utilized in the field of electrical engineering and implemented (and utilized to this day) at the Moscow Institute INFORMelectro (Pevsner & Lachuti, 1973).

Documents in the Russian and English languages were entered into this system, and two similar programs of automatic document indexing were realized giving a common search collection. For the purpose of automatic indexing, two descriptor dictionaries were compiled for the English and Russian indexing versions, where descriptors denoting classes of conditional equivalence (see Chapter 5) were the same in both dictionaries and corresponded to the same notions, though in different languages. This bilingual system was successful in finding information and formulating outputs containing documents both in the Russian and English languages.

It is clear that this method of realization is also applicable to other languages, such as German, French, or Spanish. In their work the researchers used well-known approaches in the field of IR systems development. Then what has NLP got to do with the described problem? The system proved to be unusable because an overwhelming majority of users did not know English. In other words, the system's ability to accomplish document retrieval in different languages was only the first feasible phase in the creation of a multilingual IR system. The development of a mechanism capable of presenting retrieved information in a language understandable to the user is the next phase.

Such mechanisms are commonly called machine translation, and they are directly related to NLP. NLP was developed first for machine translation (Tanimoto, 1990). In this particular case, the IR system developers hoped that this line of work would provide the required mechanism or at least significant results that would make such a mechanism feasible. Next we discuss whether this hope was realistic.

During the 1950s, researchers hoped for machines that would be able to translate a text from one language into another. Projects were funded to develop systems that could translate Russian into English and vice versa. Years later the researchers found it very difficult to make machines do quality translation. The systems used during the 1960s usually employed dictionaries, as well as morphological and syntactic analyses. They were unable to use semantic analysis, however, because no real progress had been made in this area.

Here is a well-known example that revealed the limitations of automatic translation. Researchers wanted to use the system to translate the expression "The spirit is willing but the flesh is weak" from English into Russian. However, after it had been translated into Russian and then back into English, the sentence had become "The vodka is good but the meat is rotten."

In time, the initial optimism of machine translation researchers evaporated into thin air. It became evident that the problem of programming a system to recognize the meaning of the text written in a natural language, if at all solvable, was a task of utmost complexity. Seeing no practical results, the U.S. government lost interest in this work, and after many years of waiting ceased to finance

it. In fact, the AI section of the Department of Defense manual for grants specifically states that no proposals on machine translation will be considered.

Thus, in practical terms machine translation was a flop. Nonetheless, almost two decades of intense work produced a number of interesting and useful achievements. We will mention only some of them.

First, the research showed once again that a *natural language should be considered as a whole*, and use and analysis of separate language components, such as its vocabulary or syntax, actually does not mean use and analysis of the natural language as such (Winograd, 1972).

Second, it became clear that it was impossible to analyze the meaning of what is expressed by means of natural language without a semantic analysis. Moreover, it became clear that to perform an analysis, a mechanism was needed that was not only based on formal rules for operating on meaning but also contained knowledge (Tucker, 1984).

Third, information retrieval benefited from a number of practical results. For example, morphological analysis (the stemming algorithm) was developed initially for machine translation and later found practical application in IR systems. The same applies to the lexical analysis. All of these methods are widely used for automatic indexing (see Chapter 6). In general, some semantic analysis techniques, such as the analysis of paradigmatic and syntagmatic relations, already have long been used in NLP. However, all attempts to use them in IR systems after the 1960s yielded no positive results.

Another problem frequently mentioned is the problem of the automatic abstraction of documents. Some researchers believe that solving this problem will also help in solving the problem of indexing full document texts (such as book texts). It should be noted that problems associated with automatic abstracting (by abstract we mean a coherent text reflecting in a short form the meaning of the original document) are similar in many respects to those discussed here in connection with machine translation. Indeed, successful automatic abstracting requires both meaning recognition and the ability to operate on it. This is why automatic abstracting is virtually unfeasible at present.

Also questionable is the usefulness of automatic abstracting for indexing purposes, because if the machine can recognize the meaning of the document (which is indispensable for automatic abstracting), then it will be able to recognize the meaning of the search request as well. In this case, there will be no need for indexing because after comparing the two meanings the machine will be able to decide immediately whether to include the document in the output. In other words, in this case the search will be made in natural language. Rather, automatic abstracting could be reasonably utilized in forming outputs, for in some cases it may be useful to output abstracts rather than full document texts.

Naturally, it does not follow from the previous discussion that the problem of indexing the texts of full documents is nonexistent. Researchers are trying to solve this problem through different approaches to document reduction. In do-