

[From "Wired", Issue 8.05, May 2000]

Machine Translation's Past and Future

1629 René Descartes proposes a universal language, with equivalent ideas in different tongues sharing one symbol.

1933 Russian Petr Smirnov- Troyanskii patents a device for transforming word-root sequences into their other-language equivalents.

1939 Bell Labs demonstrates the first electronic speech-synthesizing device at the New York World's Fair.

1949 Warren Weaver, director of the Rockefeller Foundation's natural sciences division, drafts a memorandum for peer review outlining the prospects of machine translation (MT).

1952 Yehoshua Bar-Hillel, MIT's first full-time MT researcher, organizes the maiden MT conference.

1954 First public demo of computer translation at Georgetown University: 49 Russian sentences are translated into English using a 250-word vocabulary and 6 grammar rules.

1960 Bar-Hillel publishes his report arguing that fully automatic and accurate translation systems are, in principle, impossible.

1964 The National Academy of Sciences creates the Automatic Language Processing Advisory Committee (Alpac) to study MT's feasibility.

1966 Alpac publishes a report on MT concluding that years of research haven't produced useful results. The outcome is a halt in federal funding for machine translation R&D.

1967 L. E. Baum and colleagues at the Institute for Defense Analyses (IDA) in Princeton, New Jersey, develop hidden Markov models, the mathematical backbone of continuous-speech recognition.

1968 Peter Toma, a former Georgetown University linguist, starts one of the first MT companies, Language Automated Translation System and Electronic Communications (Latsec).

1969 In Middletown, New York, Charles Byrne and Bernard Scott found Logos to develop MT systems.

1978 Arpa's Network Speech Compression (NSC) project transmits the first spoken words over the Internet.

1982 Janet and Jim Baker found Newton, Massachusetts-based Dragon Systems.

1983 The Automated Language Processing System (ALPS) is the first MT software for a microcomputer.

1985 Darpa launches its speech recognition program.

1986 Japan launches the ATR Interpreting Telecommunication Research Laboratories (ATR-ITL) to study multilingual speech translation.

1987 In Belgium, Jo Lernout and Pol Hauspie found Lernout & Hauspie.

1988 Researchers at IBM's Thomas J. Watson Research Center revive statistical MT methods that equate parallel texts, then calculate the probabilities that words in one version will correspond to words in another.

1990 Dragon Systems releases its 30,000-word-strong DragonDictate, the first retailed speech-to-text system for general-purpose dictation on PCs.

Darpa launches its Spoken Language Systems (SLS) program to develop apps for voice-activated human-machine interaction.

1991 The first translator-dedicated workstations appear, including STAR's Transit, IBM's TranslationManager, Canadian Translation Services' PTT, and Eurolang's Optimizer.

1992 ATR-ITL founds the Consortium for Speech Translation Advanced Research (C-STAR), which gives the first public demo of phone translation between English, German, and Japanese.

1993 The German-funded Verbmobil project gets under way. Researchers focus on portable systems for face-to-face English-language business negotiations in German and Japanese.

BBN Technologies demonstrates the first off-the-shelf MT workstation for real-time, large-vocabulary (20,000 words), speaker-independent, continuous-speech-recognition software.

1994 Free Systran machine translation is available in select CompuServe chat forums.

1997 AltaVista's Babel Fish offers real-time Systran translation on the Web.

Dragon Systems' NaturallySpeaking and IBM's ViaVoice are the first large-vocabulary continuous-speech-recognition products for PCs.

Parlance Corporation, a BBN Technologies spin-off, releases Name Connector, the first large-vocabulary internal switchboard that routes phone calls by hearing a spoken name.

1999 A televised newscast is automatically transcribed with 85 percent accuracy.

Logos releases e.Sense Enterprise Translation, the first Web-enabled multiple translator operating from a single server.

IBM releases ViaVoice for the Macintosh, the first continuous-speech-recognition Mac software.

Kevin Knight, of the University of Southern California's Information Sciences Institute (ISI), leads a multi-university team that develops Egypt, a software toolkit for building statistical MT systems. Egypt examines bilingual texts for statistical relationships, analyzes those patterns, and applies what it has "learned" to its translation functions.

2000 At MIT's Lincoln Laboratory, Young-Suk Lee and Clifford Weinstein demonstrate an advanced Korean-English speech-to-speech translation-system prototype.

USC's ISI performs backward machine-transliterations of proper nouns, which are replaced with phonetic approximations. *Southern California* translates to "Janoub Kalyfornya" in Arabic.

2001 Carnegie Mellon University's Language Technologies Institute (LTI), led by Jaime Carbonell, constructs speech-to-speech translation for "small" languages like Croatian or Mapudungun, spoken by Mapuches in Chile.

USC biomedical engineers Theodore Berger and Jim-Shih Liaw create a new Berger-Liaw Neural Network Speech Recognition System (SRS) that understands spoken languages better than humans do. Ford says the technology will be incorporated into its cars to facilitate communication at fast-food drive-thrus.

2002 NowHear offers an agent-based newsreader device that translates articles from thousands of publications worldwide, delivering them as MP3 audio files.

2003 Text of Joyce's *Ulysses* is run through Cliff's Notemaker, a new omnidirectional literary interpreter and summarizer. Program: "Your professor didn't read it either. Don't worry about what your essay says, just include the words *Dublin, pub, and fuck.*"

2004 Dragon Systems' NaturallyCursing software is added to wristwatches to ease communication at multilingual construction sites.

2005 Employee at Allstate Insurance files suit against the company, citing emotional distress from the collective chatter of coworkers using speech recognition input devices.

GeoCities pulls down 350,000 homepages for failing to use GeoCities Controlled English, a 1,000-word edictionary designed to interface with its language translation software.

2006 It's that .001 percent part that got us," moans NASA director Rafu Sanjali, after the fourth disastrous attempt to land a robot-controlled vehicle on Mars was foiled by the use of "99.999 percent accurate" MT technology.

2007 Microsoft pulls its "What do you want to think today?" campaign after reviewers unanimously trounce the company's much-anticipated Thought Recognition Interface (TRI).

2008 L&H's Travel Sunglasses offer real-time translation of road signs, marquees, and menus into a wearer's native language.

2009 CorconText introduces FinalCopy, a Japanese-to-English documentation translation program that uses AI-based semantic networks to reduce the need for human editing of output.

2012 Saruzuno embeds its Lexical Disambiguation System (LDS) into smartcards equipped with membrane microphones so travelers can converse with store clerks in dozens of languages.

2017 The Russian-made Durok II language tutor is used to train customs-and-immigrations bots (DNA-based servant-devices) employed at US points of entry.

2020 Teaching a child reading and writing is a waste of time," declares Yeo Kiah Wei, Singapore's minister of education, who cancels the subjects in schools. "Children needn't be burdened with such an onerous task as deciphering tiny markings on a page or screen. Leave it to the machines."

2021 PigLatin Furby reveals parents' plans for divorce. Dozens of toddlers are traumatized.

2043 Tower of Babel is completed in Iraq (formerly Babylonia) after a 4,000-year delay, thanks to NEC Technologies' Neutral Language.

2045 Telepathy system developed by Europeans. Users wear adhesive patches containing thought recognition and MT technology, plus a high-speed wireless transceiver.

2058 The Reformed Rifkin Institute (RRI) is awarded a patent for its invention of a symbio-parasite that feeds on the electrical impulses in the speech center of the

human brain, then excretes a translated signal that can be understood by anyone who inserts the creature in their ear. The estate of Douglas Adams files suit, claiming prior art.

2108 Procter & Gamble researchers use their newly developed Distributed Tachyon Swarm System (DTSS) to learn that diphtheria bacteria band together as a hive mind capable of communication.

2264 "Humans are dumber than bags of hair," declares Entity 296. "Only the most naive scientist would try to develop a technology to understand those smelly lumps of protoplasm," it states. "The noises they emit from the holes in their heads are ultimately less enlightening than cosmic static."

Compiled by Kristin Demos (kvdemos@yahoo.com) and Mark Frauenfelder (mark@well.com). 1629-2000: K. D.; 2001-2264: M. F.