



ÜLLE VIKS 60

Eelmise juubeli puhul tehtud kiriusutetus (KK 1998, nr 4) ütleb Ülle Viks, et erialaliselt ei oska ta end kuhugi paigutada: informaatik ei ole ta kindlasti mitte, päris keeleteadlane ka mitte, sest on liiga formaalsust armastav, ja arvutilingvistiks ka ei sobi, sest arvutitarkvara ehk programmeerimise eest hoolitsevad teised. Kas see ongi nii tähtis, aga kui tahta nime anda, siis on ta ehk meie kõige ehtsam arvutileksikograaf. Olen kindel, et arvutisõnastike loomisel oleks Eesti Keele Instituut ilma Ülle Viksita hoopis teises seisus.

Jagaksin oma mõistusega Ülle tegemised kolme rubriiki.

Avatud ehk reeglipõhine morfoloogia

Ülle Viksi esimese põhiteema võib kokku võtta väljendiga *avatud morfoloogiamudel*. Avatud morfoloogiamudel on avatud kahes mõttes. Esiteks ei sõltu see etteantud sõnade hulgast, s.t sõnastikust. Sõnastik sisaldab alati lõpliku hulga sõnu, kuid avatud morfoloogiamudelil põhinev süsteem – nagu inime-negi – tunneb ära ja oskab kasutada ka sõnu, mida ta leksikonis ei ole (pole enne kohanud): tal on reeglid, mida ta kasutab (ses mõttes on mudel reeglipõhine). Teiseks on mudel avatud ka ses mõttes, et reegleid võib ümber teha, ilma et kogu mudelit tuleks muuta. Kokku on see inimese morfoloogiapädevuse mudel. Sellel põhinevad morfoloogilised protsessid: analüüs ja sünteis. Analüüsi puhul annab süsteem suvalise tekstis oleva sõna kohta morfoloogilise info. Sünteisi puhul moodustab ta

etteantud morfoloogiliste tunnuste põhjal eesti keele korrektse sõnavormi. Arvutilingvistika seisukohalt on see puhas näide panna arvuti toimima nagu inimene. See vastandub Filosoofi lähenemisele, mille alusprintsip on efektiivsus, kiirus, sõltumata sellest, kuivõrd sarnane see inimese vastavate protsessidega on. Seejuures ei saa jätta märkimata, et Filosoofi tehtud speller põhineb Ülle Viksi „Väikesel vormisõnastikul”. Ütleksin küll, tänu jumalale, et meil mõlemad lähenemised esindatud on.

Grammatika ja sõnastikud

Teise Ülle Viksi põhiteema võib võtta kokku sõnaühendiga *sõnastikud ja grammatika*: missugust grammatilist (morfoloogilist) infot eesti keele sõnade kohta sõnastikes (näiteks kakskeelsestes sõnastikes) anda ja kuidas seda teha. Selleks on ta arendanud nn grammatilist kirjegeneraatorit kui leksikograafi töö abivahendit (vt KK 2000, nr 7). Sõnastikus on ju sõnad algvormis, ent teksti lugev või kirjutav inimene peab suutma kasutada tekstis olevaid või vajaminevaid vorme. Eriti oluline on see morfoloogiarikaste keelte puhul, nagu seda on näiteks eesti keel. Ülle Viksi idee on, et selleks, et leksikograaf saaks tegelda oma põhitööga ehk niisiis sõnadega kui leksikaalsete üksustega, mitte nende erinevate grammatiliste vormidega, tuleb luua automaatne süsteem, mis vastava rubriigi sõnaartiklis täidab. On ilmselge, et see eeldab avatud ehk reeglipõhist morfoloogiamudelit, millest eespool juttu oli.

Sõnastike haldussüsteem ehk leksikograafi töövahend/keskkond

Ülle Viksi kolmas põhitöö, mis tegelikult haakub kahe eelmisega, on üldine projekt „Leksikograafi töökeskkond” (ehk EELex). Ülle Viksi põhiosa selles on „Eesti-X-keele sõnastik”. Selle all mõeldakse kakskeelse (eesti lähtekeele) sõnaraamatu elektroonilist põhja, s.t arvutisüsteemi, mille võib aluseks võtta, kui tehakse suvalist kakskeelset sõnastikku, kus lähtekeeleks on eesti keel. EELex on taotluse järgi üldine süsteem, mida saab kasutada (minu arusaamist mööda) universaalselt, ka teiste keelte puhul. EELex-i praegune seis on järgmine.

Loodud on leksikograafi töökeskkonnatuum, sõnastike haldussüsteem, mis muudab sõnastike koostamise ja toimetamise lihtsamaks ja tõhusamaks. EELex on veebipõhine töövahend, mille olulisemad omadused on:

- andmete hoidmine XML-vormingus ja sõnastiku struktuuri esitamine XML-skeemi vahendusel;
- võimalus koostada ja toimetada sõnaartikleid kollektiivselt ja mitmel kasutajasemel;
- sõnastikuteksti esitamine mitmes eri vaates: XML-vorming, tabel, küljendusvaade;
- võimalus teha struktuuripõhiseid päringuid;
- võimalus eksportida sõnastiku tekst MS Word'i küljendusvormingusse.

EELex-i on seni kasutatud või kasutatakse kaht tüüpi sõnastike koostamisel

ja toimetamisel: 1) ükskeelsed (ÕS 2006 jt) ja 2) kakskeelsed („Eesti-vene sõnaraamat” jt). Uued rakendustüübid on 3) terminoloogiasõnastik ja 4) „Eesti-X-keele sõnastiku” andmebaas. Loodud on ka nendega seotud uued funktsioonid: illustratsioonide kasutamine, andmete import muudest EELex-i ressurssidest.

1. Koostatud on „Eesti-X-keele sõnastiku andmebaasi” (EXS) esialgne versioon:

- koostatud on märksõnastik;
- märksõnastikuga on liidetud „Eesti-vene sõnaraamatu” eesti keele materjal;
- alustatud on „Eesti-X-keele sõnastiku” toimetamist EELex-i vahenditega;
- ette on valmistatud stiili- ja erialamärgendite andmebaas (aluseks „Eesti õigekeelsussõnaraamat ÕS 2006”) „Eesti-X-keele sõnastikuga” liitmiseks.

2. Loodud on ÕS 2006 avalik veebifunktsioon ja täiendatud selle rakendusfunktsioone:

- esimene versioon (lihtpäringuga) on avalikus kasutuses aastast 2006;
- lisatud on liitsõnade täisotsing;
- lisatud on komplekspäring, mis võimaldab mitme otsitunnuse kombineerimist ja struktuuripõhist otsingut;
- loodud muude ressurssidega linkimise mehhanism.

Kokkuvõtvalt võib kindlasti öelda, et Ülle Viks on juba ära teinud suure töö ja loodetavasti jätkab ta samas vaimus. Kordaksin üle alguses öeldud mõtet: mis oleks EKI ilma Ülle Viksita?

HALDUR ÕIM