Introduction

Nowadays, it is difficult to imagine contemporary linguistics without semantics and linguistic confrontation, cognitive studies, as well as corpus linguistics and computer linguistics.

The former Semantics Group, at present the Semantics and Corpus Linguistics Group, of the Institute of Slavic Studies of Polish Academy of Sciences (ISS PAS) was the first team of scholars to develop a *semantic interlanguage for parallel comparisons of two or more languages* (the multivolume Polish-Bulgarian Contrastive Grammar). Our experience in language confrontation and semantics helped us establish our position in linguistics as a Group dealing with semantic issues rather than — traditionally — just with the grammatical (formal) aspects of the compared languages. Our practice has shown that what is important for developing electronic dictionaries and translations are first of all the so-called *parallel corpora*.

Now, we are endeavouring to not only elaborate such corpora, but also to construct on their basis bi- and trilingual dictionaries, both in the book form and electronic ones (the latter in cooperation with our colleagues — IT specialists).

At present, our Group is developing two bilingual dictionaries: *Bulgarian-Polish* and *Polish-Lithuanian*, and one trilingual dictionary: *Polish-Bulgarian-Russian*. We are trying to make use of the accessible language materials, extending them, as I have mentioned above, with our own parallel bilingual and trilingual corpora.

Parallel corpora are an important tool of contemporary contrastive linguistics. They are used, for example, in developing both electronic and traditional multilingual dictionaries (including specialist dictionaries: of technical, legal, terminology, etc.), as well as elaborating contrastive grammars and textbooks. They also enable simple, fast and easy acquisition of rich and reliable material for syntactical, lexical and semantic contrastive studies on various issues. Corpora are also applied in foreign language teaching and in everyday work of professional translators.

The Semantics Group (former name of our Group) developed (under cooperation with Institute of Mathematics and Informatics of Bulgarian Academy of Sciences (IMI BAS)) the *first bilingual electronic corpus of Bulgarian and Polish* (numbering over 5 million wordforms) (Dimitrova, Koseska). The consecutive version of the corpus — consecutive, because our corpus is being developed and enriched all the time — contains both literature (prose, journalism) and specialist vocabulary (text documents of the European Commission and the European Union).

The aim of our work (performed jointly with IT specialists in IMI BAS) was to design and construct a database handling the lexical resources of an *online*

version of the Bulgarian-Polish Dictionary (Dimitrova, Koseska-Toszewa, Satoła-Staškowiak et al.).

In our Group, we have also developed (again in cooperation with IT specialists of IMI BAS) the first *electronic Bulgarian-Polish-Lithuanian Parallel Corpus* (numbering over 3 million wordforms). The corpus is intended for experimental and comparative studies, and its authors are L. Dimitrova, V. Koseska-Toszewa, R. Roszko and D. Roszko.

At present, *we will be working — under the international Clarin Project —* on an annotated parallel corpus of Polish, Bulgarian and Russian texts equalized on the sentence level. A *parallel Polish-Bulgarian-Russian Corpus* will be constructed (M. Duškin, V. Koseska, J. Satoła-Staškowiak, A. Kisiel). The corpus will contain several subcorpora: a subcorpus of prosaic works (fiction; planned size — about 1.5 million tokens in each of the three languages), a subcorpus of poetry (about 500 thou. tokens), and a small, universal corpus (about 1 million tokens), encompassing texts from different genres (fiction, journalism, scientific literature, etc.) presented in strictly specified proportions. Within Clarin, we will also develop a *Parallel Polish-Lithuanian Corpus* (minimum 8,000,000 forms) (R. Roszko, D. Roszko). Our aim is to develop annotated parallel corpora equalized on the sentence level, which in future will be available on the Internet.

In turn, on the example of our trilingual dictionary — the *Russian-Bulgarian-Polish Dictionary* (the order of languages is irrelevant) by M. Duškin, V. Koseska-Toszewa and J. Satoła-Staškowiak — we want to show the role of a parallel corpus in developing a dictionary. We are planning to transform the book version of the dictionary into its electronic version. This task is not easy, but it is feasible if each of the languages is described with strict separation between the form and the meaning, and with help of new semantic classifiers, not used in dictionaries until now.

The extended scope of our activities is shown in the present CS|EC volume.

I would like to emphasise here our active cooperation with Polish university centres. We are also glad to be cooperating more intensively than before with Ukrainian scientific centres, and especially with Prof. V. Shyrokov, who has become yet another foreign member of our Editorial Board.

Prof. **Volodymyr Shyrokov, Kyiv, Ukraine**, (vshirokov48@mail.ru) is a well-known scientist working on mathematical, theoretical and applied linguistics. He is a founder of the Ukrainian scientific school in language technology, and the permanent director of the Ukrainian Lingua-Information Fund. Professor Shyrokov is also a Corresponding Member of the National Academy of Sciences of Ukraine, and a laureate of the State Prize of Ukraine in Science and Technology. He is the author and co-author of over 200 scientific works, including 9 monographs, 14 dictionaries, and 15 different linguistic technologies (virtual lexicographical laboratories, text analysers, linguistic corpora: <http://lcorp.ulif.org.ua>). He is also the scientific director of the 20-volume Ukrainian Language Dictionary (the first volume in 2010), as well as the author of the lexicographical systems theory and the theory of semantic states of language units.

I would like to inform all readers interested in cooperation with our journal that for the next issue we await submission of articles in the area of semantics and linguistic confrontation, cognitive studies, as well as corpus linguistics and computer linguistics, by 15 October 2012. We are also interested in fragments of doctoral theses of young scientists in the above area.

We hope that our work will be beneficial for the contemporary trends in translations and multilingual communication.

Violetta Koseska-Toszewa