

A History of NASIC

NATIONAL AIR AND SPACE INTELLIGENCE CENTER

WRIGHT-PATTERSON AIR FORCE BASE, OHIO

Machine Translation at NASIC: A 50-Year Tradition

Contributing author Dale Bostad

In 2005, NASIC celebrated the 50th anniversary of a long odyssey in the development and use of machine translation (MT), or translation by computer. MT is now pervasive throughout the world and is used extensively on the Internet. It has proven to be an increasingly valuable tool in the 21st Century environment of terrorism and proliferation of nuclear technology. MT has also proven itself as an invaluable tool for the fast turnaround of foreign language information for intelligence purposes. NASIC's MT is amazing, and the Foreign Technology Division (FTD), as NASIC was formerly called, was involved in one of the most celebrated landmark decisions in that history.

Back to the Beginning

The story starts in 1955, when NASIC was called the Air Technical Intelligence Center (ATIC). The United States was in the middle of the Cold War, and it was ATIC's mission to track technology developments in the Soviet Union. To translate Soviet science and technology articles, ATIC determined that it would need more than 20 Russian translators. But it was hard to attract Russian translators to relocate to southwest Ohio, and translator salaries were depressed at that time—between \$7,000 and \$10,000 a year. In addition, vast sums of money were being expended to pay contractors to translate Russian.

However, in 1955, there was a promising new technology on the horizon: machine translation. Based on a suggestion from the Directorate of Intelligence, Headquarters of the United States Air Force, ATIC formally requested development of an automatic

translation machine through the Aeronautical Research Laboratory and, in turn, at the Research and Development Command (RADC) in Rome, NY. After a series of development contracts with IBM, a system called the MARK II Russian Translation system was moved from the IBM Research Center in Yorktown Heights, NY, to FTD, Wright-Patterson Air Force Base, and became operational in 1963. The translation system was brought in as a test system for machine-aided translation to be evaluated by independent study of MTs of 100,000 Russian words per day. FTD was also tasked to support advanced development with several contractors to obtain higher quality MT.

In 1965, Arthur D. Little, Inc. was tasked to write an evaluation of MT at FTD and published a report entitled, "Evaluation of Machine-aided Activities at FTD." This report and other information were used as input to a report by the National Science Foundation, which had decided to evaluate the progress of MT in the United States from the 1950's through the early 1960's. It created a committee called the Automatic Language Processing Advisory Committee (ALPAC) to do this evaluation. The result was the so-called ALPAC report, published in 1966. FTD figured prominently in this report, and the repercussions of the report were far reaching.

The committee concluded that MT was more costly than human translation and produced inferior results. They saw no hope for improvement in the future. The report was widely criticized by Zbigniew Pankowicz of RADC and Colonel John Johnson, Jr. of FTD as being biased, inaccurate, misleading, hostile, and vindictive. But the damage was done. Soon after publication of the report, funding for research in MT all but stopped in the United States. In the 1950's, up to 17 institutions had benefited from MT research dollars totaling about \$25,000,000; after the ALPAC report, funding was basically reduced to zero.

Some people have characterized the period following the ALPAC report as the Dark Ages of MT, but it is here that FTD played a pivotal role in its history. FTD had the internal fortitude and confidence to flout the judgment of the esteemed scientists in the ALPAC and proceeded with the further development of MT. It was a courageous decision that history proved correct. FTD, the small organization in southwestern Ohio among the cornfields and soybeans, played a major but unappreciated role in championing MT after the devastating indictment of the ALPAC report.

Systran

By 1968, FTD was in charge of its own MT development contracts. It let its first contract with Latsec, Inc. in San Diego, to build a new Russian system based on the Mark II dictionaries. Peter Toma, the founder of Latsec,



Russian data entry, circa 1985



had been an early MT researcher on the Georgetown project. Mr. Toma was a smooth and crafty lobbyist for development money who would often come to FTD and have personal meetings with the commander. He also lobbied in Washington and was able to have Latsec with its Russian system declared a national asset. This assured continuous funding. In the ensuing years, the Russian system was developed to a very high level and became one of the most famous systems in the history of MT. Latsec was later renamed Systran Software, Inc. FTD and its successor organizations over the years bought and developed many other systems, so that NASIC now has government rights to 23 Systran systems, and Systran has become the most recognized name in MT throughout the world.

Unique Accomplishments

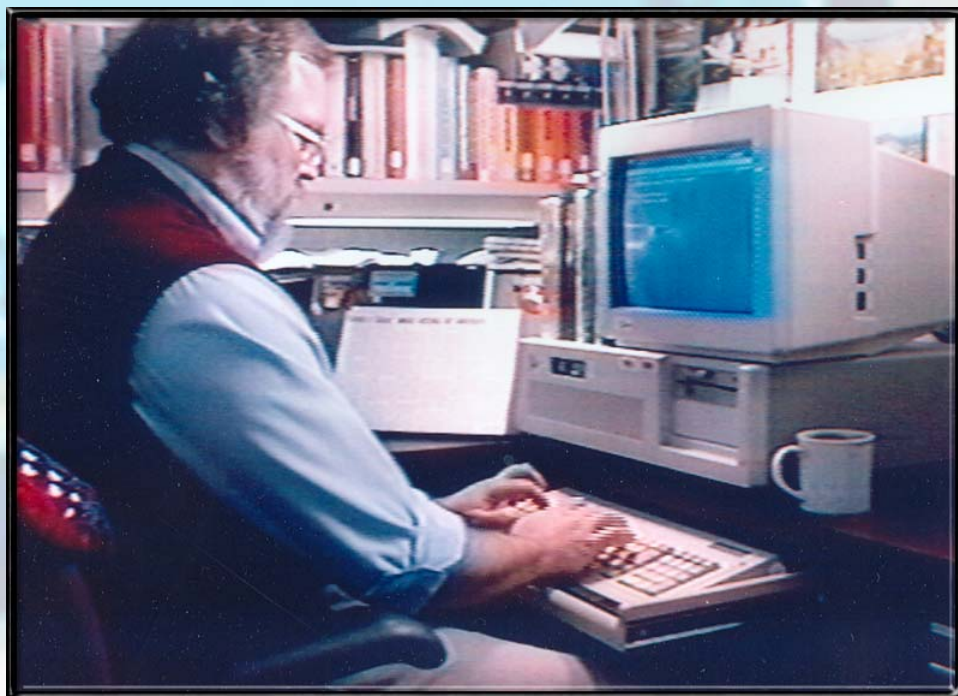
ATIC and FTD were pioneers and champions of MT in the 1950's and 1960's. But other accomplishments ensued in the last 3 decades of the 20th Century and even up to the present. First and foremost, FTD has been recognized as the first organization to implement a full-scale MT production environment. The emphasis was on MT as opposed to human translation. A culture of the use and development of MT was inculcated, and this culture was passed on to the analyst consumer. At one point there were 20 data inputers typing in Russian; 10 post-editors; a recomposition group of about 6 people; and a Russian dictionary development group of 6 linguists, called the "Z Committee," which provided feedback to Systran developers.

The Translation Division in 1976 was also the first group to develop a program for semiautomatic post-editing of raw MT. The program provided for high productivity of the MT product, improving both accuracy and turnaround time. The term used by the Translation directorate was "partially edited MT." This term became part of the literature of MT and is now often called fast or rapid post-editing. These are just some of the breakthrough techniques developed at FTD that have assured FTD a secure position in the history books on MT.

Publicity of All Kinds

Over time, both NASIC and its predecessor organizations received a lot of publicity. Systran was also featured in many technical articles and journalistic stories throughout the world. For example, on 5 August 1987, the morning edition of the *Dayton Daily News* ran a front page story entitled, "Base Division Low-Profile, High Tech," with the subtitle, "Secretive FTD Sifts, Analyzes Foreign Data." The writer, Tim Gaffney, interviewed Colonel Gary Culp, the commander of FTD. On his desk, Colonel Culp had a photo of a fighter plane of an exotic design with a Russian title below it. It was a photo of a plastic model, developed by the Testor Corp., of what a Soviet stealth fighter might look like. The Russian text was taken to the Translation Division, and a typist typed in "protivoradarny1 sovetski1 istrebitel6" (the English character equivalent of the Cyrillic characters). Thirty seconds later, the translation appeared on the screen, "Antiradar Soviet Fighter." Such was the state of FTD's MT system in 1987, with the Russian system accessible on 1,400 personal computers within FTD.

In 1983, a Soviet MT researcher named Yu. N. Marchuk wrote a book entitled, *Problems of Machine Translation*, which was published in Moscow. FTD obtained a copy of this book, which contains a very large section on Systran. This section of the book was translated by the very Systran Russian system described in detail in the book. The translation was then sent to FTD's Chief Scientist to read. Here is part of the translation: "Since at least 1970, the system has been used for technical translations from Russian into English at the Foreign Technology Division of the USAF at Wright-Patterson Air Force Base in San Diego, California." Well, nobody ever said that intelligence was a perfect science!



Russian post-editor, circa 1985