

# **PARS/U FOR WINDOWS: THE WORLD'S FIRST COMMERCIAL ENGLISH-UKRAINIAN AND UKRAINIAN-ENGLISH MACHINE TRANSLATION SYSTEM**

Michael S. Blekhman, Alia Rakova, Andrei Kursin.

Vice-President, POLYGLOSSUM, Inc.; Director, Lingvistica '93 Co  
94a Prospekt Gagarina, ap.1 11, Kharkov 310140, Ukraine.

Tel.: (0572) 27-71-35.

E-mail address: blekhman@lotus.kpi.kharkov.ua

## **Abstract**

The paper describes the PARS/U Ukrainian-English bidirectional MT system by Lingvistica '93 Co. PARS/U translates MS Word and HTML files as well as screen Helps. It features an easy-to-master dictionary updating program, which permits the user to customize the system by means of running subject-area oriented texts through the MT engine. PARS/U is marketed in Ukraine and North America.

## **1. Background**

Ukrainian is the state language in Ukraine, the second largest republic in the former Soviet Union. The country gained its independence in 1991. Before that, Russian had been mainly used in science, technology, business, and other branches of socio-economic life in Ukraine. At the same time, Ukrainian is the language of a very rich culture of European orientation, and it's only natural that Ukrainians want to revive it after a long period of a certain stagnation.

The difficulties are obvious: lack of dictionaries and standardized terminology. We are deeply convinced that it is impossible to solve this task without including Ukrainian in the orbit of language engineering.

The first Ukrainian-based machine translation engine developed in Ukraine was RUMP, a Russian-Ukrainian bidirectional system by Lingvistica '93 Co. We started marketing it as early as in 1990. RUMP runs on dozens of hundreds of computers in Ukraine.

Inspired with the success, we developed a system for translating between Ukrainian and English.

## **2. Lexicography**

The first problem to be solved was of lexicographic nature. We developed a bidirectional English-Ukrainian basic dictionary of 35,000 words and phrases, which contains not only general-usage, but also some business terms. When compiling the dictionary, we made use of some English and English-Ukrainian printed dictionaries, and we have also run numerous original texts through PARS/U to enlarge the dictionary.

We are also developing terminological dictionaries for PARS/U. Our company makes agreements with leading Ukrainian lexicographers to convert their dictionaries into PARS/U data bases. The dictionaries are automatically made bidirectional, English-Ukrainian-English.

A dictionary of 20,000 computer terms has been compiled, based on the English-Ukrainian dictionary by Dr.O.Kossak. The other specialist dictionaries for PARS/U are ecological, by Prof.V.Lozauskyy,

V.Beliavsky, and V.Borovitsky, 18,000 terms, telecommunications, by Dr.O.Kossak and Dr.R.Kravets, 3,500 terms, and automatic measuring systems, by Prof.M.Ginzburg, 6,000 terms.

1997 is supposed to be a very important year in PARS/U history. Lingvistica '93 Co. is launching an ambitious project of setting up what is called «Ukrainian-English lexicography». Dozens of Ukrainian-English dictionaries will be made for PARS/U, such as a large aviation dictionary created at Kharkov Aviation University.

### 3. Grammar

Another problem was working out a formal grammar of Ukrainian and a transformation grammar. We have made a thorough description of Ukrainian morphology and developed a Ukrainian grammatical dictionary. Dozens of rules for analysis and synthesis of Ukrainian texts have also been written and included in PARS/U.

It turned out that translating between English, a vividly-expressed analytic language, and Ukraine, a Slavonic language of expressed synthetic kind, is even more difficult than translating between English and Russian. The set of transfer rules in PARS/U is even more sophisticated than in PARS, the Russian-English bidirectional system. One of the examples is the rare use of participles in Ukrainian: English active participles are to be translated using Ukrainian participial clauses.

### 4. How PARS/U translates

PARS/U runs on IBM PCs under Windows 3.1, Windows 95, and Windows NT.

The system is linked with MS Word 6 and 7 for Windows. When the latter is activated, the «Translate» item is displayed in its main menu, which lets the user activate PARS/U. The user then opens an English or Ukrainian text in WinWord and has it translated by one keystroke. The target text appears under the source one, in a separate window created automatically by WinWord. The source text format is preserved in the target text, including fonts, styles, paragraphs, tables, etc. (see Fig. 1).

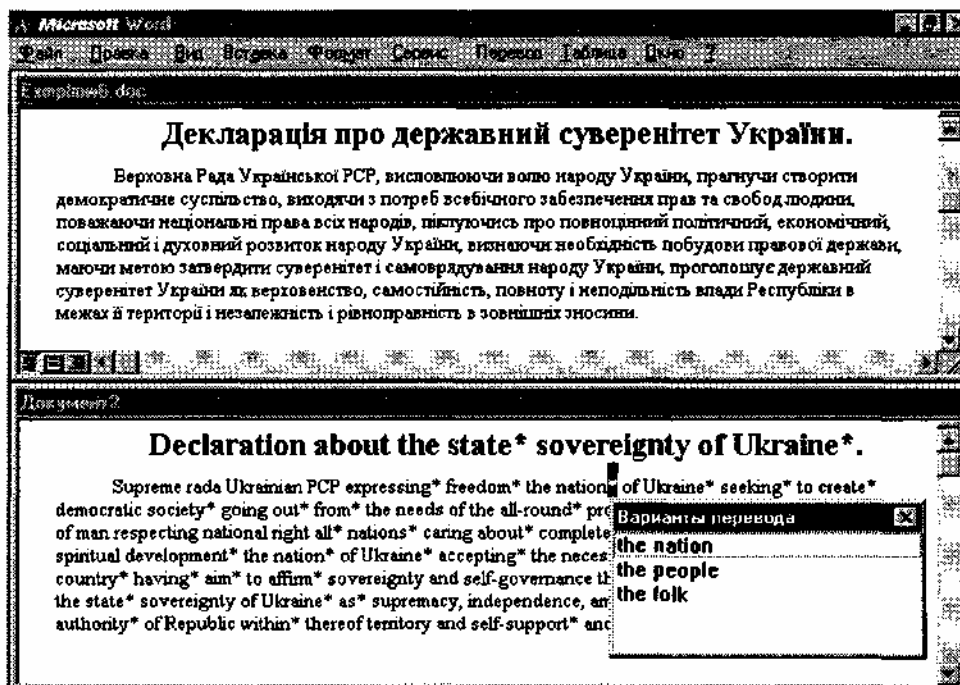


Fig.1 PARS/U has translated in WinWord the Ukrainian Declaration on Independence.

PARS/U also translates HTML files, as well as Windows screen Helps. To perform translation, the user, working, for example, in Netscape Navigator, marks the portion to be translated, activates PARS/U, and the translation pops up in a separate window under the source text ( see Fig. 2). The target file can be saved for subsequent editing, if necessary.

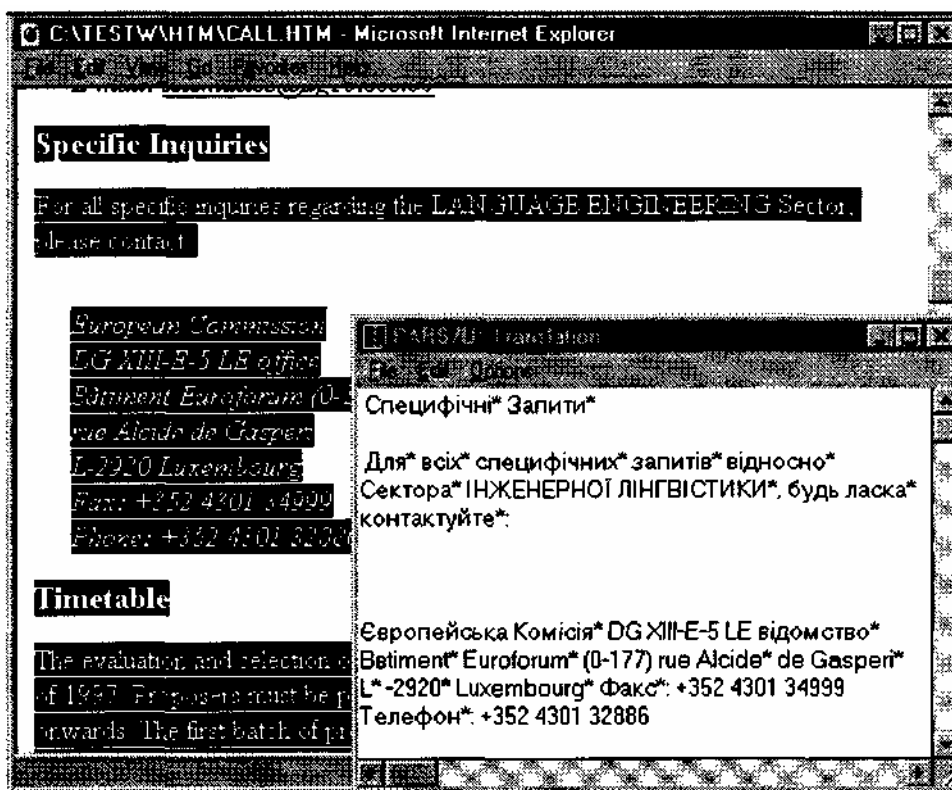


Fig.2. PARS/U has translated a WWW text on-line.

To compensate for lack of deep semantic analysis, PARS/U features marking polysemantic words with asterisks in the target text, so that the user could select a more appropriate translation if necessary. A double click on an asterisk displays a panel of translation options; another double click substitutes the initial word in the text with the word (phrase) selected by the user.

The last but not the least feature to mention is that PARS/U runs both in stand-alone and network modes. This, in particular, has made it possible to use PARS/U as a tutorial facility at the Department for intelligent information systems at Kharkov State Polytechnical University.

## 5. How PARS/U dictionaries can be updated

Dictionary updating facilities are very important for an operational MT system. They make it possible for the users to customize the system to their subject areas and raise translation efficiency substantially.

In PARS/U, the user may enter a word or phrase into the dictionary directly from the screen by marking it in the WinWord window and clicking on the "New word" item in the PARS/U menu; the word (phrase) will be immediately inserted into the dictionary.

This has proven to be a reliable and convenient way of customizing PARS/U. The user runs a representative corpus of his/her texts through PARS/U and creates a dictionary of specific terminology.

PARS/U features the "Auto-encoding" routine. It assigns grammatical characteristics to the Ukrainian words entered into the dictionary. When analyzing a word, PARS/U uses the Ukrainian grammatical

dictionary. Taking into account the high morphological ambiguity of Ukrainian words, PARS/U displays variants in those situations when it cannot choose between two variants, for example, when the noun/adjective ambiguity is encountered. In such situations, it is up to the user to make the final choice.

PARS/U also features so-called batch operations with dictionaries. In particular, the user may have 2 dictionaries merged, which lets him/her share dictionary compilation between several dictionary officers, and make one dictionary out of several ones.

PARS/U dictionaries are bidirectional: creating, for example, a Ukrainian-English dictionary, the user automatically obtains the Ukrainian-English-Ukrainian correspondence.

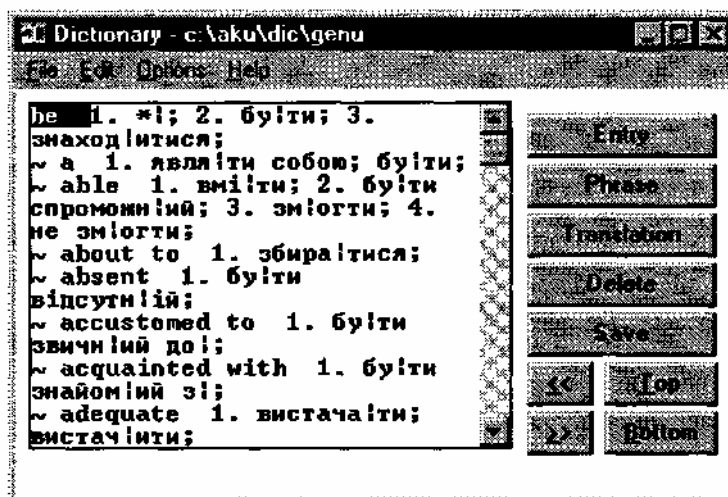


Fig.3. PARS/U dictionary entry

## 6. Where PARS/U is distributed

There are two large areas where PARS/U is commercially available, and these areas are the places where large Ukrainian communities exist: Ukraine itself, and the North America. In the latter, the system is marketed by Erudite Corporation Ltd., a company based in Toronto, Canada.