

Computer-Aided Translation Tools for Russian, Ukrainian, and English

Dr. Michael S. Blekhman	Maryna	Andrei	Dr. Igor	Alexander	Vladimir
	Byezhanova	Kursin	Backlan	Kazakov	Petrov
President, <i>Lingvistica '98</i> <i>Inc.</i>	Lingvist, <i>Lingvistica</i>	Program- mer,	Program- mer,	Program- mer, <i>Ling-</i>	Vice- president,
<i>E-mail: 924592@ican.net</i>	<i>'98 Inc.</i>	<i>Ling-</i>	<i>Ling-</i>	<i>vistica'98</i>	<i>Polyglossum,</i>
<i>Web:</i>		<i>vistica</i>	<i>vistica</i>	<i>Inc.</i>	<i>Inc.</i>
<i>members.tripod.com ~ling98</i>		<i>'98 Inc.</i>	<i>'98 Inc.</i>		

Abstract

The paper presents a new development by *Lingvistica '98 Inc.*: the PG-PARS computer-assisted translation system and a series of professional bidirectional dictionaries. PG-PARS was designed as a Windows 95 and Windows 98 application to support English-Russian-English and English-Ukrainian-English dictionaries. PG-PARS dictionaries are all bidirectional, i.e. they include, for example, both the English-Russian and Russian-English parts. Each part has its alphabetical index of entries displayed on a separate tab in the PG-PARS main window. The word entries are displayed for both parts of the dictionary, and translations of translations can be found easily, which is useful for a professional translator. One of the main peculiarities of PG-PARS is the **Smart search** mode based on morphological analysis of Slavic and English words looked up in the dictionaries. This is especially beneficial if the user is not a native speaker of Russian and Ukrainian. Another important feature is the **Selection option** which allows the user to mark a portion of the dictionary entry and paste it into the text. The presentation will show professional applications of the PG-PARS system for translating Russian and English texts.

1 Introduction

This paper presents a new development by *Lingvistica '98 Inc.*: the PG-PARS machine-aided translation system and a series of professional bidirectional dictionaries. It was our goal to:

- develop user-friendly dictionary look-up programs;
- compile representative bidirectional dictionaries for two language pairs: Russian and Ukrainian to and from English;
- make these translation tools available for professional translators as well as for language students all over the world, either separately or together with the PARS automatic translation systems.

PG-PARS was designed as a Windows 95 and Windows 98 application to support English-Russian-English and English-Ukrainian-English dictionaries. Two unique dictionaries have been developed:

- The Comprehensive English-Russian Bidirectional Dictionary of the XX Century, and a student-oriented dictionary based on it.
- The Comprehensive English-Ukrainian Bidirectional Dictionary of the XX Century, and a student-oriented dictionary based on it.

2 Linguistic aspects

As was mentioned above, the present PG-PARS version supports 2 unique bidirectional dictionaries, English-Russian and English-Ukrainian. Besides, a number of specialist dictionaries have been and are being compiled, covering such subject areas as aviation, military, engineering, electronics, medicine, social-political, and others. To compile them, we have used several sources, in particular:

- Large English-Russian Dictionary by Prof. Yu. Apresian et al.
- PARS MT Social-Political English-Russian-English Dictionary.
- PARS as well as printed English-Russian and Russian-English technical dictionaries.
- PARS as well as printed Russian-Ukrainian dictionaries.
- Numerous modern newspaper texts.

When analyzing existing English-Russian and Russian-English terminological dictionaries published in the Soviet Union and Russia within the 1970s-1990s, we discovered a number of typical mistakes often made by dictionary authors. Our conclusions may be of interest to computer lexicographers developing MT and MAT oriented professional dictionaries. It turned out that, rather often, the authors of the dictionaries have poor engineering background. And the lack of expertise, in its turn, results in erroneous word-entries. Here are some typical examples.

- (1) A slang expression is used instead of a technical term, for example, *mating area* instead of *assembly area* or *erecting area*.

(2) A general usage word is used instead of a technical term, such as *light-off velocity* instead of *ignition velocity*.

(3) Misuse of terms: *traffic-handling capacity* instead of *guidance ability* in a dictionary on military aviation (a civil aviation term was used instead of a military one).

(4) Illogical expressions, for example: *manned antitank missile*, which means an antitank missile controlled by a person sitting in it: of course, a missile can only be controlled from outside, so it cannot be manned.

Besides, numerous mistakes have been observed in Russian translations of English technical terms, again due to poor understanding of technical peculiarities.

All these problems were taken into account when developing PG-PARS professional dictionaries.

We also face the necessity of «inventing» new terms when developing Ukrainian-English specialist dictionaries as well as making social-political dictionaries for both language pairs. The latter is, in particular, due to wide use of criminal slang in modern political life in Russia.

dictionaries. This is especially vital if the user is not a native speaker of Russian and Ukrainian.

Another important feature is the **Selection** option which allows the user to mark a portion of the dictionary entry and paste it into the text.

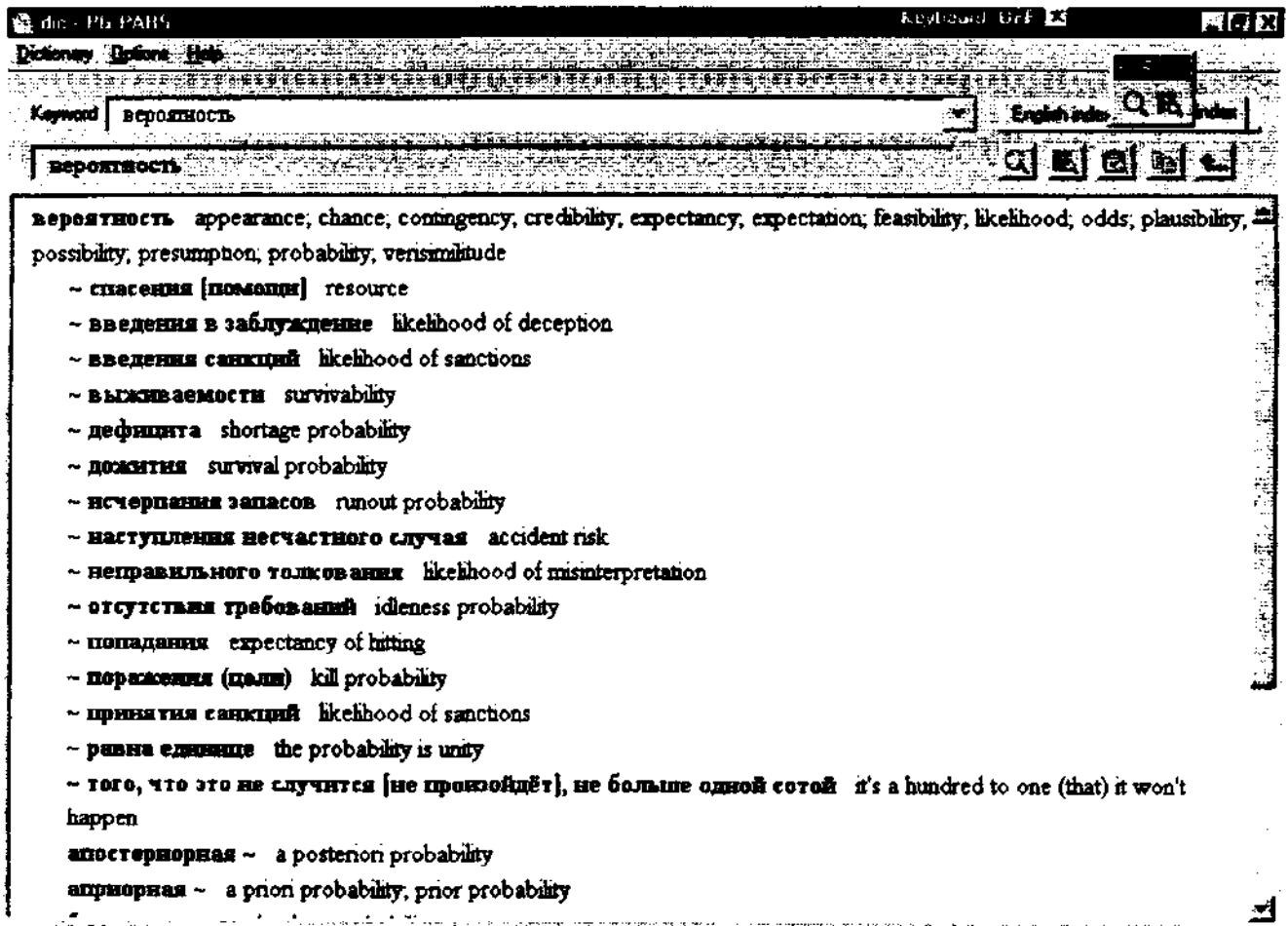
3.1 Dictionary structure

PG-PARS dictionaries are organized like traditional printed-on-paper dictionaries.

A dictionary entry consists of the head word, its translations, phrases comprising the head word, their translations, comments and examples.

The word entry structures may vary slightly from dictionary to dictionary due to the preferences of their authors.

PG-PARS dictionaries are all bidirectional, i.e. they include, for example, both the English-Russian and Russian-English parts. Each part has its alphabetical index of entries displayed on a separate tab in the PG-PARS main window. The word entries are displayed for both parts of the dictionary, and translations of translations can be found easily, which is sometimes very useful for professional translators.



3 Technological aspects

One of the main peculiarities of PG-PARS is the **Smart search** mode. It is based on morphological analysis of Slavic and English words looked up in the

The traditional search method can be used, which consists in looking up a single word or the phrase keyword in the index and then examining the dictionary entry. This is called **Simple search** in PG-PARS. Besides, the **Smart search** mode can be used (see below).

To see the word entry including the translations, the user selects a word in the index and makes a double click on it

3.2 Smart search

The smart search mode is one of the most important features in PG-PARS. In this mode, a word or phrase is searched regardless of the inflection form in which the query is specified. This allows the user to start searching by simply dragging a word or a phrase from the source text to the special icon above the dictionary window.

If the smart search succeeds, the corresponding entry is displayed in the dictionary entry window. Otherwise, simple search on the query is performed, i.e. the text of the query is placed in the keyword box and the index is positioned accordingly. The smart search mode may fail because the word/phrase is absent in the dictionary or due to certain homonymy of word forms or endings. However, our experiments show that search precision in the smart search mode is rather high: 97% for Russian and 98% for Ukrainian, that is purely linguistic mistakes are rather rare. And, what is especially important, we do not assign grammatical characteristics to the words in PG-PARS dictionaries: the search engine makes use of the PARS MT system morphological tables, so the lexicographer doesn't have to spend his/her time tagging hundreds of thousands of words.

To initiate smart search, the user drags the word/phrase to be found from another application (text editor) and drops it onto the special PG-PARS icon above the PG-PARS window.

3.3 Selecting text in the dictionary entry

Another peculiarity is the option for flexible selection of elements in a dictionary entry, such as translation, source phrase, etc., by double-clicking it.

A dictionary entry element may have variants given in square or curly brackets. To select the variant including text in brackets, the user double-clicks inside, otherwise outside them. Here is an example:

- bring to true [uniform] surface

double-click on **bring** selects **bring to true surface**;
double-click on **uniform** selects **bring to uniform surface**.

- tow(ing) rope

double-click on **tow** selects **tow rope**;
double-click on **ing** selects **towing rope**.

It is also possible to select a portion of an entry element by pressing the left mouse button and then dragging the mouse. In this mode, the user can select italic text (all kinds of comments). Selecting an element is performed for copying the translation to another application or another text. That is, when typing the target text, the user can select a (the) translation of the word in question and pastes it into the text. This is how the user does that. When in the dictionary entry window, he/she selects an entry element and clicks on the *Copy* button to

copy the selected text into the Clipboard, or drags the selection to the target application.

If a tilde is present in the selection, it will be automatically replaced with the word it substitutes. Brackets, semicolons in the ends of the translations, etc., are not copied.

It is also possible to drag the text selected in the dictionary entry windows to another application.

Many professionals find it useful to examine back translations of translation variants before choosing one of them. When in the dictionary entry window, the user selects the translation and clicks on the *Show entry* button, after which the entry for the text selected is displayed. The user can also return to the previously displayed entry by clicking on the *Back* button.

The last but not the least, PG-PARS supports word look-up in several dictionaries. The user may set up the dictionaries to be used in the translation session as well as their priorities.

The presentation will show professional applications of the PG-PARS system for translating Russian and English texts.