conventional design is able to refer to a collection of existing translations. A much more radical approach, championed by IBM (Brown, Cocke, et al., 1990), is the one in which virtually the entire body of knowledge that the system uses is acquired automatically from statistical properties of a very large body of existing translation.

In recent years, work on machine translation has been most vigorously pursued in Japan and it is also there that the greatest diversity of approaches is to be found. By and large, the Japanese share the general perception that the transfer approach offers the best chance for early success.

Two principal advantages have always been claimed for the interlingual approach. First, the method is taken as a move towards robustness and overall economy in that translation between all pairs of a set of languages in principle requires only translation to and from the interlingua for each member of the set. If there are n languages, n components are therefore required to be translated into the interlingua and n to translate from it, for a total of $2n$. To provide the same facilities, the transfer approach, according to which a major part of the translation system for

a given pair of languages is specific to that pair, requires a separate device to translate in each direction for every pair of languages for a total of $n(n-1)$.

The PIVOT system of NEC (Okumura, Muraki, et al., 1991; Muraki, 1989) and ATLAS II of Fujitsu (Uchida, 1989) are commercial systems among a number of research systems based on the two-step method according to which texts are translated from the source language to an artificial interlingual representation and then into the target language. The Rosetta system at Philips (Landsbergen, 1987), and the DLT system at BSO (Witkam, 1988; Schubert, 1988) in the Netherlands also adopted this approach. In the latter, the interlingua is not a language especially designed for this purpose, but Esperanto.

According to the majority transfer view of machine translation, a certain amount of analysis of the source text is done in the context of the source language alone and a certain amount of work on the translated text is done in the context of the target language. But the bulk of the work relies on comparative information about the specific pair languages. This is argued for on the basis of the sheer difficulty of designing a single interlingua that can be all things for all languages and on the view that translation is, by its very nature, an exercise in comparative linguistics. The massive Eurotra system (Schutz, Thurmair, et al., 1991; Arnold & des Tombes, 1987; King & Perschke, 1987; Perschke, 1989), in which groups from all the countries of the European Union participated, was a transfer system, as is the current Verbmobil system sponsored by the German Federal Ministry for Research and Technology (BMFT).

A transfer system, in which the analysis and generation components are large relative to the transfer component and where transfer is therefore conducted in terms of quite abstract entities, takes on much of the flavor of an interlingual system, while not making the commitment to linguistic universality that many see as the hallmark of the interlingual approach. Such semantic transfer systems are attracting quite a lot of attention. Fujitsu's ATLAS I (Uchida, 1986) was an example, and Sharp's DUET system is another. The approach taken by SRI (Cambridge) with the Core Language Engine (Alshawi, Carter, et al., 1991) also falls in this category.

Just as these systems constitute something of an intermediate position between interlingua and transfer, they can also be seen to some extent as a compromise between the mainly linguistically based approaches we have been considering up to now and the so-called knowledge-based systems pursued most notably at Carnegie Mellon University (Nirenburg, Raskin, et al., 1986; Carbonell & Tomita, 1987), and at the Center for Research in Language at New Mexico State University (Farwell & Wilks, 1990). The view that informs these efforts, whose most forceful champion was Roger Shank, is that translation relies heavily on information and abilities that are not specifically linguistic. If it is their linguistic knowledge that we often think of as characterizing human translators, it is only because we take their common sense and knowledge of the everyday world for granted in a way we clearly cannot do for machines.

Few informed people still see the original ideal of fully automatic high-quality translation of arbitrary texts as a realistic goal for the foreseeable future. Many systems require texts to be pre-edited to put them in a form suitable for treatment by the system, and post-editing of the machine's output is generally taken for granted. The most successful systems have been those that have relied on their input being in a sublanguage (Kittredge, 1987), either naturally occurring, as in that case of weather reports, or deliberately controlled. The spectacular success of the METEO system (Chevalier, Dansereau, et al., 1978) working on Canadian weather reports encouraged the view that sublanguages might be designed for a number of different applications, but the principles on which such languages should be designed have failed to emerge and progress has been very limited.

Future Directions

Research in machine translation has developed traditional patterns which will clearly have to be broken if any real progress is to be made. The traditional view that the problem is principally a linguistic one is clearly not tenable, but

the alternatives that require a translation system to have a substantial part of the general knowledge and common sense that humans have also seems to be unworkable. Compromises must presumably be found where knowledge of restricted domains can facilitate the translation of texts in those domains. The most obvious gains will come from giving up, at least for the time being, the idea of machine translation as a fully automatic batch process in favor of one in which the task is apportioned between people and machines. The proposal made in Kay (1980), according to which the translation machine would consult with a human speaker of the source language with detailed knowledge of the subject matter, has attracted more attention in recent times. A major objection to this approach, namely that the cost of operating such a system would come close to that of doing the whole job in the traditional way, will probably not hold up in the special, but widespread situation in which a single document has to be translated into a large number of languages.

8.3 Chapter References

ACL (1991). *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, Berkeley, California. Association for Computational Linguistics.

Alshawi, H., Carter, D., et al. (1991). Translation by quasi logical form transfer. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, Berkeley, California. Association for Computational Linguistics.

ANLP (1994). *Proceedings of the Fourth Conference on Applied Natural Language Processing*, Stuttgart, Germany. ACL, Morgan Kaufmann.

Arnold, D. and des Tombes, L. (1987). Basic theory and methodology in EUROTRA. In Nirenburg, S., editor, *Machine Translation: Theoretical and Methodological Issues*, pages 114–135. Cambridge University Press.

ARPA (1993). *Proceedings of the 1993 ARPA Human Language Technology Workshop*, Princeton, New Jersey. Advanced Research Projects Agency, Morgan Kaufmann.

Bamberg, P., Demedts, A., Elder, J., Huang, C., Ingold, C., Mandel, M., Manganaro, L., and van Even, S. (1991). Phonem-based training for large-vocabulary recogntition in six european languages. In *Eurospeech '91, Proceedings of the Second European Conference on Speech Communication and Technology*, volume 1, pages 175–181, Genova, Italy. European Speech Communication Association.

Batchelder, E. O. (1992). A learning experience: Training an artificial neural network to discriminate languages. Technical Report.

Beesley, K. R. (1988). Language identifier: A computer program for automatic natural-language identification on on-line text. In *Proceedings of the 29th Annual Conference of the American Translators Association*, pages 47–54.

Belogonov, G., Khoroshilov, A., Khoroshilov, A., Kuznetsov, B., Novoselov, A., Pashchenko, N., and Zelenkov, Y. (1993). An interactive system of Russian-English and English-Russian machine translation of polythematic scientific and technical texts. Technical report, VINITI internal Report, Moscow.

Berkling, K. M. and Barnard, E. (1994). Language identification of six languages based on a common set of broad phonemes. In *Proceedings of the 1994 International Conference on Spoken Language Processing*, volume 4, pages 1891–1894, Yokohama, Japan.

Boitet, C. (1986). The French national MT-project: technical organization and translation results of CALLIOPE-AERO. *Computers and Translation*, 1:281.

Boitet, C. and Blanchon, H. (1993). Dialogue-based machine translation for monolingual authors and the LIDIA project. In Nomura, H., editor, *Proceedings of the 1993 Natural Language Processing Rim Symposium*, pages 208–222, Fukuoka. Kyushu Institute of Technology.

Bourbeau, L., Carcagno, D., Goldberg, E., Kittredge, R., and Polguere, A. (1990). Bilingual generation of wheather forecasts in an operations environment. In Karlgren, H., editor, *Proceedings of the 13th International Conference on Computational Linguistics*, volume 3, pages 318–320, Helsinki. ACL.

Brown, P., Cocke, J., Pietra, S. D., Pietra, V. J. D., Jelinek, F., Lafferty, J. D., Mercer, R. L., and Roossin, P. S. (1990). A statistical approach to machine translation. *Computational Linguistics*, 16(2):79–85.

Carbonell, J. G. and Tomita, M. (1987). Knowledge-based machine translation, the CMU approach. In Nirenburg, S., editor, *Machine Translation: Theoretical and Methodological Issues*, pages 68–89. Cambridge University Press.

Cavner, W. B. and Trenkle, J. M. (1994). N-gram based text categorization. In *Proceedings of the Third Annual Symposium on Document Analysis and Information Retrieval*, pages 261–169.

Cerf-Danon, H., DeGennaro, S., Ferreti, M., Gonzalez, J., and Keppel, E. (1991). Tangora—a large vocabulary speech recognition system for five languages. In *Eurospeech '91, Proceedings of the Second European Conference on Speech Communication and Technology*, volume 1, pages 183–192, Genova, Italy. European Speech Communication Association.

Chandioux, J. (1989). Meteo: 1000 million words later. In Hammond, D. L., editor, *American Translators Association Conference 1989: Coming of Age*, pages 449–453. Learned Information, Medford, New Jersey.

Chandler, B., Holden, N., Horsfall, H., Pollard, E., and McGee, M. W. (1987). N-tran final report, Alvey project. Technical Report 87/9, CCL/UMIST, Manchester.

Chevalier, M., Dansereau, J., et al. (1978). *TAUM-METEO: Description du Système*. Université de Montréal.

Cole, R. A., Hirschman, L., Atlas, L., Beckman, M., Bierman, A., Bush, M., Cohen, J., Garcia, O., Hanson, B., Hermansky, H., Levinson, S., McKeown, K., Morgan, N., Novick, D., Ostendorf, M., Oviatt, S., Price, P., Silverman, H., Spitz, J., Waibel, A., Weinstein, C., Zahorian, S., and Zue, V. (1995). The challenge of spoken language systems: Research directions for the nineties. *IEEE Transactions on Speech and Audio Processing*, 3(1):1–21.

COLING (1986). *Proceedings of the 11th International Conference on Computational Linguistics*, Bonn. ACL.

COLING (1988). *Proceedings of the 12th International Conference on Computational Linguistics*, Budapest.

COLING (1992). *Proceedings of the 14th International Conference on Computational Linguistics*, Nantes, France. ACL.

Dalsgaard, P. and Andersen, O. (1994). Application of inter-language phoneme similarities for language identification. In *Proceedings of the 1994 International Conference on Spoken Language Processing*, volume 4, pages 1903–1906, Yokohama, Japan.

Debili, F., Fluhr, C., and Radasao, P. (1989). About reformulation in full-text IRS. *Information Processing and Management*, 25:647–657.

Delin, J., Hartley, A., Paris, C., Scott, D., and Vander Linden, K. (1994). Expressing procedural relationships in multilingual instructions. In *Proceedings of the Seventh International Workshop on Natural Language Generation*, pages 61–70, Kennebunkport, Maine. Springer-Verlag, Berlin.

EMIR (1994). Final report of the EMIR project number 5312. Technical report, European Multilingual Information Retrieval Consortium For the Commission of the European Union, Brussels.

Eurospeech (1991). *Eurospeech '91, Proceedings of the Second European Conference on Speech Communication and Technology*, Genova, Italy. European Speech Communication Association.

Farwell, D. and Wilks, Y. (1990). *Ultra: A Multi-lingual Machine Translator*. New Mexico State University.

Fluhr, C. (1990). Multilingual information. In *AI and Large-Scale Information*, Nagoya.

Furuse, O. and Iida, H. (1992). Cooperation between transfer and analysis in example-based framework. In *Proceedings of the 14th International Conference on Computational Linguistics*, Nantes, France. ACL.

Gauvain, J.-L. and Lamel, L. F. (1993). Identification of non-linguistic speech features. In *Proceedings of the 1993 ARPA Human Language Technology Workshop*, page Session 6, Princeton, New Jersey. Advanced Research Projects Agency, Morgan Kaufmann.

Glass, J., Goodine, D., Phillips, M., Sakai, S., Seneff, S., and Zue, V. (1993). A bilingual voyager system. In *Proceedings of the 1993 ARPA Human Language Technology Workshop*, Princeton, New Jersey. Advanced Research Projects Agency, Morgan Kaufmann. Session 6.

Harman, D., editor (1993). *National Institute of Standards and Technology Special Publication No. 500-207 on the The First Text REtrieval Conference (TREC-1)*, Washington, DC. National Institute of Standards and Technology, U.S. Department of Commerce, U.S. Government Printing Office.

Hazen, T. J. and Zue, V. W. (1994). Recent improvements in an approach to segment-based automatic language identification. In *Proceedings of the 1994 International Conference on Spoken Language Processing*, volume 4, pages 1883–1886, Yokohama, Japan.

Heinrich, P. (1989). Language identification for automatic grapheme-to-phoneme conversion of foreign words in a german text-to-speech system. In *Speech-89*, pages 220–223.

House, A. S. and Neuberg, E. P. (1977). Toward automatic identification of the language of an utterance. I. Preliminary methodological considerations. *Journal of the Acoustical Society of America*, 62(3):708–713.

Huang, X. M. (1990). A machine translation system for the target language inexpert. In Karlgren, H., editor, *Proceedings of the 13th International Conference on Computational Linguistics*, volume 3, pages 364–367, Helsinki. ACL.

ICASSP (1994). *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing*, Adelaide, Australia. Institute of Electrical and Electronic Engineers.

ICSLP (1994). *Proceedings of the 1994 International Conference on Spoken Language Processing*, Yokohama, Japan.

Iida, E. S. and Iida, H. (1991). Experiments and prospects of example-based machine translation. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics*, pages 185–192, Berkeley, California. Association for Computational Linguistics.

Ingle, N. C. (1991). A language identification table. *The Incorporated Linguist*, 15(4):98–101.

IWNLG (1994). *Proceedings of the Seventh International Workshop on Natural Language Generation*, Kennebunkport, Maine. Springer-Verlag, Berlin.

JEIDA (1989). A Japanese view of machine translation in light of the considerations and recommendations reported by ALPAC, USA. Technical report, Japanese Electronic Industry Development Association, Tokyo.

Kadambe, S. and Hieronymus, J. L. (1994). Spontaneous speech language identification with a knowledge of linguistics. In *Proceedings of the 1994 International Conference on Spoken Language Processing*, volume 4, pages 1879–1882, Yokohama, Japan.

Karlgren, H., editor (1990). *Proceedings of the 13th International Conference on Computational Linguistics*, Helsinki. ACL.

Kay, M. (1973). The MIND system. In Rustin, R., editor, *Courant Computer Science Symposium 8: Natural Language Processing*, pages 155–188. Algorithmics Press, New York.

Kay, M. (1980). *The Proper Place of Men and Machines in Language Translation*. Xerox Palo Alto Research Center, Palo Alto, California.

Kay, M., Gawron, J. M., and Norvig, P. (1991). Verbmobil: A translation system for face-to-face dialog. Technical report, Stanford University.

King, M. and Perschke, S. (1987). *Machine Translation Today: The State of the Art*. Edinburgh University Press. EUROTRA.

Kittredge, R. I. (1987). The significance of sublanguage for automatic translation. In Nirenburg, S., editor, *Machine Translation: Theoretical and Methodological Issues*, pages 59–67. Cambridge University Press.

Kulikowski, S. (1991). Using short words: a language identification algorithm. Unpublished technical report.

Lamel, L. F. and Gauvain, J.-L. S. (1994). Language identification using phone-based acoustic likelihoods. In *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 293–296, Adelaide, Australia. Institute of Electrical and Electronic Engineers.

Landauer, T. K. and Littman, M. L. (1990). Fully automatic cross-language document retrieval using latent semantic indexing. In *Proceedings of the Sixth Annual Conference of the UW Centre for the New Oxford English Dictionary and Text Research*, UW Centre for the New OED and Text Research, Waterloo Ontario.

Landsbergen, J. (1987). Isomorphic grammars and their use in the ROSETTA translation system. In *Machine Translation Today: The State of the Art*. Edinburgh University Press, Edinburgh.

Lehrberger, J. and Bourbeau, L. (1988). *Machine translation: linguistic characteristics of MT systems and general methodology of evaluation*. John Benjamins, Amsterdam, Philadelphia.

Levin, L., Suhm, B., Coccaro, N., Carbonell, J., Horiguchi, K., Isotani, R., Lavie, A., Mayfield, L., Rose, C. P., Van Ess-Dykema, C., and Waibel, A. (1994). Speech–language integration in a multi-lingual speech translation system. In *Proceedings of the 1994 AAAI Conference*, Seattle. American Association for Artificial Intelligence.

Li, K.-P. (1994). Automatic language identification using syllabic features. In *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 297–300, Adelaide, Australia. Institute of Electrical and Electronic Engineers.

Maruyama, H., Watanabe, H., and Ogino, S. (1990). An interactive Japanese parser for machine translation. In Karlgren, H., editor, *Proceedings of the 13th International Conference on Computational Linguistics*, volume 2, pages 257–262, Helsinki. ACL.

Melby, A. K. (1982). Multi-level translation aids in a distributed system. In *Proceedings of the 9th International Conference on Computational Linguistics*, volume 1 of *Ling. series 47*, pages 215–220, Prague. ACL.

Miller, G. (1990). Wordnet: An on-line lexical database. *International journal of Lexicography*, 3(4):235–312.

Morimoto, T., Takezawa, T., Yato, F., Sagayama, S., Tashiro, T., Nagata, M., and Kurematsu, A. (1993). ATR's speech translation system: ASURA. In *Proceedings of the Third Conference on Speech Communication and Technology*, pages 1295–1298, Berlin, Germany.

MTS (1989). *Proceedings of the Second Machine Translation Summit*, Tokyo. Omsha Ltd.

MTS (1991). *Proceedings of the Third Machine Translation Summit*, Carnegie Mellon University.

Muraki, K. (1989). PIVOT: Two-phase machine translation system. In *Proceedings of the Second Machine Translation Summit*, Tokyo. Omsha Ltd.

Mustonen, S. (1965). Multiple discriminant analysis in linguistic problems. In *Statistical Methods in Linguistics*. Skriptor Fack, Stockholm. Number 4.

Muthusamy, Y. K. (1993). *A Segmental Approach to Automatic Language Identification*. PhD thesis, Oregon Graduate Institute of Science & Technology, P.O.Box 91000, Portland, OR 97291-1000 USA.

Muthusamy, Y. K., Barnard, E., and Cole, R. A. (1994). Reviewing automatic language identification. *IEEE Signal Processing Magazine*, 11(4):33–41.

Muthusamy, Y. K., Cole, R. A., and Oshika, B. T. (1992). The OGI multi-language telephone speech corpus. In *Proceedings of the 1992 International Conference on Spoken Language Processing*, volume 2, pages 895–898, Banff, Alberta, Canada. University of Alberta.

Muthusamy, Y. K., Jain, N., and Cole, R. A. (1994). Perceptual benchmarks for automatic language identification. In *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 333–336, Adelaide, Australia. Institute of Electrical and Electronic Engineers.

Nagao, M. (1992). Some rationales and methodologies for example-based approach. In *Fifth Generation Natural Language Processing*. Publisher Unknown.

Nakayama (1994). Modeling content identification from document images. In *Proceedings of the Fourth Conference on Applied Natural Language Processing*, pages 22–27, Stuttgart, Germany. ACL, Morgan Kaufmann.

Newman, P. (1987). Foreign language identification: First step in the translation process. In *Proceedings of the 28th Annual Conference of the American Translators Accociation*, pages 509–516.

Ney, H. and Billi, R. (1991). Prototype systems for large-vocabulary speech recognition: Polyglot and Spicos. In *Eurospeech '91, Proceedings of the Second European Conference on Speech Communication and Technology*, volume 1, pages 193–200, Genova, Italy. European Speech Communication Association.

Nirenburg, S., editor (1987). *Machine Translation: Theoretical and Methodological Issues*. Cambridge University Press.

Nirenburg, S., Raskin, V., et al. (1986). On knowledge-based machine translation. In *Proceedings of the 11th International Conference on Computational Linguistics*, Bonn. ACL.

Okumura, A., Muraki, K., and Akamine, S. (1991). Multi-lingual sentence generation from the PIVOT interlingua. In *Proceedings of the Third Machine Translation Summit*, Carnegie Mellon University.

Paris, C. and Scott, D. (1994). Stylistic variation in multilingual instructions. In *Proceedings of the Seventh International Workshop on Natural Language Generation*, pages 45–52, Kennebunkport, Maine. Springer-Verlag, Berlin.

Perschke, S. (1989). EUROTRA project. In *Proceedings of the Second Machine Translation Summit*, Tokyo. Omsha Ltd.

Ramesh, P. and Roe, D. B. (1994). Language identification with embedded word models. In *Proceedings of the 1994 International Conference on Spoken Language Processing*, volume 4, pages 1887–1890, Yokohama, Japan.

Rau, M. D. (1974). Language identification by statistical analysis. Master's thesis, Naval Postgraduate School.

Rayner, M. et al. (1993). A speech to speech translation system built from standard components. In *Proceedings of the 1993 ARPA Human Language Technology Workshop*, Princeton, New Jersey. Advanced Research Projects Agency, Morgan Kaufmann.

Reyes, A. A., Seino, T., and Nakagawa, S. (1994). Three language identification methods based on HMMs. In *Proceedings of the 1994 International Conference on Spoken Language Processing*, volume 4, pages 1895–1898, Yokohama, Japan.

Roe, D. B., Pereira, F. C., Sproat, R. W., and Riley, M. D. (1992). Efficient grammar processing for a spoken language translation system. In *Proceedings of the 1992 International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 213–216, San Francisco. Institute of Electrical and Electronic Engineers.

Rösner, D. and Stede, M. (1994). Techdoc: Multilingual generation of online and offline instructional text. In *Proceedings of the Fourth Conference on Applied Natural Language Processing*, pages 209–210, Stuttgart, Germany. ACL, Morgan Kaufmann.

Sadler, V. (1989). Working with analogical semantics: Disambiguation technics in DLT. In Witkam, T., editor, *Distributed Language Translation (BSO/Research)*. Floris Publications, Dordrecht, Holland.

Salton, G. and McGill, M. (1983). *An Introduction to Modern Information Retrieval*. McGraw-Hill, New York.

Sato, S. (1992). CTM: An example-based translation aid system using the character-based best match retrieval method. In *Proceedings of the 14th International Conference on Computational Linguistics*, Nantes, France. ACL.

Schmitt, J. C. (1991). Trigram-based method of language identification. U.S. Patent number: 5062143.

Schubert, K. (1988). The architectre of DLT—interlingual or double direct. In *New Directions in Machine Translation*. Floris Publications, Dordrecht, Holland.

Schutz, J., Thurmair, G., et al. (1991). An architecture sketch of Eurotra-II. In *Proceedings of the Third Machine Translation Summit*, Carnegie Mellon University.

Sibun, P. and Spitz, L. A. (1994). Language determination: Natural language processing from scanned document images. In *Proceedings of the Fourth Conference on Applied Natural Language Processing*, pages 15–21, Stuttgart, Germany. ACL, Morgan Kaufmann.

Sigurdson, J. and Greatex, R. (1987). *Machine Translation of on-line searches in Japanese Data Bases*. RPI, Lund University.

Somers, H. L., Tsujii, J.-I., and Jones, D. (1990). Machine translation without a source text. In Karlgren, H., editor, *Proceedings of the 13th International Conference on Computational Linguistics*, volume 3, pages 271–276, Helsinki. ACL.

Spitz, L. A. (1993). Generalized line word and character finding. In *Proceedings of the International Conference on Image Analysis and Processing*, pages 686–690.

Tomita, M. (1986). Sentence disambiguation by asking. *Computers and Translation*, 1(1):39–51.

Tong, L. C. (1987). The engineering of a translator workstation. *Computers and Translation*, 2(4):263–273.

Uchida, H. (1986). Fujitsu machine translation system: ATLAS. In *Future Generations Computer Systems 2*, pages 95–100. Publisher Unknown.

Uchida, H. (1989). ATLAS-II: A machine translation system using conceptual structure as an interlingua. In *Proceedings of the Second Machine Translation Summit*, Tokyo. Publisher Unknown.

Ueda, Y. and Nakagawa, S. (1990). Prediction for phoneme/syllable/word-category and identification of language using HMM. In *Proceedings of the 1990 International Conference on Spoken Language Processing*, volume 2, pages 1209–1212, Kobe, Japan.

Vasconcellos, M. and Len, M. (1988). SPANAM and ENGSPAM: Machine translation at the Pan American Health Organization. In Slocum, J., editor, *Machine Translation systems*, pages 187–236. Cambridge University Press.

Velho Lopes, R. R. (1989). Automated access to multilingual information: a Brazilian case study. *Information Development*, 5(3).

Wahlster, W. (1993). Verbmobil, translation of face-to-face dialogs. In *Proceedings of the Fourth Machine Translation Summit*, pages 127–135, Kobe, Japan.

Waibel, A., Jain, A., McNair, A., Saito, H., Hauptmann, A., and Tebelskis, J. (1991). JANUS: a speech-to-speech translation system using connectionist and symbolic processing strategies. In *Proceedings of the 1991 International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 793–796, Toronto. Institute of Electrical and Electronic Engineers.

Wehrli, E. (1992). The IPS system. In *Proceedings of the 14th International Conference on Computational Linguistics*, volume 3, pages 870–874, Nantes, France. ACL.

Whitelock, P. J., Wood, M. M., Chandler, B. J., Holden, N., and Horsfall, H. J. (1986). Strategies for interactive machine translation: The experience and implications of the UMIST Japanese project. In *Proceedings of the 11th International Conference on Computational Linguistics*, pages 25–29, Bonn. ACL.

Winsoft (1987). *Manuel d'utilisation de WinTool*. Winsoft Inc., Grenoble. Version 1.1.

Witkam, T. (1988). DLT—an industrial R&D project for multilingual machine translation. In *Proceedings of the 12th International Conference on Computational Linguistics*, Budapest.

Wood, M. M. and Chandler, B. (1988). Machine translation for monolinguals. In *Proceedings of the 12th International Conference on Computational Linguistics*, pages 760–763, Budapest.

Woszczyna, M., Aoki-Waibel, N., Buo, F. D., Coccaro, N., Horiguchi, K., Kemp, T., Lavie, A., McNair, A., Polzin, T., Rogina, I., Rose, C. P., Schultz, T., Suhm, B., Tomita, M., and Waibel, A. (1994). Towards spontaneous speech translation. In *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 345–349, Adelaide, Australia. Institute of Electrical and Electronic Engineers.

Ziegler, D. V. (1991). *The automatic identification of languages using linguistic recognition signals*. PhD thesis, SUNY Buffalo.

Zissman, M. A. and Singer, E. (1994). Automatic language identification of telephone speech messages using phoneme recognition and n-gram modeling. In *Proceedings of the 1994 International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 305–308, Adelaide, Australia. Institute of Electrical and Electronic Engineers.