NKFP-5/135/01: Komplex uráli nyelvészeti adatbázis

A projektum célja három komponensből álló számítógépes adatbázis, illetőleg elemző létrehozása volt: Uráli Etimológiai Adatbázis: az uráli nyelvek közös szókincsének szótári adatbázisa; Uráli Morfológiai Elemző: kilenc uráli nyelv egységes szempontú morfológiai leírásán alapuló számítógépes morfológiai elemzése (a nyelvi szövegeket a program elemzi, s megmutatja, hogy a szöveget alkotó szavaknak melyek az alaktani összetevői); Osztják (hanti) értelmezőszótár és szöveggyűjtemény: az osztják nyelv szinjai nyelvjárásának értelmező szótára, valamint szurguti nyelvjárásának folklór szöveggyűjteménye (hangzó anyaggal) elektronikus adathordozón. Mindhárom résztéma hiányzott nemcsak a hazai, de a nemzetközi finnugrisztikai kutatásokból is. A leírt, feldolgozott nyelvek veszélyeztetettek.

A CD-n terjesztendő számítógépes etimológiai adatbázis egyáltalán nem létezik a világon. Az adatbázis nem egy elektronikusan olvasható szótárt kínál, hanem olyan anyagot, amely különféle szempontok szerint teszi lehetővé a tartalom lekérdezését, bővítését. A morfológiai elemző legfőbb tudományos eredménye kilenc kisebb uráli nyelv – az északi számi (lapp), a(z erza) mordvin, a mari (cseremisz), az udmurt (votják), a komi-zürjén, az északi manysi (vogul), a szurguti hanti (osztják), a nyenyec és a nganaszan – morfonológiai és morfológiai tulajdonságairól való ismereteinknek bővülése, a számítógépes elemző program hátterét képező nyelvi leírások elkészítése és hat nyelv elemzőjének a létrehozása. A hanti értelmezőszótár és szöveggyűjtemény két, eleddig csak szerényebben ismert, veszélyeztetett nyelvjárás feltárása. A szinjai értelmezőszótár egyetlen közlő teljes mentális lexikonát tartalmazza elektronikus formában. A szurguti gyűjtések e nyelvjárás folklór anyagát ölelik fel, s teszik közzé hangzó-anyaggal együtt CD-n.

A szerződésben vállalt főbb célkitűzéseket a konzorcium teljesítette: működőképesen létrehozta az Uráli Etimológiai Adatbázist, elkészítette működőképes állapotban a komi-zürjén, az udmurt, a mari, a manysi, a nyenyec és a nganaszan morfológiai elemzőt, továbbá kidolgozta az északi-számi, az erza-mordvin és a szurguti hanti elemzők nyelvészeti hátterét; szerkesztett állapotba hozta és elektronikus szótári konvertálásra előkészítette a szinjai osztják értelmező szótárat, továbbá teljesítette a kb. 12 ív terjedelmű folklóranyag lejegyzését és fordítását valamint elkezdte a hangzó anyaggal együtt annak CD-n való közzétételét. A morfológiai elemzők kidolgozása során egy sor, eddig tisztázatlan vagy a korábbi nyelvészeti feldolgozásokból teljes mértékben hiányzó grammatikai részkérdésre derült fény, ami azt jelenti, hogy nem csupán az elemzők, illetőleg azok nyelvi háttere alakult ki, hanem e nyelvek pontosabb megismerése is lehetővé vált. Ugyanez mondható el a szinjai hanti értelmezőszótár munkálatairól: egy sor hangtani és jelentéstani kérdés tisztázásához segített hozzá. A projekt legfőbb hozadéka az erősen vagy kevésbé erősen veszélyeztetett nyelvek, illetőleg egyes esetekben a teljes kihalás szélén álló nyelvek (pl. manysi vagy nganaszan) nyelvi anyagának feldolgozása, mentése, a világhálón közzéteendő elemzők révén pedig e beszélők nyelvi öntudatának megtámogatása.