

Pórdís Úlfarsdóttir

ISLEX – norræn margmála orðabók

1 Inngangur

ISLEX er samnorrænt, rafrænt orðabókarverk sem er unnið hjá orðfræðisviði Stofnunar Árna Magnússonar í íslenskum fræðum (SÁ) (áður Orðabók Háskólans) í samstarfi við fjórar aðrar háskóla- og fræðastofnanir á Norðurlöndum, í Danmörku, Noregi, Svíþjóð og Færeyjum. Samstarfsaðilar verkefnisins eru *Det Danske Sprog- og Litteraturselskab* (DSL) í Kaupmannahöfn, *Institutt for lingvistiske, litterære og estetiske studier* við Háskólann í Bergen, *Institutionen för svenska språket* við háskólann í Gautaborg og *Fróðskaparsetur Føroya* í Pórhöfn.

ISLEX er veforðabók milli íslensku og fimm annarra norrænna mála. Hún hefur að geyma tæplega 50.000 íslensk uppflettiorð en markmálin eru danska, norska (bókmál og nýnorska), sænska og færeyska. ISLEX er eina norræna orðabókin hingað til sem sameinar svo mörg mál í einu verki. Hún er fyrst og fremst orðabók yfir nútímaíslensku og taka uppflettiorð, notkunardæmi og orðasambönd einkum mið af málnotkun samtímans. Verkinu er ætlað að svara brýnni þörf fyrir handhægar orðabækur milli íslensku og hinna norrænu málanna þar sem þær bækur sem til eru hafa flestar úrelst og uppfylla ekki lengur kröfur tímans. En um leið er markmið aðstandenda verkefnisins að búa til fjölpætta, nýstárlega og vandaða orðabók sem nýst getur sem flestum, Íslendingum og útlendingum, tungumálanemendum, kennurum og þýðendum, og raunar öllum þeim sem þörf hafa fyrir slíkt verk. Vinnan við verkið stóð að mestu leyti yfir á árunum 2006–2011 og var ISLEX

opnað fyrir almennt aðgang í nóvember 2011, enda þótt ýmsum verkþáttum væri þá ekki fulllokið.

Í þessari grein er fjallað um ISLEX-verkefnið og er megináhersla lögð á þá þætti sem snúa að íslenska hluta orðabókarinnar. Efninu er skipað á eftirfarandi hátt: 2. kafli lýsir forsögu og upphafi verkefnisins, 3. kafli fjallar almennt um rafrænu orðabókina, 4. kafli greinir frá verklagi við ritstjórn í fjölþjóðlegu samstarfi, í 5. kafla segir frá efni og uppbyggingu orðabókarinnar, í 6. kafla er lýst notendaviðmóti hennar á vefnum og í 7. kafla er fjallað um ýmsa þá möguleika sem búa í svona verki og hugað að framhaldi þess. Lokaorð eru í 8. kafla.

2 Upphaf verkefnisins

Upphaf ISLEX-verkefnisins má rekja til tíunda áratugarins, um það bil áratug áður en byrjað var á sjálfri orðabókarvinnunni.¹ Árið 1994 réðst Orðabók Háskólans (OH) í samvinnu við Norræna málstöð í Osló í að semja íslenskan orðabókarstofn sem orðið gæti grunnur að tvímála orðabókum milli íslensku og annarra Norðurlandamála. Til verksins fékkst styrkur til tveggja ára frá Norræna menningarsjóðnum. Orðabókarstofninn skyldi vera um 50–70 þúsund orð. Stofninn var listi yfir valin orð í íslensku með áherslu á nútímamálið, en hverju orði fylgdu upplýsingar um orðflokk, formgerð (grunnorð eða samsett orð) og heimild um orðið. Kristín Bjarnadóttir hafði umsjón með þessu verki (sjá Kristín Bjarnadóttir 1998).

Lengi hafði verið vilji til þess að setja saman nútímalega íslensk-sænska orðabók, bæði á Íslandi og í Svíþjóð, og árið 2002 var leitað samstarfs við *Institutionen för svenska språket* við Gautaborgarháskóla um gerð slíkrar orðabókar. Áhugi var á því að koma á sams konar samvinnuverkefni við Norðmenn um íslensk-norska orðabók, og árið 2003 fékkst fjárstyrkur frá Det kongelige kultur- og kirkedepartement í Noregi til undirbúnings þess. Norski samstarfsaðilinn í verkefninu varð *Nordisk institutt* við háskólann í Bergen. Norðmenn lögðu áherslu á að bæði norskt bókmál og nýnorska skyldu vera með í orðabók af þessu tagi en ekki bókmálið eingöngu. Frá upphafi var jafnframt ljóst að það yrði á allan hátt hentugt að bæði tungumálin, sænska og norska, yrðu unnin samhliða í sama tölvukerfinu.

Haldin var frumkynning á orðabókinni í Kaupmannahöfn 2005, og

1 Sjá nánar um forsögu ISLEX: Þórdís Úlfarsdóttir 2011, Svavar Gestsson 2011 og Anders Ljuggren 2011.

í kjölfarið lýstu fulltrúar frá *Det Danske Sprog- og Litteraturselskab* (DSL) yfir vilja sínum til þess að taka þátt í ISLEX með dönsku sem markmál. Eftir frekari viðræður varð það úr að DSL sótti um fjárveitingu til að geta hafið verkið, og fékkst hún. Þar með voru fjögur lönd orðin aðilar að verkefninu.

Þátttaka Færeyja kom til sögunnar í ársbyrjun 2011 en þá var verkið komið langt á veg í hinum löndunum.

Hvert land hefur séð um kostnaðarhliðina við sitt tungumál. Verkið hefur að mestu verið fjármagnað með opinberu fé í öllum þátttöku-löndunum, auk styrkja úr samnorrænum og öðrum sjóðum.

3 Rafræn orðabók

Þegar byrjað var á verkefninu höfðu runnið upp nýir tímar í íslenskri orðabókargerð, tölvutæknin var stöðugt í þróun og ekki síst hafði veraldarvefurinn og allt umhverfi hans tekið stórfelldum framförum á liðnum áratug. Á fáum árum höfðu einnig kröfur umhverfisins aukist um að sem mest af upplýsingum væru aðgengilegar á vefnum, gjarnan ókeypis. Það er ekki að öllu leyti einfalt í orðabókarvinnu að svara þessum kröfum vegna þess að hvort sem birtingarformið er rafrænt eða ekki er starfið í eðli sínu afar seinlegt, og hætt er við að tæknileg framsetning sem hönnuð er utan um orðabókarverk snemma á ferli þess verði óþægilega fljótt úrelt eða gamaldags. Reynt hefur verið að mæta þessum vanda í ISLEX með því að endurbæta og byggja utan á gagnagrunn og notendaviðmót orðabókarinnar eftir því sem aðstæður og fjárhagur hafa leyft.

ISLEX-orðabókin er aðeins til sem rafrænt verk en hún var frá upphafi ætluð til birtingar á vefnum. Að sumu leyti gilda önnur lögmál um rafræna orðabók en prentaða bók. Rafrænt verk býður upp á meiri hraða í upplýsingaöflun, meira myndefni, hreyfimyndir og hljóð, tengla innan verksins og út fyrir það (í aðrar vefsíður). Einnig hefur rafræn bók þann kost að stöðugt er hægt að gera endurbætur og breytingar sem skila sér strax til notandans. Rafræn bók hefur enn fremur möguleika á miklu meiri útbreiðslu en prentuð bók. En fyrir utan sjálfa orðabókina hefur starfið við ISLEX skilað dýrmætu gagnasafni með skipulegum upplýsingum um vensl norrænna tungumála sem nýta má í öðrum tilgangi, bæði hagnýtum og fræðilegum.

ISLEX-verkefnið er unnið í sérsmíðuðum gagnagrunni með viðmót fyrir vefbirtingu og er tölvukerfið rekið hjá SÁ. Sérstakt viðmót er

fyrir ritstjórnarvinnu og annað sem er ætlað til uppflettinga, þ.e. notendaáýnd orðabókarinnar. Þessi tvö viðmót eru gjörólik en bæði sækja efnið í einn og sama gagnagrunninn.

4 Verklag við ritstjórn

Fjórar til fimm ritstjórnir í jafnmörgum löndum hafa að jafnaði starfað við ISLEX-verkefnið. Hver hópur er sjálfstæð starfseining sem hefur séð um sitt tungumál í orðabókinni. Eðli málsins samkvæmt er starf íslensku ritstjórnarinnar allólíkt starfi hópanna í Skandinavíu en þeirra verkefni er fyrst og fremst að þýða íslenskt efni orðabókarinnar á hin málin. Íslenski hópurinn hefur séð um að gera viðfangsmálinu skil og búa það í hendur ritstjóra markmálanna. Vinnan við verkefnið hefur ætíð farið fram samtímis í öllum löndunum, með þeirri undantekningu að Færeyingar bættust ekki í hópinn fyrr en árið 2011 eins og áður segir.

Starfandi er samráðsnefnd allra þátttökulanda sem hefur það hlutverk að annast formleg tengsl milli samstarfsstofnananna og eru meðlimir hennar jafnframt opinberir fulltrúar verkefnisins fyrir hönd sinna stofnana.

4.1 Verklag á Íslandi

Íslenska er viðfangsmál ISLEX-orðabókarinnar og er úrvinnsla þess margslungið starf en í grófum dráttum felst hún einkum í eftirfarandi þáttum. Orðum er skipt upp í merkingarliði þar sem það á við, samin er orðskýring á íslensku við hvern lið, notkunardæmi eru búin til eftir þörfum, orðasambönd eru færð í staðlað form og samin við þau skýring. Sérmerkingum og millivísunum er komið fyrir þar sem þurfa þykir, og sums staðar er mynd bætt inn í flettugreinina. Ástæður þess að íslenskum skýringum er komið fyrir í flettugreinunum eru ýmsar, ein er sú að með því fæst skýrari sundurgreining á merkingarliðunum þar sem þeir eru fleiri en einn. Það hefur líka komið á daginn að íslenskar skýringar eru víða nauðsynlegar til að þýðandinn hafi betri tilfinningu fyrir því sem uppflöttiorðið felur í sér og geta jafnvel komið í veg fyrir misskilning, þar sem ritstjórar íslenska hlutans og markmálanna eru ekki þeir sömu.

grein	no kvk
→	BEYGING
1	(= trjágrein) langur, trékenndur útvöxtur á tré, trjágrein
2	(= ritsmíð) ritsmíð, lítil ritgerð, einkum í blaði eða tímariti, sem fjallar um tiltekið efni <i>grein um bókmenntir</i>
3	(= námsgrein) nám, kennsla og fræði um tiltekið efni, námsgrein, fag <i>hún var hæst í þremur greinum á lokaprófinu</i>
4	(= lagagrein o.þ.h.) eitt atriði í lista með samfelldum texta, t.d. í reglugerðum og lögum (lagagrein) <i>20. grein fjallar um tollamál</i>
gera grein fyrir <málinu>	
útskýra <málið>	
gera sér grein fyrir <ástandinu>	
skilja og þekkja ástandið	
taka <afsökunina> til greina	
samþykkja <afsökunina>	

Mynd 1. Uppflettiorðið grein. *Búið er að vinna orðið á íslensku en engar þýðingar eru komnar inn. Flettan er í fjórum merkingarliðum og eru íslenskar skýringar við þá alla. Notkunardæmi eru í lið 2, 3 og 4. Fyrir neðan strikið eru orðasambönd sem einnig eru skýrð á íslensku.*

Ekki er farið eftir stafrófsröð í vinnu við flettugreinarnar heldur er að langmestu leyti unnið eftir merkingarflokkum (því er nánar lýst í kafla 5.6). Stærstu orðflokkanir þrír, nafnorð, lýsingarorð og sagnir, hafa nær eingöngu verið unnir á þann hátt og auk þess stór hluti atviksorða. Þannig hafa t.d. öll orð innan merkingarflokksins DÝR eða LÖGFRÆÐI verið unnin í einu. Með því að vinna eftir merkingarflokkum fæst mun betra samræmi í framsetningu orðabókarinnar.² Aðrir orðflokkar, fornöfn, forsetningar og samtengingar (þ.e. kerfisorð eða málfræðileg orð), hafa þó að mestu verið unnir sem ein heild, enda eru þeir orðflokkar tiltölulega litlir.

Af öðrum verkþáttum má nefna umræður og samvinnu um flókin og vandmeðfarin orð og gagnrýninn yfirlestur á notkunardæmum sem aðrir ritstjórar hafa búið til. Loks hefur ætíð þurft að sinna samskiptum við samstarfsmennina á hinum Norðurlöndunum, bregðast við spurningum og ábendingum og leysa úr ýmsum málum sem upp hafa komið.

2 Í eldri útgáfum af *Íslenskri orðabók* (t.d. 2. útg. 1993) er víða afar mikið ósamræmi í skýringum á merkingarskyldum flettum þar sem langt er á milli þeirra í stafrófinu, sbr. t.d. skýringarnar á orðunum *harðsjóða* og *linsjóða*.

4.2 Verklag í Skandinavíu

Fjórar ritstjórnir markmálanna starfa við ISLEX-orðabókina í Skandinavíu. Starf þeirra er einkum fólgið í því að þýða íslenska efnið (upp-flettiorð, orðastæður, orðasambönd og notkunardæmi) yfir á sitt eigið mál, dönsku, norsku (bókmál og nýnorsku), sænsku og færeysku, og bæta við frekari upplýsingum eða skýringum þar sem þörf krefur. Í Bergen þýðir sami ritstjórinn ætíð bæði á norskt bókmál og nýnorsku. Eins og ritstjórnarnir við íslenska hluta verkefnisins hafa kollegar þeirra í Skandinavíu tekið mið af merkingarflokkunum í þýðingum sínum. Þetta fyrirkomulag tryggir betra samræmi í frágangi greinanna en einnig fær þýðandinn víðari yfirsýn yfir úrval orða innan tiltekins merkingarsviðs sem hann á kost á að nota þegar hann þarf að velja sem nákvæmast jafnheiti við íslenska orðið.

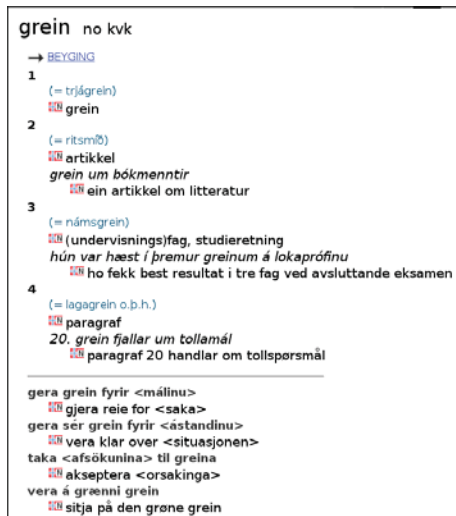
Pótt ekki sé unnið eftir stafrófsröð geta þýðendur hvers tungumáls ávallt séð á hvaða stigi tiltekið upp-flettiorð er í vinnslunni þar sem merkt er í sérstakan reit (sem ber vinnuheitið „stig“) þegar búið er að vinna flettugreinina. Ef orð er t.d. fullþýtt á dönsku er sett merki um það í reitinn og þurfa þá ritstjórar danska hlutans ekki að koma aftur að því orði í bráð.

Ritstjórar markmálanna hafa tekið virkan þátt í uppbyggingu ISLEX-orðabókarinnar og veitt ritstjórninni á Íslandi mikilvægt aðhald og er þar því um að ræða ákveðið gæðaeftirlit. Þýðendur hafa m.a. oft komið með tillögur að nýjum upp-flettiorðum eða merkingarliðum, þeir hafa óskað eftir að sett yrði inn notkunardæmi við ákveðin orð og lagt til breytingar á dæmum sem þeir hafa talið ófullnægjandi. Einnig hafa þeir komið auga á efnislegar og kerfislegar misfellur, bæði hverjir hjá öðrum og í íslenska hluta verksins.

4.3 Samskipti

Þar sem unnið er við ISLEX-orðabókina í mörgum löndum samtímis er algerlega nauðsynlegt að samskipti milli manna geti farið fram á skjótan og þægilegan hátt. Nú eru til ótal samskiptaleiðir um rafræna miðla svo að það hefur ekki háð verkefninu þótt langt sé á milli þátttakenda landfræðilega séð. Dagleg samskipti hafa að miklu leyti farið fram um tölvupóst en einnig er notast við sérstakt skilaboðakerfi sem er hluti af vinnuumhverfi gagnagrunnsins. Í tengslum við ISLEX-verkefnið var auk þess búið til sérstök lokuð vefsíða (forum-síða) þar sem birtar eru upplýsingar um merkingarflokka sem eru tilbúnir frá

íslensku ritstjórninni og sem þýðendur geta þá byrjað að þýða, og einnig eru þar ýmsar töflur, fundargerðir o.þ.h. og annað sem tengist verkefninu.



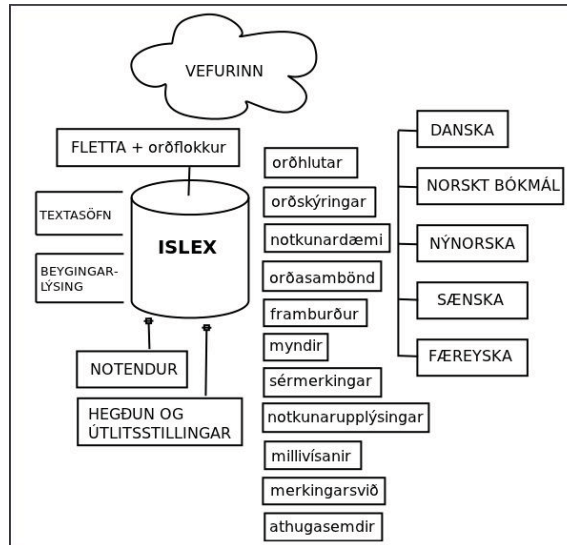
Mynd 2. Uppflettið grein ásamt þýðingum á nýnorsku. Þetta er það sem notandinn sér þegar nýnorsk orðabók er valin.

Ásamt þessum rafrænu samskiptaleiðum hafa að jafnaði verið haldnir tveir þriggja daga vinnufundir ár hvert þar sem allir þátttakendur hafa hist í eigin persónu. Þessir fundir hafa gagnast til upplýsingagjafar, umræðna og áætlana og eru menn sammála um að samskipti um tölvur geta ekki að öllu leyti komið í stað persónulegra samtala. Fundirnir hafa því verið mjög mikilvægur þáttur í samstarfinu. Þýðendur verkefnisins í Skandinavíu hafa einnig hist reglulega til að bera saman bækur sínar.

Eftir að orðabókin var opnuð á vefnum 2011 hafa samskipti við notendur verið hluti af ritstjórnarstarfinu. Á vefsíðu ISLEX er sérstakt kerfi fyrir athugasemdir notenda (undir tenglinum 'Hafa samband') sem hægt er að koma á framfæri við ritstjórnina og hafa margir nýtt sér það. Sagt hefur verið frá villum og öðrum ágöllum, bent hefur verið á orð sem vantar o.s.frv. Allar slíkar ábendingar eru skoðaðar vandlega.

5 Efniviður ISLEX

ISLEX er meðalstór orðabók að því er varðar fjölda uppflettiorða, með tæplega 50.000 orð. Til samanburðar er fjöldi uppflettiorða í *Íslenskri orðabók* (4. útg. 2007) um 94.000, en sú bók telst vera fremur stór. Í gagnagrunni ISLEX eru fólgin margvísleg gögn sem öll eru nákvæmlega skilgreind. Mynd 3 er e.k. flæðirit sem sýnir helstu þætti sem snerta verkefnið.



Mynd 3. Skipan ISLEX-orðabókarinnar. Gagnagrunnurinn er táknður með stóra sívalningnum. Fyrir ofan hann er reitur með flettu og orðflokki, sem gegna veigamiklu hlutverki í gagnagrunninum. Til vinstri eru mikilvæg gagnasöfn, textasöfnin og beygingar orða. Hægra megin eru reitir með helstu gagnaeyingum í ISLEX, og lengst til hægri eru markmálin, danska, norska, sænska og færeyska. Neðan við ISLEX eru töflur sem geyma upplýsingar um notendur, og stillingar á hegðun og útliti gagnategunda.

Á þeim tíma sem liðinn er frá því að vinna hófst við ISLEX hafa safnast saman miklar upplýsingar í gagnagrunninn, bæði um íslenska orðaförðann og markmálin fimm. Afar mikill efniviður er fólgin inni í sjálfum flettugreinunum, þ.e. því efni sem birt er undir hverri einstakri flettu. Allt efnið byggist á kerfislega flokkuðum upplýsingum ("tegundum"), jafnt íslenski hlutinn og markmálin, og eru t.d. notkunardæmi ein slík tegund og millivísun önnur tegund. Einingarnar hafa sérstakt heiti eftir því hlutverki sem þær gegna í orðabókinni. Alls eru skilgreindar um 80 tegundir gagna, þar af er um helmingurinn

ætlaður fyrir upplýsingar um íslensku auk kerfislegra atriða, en hinn helmingurinn er fyrir ýmis gögn er varða markmálin fimm. Þar eru sjálfar þýðingarnar vitaskuld fyrirferðarmestar en einnig eru þar notkunarupplýsingar af ýmsu tagi og beygingarendingar jafnheita, allt sundurgreint eftir tungumálum.

5.1 Textasöfn og önnur hjálpargögn

Rafræn textasöfn hafa á undanförunum áratugum leyst af hólmi hefðbundna orðtöku sem öll viðamikil orðabókarvinna hefur óhjákvæmlega útheimt í gegnum tíðina, þar sem tilgangurinn hefur verið að leita uppi sem mest af heimildum um tiltekið orð, s.s. merkingu þess, notkun, aldur, útbreiðslu, beygingu o.s.frv. (Um kosti og galla textasafna við orðabókargerð má t.d. benda á Svensén 2004:56-59).

Vitnisburður textasafna hefur fjölþætt gildi í orðabókarvinnu:

- Textarnir gefa vísbendingu um hversu algengt eða sjaldgæft orðið er í málinu.
- Upplýsingar fást um aldursdreifingu orðsins og notkunarvið þess.
- Hægt er að ákvarða merkingu orðs (stundum fleiri en eina) út frá samhengi þess í setningum.
- Ákveðin mynstur orða geta komið í ljós í röðuðum niðurstöðum textasafnsleitar, þ.e. mynstur um fasta orðanotkun sem gera þarf skil í orðabók en auðvelt væri að yfirsjást með öðrum hætti.
- Fjölbreyttir textar eru mikilvæg hjálp þegar búa þarf til notkunar-dæmi um orðin.

ISLEX telst til þeirra orðabóka sem eru að miklu leyti grundvallaðar á textasöfnum, og búum við svo vel að hafa aðgang að fjölbreytilegum, rafrænum íslenskum textum. Einkum hafa tvö stór textasöfn verið mikilvæg hjálpartæki í starfinu.

Um svipað leyti og starfið við ISLEX hófst var *Íslensku textasafni* komið fyrir á vef stofnunarinnar. Textasafnið hefur aukist jafnt og þétt með árunum og er stærð þess um mitt ár 2012 um 67 milljónir lesmálsorða. Það hefur að geyma efnisflokkaða og að nokkru leyti aldursflokkaða texta þar sem leitað er að stökum orðum, í einum efnisflokki eða fleirum í senn. Efnisflokkar textasafnsins eru t.d. SKÁLDRIT, FORNRIT, SAGA og HEIMSPEKI, LÍFFRÆÐI, RAUNGREINAR og MATREIÐSLA. Mynd 4 sýnir búið af niðurstöðu leitar í *Íslensku textasafni*.

og breitt bros á tala við tvo því að hún var svo upprifin og yssu í beltissstað. Norsararnir Guðrún Hauksdóttir, falleg, áttu það sameiginlegt að vera þegar í veitingabúð Gerðubergs. "Þaðbúð Gerðubergs. "Kátir dagar, þar á eftirlaunum. Karlakörinn ganga að Kattholti. Karlakörinn setta: Annar þeirra snarði sér öll hjónaefni jöfn; feru fyrst við verum að spjalla saman í undirfyrirsögn: ~ UNN SÖHUNDURÖÐ einn daginn er hann syngjandi, var frændi hans Þórbörn kakali degi. Þegar allir voru nú bara hafði svosem ástaðu til að vera er og segir að verið hafi honum líkaði. Öttar kveðst verið hafa til lögmansins. Gísur var hinn í og veifaði hattinum. -Hann er sem honum voru settir þeim mun ni eftir auðfari sín. Þeir verbað í eyra, sagði Daði og horfði var gaman að heilsa þeim. allir erf. Maður verður bái betri og þessu tíðindi. Og varð feiti í sem ég lék í Stubbnum, feiti, egri en Lüðvík. Lüðvík er ekki svoglumalgi á leiðinni en allir ín Óskarsdóttir voru hressar og sunna og starfsfólk. "segir Bára	kata sjóliða. Tvar vörpulegar löggur gengu fyrir gluggann með hendur fyrir aftan kat. Reyndi að láta alls ekki á því bera að mér var verulega illa brugðið. Hann katu gáfust upp á að biða eftir afgreiðslu og fóru út. Eg horfði á eftir þeim þa kat og hress Kópavogsmér sem innan tíðar gengur í kristinna manna tölu. Draumuri kátir og yndislegir og greinilega vanir mannanna meðhöndlun. Eg hef einnig heyrt kátir dagar, kát fólk": Föstudaginn 12. okt. kl. 19 haustfagnaður á Hótel Sögu, miðar til só kát fólk": Föstudaginn 12. okt. kl. 19 haustfagnaður á Hótel Sögu, miðar til só kátir karlar. Hefur áfingar þriðjudaginn 17. sept. kl. 13 í Félags- og þjónustumi kátir karlar hefur áfingar þriðjudaginn 17. sept. kl. 13 í Félags- og þjónustumi kátur inn á sýslumannsskrifstofu og sagði við afgreiðslukonuna: Jája, heldurðu á kát á sýslumannsskrifstofu og svo - vildu þau helga sambandið - í kirkju. Að min kátum kunningjahöpi. Mér dettur í hug saga sem mér finnst passa vel inn í umráðu kátir dagarnir, kát fólk": Föstudaginn 12. okt. kl. 19 haustfagnaður á Hótel Sögu, miðar til só kátir dagar, kát fólk": Föstudaginn 12. okt. kl. 19 haustfagnaður á Hótel Sögu, miðar til só kátir karlar. Hefur áfingar þriðjudaginn 17. sept. kl. 13 í Félags- og þjónustumi kátir og glaðlegir vindar blésu. Hún sagði við Eyjólf: -Það er vegna þess að ha kátir því að hann er sannanlega gleðiefni giftingardagur dótturinnar, þannig er kátara. Síghvatur það hann þá kveða kveðið það er hann hafði ort um Ástriði. Ött kátari. Síghvatur það hann kveða sér kveðið það er hann hafði ort um Ástriði drot kátasti. -Það þýðir lítið fyrir Erlend að þratta við mig um jarðir, sagði hann. kátur, karlinn, sagði biskupsveinninn. -Hann hefur lítið til að kátast yfir, þa kátari varð hann. Átti hann í sálarskaska flíkaði hann inni manni með því að snú kátir sem ná undir sig jörðum og skreið. Aularnir munu sitja volandi yfir aulask kátur á Ára. -Það er ekki svo lítið sem hún og börnin hafa þjáðst vegna skáldska kátir, brún mín léttest og lundin lifnaði. Margir komu í bílinn á Afلاغrandanum, kátari að vera með þessu góða folki. Þökk sé skórnarféndi Neskirku fyrir að bjóð kátí strákurinn, feiti trúðurinn, og það var einmitt hann sem ég lék í Stubbnum, kátí strákinn, karkvissa þrákarann, trúðinn. Þannig sló ég í gegn. Sem ég, einn kátur, hann er ekki skemtilegur, en hann er ekki vitlaus. Lüðvík er langafabás kátir og hressir. Þetta var fjölskylduból, alveg ágæt. Frændsystkini mín eru al kátar, en Guðrún kennir ballett fyrir skólann og Lovisa jazz og leikfimi. Margir kát sem kennt hefur jazzballett og líkamsrækt í tugi ára. -Við erum búin að stæk	(mann) (mann) (mogg-bladauki) (mogg-brjéftilbi) (mogg-dagbok) (mogg-dagbok) (mogg-dagbok) (mogg-dagbok) (mogg-lesbok4) (mogg-lesbok4) (mogg-lesbok4) (mogg-sudurnes) (ofsi) (ofsi) (ofsi) (ofsi) (ottav_fl) (ottav_to) (oxinogjordin) (oxinogjordin) (oxinogjordin) (oxinogjordin) (oxinogjordin) (prestabl) (prestabl) (riklarhugans) (riklarhugans) (riklarhugans) (riklarhugans) (sedogheyr2006) (sedogheyr2006)
---	--	---

Mynd 4. *Hluti af niðurstöðu leitar að orðinu kátur í Íslensku textasafni. Orðið finnst í öllum beygingarmyndum sínum. Þarna er niðurstöðum raðað eftir textaskrá. Aftast eru tenglar í upplýsingar um viðkomandi ritverk.*

Nokkru áður hafði verið opnaður aðgangur að vefnum *Tímarit.is* sem var samstarfsverkefni Landsbókasafns Íslands - Háskólabókasafns og bókasafna í Færeyjum og á Grænlandi. Tímaritavefurinn er stafrænt safn sem veitir aðgang að milljónum myndaðra eða skannaðra blaðsíðna úr blöðum og tímaritum gefnum út í þessum löndum, en þar er íslenska efnið langfyrirferðarmest. Blöðin hafa að geyma almennt fréttafefni, auglýsingar, efni á sviði bókmennta, sagnfræði, ættfræði, þjóðlífs, menningar, atvinnuvega og viðskipta (sbr. <http://timarit.is> – *Um vefinn*). Með árunum hefur leitaraðgangurinn orðið greiðari og blaðsíðum fjölgað, og hefur mikið verið stuðst við þetta einstæða gagnasafn við uppbyggingu ISLEX-orðabókarinnar. Mynd 5 sýnir nút af leitarniðurstöðum á *Tímarit.is*.

Þessi tvö textasöfn eru ómetanlegar ritheimildir um íslensku og hafa þau reynst ómissandi við orðabókarstarfið. Nú er að vísu hægt að leita að orðum á öllum veraldarvefnum og fá fjölmargar niðurstöður úr þeirri leit, en oft er samhengið sem þannig fæst mjög ófullnægjandi heimild um orðið og mikið er um síendurtekin tilvik. Að auki koma mörg orð afar sjaldan fyrir í því efni sem birtist almennt á vefnum og þarf að leita í verk rithöfunda eða fræðimanna til að finna þau. Mörg slík verk er einmitt að finna í *Íslensku textasafni*.

En þrátt fyrir mikla kosti textasafna eins og þeirra sem hér hefur verið lýst hafa þau sín takmörk og ekki er hægt að sækja í þau allan fróðleik. Ef afla þarf ákveðinnar vitneskju um sértækan orðaforða er venjulega betra að leita hennar annars staðar. Með þessu er átt



Mynd 5. Hluti af niðurstöðu leitar að orðinu kátur á vefnum Tímarit.is. Hægt er að smella á greinarnar og sjá stærra samhengi orðanna.

við t.d. latnesk tegundarheiti fugla og skilgreiningar á hugtökum eðlisfræðinnar. Við vinnslu orðabókarinnar hefur því stundum verið stuðst við handbækur og orðabækur um ýmis efni, hinar og þessar vefsíður, og auk þess þekkingu fróðra manna á SÁ og annars staðar. Loks má nefna *Ritmálssafn Orðabókar Háskólans* sem er heljarmikið gagnasafn um íslenskt mál síðustu fimm alda. Það hefur komið að góðum notum í orðabókarstarfi hér á landi síðustu áratugina og hefur líka verið notað við vinnuna við ISLEX.

5.2 Uppflettiorðin

Eins og fyrr greinir er ISLEX meðalstór orðabók. Hlutfall orðflokka innan orðaforðans er á þá leið að nafnorð eru í yfirgnæfandi meirihluta, um 68% orðaforðans. Lýsingarorð eru um 17% orðaforðans, sagnir um 9%, atviksorð um 4% og orð af öðrum orðflokkum eru miklu færri. Orðaforðinn skiptist eftir orðflokkum samkvæmt töflu 1:

Orðflokkur	Fjöldi orða	Hlutfall orðaforðans
Nafnorð	33.136	68%
Lýsingarorð	8362	17%
Sagnir	4234	9%
Atviksorð	1815	4%
Töluorð	71	< 1%
Fornöfn	126	< 1%
Forsetningar	235	< 1%
Samtengingar	93	< 1%
Upphrópanir	62	< 1%
Forliðir	136	< 1%
Greinir	1	< 1%
Alls	48.432	100%

Tafla 1. Skipting orða í ISLEX eftir orðflokkum.

5.2.1 Orðaforðinn: val og viðbætur

Eins og greint var frá í kafla 2 var settur saman sérstakur orðabókarstofn hjá Orðabók Háskólans á síðasta áratug 20. aldar (nefndur *norræni orðabókarstofninn*) og varð hann grundvöllur orðaforðans í ISLEX. Í upphaflega orðabókarstofninum er kjarni orðaforðans: algengustu nafnorð, lýsingarorð og sagnir og öll kerfisorð. Þar er áhersla lögð á nútímaíslensku en líka er dálítill orðaforði úr fornmáli, eldra máli og skáldamáli. Einnig er nokkuð um fagorð af ýmsum sviðum s.s. líffræði, lögfræði og viðskiptum.

Á þeim árum sem orðabókin hefur verið í smíðum hefur ýmsu verið bætt við upprunalega flettiorðalistann og annað hefur verið fellt brott. Frá því að verkið hófst standa eftir um 41.600 orð úr upphaflega norræna orðabókarstofninum (um 86% orðaforða ISLEX) en um 7.000 orðum hefur verið bætt við (sem svarar til um 14% af orðaforðanum).

Nýrri orðin endurspegla alloft aðstæður og umræðu líðandi stundar en ekki er það þó algilt. Dæmi um orð sem bæst hafa við eru nafnorðin *auglýsingaherferð*, *áfallahjálp*, *átröskun*, *bílsprengja*, *bökunarplata*, *dagmóðir*, *eftirlitsmyndavél*, *flisteppi* og *hnattvæðing*; lýsingarorðin *alsaklaus*, *áralangur*, *bensínknúinn*, *eldspúandi*, *enskuskotinn*, *glæraugnalaus*, *grænblár*, *heimaprjónaður* og *hjartaveill*; og sagnirnar *aftengja*, *djúpsteikja*, *endurlifa*, *fúaverja* og *strádrepa*. Flettulistinn í ISLEX er þannig talsvert sveigjanlegur og opinn fyrir viðbótum.

Um viðbótarorð hafa einkum gilt þessi sjónarmið:

- Orðið er algengt í málinu og hafði ekki ratað inn í upphaflega orðabókarstofninn.
- Orðið er í umræðu líðandi stundar og líklegt virðist að það eigi sér framtíð í tungumálinu (*efnahagskreppa, samkeppnisstaða*). Þetta getur þó stundum verið erfitt að meta.
- Orðið heyrir eða kemur fyrir víða í umhverfinu þótt það sjáist ekki endilega mikið á prenti (*pressukanna, mátunarklefi, slökun-artónlist*).
- Ýmis svið orðaforðans hafa verið augin með orðtöku úr tiltækum heimildum, t.d. tegundir fugla og plantna, landaheiti og fatnaður. Einnig hafa nokkur fagorðasvið verið augin, t.d. orð sem tilheyra jarðfræði (*kvikuþró* og *megineldstöð* hafa t.d. bæst við þann flokk).

Annars er ekki að öllu leyti auðvelt að skilgreina reglur um val orða í orðabók. Það hefur komið til tals að útbúa tíðniraðaða orðalista úr rafrænum textum og ákveða að “toppurinn” skuli verða uppflattiorð og hefur það raunar verið prófað lauslega í tengslum við ISLEX-verkefnið. Það er þó ekki einfalt verk þegar á hólminn er komið og margar hindranir í veginum. Þættir sem valda örðugleikum eru t.d. röng lemmun (eða vöntun á lemmun), mjög mikið af óviðkomandi orðum s.s. nöfnum, örnefnum, erlendum orðum, tölustöfum og táknum, og eðlislægar takmarkanir textasafna hvað snertir orðaforða málsins. Þegar upp er staðið er það því að miklu leyti máltilfinning ritsjáranna sem ræður vali orðanna, og er stuðnings gjarnan leitað í tiltækum heimildum.

5.2.2 Fleiryrtar flettur

Sum uppflattiorðin í ISLEX-orðabókinni eru fleiryrtar flettur, þ.e. tvö orð eða fleiri. Þar eru atviksliðaflettur áberandi (*að fullu og öllu, alla jafna, af og til, sitt á hvað, smátt og smátt, sums staðar*) en einnig forsetningar (*á milli, fram eftir, í kringum, niður í, upp úr, út undan*), samtengingar (*af því að, né heldur, til þess að*) og fornöfn (*annar hver, sinn hvor*). Að auki er að finna meðal fleiryrttra flettna nokkra nafnliði (*grafísk hönnun, háls-, nef- og eyrnalæknir, heitur pottur*) og lýsingarorðsliði (*annars hugar, frá sér numinn, önnum kafinn*).

Fleiryrtar flettur hafa fengið orðflokksmerkingu í samræmi við setningarstöðu þeirra og hlutverk. Atviksliðafletturarnar eru sambönd tveggja eða fleiri orða sem mynda merkingarlega heild og hafa sömu

stöðu og atviksorð í setningum. Þau fá orðflokksmerkinguna *ao*. Sama gildir um fleiryrtta nafnliði (nefndir *no*) og lýsingarorðsliði (*lo*), en fleiryrtar samtengingar eiga sér lengri hefð í málinu og jafnvel samsettar forsetningar.

Í ISLEX eru alls um 860 uppflöttiorð fleiryrtar einingar, og er með þessu farið nokkuð út fyrir hefðina í orðabókum. Hugmyndina að því að taka með svo margar fleiryrtar flettur í ISLEX má að miklu leyti rekja til verkefnisins *Íslenskt orðanet* (www.ordanet.is), sem unnið hefur verið að hjá SÁ frá árinu 2004 undir stjórn Jóns Hilmars Jónssonar. Flest fleiryrttra atviksorða í ISLEX eiga uppruna sinn í því verkefni (sjá Jón Hilmar Jónsson 2009:165–170 og 2012).

Nokkur rök eru með því að hafa fleiryrt uppflöttiorð í orðabókinni:

- Fleiryrt sambönd geta oft staðið jafnfætis stökum orðum sem sjálfstæðar, merkingarbærar einingar. Í tvímála samhengi vill svo til að margar fleiryrtar einingar eiga sér einyrt jafnheiti á öðrum málum. Dæmi um það er atviksorðið *allt í einu* sem á sér danska jafnheitið *pludselig* og nýnorska jafnheitið *plutselig*; að innanverðu á sænska jafnheitið *invändigt* og *aftan frá* samsvarar sænsku *bakifrån*. (Þetta á raunar ekki aðeins við milli tungumála heldur á t.d. þríyrta íslenska sambandið *allt í einu* sér einyrta samheitið *skyndilega*.)
- Mörg slík sambönd eru ákaflega algeng í málinu. Aðgangur að þeim í hefðbundnum orðabókum er oft óskýr og ógreiður og þau vilja oft tynast inni í greinunum. Það er þægilegt fyrir notandann að geta gengið beint að þeim.
- Sum orð eru ýmist rituð í einu orði eða tveimur: *ýmiss konar* eða *ýmiskonar*, *sums staðar* eða *sumstaðar*, *báðum megin* eða *báðumegin*. Spyrja má hvort eðlilegt sé að stafsetning ráði því hvernig orði er komið fyrir í orðabók og með hvaða orðflokksmerkimiða. Í ISLEX eru tekin með bæði ritform slíkra uppflöttiorða og vísað á milli, en orðflokkurinn er hafður sá sami hvort sem rituð eru eitt orð eða tvö. Þannig eru bæði *sums staðar* og *sumstaðar* greind sem atviksorð.

Við þetta má bæta að orðasambönd í ISLEX-orðabókinni þar sem kjarninn er nafnorð (*fara á fætur*, *hafa hjartað á réttum stað*, *taka til hendinni*) eða lýsingarorð (*fá sér einn gráan*, *vera harður af sér*) eru að jafnaði tilfærð undir kjarnaorðunum en ekki sem sjálfstæðar flettur. Að þessu leyti er farin leið hefðbundinna orðabóka.

5.3 Orðasambönd og orðastæður

Orðasambönd eru áberandi hluti af ISLEX-orðabókinni og er lögð meiri áhersla á að lýsa merkingu þeirra og notkun en hingað til hefur tíðkast í íslensk-erlendum orðabókum. Ástæður þess eru einkum þær að orðasambönd eru mjög stór hluti af tungumálinu og að mörg föst orðasambönd geta verið ógegnisæ og erfið viðureignar fyrir erlenda notendur. Í ISLEX eru öll orðasamböndin þýdd á markmálin og hefur það hagnýtt gildi í báðar áttir í tveggja mála (eða margmála) orðabók þar sem íslenskir notendur geta notfært sér þennan þátt ekki síður en hinir erlendu.

Hugtakið *orðasamband* hefur fremur víða skilgreiningu (sjá t.d. Jón Hilmar Jónsson 2005:xix–xx) en í einfaldasta skilningi er það fast samband tveggja eða fleiri orða. Í ISLEX er gerður greinarmunur á tvenns konar orðasamböndum. *Orðasamband* er notað um orðtak eða myndhverft samband þar sem kjarni þess hefur oft lítið eða ekkert að gera með merkingu sambandsins í heild: *<biða> með öndina í hálsinum, <þeir> fara í hár saman, vera á höttunum eftir <upplýsingum>, hafa <lítið> upp úr krafsinu*.

Orðastæða samsvarar því sem nefnist á ensku *collocation* og *kollokation* á dönsku og sænsku. Í slíkum samböndum kemur orðið fyrir sem kjarnaliður með föstum fylgdarorðum, einu eða fleirum (*franskar kartöflur, komandi kynslóðir, kaffi og með því, sólin er hátt/lágt á lofti*). Stundum er sagnorð kjarnaliður (*brjóta af sér, ráða <fólk> til starfa, vilja <honum> vel*).

Fjöldi myndhverfra orðasambanda í ISLEX er um 4400 og orðastæðna um 9600. Mikill meirihluti beggja gerða þessara sambanda er fenginn úr gagnagrunni *Stóru orðabókarinnar um íslenska málnotkun* (hér eftir skammstafað SOB) (Jón Hilmar Jónsson 2005) með leyfi höfundar, sem raunar annaðist sjálfur val orðasambanda fyrir nafnorðaflettunar í ISLEX. Sum orðasamböndin hafa þó komið inn í verkið eftir öðrum leiðum.

Framsetning orðasambanda er með svipuðum hætti og í SOB (sjá ítarlega greinargerð um það í inngangi bókarinnar; *ibid.*:xix–xxii). Eitt megineinkenni orðasambanda í SOB er að breytilegir liðir eru afmarkaðir með oddklofum: *krús af <öli>, rissa upp <mynd af fugli>, vera um kyrrt <í sveitinni>*. Þetta fyrirkomulag hefur verið tekið upp í ISLEX-orðabókinni með dálitlum breytingum sem til eru komnar vegna ólíks hlutverks þessara tveggja orðabóka. Þar sem öll orðasambönd í ISLEX eru þýdd á markmálin, líka þau orð sem eru umlukin oddklofum, verður að gera þau þannig úr garði að hægt sé með góðu móti að þýða

þau á önnur mál. Breytilegir liðir hafa því víða verið einfaldaðir og eru mörg orðasambönd í ISLEX talsvert styttri og einfaldari en í SOB.

Í ISLEX hafa eftirfarandi nafnorð flest orðasambönd (fjöldi þeirra er á eftir nafnorðinu og í sviganum er dæmi um orðasamband. Orðastæður eru ekki taldar með):

- (1) hönd 75 (*hafa hendur í hári* <þjófanna>)
- (2) fótur 55 (*setja á fót* <fyrirtæki>)
- (3) mál 53 (*vekja máls á* <þessu>)
- (4) orð 43 (*fá orð í eyra*)
- (5) auga 31 (*koma auga á* <snjalla lausn>)
- (6) maður 30 (*vera maður að meiri*)
- (7) vegur 26 (<ferðast> *á eigin vegum*)
- (8) dagur 25 (<þannig var lífið> *forðum daga*)
- (9) líf 24 (*komast lífs af*)
- (10) ráð 24 (*vera ekki með réttu ráði*)

5.4 Notkunardæmi

Mörg notkunardæmi eru í ISLEX-orðabókinni. Þeim er ætlað að sýna uppflöttiorðið í eðlilegu samhengi og einnig að skerpa greinarmun á mismunandi merkingum orðsins, sé um fleiri en einn merkingarlið að ræða. Nafnorðið *meðferð* skiptist t.d. í fjóra liði. Einn liðurinn merkir 'meðhöndlun' og þar er dæmið *leikföngin hafa fengið slæma meðferð*. Annar liður merkir 'læknismeðferð' og þar er dæmið *hún fékk meðferð við sjúkdómnum*.

Þarna er ekki um að ræða raunveruleg ritdæmi með heimildarvísun eins og tíðkast í sumum orðabókum heldur eru þau samin af ritstjórum verksins. Dæmin eru viðkvæm og vandmeðfarin. Oft er auðveldara að láta lengri dæmi hljóma eðlilegar og betur en mjög stutt dæmi, en í orðabók er samt æskilegt að hafa þau eins stutt og hnitmiðuð og kostur er. Þarna þarf að finna milliveg.

Textasöfn (sbr. kafla 4.1) hafa komið að góðum notum við að búa til dæmin enda eru textar þeirra óþrjótandi uppspretta samhengis orða í tungumálinu. Ekki hefur tíðkast að taka heilar setningar úr textum og gera þær óbreyttar að notkunardæmum heldur eru fengnar þaðan hugmyndir að dæmum þegar innblásturs er þörf. Dæmin eru á eðlilegri nútímaíslensku og eru þau þýdd á öll markmálin eins og orðasamböndin.

Dæmi koma fyrir í öllum orðflokkum. Þótt nafnorð séu langtum fleiri en sagnir eru samanlagt fleiri notkunardæmi í sagnagreinunum heldur en í nafnorðagreinunum (í sögnum eru 10.000 dæmi en 9000 í nafnorðunum). Það er vegna þess að sagnir eru þess eðlis að lýsing þeirra í orðabókinni hvílir mun meira á dæmum en lýsing flestra nafnorðanna. Hér er sýnishorn af dæmum í sagnaflettum þar sem merkingarviðið er HREYFING/GÖNGULAG. Uppflettiorðið er í sviga fyrir aftan:

áin hlykkjast eftir dalnum (*hlykkjast*)
 þeir bröltu yfir grjótið í fjörunni (*brölta*)
 blaðamaðurinn skálmaði beint að ráðherranum (*skálma*)
 hún ranglaði ein um gamla bæjarhlutann (*rangla*)
 ég reikaði niður að höfninni (*reika*)
 ég skreiddist fram úr rúminu kl. 12 (*skreiddast*)
 gamla konan staulaðist niður stigann (*staulast*)
 hann stikaði löngum skrefum út ganginn (*stika*)

Í dæmunum er leitast við að sýna orðið í örstuttu textasamhengi sem er lýsandi fyrir notkun þess. Sum dæmin eru aðeins lýsingarorð með nafnorði: *adkeypt þjónusta*, *auðfenginn gróði*; undir uppflettiorðunum *adkeyptur* og *auðfenginn*. Í slíkum tilfellum getur munurinn á dæmi og orðastæðu stundum verið óljós.

5.5 Beygingar

ISLEX-orðabókin gefur ekki upp beygingarendingar eða beygingarflokk við uppflettiorðin eins og gjarnan er gert í prentuðum orðabókum, heldur er við hverja flettu tengill í vef verkefnisins *Beygingarlýsing íslensks nútíðsmáls* (á <http://bin.arnastofnun.is>). Þá opnast beygingardæmi orðsins í heild sinni. Í ISLEX eru tenglar í beygingarlýsinguna við nær öll nafnorð, lýsingarorð og sagnir.

Í orðabókinni er einnig að finna upplýsingar um beygingarendingar jafnheita markmálanna, og er því lýst í kafla 5.8.3.

5.6 Merkingarflokkar

Öllum orðaforðanum í ISLEX er skipt upp í merkingarflokka. Í köflum 3.1 og 3.2 kom fram að bæði íslensku flettugreinarnar og þýðingarnar hafa verið unnar eftir þessum merkingarflokkum. Þeir eru tæki fyrir

ritstjórnarvinnuna og eru ekki sýnilegir notendum. Flokkarnir bjóða þó upp á ýmsa möguleika til frekari birtingar á orðabókarefninu, t.d. væri með áframhaldandi þróun hægt að nota þá til að útbúa samheita- og skyldheitabálka innan orðaforðans.

Merkingarflokkarnir eru hópar merkingarskyldra orða sem eru allbreytilegir að stærð. Þeir geta verið fáein orð upp í að vera mörg hundruð, en dæmigerður merkingarflokkur er nokkrir tugir orða. Dæmi um merkingarflokka eru DÝR, FATNAÐUR, HÚSBÚNAÐUR, BÍLAR (mest hlutbundin nafnorð); HLJÓÐ, ÓVEÐUR, SAMKOMA, ÞJÓÐARHEITI, STAÐARHEITI, STARFSHEITI, ÆTTARTENGL (óhlutbundnari orð); EFNAFRÆÐI, JARÐFRÆÐI, LÍFRÆÐI, LÆKNISFRÆÐI, MÁLFRÆÐI, STÆRÐFRÆÐI (sérsvið innan orðaforðans). Áberandi merkingarflokkar meðal lýsingarorða eru ALDUR, ÁNÆGJA, ÁSTAND, DUGLEYSI, FJÁRHAGUR, HEILSA, HOLDAFAR, LITUR, LÖGUN og ÚTLIT. Merkingarflokkar sagna eru t.d. BREYTING, DAUÐI, DVÖL, HREYFING, SNERTING, SVEFN, ÚRKOMA og VINDUR. Merkingarsviðin skarast mjög víða, t.d. tilheyrir orðið *gul-rót* bæði flokknum GRAS og MATUR, og *kennari* er bæði í flokkunum STARFSHEITI og SKÓLI.

Til að gefa skýrari mynd af þessu má taka tvö dæmi um orðahópa sem tilheyra sama merkingarflokknum. Í flokknum STJÓRNARFAR eru einkum nafnorð, m.a.

alræði, bananalýðveldi, einræði, einræðisríki, einveldi, flokksveldi, flokksræði, fulltrúalýðræði, harðstjórn, hægristjórn, konungdæmi, lénsskipulag, lögregluríki, ógnarstjórn, skrifræði, vinstristjórn, þingkjörinn, þjóðveldi

Í flokknum GANGA/GÖNGULAG eru sagnir mest áberandi, m.a.

arka, feta, ganga, haltra, hökta, kjaga, keifa, labb, labba, léttstígur, læðast, lötra, marsera, paufast, plampa, rangla, ráf, ráfa, ráp, rápa, reika, röлта, silalegur, skálma, skjögra, slaga, spígspora, staulast, stjákl, stjákla, stórstígur, storma, strunsa, tipla, trítla, tölta, þungstígur

Með því að einbeita sér að einum merkingarflokki í einu fær ritstjórinn yfirsýn yfir heilt svið orðaforðans, og hefur það reynst heppilegri aðferð en að vinna fletturnar eftir stafrófsröð.

5.7 Myndir og hljóð

Fjöldinn allur af myndum eru í ISLEX. Þeim er ætlað að vera til prýði

og gleði en þær hafa einnig víðtækara hlutverk. Mynd getur skýrt merkingu sumra orða betur en hægt er að gera með orðum (*blævængur*, *gleraugu*, *íhorni*, *skæri*, *tannhjól*, *tvinnakefli*, *þvottaklemma*), og stundum er hentugt að hafa myndir til að greina í sundur mismunandi merkingar orða, séu þær fleiri en ein. Myndirnar gera auk þess orðabókina meira aðlaðandi í augum notenda. Í ISLEX eru um 3200 ljósmyndir og teikningar úr ýmsum áttum. Þar af eru um 700 myndir eftir Jón Baldur Hlíðberg af fyrirbærum úr ríki náttúrunnar: plöntum, fuglum, fiskum og fleiri dýrum.



Mynd 6. Myndskreytingar með dýrum og plöntum eftir Jón Baldur Hlíðberg.

Framburður uppflöttiorða í ISLEX er gefinn með hljóðdæmum en ekki með hljóðritunartáknum. Hægt er að smella á hljóðtengil til þess að hlusta á orðið með eðlilegum íslenskum framburði.³ Sum orð eru þannig að málhljóð þeirra eru breytileg eftir beygingarmynd (t.d. *dagur* – *degi*, *saga* – *sagna*, *trefill* – *treflar*), og í slíkum tilvikum er ætlunin að hafa ekki aðeins hljóðdæmi með grunnformi orðsins heldur valdar orðmyndir þar sem tilbrigði í framburði koma fram.

5.8 Markmálin

Markmálin í ISLEX-orðabókinni eru danska, norskt bókmál, nýnorska, sænska og færeyska eins og áður segir. Hér verða markmálunum ekki gerð tæmandi skil en aðeins nefnd nokkur atriði (sjá nánar Aldís Sigurðardóttir *et al.* 2008:784–786 og Margunn Rauset *et al.* 2012).

Í hverju markmálslandi hafa að jafnaði starfað 1–4 ritstjórar, þ.e. þýðendur. ISLEX-orðabókin gerir miklar kröfur til þeirra. Þýðendurnir þurfa að vera mjög vel að sér bæði í íslensku og markmálinu, þeir þurfa að vera næmir á báðar tungurnar og fundvísir á réttu jafnheitin við íslensku flettarnar. Bakgrunnur þýðenda ISLEX er mjög mismunandi. Sumir þeirra eru Íslendingar, aðrir Skandínavar, nokkrir eru af blönduðu þjóðerni. Sumir Skandínavanna hafa starfað sem sendikennarar um tíma á Íslandi og kennt móðurmál sitt við Háskóla

3 Þegar þetta er ritað er þessi verkþáttur enn á vinnslustigi.

Íslands og lært íslensku í leiðinni. Aðrir eru Íslendingar sem hafa búið áratugum saman í einhverju öðru Norðurlandanna.

5.8.1 Þýðingarnar: helstu áskoranir

Starf þýðandans er að finna og skrá á sínu tungumáli sem nákvæmust jafnheiti við íslensku uppflettiorðin, í öllum merkingarliðum þegar þeir eru fleiri en einn. Einnig þarf að þýða orðastæður, orðasambönd og notkunardæmi. Víða þarf að bæta frekari upplýsingum við jafnheitin sem valin eru til nánari skýringar.

Engin sérstök regla gildir um það í hvaða röð markmálin koma inn í gagnagrunninn. Stundum geta sænskar þýðingar orðið fyrstar, stundum þær dönsku o.s.frv. Markmálin styðja hvert annað í mörgum tilfellum en þau geta líka haft truflandi áhrif hvert á annað við þýðingarnar. Í vinnuumhverfi gagnagrunnsins blasa öll málin samtímis við ritstjórunum. Þá getur komið sér vel fyrir þýðandann að horfa á notendaviðmót orðabókarinnar, en þar er hægt að sjá eitt tungumál í einu án áreitis frá hinum málunum. Oft finnst þýðendunum gott að koma að hreinni og ósnertri flettu en í öðrum tilvikum geta þeir nýtt sér rannsóknarvinnu annars þýðanda ef hann hefur orðið á undan til þess að þýða t.d. viss fræðiorð á sitt mál. Mörg fræðiorð og alþjóðleg orð eru að stofni til eins á norðurlandamálunum öðrum en íslensku (og e.t.v. færeysku), t.d. er orðið *veðurfræði* nefnt *meteorologi* á dönsku, norsku og sænsku, og *hnattvæðing* er *globalisering* á hinum málunum.

Norrænu tungumálin eru oft sviplík en stundum standast þau ekki á. Það kemur fyrir að sum málin eiga jafnheiti við íslenska orðið á meðan önnur þurfa að grípa til skýringa:

afréttur no kk

Norska: *beite, fjellbeite*

Danska: *græsgang* (*i vildmarken; fælles græsning for et områdes færehold*)

Færeyska: *fjallhagi*

Oft koma upp slík tilfelli þar sem er aðeins í sumum málunum hægt að þýða beint með orði á móti orði. *Afréttur* er fyrirbæri sem þarfnast lengri útskýringar á dönsku, væntanlega vegna þess að landslag og landshættir í Danmörku eru þess eðlis að nákvæmlega sambærilegt hugtak fyrirfinnst ekki í dönsku. Danska þýðingin verður því útskýringarkennd. Aftur á móti er *afréttur* til í Noregi og Færeyjum líkt og á Íslandi og þar getur þýðandinn fært inn jafnheiti á sínu máli.

Sveitungi er orð þar sem jafnheiti er til á norsku en ekki á dönsku og sænsku:

sveitungi no kk

Norska: *sambygding*

Danska: *person som er fra samme område som nogen*

Sænska: *person som är från samma ort eller trakt som angiven person*

Sumt er vandasamara en annað í þýðingum á efninu í ISLEX. Fag-orðaforði íslenskunnar, þ.e. orð sem heyra til sérsviðum, getur verið mjög erfiður viðureignar fyrir þýðendurna og í sumum löndunum hafa þeir fengið aðstoð fræðimanna við þann orðaforða. Sum orð innan sérsviða eins og bragfræði, jarðfræði og lögfræði geta verið illskiljanleg mörgum íslenskum málnotendum og er ekki erfitt að gera sér í hugarlund hve tímafrekt getur verið að þýða þau á önnur mál, sérstaklega þar sem handbækur um þessi fög með erlendum atriðisorðum eru ekki ýkja margar.

Enda þótt leitast hafi verið við að hafa málsnið notkunardæmanna í ISLEX sem hlutlausast hefur einn þýðendanna getið um gildru sem auðvelt er að falla í við þýðingarnar og sem þurfi að varast, en það er að íslenskan virkar oft fremur hátíðleg á Skandínava. Það getur bæði átt við um einstök orð og byggingu setninga, og það þarf að gæta þess að hafa þýðingarnar ekki í of hátíðlegum stíl þar sem það á ekki við.

5.8.2 Upplýsingar um jafnheiti

Þegar tvö eða fleiri erlend jafnheiti eru valin sem þýðing við tiltekið íslenskt orð er oft bætt við fróðleik um rétta notkun hvers jafnheitis. Þetta kemur sér vel fyrir þá málnotendur sem hafa það markmál ekki að móðurmáli. Sænsku þýðingarnar á so. *rymja* og no. *fimmkall* lýsa þessu:

rymja, v: *bröla (om tjur), råma (om ko), skria (om åsna)*

fimmkall, n m: *femna (om myntet); fem spänn (om beloppet)*

Orðið *amma* er ekki til í sænsku í sama skilningi og í íslensku, dönsku og norsku. Í sænsku þarf að taka fram hvort um er að ræða föðurömmu eða móðurömmu:

amma no kvk

Danska: *bedstemor*

Norska: *bestemor*

Sænska: *farmor (móðir föður); mormor (móðir móður)*

Sums staðar er að finna upplýsingar um íslenska uppflöttiorðið sem skýra nánar merkingu þess eða notkun fyrir erlendum málhöfum. Eftirfarandi danskar þýðingar sýna þetta:

kvæntur, adj: *gift (om mand)*

skólasystir, n f: *skolekammerat (om pige)*

Íslensku orðin að ofan eru takmörkuð við ákveðið kyn og hafa þannig þrengri merkingu en dönsku jafnheitin *gift* og *skolekammerat*. Í slíkum tilvikum er nauðsynlegt að bæta við frekari skýringu sem er látin birtast innan sviga.

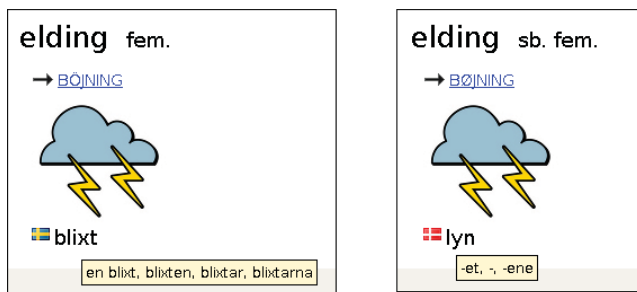
5.8.3 Málfræðiupplýsingar um markmálin

Þegar Íslendingur notar ISLEX-orðabókina getur verið gagnlegt fyrir hann að vita hvort danska jafnheitið er í samkyni eða hvorugkyni eða hvernig þátíð er af sænskri sögn. Þessar upplýsingar er að finna í gagnagrunni orðabókarinnar. Þær eru ekki sýnilegar á yfirborðinu en birtast þegar mús er rennt yfir erlenda jafnheitið, þá sprettur fram rammi með beygingarendingum viðkomandi orðs. Markmálin fara dálítið mismunandi leiðir til að birta slíkar málfræðiupplýsingar eins og sést á næstu fjórum myndum. Myndir 7 og 8 sýna beygingarendingar sænska orðsins *fjäril* og danska orðsins *sommerfugl* 'fiðrildi'.



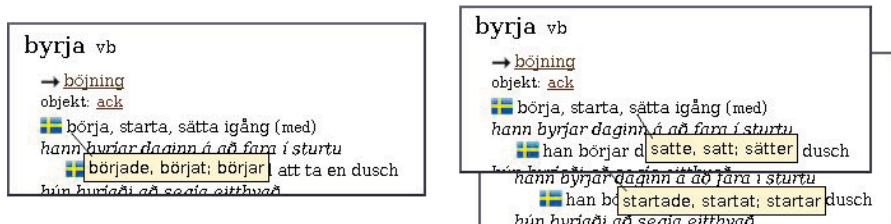
Mynd 7 og 8. Uppflöttiorðið fiðrildi með sænskri og danskri þýðingu. Í sænsku gerðinni birtist rammi með beygingarmyndum jafnheitisins þegar mús er rennt yfir það: orðið með óákveðnum greini, með viðskeyttum ákveðnum greini, fleirtala orðsins og fleirtalan með greini. Í dönsku gerðinni birtast beygingarendingar jafnheitisins á sama hátt: viðskeyttur ákveðinn greinir, fleirtöluending og ákveðinn greinir í fleirtölu.

Myndir 9 og 10 sýna beygingarendingar sænska orðsins *blíxt* og danska orðsins *lyn* 'elding'.



Mynd 9 og 10. Uppflettiorðið elding með sænskri og danskri þýðingu. Músinni hefur verið rennt yfir jafnheitin svo að beygingarendingar þeirra koma í ljós.

Málfræðiupplýsingar af þessu tagi eru hafðar við nafnorð, lýsingarorð og sagnir. Mynd 11 sýnir beygingarendingar þriggja sagna í sænsku. Í sérhverju af jafnheitunum þremur eru fólgnar upplýsingar um beygingu þess.



Mynd 11. So. byrja með þýðingum á sænsku. Beygingarendingar sænsku jafnheitanna börja, starta og sætta igång sjást þegar mús er rennt yfir viðkomandi sænskt orð.

Þótt sænskir og danskir notendur þurfi varla á slíkum fróðleik að halda um sín eigin mál geta þær komið sér vel fyrir Íslendinga (og aðra) sem ætla sér að tala eða skrifa sænsku eða dönsku.

6 Notendaaðgangur að ISLEX

Aðgangur að ISLEX-orðabókinni er á vefnum. Það form býður vitanlega upp á aðrar og á margan hátt líflegri aðferðir við framsetningu efnisins en kostur er á í hefðbundinni, prentaðri bók, t.d. hljóddæmi, myndir og hreyfimyndir, svo og tengla til að ferðast á milli flettna. Gefinn er kostur á nokkrum mismunandi leitarmöguleikum. Mynd 12 sýnir orðið *eldgos* eins og það birtist í notendaviðmóti orðabókarinnar.

- íslensku uppfllettiorði
- orðmynd í íslensku (þ.e. hvaða beygingarmynd sem er af uppfllettiorðunum)
- jafnheiti á völdu markmáli

Á mynd 14 hefur verið slegið inn orðið *skál* og vinstra megin á myndinni sjást niðurstöður úr þeirri leit. Í íslenskum notendaham er ætíð leitað að jafnheitum í öllum markmálunum og því sjást margir tungumálafánar við orðin í listanum. Fyrsta orðið í listanum er *skál*, en hún er uppfllettiorð og fær því mest gildi í leitinni. Næsta orð er *skal* sem er beygingarmynd af so. *skulu*. Í þriðja sæti listans og niður úr eru jafnheiti *skal*, *skål* og *skäl* á fjórum málum. “Loðin leit” (e. *fuzzy search*) er sjálfgild leitarhegðun. Með því er átt við að leyfð er ákveðin ónákvæmni í innsætti til hagræðis fyrir notandann, og er þannig ekki gerður greinarmunur á bókstöfunun *a*, *á*, *å* og *ä*; *d* og *ð* eru jafngildir stafir í leit o.s.frv.

Leitarniðurstöður fyrir 'skál'	
o skál no kvk	skál
o skal = skulu so	Leita
o skal : börkur no kk	á ö é í ó ú ý þ æ ö å ä ø Flóknari leit
o skal : hismi no hk	
o skal : hýði no hk	
o skal : skel no kvk	
o skal : skræl no hk	
o skal : skurn no kvk/hk	
o skål : dallur no kk	
o skål : full no hk	
o skål : matarílát no hk	
o skål : skál no kvk	
o skäl : ástæða no kvk	
o skäl : orsök no kvk	
o skäl : rök no hk ft	
o skäl : röksemd no kvk	

Mynd 14. Íslenskt notendaviðmót ISLEX. Leit að *skál* skilar mörgum ólíkum niðurstöðum, fyrst uppfllettiorðinu *skál*, svo orðmyndinni *skal* (af *skulu*) og því næst ýmsum jafnheitum markmálanna. Hægra megin við tungumálafánana eru íslensku uppfllettiorðin (feitletruð) sem jafnheiti koma fyrir í.

Hægt er að leita að skandínavískum orðum sem ekki eru íslensk uppfllettiorð en óvíst er hvaða árangur slík leit ber. Mynd 15 sýnir niðurstöðu leitar að orðinu *monster* í sænsku notendaviðmóti. *Monster* er hvorki til sem uppfllettiorð né beygingarmynd, heldur eru í niðurstöðulistanum flettur þar sem orðið kemur fyrir sem sænskt jafnheiti. Neðsta orðið *mönster* finnst líka vegna þess að leitin er loðin.



Mynd 15. Sænskt notendaviðmót ISLEX. Leitað hefur verið að monster. Niðurstöður leitarinnar birtast í tveimur dálkum: vinstri megin eru sænsku jafnheitið og hægra megin við þau eru viðkomandi íslensk uppflettiorð. Næsta skref er að velja eitthvert af orðunum úr listanum og fara inn í viðkomandi flettugrein.

Þegar flóknari leit er valin eru fleiri möguleikar í boði. Hægt er að:

- Slökkva á leit að uppflettiorði (þá er bara leitað að erlendu jafnheiti).
- Slökkva á leit að jafnheiti (þá er aðeins leitað að íslensku uppflettiorði og orðmynd þess).
- Kveikja á leit í *texta*, en þá finnast orð eða stafastrengir sem koma fyrir í notkunardæmum og orðasamböndum, jafnt á íslensku sem markmálunum.
- Slökkva á loðinni leit.
- Tilgreina nánar þau markmál sem leitin nær til með því að haka við eða slökkva á reitunum *danska*, *norska*, *nýnorska* og *sænska*.

ISLEX hefur einnig verið aðlagð að helstu stýrikerfum snjallsíma. Þá er farið á vefsíðuna á venjulegan hátt og sérstakur hugbúnaður skynjar hvaða tæki verið er að nota og sér um að birta vefinn í samræmi við það.



Mynd 16. Flettan eldgos í snjallsíma.

7 Frekari þróun og notkunarmöguleikar

Verkefni eins og ISLEX felur í sér marga spennandi möguleika til frekari þróunar. Hægt er til dæmis að bæta við fleiri tungumálum í gagnagrunninn, og í undirbúningi er að finnska verði sjötta markmál orðabókarinnar. Einnig hafa enska, franska og þýska verið nefndar til sögunnar sem vænlegir kostir.

ISLEX er nú gefin út sem veforðabók en útgáfuformið getur breyst með tímanum. Þegar upphaflega var gengið til samstarfs við sænska og norska aðila stóð til að ISLEX yrði prentaðar orðabækur milli íslensku og sænsku annarsvegur og íslensku og norsku hinsvegar. Það fyrirkomulag getur vel orðið að veruleika en hafa þarf í huga að talsverður kostnaður fylgir umbroti og prentun orðabóka. Ennfremur er þá lokað fyrir möguleikann á viðbótum og leiðréttingum sem svo fyrirhafnarlítið er að bæta inn í rafrænt verk. Á móti kemur að prentuð orðabók hefur ætíð nokkurn hóbýlapokka í bókahillu eða á borði og sumir vilja frekar nota orðabækur á pappír en í tölvu.

Með tilkomu rafbóka og lestölva (lesbretta) er orðið til þriðja sniðið sem hvorki er á vef né pappír. Vel má vera að slík útgáfa á ISLEX líti dagsins ljós innan fárra ára. Orðabækur á geisladiskum hafa áður gegnt svipuðu hlutverki en útgáfa efnis á geisladiskum hefur nú að miklu leyti vikið fyrir öðrum miðlum.

ISLEX er líka kjörinn efniviður í nýja íslensk-íslenska orðabók, en slíka bók af hentugri stærð hefur sárlega vantað í nokkra áratugi. Samhliða vinnunni við ISLEX hefur orðið til stofn í nýja íslenska orðabók yfir nútíamálið með vandlega völdum uppflettiorðum, orðskýringum og notkunardæmum. Orðskýringar á íslensku í ISLEX eru upphaflega tilkomnar til þess að auðvelda þýðendunum verk sitt en þær hafa gert það að verkum að vísir að íslensk-íslenskri orðabók hefur myndast sem hliðarafurð. Skýringarnar eru þó ekki alltaf nægilega vandaðar til að vera birtingarhæfar eins og sakir standa. Ekki þyrfti þó mjög mikla viðbótarvinnu til að hægt væri að gefa út handhæga, meðalstóra íslenska orðabók.

Eins og hún er hönnuð hentar ISLEX-orðabókin vel til að nota við tungumálakennslu í skólum þar sem hún er mjög þægilegt verkfæri. Einnig er hún kjörin til þess að búa til kennsluefni fyrir grunnskóla og menntaskóla, og hefur nú þegar verið unnið slíkt verkefni fyrir menntaskóla (sjá www.flensborg.is/islex).

ISLEX er ekki aðeins orðabók heldur hefur sprottið af verkefninu stórt gagnasafn um norrænu tungumálin. Sérhvert íslenskt uppfletti-

orð er tengt jafnheitum sínum á dönsku, norsku, sænsku og færeysku í gegnum færslur í gagnagrunninum. Allur þessi efniviður með vandlega skilgreindum upplýsingum gerir það að verkum að ISLEX er gott hráefni í ýmiss konar máltæknileg verkefni þar sem fengist er við norðurlandamálin. Geta má þess að gert hefur verið samkomulag við forráðamenn META-NORD-verkefnisins um að hagnýta megi ákveðin gögn úr ISLEX. META-NORD er verkefni sem Norðurlöndin og Eystrasaltslöndin standa að, og er markmið þess að efla gagnsöfn um viðkomandi tungumál sem nýst geti í margvísleg máltækniverkefni (sjá nánar á www.meta-nord.eu).

Í vinnunni við verkefnið hefur safnast upp mikill fjöldi jafnheita á markmálunum. Svo dæmi sé tekið eru um 68.000 sænsk jafnheiti í ISLEX (41.000 jafnheiti ef hvert er talið einu sinni). Svipað má segja um hin markmálin. Vegna þessa auðuga orðaforða gæti efnið í ISLEX hentað sem grunnur í orðabækur sem snúa öfugt, þ.e. í sænsk-íslenska orðabók o.s.frv.

Af ofangreindu er ljóst að ISLEX felur í sér margvíslega möguleika til stækkunar, þróunar, útgáfu og rannsóknarverkefna. Tíminn leiðir svo í ljós hvað á eftir að spretta af þessu orðabókarverki.

8 Lokaorð

Hér hefur verið fjallað vítt og breitt um veforðabókina ISLEX en einkum þá þætti sem snúa að íslenska viðfangsmálinu. Fyrst var sagt frá bakgrunni verkefnisins og upphafi þess, og því næst var farið almennum orðum um hina rafrænu orðabók. 4. kafli fjallaði um verklagið við ritsstjórn orðabókarinnar, fyrst um íslenska hlutann og síðan um starfið við markmálin í nágrannalöndunum. Í 5. kafla sagði frá efni og uppbyggingu orðabókarinnar og var þar m.a. fjallað um notkun textasafna í starfinu, orðaforðann í ISLEX, uppbyggingu flettugreina, orðasambönd og notkunardæmi. Greint var frá merkingarflokkum sem hafa verið nýttir við að gera flestum flettugreinunum skil, bæði við vinnslu íslenska hlutans og markmálanna. Rætt var um markmálin og helstu áskoranir þar, en hafa ber í huga eins og fyrr er getið að þessi grein er rituð út frá sjónarhóli ritsstjóra íslenska hluta ISLEX. Í 6. kafla var lýst notendaviðmóti orðabókarinnar á vefnum og í 7. kafla var fjallað um ýmsa möguleika sem búið geta í svona verki og hugað að framhaldi þess.

Fram kom að orðabókin hefur verið unnin af mörgum ritstjórnunum samtímis í mörgum löndum og hefur íslenska ritstjórnin leitt verkið. Þetta verklag hefur óhjákvæmilega verið nokkuð þungt í vöfum og gert samhæfingu orðabókarstarfsins allkrefjandi. Þrátt fyrir það hefur starfið við ISLEX almennt gengið greiðlega en svo margslungið verkefnið útheimtir talsverða lipurð í samskiptum, bæði á milli landa og innan ritstjórnarhópanna innbyrðis. Einnig hefur þurft að halda mjög vel á spöðunum varðandi nýtingu á sameiginlegum styrkjum sem verkefnið hefur fengið, skipulag starfsins, skipulag funda o.s.frv.

ISLEX-orðabókin var opnuð almenningi þann 16. nóvember 2011, á degi íslenskrar tungu. Viðtökur hennar hafa verið afar jákvæðar og orðabókin fékk fljótt fjölmarga notendur. Verkinu var þó ekki fulllokið þegar það var opnað, t.d. vantaði talsvert upp á að norsku þýðingarnar væru tilbúnar, margar myndir vantaði og upptökur af framburði orðanna voru ekki tilbúnar. Notendur hafa verið þolinmóðir gagnvart þessum annmörkum enda kemur fram fyrirvari um slíkt á vefsíðu orðabókarinnar þar sem segir frá því að efnið sem vantar sé á leiðinni. Margir notendur hafa haft samband og lýst yfir ánægju sinni með verkefnið, sem vonandi auðnast að vaxa áfram og dafna.

Heimildir

- Aldís Sigurðardóttir, Anna Hannesdóttir, Halldóra Jónsdóttir, Hákan Jansson, Lars Trap-Jensen, Þórdís Úlfarsdóttir. 2008. ISLEX – An Icelandic-Scandinavian Multilingual Online Dictionary. *Proceedings of the XIII Euralex International Congress*, bls. 779–789. (Á geisladiski).
- Beygingarlýsing íslensks nútíðsmáls*. Stofnun Árna Magnússonar í íslenskum fræðum: <http://bin.arnastofnun.is>.
- Halldóra Jónsdóttir og Þórdís Úlfarsdóttir. 2010. ISLEX – Íslensk-Skandinavisk webordbog. *Málfríður* 2. tbl. 2010: 11–15.
- ISLEX-orðabókin*. Stofnun Árna Magnússonar í íslenskum fræðum: www.islex.is, www.islex.dk, www.islex.no og www.islex.se.
- Íslensk orðabók*. 1963. Ritstjóri Árni Böðvarsson. Reykjavík: Bókaútgáfa Menningarsjóðs.
- Íslensk orðabók*. 1993. 2. útg. aukin og bætt, 7. prentun. Ritstjóri Árni Böðvarsson. Reykjavík: Mál og menning.
- Íslensk orðabók*. 2007. 4. útg. Ritstjóri Mörrður Árnason. Reykjavík: Edda.
- Íslenskt orðanet*. Stofnun Árna Magnússonar í íslenskum fræðum: www.ordanet.is.
- Íslenskt textasafn*. Stofnun Árna Magnússonar í íslenskum fræðum: www.lexis.hi.is/corpus/leit.pl.

- Jón Hilmar Jónsson. 2005. *Stóra orðabókin um íslenska málnotkun*. Reykjavík: JPV.
- Jón Hilmar Jónsson. 2009. Lemmatisation of Multi-word Lexical Units: Motivation and benefits. Í: H. Bergenholtz, S. Nielsen og S. Tarp (ritstj.). *Lexicography at Crossroads*, bls. 165–194. Bern, Berlin o.v.: Peter Lang.
- Jón Hilmar Jónsson. 2012. Að fanga orðaforðann: orðanet í þágu orðabókar. *Orð og tunga* 14:39–65.
- Kristín Bjarnadóttir. 1998. *Norræna verkefnið*: www.lexis.hi.is/kristinb/norr.pdf.
- Ljunggren, Anders. 2011. *Några skärvor som kan fogas till en bild av vägen fram till ISLEX tillkomst*. <http://islex.is/islex/greinar/Ljunggren-nov2011.pdf>.
- Margunn Rauset, Anna Helga Hannesdóttir og Aldís Sigurðardóttir. 2012. Ein-, to- eller fleirspråklig ordbok? Í: Eaker, B., L. Larsson og A. Mattisson (ritstj.). *Nordiska studier i lexikografi* 11, bls. 512–523. Lundur.
- META-NORD-verkefnið. www.meta-nord.eu.
- Ritmálssafn Orðabókar Háskólans. Stofnun Árna Magnússonar í íslenskum fræðum: www.arnastofnun.is/page/arnastofnun_gagnasafn_ritmal.
- Svavar Gestsson. 2011. *Ævintýri handa öllum heiminum*. http://islex.is/islex/greinar/avarp_SG-nov11.pdf.
- Svensén, Bo. 2004. *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Stockholm: Norstedts Akademiska Förlag.
- Tímarit.is. Landsbókasafn Íslands – Háskólabókasafn: <http://tímarit.is>.
- Þórdís Úlfarsdóttir. 2011. *Upphaf ISLEX-verkefnisins*. http://islex.is/islex/greinar/Islex_upphaf.pdf.

Lykilorð

tvímála orðabók, margmála orðabók, veforðabók

Keywords

bilingual dictionary, multi-lingual dictionary, online dictionary

Abstract

The Icelandic-Scandinavian dictionary ISLEX was opened on the web in November 2011. The outcome of the project is twofold. Firstly, a long awaited Icelandic-Scandinavian dictionary has been produced using the web-medium and its many possibilities. Secondly, the project has resulted in the compilation of a large amount of data on the Nordic languages, and provided important experience in jointly building such work in five countries simultaneously, as the Icelandic source language has been processed in Reykjavík and the target languages in Copenhagen, Bergen, Göteborg and Tórshavn, respectively. ISLEX has thus contributed to Nordic language collaboration.

The ISLEX-dictionary describes modern Icelandic, including a range of phrases, collocations and examples of use, all of which are translated into the five TLs, Danish, Norwegian (Bokmål og Nynorsk), Swedish and Faroese. ISLEX is thus a multi-lingual dictionary within which are contained five bilingual dictionaries. The electronic form offers direct search both in Icelandic and the TLs and is therefore ideal for comparative studies between the languages. The project also offers other possibilities for research. The material is a good base for various internordic projects within education and language technology. Moreover, the whole Icelandic vocabulary is semantically classified, a classification which extends to the TLs too up to a point.

Þórdís Úlfarsdóttir

Stofnun Árna Magnússonar í íslenskum fræðum, orðfræðisvið

Neshaga 16, 107 Reykjavík

disa@hi.is