NKFP-5/135/01: Complex Uralic Linguistic Database

The aim of the three module project was to set up electric databases and morphological analysers for Uralic languages. Uralic Etymological Database: electric database of the common Uralic vocabulary, Uralic morphological analyser: computer-based morphological analysis of nine smaller Uralic languages with a unified morphological and morphonological description as background (texts are analysed by a program showing the morphological components of the words), Ostyak (Khanty) monolingual electronic dictionary for the Sinya dialect, and Surgut-Khanti folklore texts with translations and recorded sound materials on CD. All three moduls were missing not only from the Hungarian but from the international Finno-Ugric researches as well. Languages/dialects involved are endangered ones. Electric etymological database did not exist according to our information. The Uralic Etymological database is not a simple electronic dictionary; it offers the function of multiple database querries and ensures the possibility of its expension too. The main scientific result of the morphological analyser is new informations on the morphological and morphonological characteristics of nine Uralic languages (Northern Saami, Erzya-Mordva, Mari, Udmurt, Komi, Northern Mansi, Surgut-Khanti, Nenets and Nganasan), compilation of the linguistic background of the analysis-program for nine languages, and setting up of the analysers for six languages. The Khanty monolingual dictionary and the folklore texts are giving new informations on two, less revealed, endangered dialects. The Sinya-Khanti monolingual dictionary provides us with the mental lexicon of a single speaker, whereas the Surgut material is the compendium of folklore traditions.

The main objectives of the contract have been fulfilled by the consortium, i.e. the operability of the Uralic Etymological Database has been created, the morphological analysers of Komi, Udmurt, Mari, Northern Mansi, Nenets, and Nganasan languages/dialects have also been produced, and the linguistic background of Northern Sami, Erzya-Mordva, and Surgut Khanty analysers have been elaborated, the material of the Synya Khanty monolingual dictionary has been edited, and prepared for electronic conversion. In addition, the folklore texts running to twelve printed sheets were taken down and translated, furthermore, with the recorded texts, the publishing has also commenced. In the course of the elaboration of the morphological analysers, a number of points, which had been problematic or absent from former descriptive grammars, have been cleared up, which means at the same time, that it is not only the analysers that resulted from this work but also it led to a more precise knowing of the languages in question. The same applies to the Synya Khanty monolingual dictionary: in the course of the work several phonological and semantic problems were solved. The most important result of the project is that it is a number of endangered, or moribund (e.g. Mansi, Nganasan) languages whose material was processed, saved, furthermore, by publishing the above mentioned texts via internet we may enhance the linguistic identity of the speakers of the languages concerned.