

Eiríkur Rögnvaldsson

Tungutækni-verkefni sem Orðabók Háskólans tekur þátt í

Frá árinu 2001 hefur Orðabók Háskólans rekið *Íslenskt upplýsingasetur um tungutækni*, sem kostað var af norrænu tungutækniáætluninni fram á mitt ár 2005. Setrið hefur tekið þátt í samstarfsneti norrænna upplýsingasetra, *NorDokNet*, www.nordoknet.org, en starfstíma þess lauk einnig á árinu. Upplýsingasetrið beitti sér fyrir samningu íslensks iðorðasafns í tungutækni, en fé fékkst til þess úr norrænu tungutækniáætluninni. Valdís Ólafsdóttir, meistaranemi í tungutækni, var ráðin til að leggja drög að safninu og vann hún við það um þriggja mánaða skeið (í 75% starfi). Verkið var unnið í samvinnu við Orðabanka Íslenskrar málstöðvar, og var iðorðasafnið opnað sem hluti bankans 1. september.

Þegar tungutækniáætlun menntamálaráðuneytisins lauk og verkefnisstjórn í tungutækni var lögð niður um áramótin 2004-5 yfirtók upplýsingasetrið vef hennar, www.tungutaekni.is. Til að halda utan um og ýta undir starfsemi á sviði íslenskrar tungutækni var ákveðið að stofna *Tungutækni-setur*, sem er samstarfsverkefni Orðabókarinnar, Málvísindastofnunar Háskóla Íslands og tækni- og verkfræðideildar Háskólans í Reykjavík. Í stjórninni situr einn fulltrúi frá hverjum samstarfsaðila, og er Sigrún Helgadóttir fulltrúi Orðabókarinnar. Meðal markmiða setursins er að halda árlega ráðstefnu um íslenska tungutækni, og verður sú fyrsta haldin nú í maí.

Norræna nýsköpunarmiðstöðin (Nordisk Innovationscenter) og Norræna rannsóknarmiðstöðin (NordForsk) veittu norrænum upplýsingasetrum um tungutækni og samtökum iðnaðarins á Norðurlöndum styrk til forverkefnis sem nefnt var *NLTNet*. Styrknum var varið til að halda sameiginlegan fund í Kaupmannahöfn þar sem ræddar voru forsendur fyrir samvinnu um kynningu og eflingu tungutækni á Norðurlöndum. Þar var samþykkt að stefna að áframhaldandi samstarfi, og var gefin út skýrsla um niðurstöður fundarins – sjá www.nlt.net.org. Snemma árs 2006 sendi sami hópur svo umsókn til Norrænu nýsköpunarmiðstöðvarinnar um nýtt tveggja ára verkefni, *Nordic ICT and Language*, þar sem ætlunin er að fylgja eftir þeirri stefnu sem mörkuð var í forverkefninu. Þegar þetta er ritað er ekki komið í ljós hvort styrkur fæst til þessa verkefnis.

Orðabókin og Háskóli Íslands tóku þátt í *Nordisk Netordbog*, samstarfsneti um rannsóknir og þróun í tungutækni, einkum margmála leit á netinu og í gagnabönkum. Aðrir þátttakendur í verkefninu eru Háskólinn í Bergen, Kungliga tekniska höskolan (KTH) í Stokkhólmi, Háskólinn í Helsinki og Center for sprogteknologi (CST) í Kaupmannahöfn, sem stýrir verkefninu (verkefnisstjóri Bente Maegaard). Norræna ráðherranefndin átti frumkvæði að þessu verki og hefur veitt fé til þess, ásamt Málráði Norðurlanda. Vinnan í verkefninu á árinu fólst einkum í því að afla ýmissa tví- eða margmála orðalista til að nýta í margmála leit. Þrjú fundir hafa verið haldnir í verkefninu og hefur Eiríkur Rögnvaldsson sótt þá. Framhald þessa verkefnis, *Tværsproglig søgning på tekster og ordbøger*, fékk svo styrk frá Nordplus Sprog í lok ársins. Þessi verkefni halda því áfram á árinu 2006.

Auk þessa átti Orðabókin aðild að þremur fjölbjóðlegum umsóknum um styrki til tungutækni-verkefna á árinu 2005. Þar var í fyrsta lagi um að ræða verkefnið *Nordic Multilingual Technologies*, umsókn til NOS-HS um norrænt öndvegissetur í margmála upplýsingatækni. Forsvarsmaður umsóknarinnar var Koenraad de Smedt prófessor í Bergen. Í öðru lagi var verkefnið *NorPar*, umsókn til NORA, Norrænu Atlantshafsnefndarinnar, um gerð samskipaðrar málheildar (parallel korpus) fyrir dönsku, norsku, sænsku, færeysku og íslensku, með áherslu á tvö síðastnefndu málin. Forsvarsmaður umsóknarinnar var Janne

Bondi Johannessen prófessor í Osló. Í þriðja lagi var verkefnið *EuroDocNet*, umsókn í 6. rammaáætlun Evrópusambandsins. Hans Uzkoreit prófessor í Saarbrücken var forsvarsmáður umsóknarinnar, en markmið verkefnisins var að koma á evrópsku samstarfsneti upplýsingasetra í tungutækni, koma upp upplýsingasetrum í Eystrasaltslöndunum, koma upp sameiginlegum gagnagrunni með upplýsingum um tungutækni, þróa aðferðir við margmála leit, þróa og samhæfa íðorðaförða greinarinnar á mörgum tungumálum o.fl.

Engin þessara þriggja umsókna hlaut brautargengi, en þátttaka í þeim var samt mjög gagnleg fyrir Orðabókina. Með henni fékkst dýrmæt reynsla af samningu umsókna af þessu tagi, auk þess sem Orðabókin komst í tengsl við ýmsar erlendar stofnanir og einstaklinga sem starfa á sviði tungutækni. Síðastnefnda umsóknin hefur nú verið endurskoðuð og send aftur til Evrópusambandsins en niðurstöðu er ekki að vænta fyrr en í haust.