

Veffregnir

Á árunum 2018 og 2019 hafa margir vefir Stofnunar Árna Magnússonar í íslenskum fræðum verið uppfærðir auk þess sem nýir vefir hafa litið dagsins ljós. Í þeirri vinnu hefur verið leitast við að bæta aðgengi að gögnum á vefunum, auðvelda notkun, samræma útlit og hönnun og gera þá farsímavænni. Margir af nýju eða endurbættu vefunum eru vefir fyrir tiltekin orðasöfn eða orðabækur, en auk þeirra hafa verið settir upp nýir málheildavefir, vefur fyrir *Orð og tungu* og nokkrir minni vefir með sérefni. Við lýsum stuttlega þessum nýjungum og segjum frá virkni vefanna.

Beygingarlýsing íslensks nútímamáls

Ný útgáfa *Beygingarlýsingar íslensks nútímamáls (BÍN)* var kynnt haustið 2019 (Kristín Bjarnadóttir o.fl. 2019). Í nýju útgáfunni var gagnagrunnurinn sem *BÍN* keyrir á skrifaður upp á nýtt svo hægt væri að skrá alls kyns nýjar upplýsingar. Þannig er nú hægt að gera greinarmun á lýsandi og vísandi málfræði í gagnagrunninum og tekið hefur verið saman vísandi úrtak með yfir 50.000 algengum orðum, *BÍN*-kjarninn. *BÍN*-kjarninn er aðgengilegur í forritaskilum, sem gefur öðrum vefum kost á að nýta efnið í honum. Auk þess eru nú aðgengileg ýtarlegri gögn úr *BÍN* til nota í máltækni og *BÍN*-vefurinn var hannaður upp á nýtt til að verða farsímavænni og aðgengilegri fyrir notendur.

Ritmálssafn Orðabókar Háskólans

Í *Ritmálssafni Orðabókar Háskólans* má finna dæmi um orðnotkun úr íslenskum textum frá siðaskiptum til loka 20. aldar. Í safninu eru orðtekin dæmi, auk upplýsinga um heimildir. Haustið 2018 lauk þróun á nýju viðmóti fyrir ritmálssafnið á vefslóðinni *ritmalssafn.arnastofnun.is*. Um er að ræða endurhönnun á eldri vef með mjög auknum leitar-möguleikum.

Nýja útgáfan býður ekki aðeins upp á leit í uppflettiorðum eða ritmyndum, heldur einnig í heimildaskrá og þannig er auðvelt að nálgast öll dæmi sem orðtekin hafa verið í ákveðinni heimild. Þegar leitað er eftir orðum er boðið upp á að þrengja leitina svo aðeins komi fram dæmi frá ákveðnu tímabili. Í heimildaleit er hægt að leita eftir aldri heimildar og nöfnum höfunda og þýðenda. Í dæmalista geta notendur hlaðið niður dæmum í töfluskrá sem opna má í töflureikni. Loks geta notendur safnað einstökum dæmum í sinn eigin dæmalista og hlaðið þeim lista niður.

Nýyrðavefurinn

Á degi íslenskrar tungu 2018 var *Nýyrðavefur Stofnunar Árna Magnússonar í íslenskum fræðum* opnaður á vefslóðinni *nyyrði.arnastofnun.is*. Þeim vef er ætlað að vera vettvangur fyrir almenning til að koma nýyrðum á framfæri. Vefnum er ekki ætlað að vera uppflettivefur með samþykktum eða viðurkenndum nýyrðum heldur frekar að vera staður til að birta tillögur að nýyrðum og samræðuvettvangur um þau. Þau nýyrði sem notendur senda inn safnast fyrir í ritstjórnarkerfi þar sem þau bíða samþykkis ritstjóra fyrir birtingu. Hverju nýyrði geta fylgt skýringar, ein eða fleiri, ásamt dæmum og vísun í heimild ef orðin hafa birst á prenti eða á netmiðli.

Notendur geta skráð athugasemdir um birt nýyrði á vefnum og gefið þeim einkunn í formi þumals sem annaðhvort vísar upp eða niður og er þar um beina skírskotun til samfélagsmiðla að ræða. *Nýyrðavefurinn* sker sig því þannig frá hefðbundnum orðasöfnum stofnunarinnar þar sem á honum fær talsverður léttleiki að ríkja. Fyrsta árið voru um 400 nýyrði frá notendum birt á vefnum.

Íðorðabankinn

Íðorðabanki Stofnunar Árna Magnússonar í íslenskum fræðum, áður *Orðabanki Íslenskrar málstöðvar*, var opnaður í október 2019 í nýrri mynd. Í *Íðorðabankanum* má fletta upp um 178.000 hugtökum í 66 orðasöfnum sem hvert fyrir sig tengist ákveðnu sérsviði. Nýtt viðmót *Íðorðabankans* býður upp á leit í öllum hugtökum án þess að þörf sé á að velja leitarmál. Þegar notendur leita, hvaða tungumál sem þeir nota, fá þeir strax tillögur að niðurstöðum í sprettiglugga þar sem sést í hvaða orðasöfnum niðurstöður er að finna og á hvaða tungumálum. Í niðurstöðulista má síðan sía eftir einstökum orðasöfnum og raða eftir tungumálum. Notendur geta þannig smellt á orðasafnshnapp og

einskorðað sig þannig við það safn. Í niðurstöðulista getur notandi valið á milli þess að fá knappan lista með grunnupplýsingum sem sýnir meira á skjánum eða ítarlegri lista þar sem auknar upplýsingar um hvert hugtak sést. Vef Íðorðabankans má finna á vefslóðinni *idord.arnastofnun.is*.

Málfarsbankinn

Vefur *Málfarsbankans* var endurskoðaður haustið 2018 og hannaður upp á nýtt. Vefurinn inniheldur stuttar greinar um málfar og málnotkun og er efninu þar skipt í um tuttugu flokka. Hægt er að leita inn í textum greinanna með textastrengjum eða eftir sérstökum leitarorðum sem tengd eru hverri grein og eiga að stuðla að því að auðveldara sé að finna efni í bankanum. Leitin virkar eins og á flestum hinna nýju vefa, tillögur birtast í lista um leið og ritað er í leitarreit og auðvela þannig notandanum að finna efni. Á vormisseri 2019 var málfarspistlum Jóns G. Friðjónssonar bætt við bankann. Þar er um að ræða pistla um íslenskt mál sem Jón hefur skrifað frá 2015 og birt á vikulegum tölvupóstlista sínum.

ISLEX og Íslensk nútímamálsorðabók

Margmála orðabókin *ISLEX* (Halldóra Jónsdóttir og Þórdís Úlfarsdóttir 2016) og *Íslensk nútímamálsorðabók* (Halldóra Jónsdóttir og Þórdís Úlfarsdóttir 2019) eru nátengdar. *ISLEX*, samvinnuverkefni sex stofnana á Íslandi, í Danmörku, Noregi, Svíþjóð, Finnlandi og Færeyjum, hefur íslensku sem viðfangsmál en hin Norðurlandamálin eru markmál. *Íslensk nútímamálsorðabók* byggir svo á grunni *ISLEX*-orðabókarinnar.

Í nóvember 2019 opnaði nýr vefur *ISLEX* á vefslóðunum *islex.is* og *islex.arnastofnun.is*. Nýr vefur *Íslenskrar nútímamálsorðabókar* fór svo í loftið í apríl 2020 á vefslóðinni *islenskordabok.arnastofnun.is*. Grunnvirkni nýju vefanna er í flestu sú sama og í hinum eldri. Leitin er þó orðin mun hraðari og leitarglugginn býður upp á sprettiniðurstöður eins og aðrir nýir vefir og kemur þannig með tillögur að orðum um leið og ritað er. Þá er viðmótið hannað þannig að það gagnast notendum flestra gerða af tölvum og öðrum tækjum.

Stærsta nýjungin er möguleikinn á að sjá samheiti við flestar flettur orðabókanna. Samheitin eru fundin sjálfvirkt með því að bera saman öll markmál *ISLEX* nema færeysku. Til að finna möguleg samheiti íslensks orðs eru jafnheiti orðanna á hinum tungumálunum skoðuð og reynt að finna önnur íslensk orð sem hafa sömu jafnheiti. Þannig fæst nokkuð góður listi líklegra samheita en þó er sá fyrirvari gefinn á vefsíðunni að um sjálfvirkan lista sé að ræða og því geti verið að sumar samheitatillögurnar eigi ekki við. Til að tryggja að samheitin séu sem réttust er færeysku sleppt úr samanburðinum því í færeysku og íslensku eru mörg orðin eins en merkingin ekki alltaf sú sama. Þessi möguleiki er notaður bæði í *ISLEX* og *Íslenskri nútímamáls-orðabók* en í báðum tilfellum er gagnagrunnur *ISLEX* notaður til að finna samheitin.

Skrá um orðasambönd

Nýr vefur með skrá um orðasambönd fór í loftið í lok árs 2019. Skráin er unnin upp úr dæmum úr *Ritmálssafni Orðabókar Háskólans* og byggist því á textum frá miðri 16. öld og fram til loka 20. aldar. Í skránni má finna um 130.000 orðasambönd. Hægt er að leita í skránni eftir lykilorðum, sem eru veigamestu orðin í hverju orðasambandi, eða með strengjaleit í texta. Við leit birtist listi yfir þau orðasambönd sem tengjast viðkomandi lykilorði og í listanum má sjá önnur lykilorð hvers orðasambands. Þau lykilorð er svo hægt að smella á til að fíkra sig áfram um vefinn. Orðasambandaskráin er aðgengileg á vefslóðinni ordasambond.arnastofnun.is.

Málheildarvefir

Vorið 2018 var ný textamálheild, *Risamálheildin*, opnuð á vefnum eins og gerð var grein fyrir í 21. hefti *Orðs og tungu* (Steinþór Steingrímsson 2019). Í tengslum við málheildina hafa þrír vefir verið opnaðir. Fyrst ber að telja leitarvélina á malheildir.arnastofnun.is. Þar er hægt að leita í textunum með fyrirspurnum sem byggja á málfræði og skoða þannig breytileika í setningagerð, beygingar, orðmyndun, merkingarbreytingar og svo má áfram telja. Málheildin og leitarvefurinn gerbreytir aðstöðu málfræðinga til rannsókna á því hvernig íslenskt mál er notað í ritmáli.

Í tengslum við *Risamálheildina* rekur stofnunin líka tvo aðra vefi. *ord-tidni.arnastofnun.is* veitir upplýsingar um tíðni orða í málheildinni allri eða tilteknum undirmálheildum. Þar er hægt að skoða tíðni orðmynda

eða lemma, tiltekinna orðflokka eða jafnvel ákveðinna greiningarstrengja í *Risamálheildinni*. *n.arnastofnun.is* er n-stæðuskoðari. Hann sýnir orðnotkun yfir tíma og í honum er auðveldlega hægt að sjá hvort notkun hafi breyst, aukist eða minnkað, á orðum, orðmyndum eða orðastæðum með allt upp í þremur orðum. Gögnin á öllum þessum þremur vefum eru uppfærð árlega.

Tímaritavefir

Unnið hefur verið að því að koma tímaritunum tveim sem gefin eru út af Stofnun Árna Magnússonar í íslenskum fræðum, *Gripla* og *Orði og tungu*, á vefinn í rafrænni útgáfu. *Gripla* er ekki enn komin í rafræna útgáfu en haustið 2019 fór í loftið sérstakur vefur fyrir *Orð og tungu* þar sem ritið er birt rafrænt. Þegar ákveðið var hvernig best væri að setja upp vef fyrir tímaritið var leitað í smiðju þeirra sem gefa út *Ritið: Tímarit Hugvísindastofnunar* og *Stjórnmal og stjórnsýslu*, sem gefið er út af Stofnun stjórnsýslufræða og stjórn mála við Háskóla Íslands. Þau tímarit eru gefin út rafrænt og nota Open Journal System sem notað er víða um heim fyrir rafræn tímarit. Stofnunin ákvað að gera það líka, því auk þess að vera aðgengilegt fyrir lesendur tímaritsins, þá fullnægir kerfið helstu skilyrðum um opinn aðgang, styður við varanleg auðkenni (e. *digital object identifier*, DOI) og er þannig úr garði gert að auðveldara er að keyra lýsigögn úr því yfir í fræðigreina veitur á borð við Google Scholar og CrossRef (Haukur Arnþórsson 2013). Það fylgir því þess vegna mikill ávinningur að nota þetta kerfi. Slóðin fyrir *Orð og tungu* er *ordogtungu.arnastofnun.is*.

Aðrir vefir

Af öðrum vefum sem birst hafa síðustu tvö árin ber helst að nefna tvo sem tengjast orðfræðivefum stofnunarinnar en eru þó ekki hefðbundnir orðfræðivefir, heldur frekar til gamans og jafnvel gagns. Fyrst er að nefna *Lykilorðasmiðinn* en hlutverk hans er að búa til lykilorð sem fólk getur notað fyrir aðgang að tölvukerfum. Það er þekkt vandamál að mörg tölvukerfi krefjast þess að lykilorð innihaldi ákveðið marga bókstafi og þá getur verið erfitt að muna þau ef þau eru flókin. *Lykilorðasmiðurinn* sækir tvö orð af handahófi úr gagnagrunni *Ritmálssafns Orðabókar Háskólans* og skeytir þeim saman með punkti á milli. Þannig fæst í flestum tilfellum mjög sterkt lykilorð sem þó er tiltölulega auðvelt að muna. *Lykilorðasmiðinn* má nálgast á slóðinni *lykilorð.arnastofnun.is*.

Þá er að nefna vefinn *Óravíddir – orðaforðinn í nýju ljósi* en þar er um að ræða vef sem settur var upp samhliða sýningu í Safnahúsinu við Hverfisgötu sem var opnuð 7. júní 2019. Sýningin byggir á myndrænni framsetningu á gögnum úr Íslensku orðaneti. Í Íslensku orðaneti eru vensl orða skráð þannig að hvert orð er tengt fleiri orðum í gegnum pör sem finnast í rituðu máli. Fyrir sýninguna voru skyldheiti tekin fyrir og teiknað upp myndrænt hvernig þau tengjast saman. Með þessu móti birtast orðin í netum sem minna á stjörnakerfi og á sýningunni gátu gestir fylgst með því þegar flogið var um orðageiminn og staldrað við eitt og eitt orð af handahófi. Á vef *Óravídda* má lesa um sýninguna en einnig er þar hægt að fletta upp tveim orðum og sjá hvernig þau tengjast í Íslensku orðaneti í gegnum parasambönd. Vefurinn er á oraviddir.arnastofnun.is.

Lokaorð

Notkun á vefum Stofnunar Árna Magnússonar í íslenskum fræðum hefur aukist mikið síðustu ár. Fjöldi heimsóknna á fjóra vinsælustu vefina, BÍN, [málið.is](http://malið.is), *Íslensk nútímamálsorðabók* og ISLEX, ríflega tvöfaldaðist frá apríl 2018 til apríl 2020. Í upphafi tímabilsins voru heimsóknirnar á þessa fjóra vefi um 154.000 á mánuði, en í lok þess um 323.000.

Árnastofnun rekur nú um þrjátíu vefi sem tengjast öllum rannsóknarsviðum stofnunarinnar. Undanfarin tvö ár hefur verið unnið að því að samræma yfirbragð og útlit vefanna, bæta aðgengi að gögnum og auðvelda notkun í mismunandi tækjum. Nýir vefir hafa bæst við, þrír sem tengjast málheildaverkefnum, nýyrðavefur, lykilorðavefur og sýningarvefurinn *Óravíddir*. Aðrir vefir hafa verið endurskoðaðir og sumir smíðaðir aftur alveg frá grunni. Uppfærslur og endurbætur halda áfram, með það að leiðarljósi að auðvelda notkun og bæta þjónustu, og áfram munu fleiri vefir bætast við í tengslum við rannsóknarverkefni og önnur verkefni sem starfsmenn stofnunarinnar sinna.

Heimildir

- Halldóra Jónsdóttir og Þórdís Úlfarsdóttir. 2016. Hvem bruger ISLEX og hvordan? *LexicoNordica* 23:89–104.
- Halldóra Jónsdóttir og Þórdís Úlfarsdóttir. 2019. Íslensk nútímamálsorðabók. Kjarni tungumálsins. *Orð og Tunga* 21:1–26. <https://doi.org/10.33112/ordogtunga.21.2>.

- Haukur Arnþórsson. 2013. Útgáfa fræðitímarita á netinu. *Stjórnsmál og stjórnsýsla* 9/2:2–7. http://www.irpa.is/article/view/b.2013.9.2.1/pdf_284 (maí 2020).
- Kristín Bjarnadóttir, Kristín Ingibjörg Hlynsdóttir og Steinþór Steingrímsson. 2019. DIM: The Database of Icelandic Morphology. Í: Mareike Hartmann og Barbara Plank (ritstj.), *Proceedings of the 22nd Nordic Conference on Computational Linguistics*. Linköping University Electronic Press. <https://www.aclweb.org/anthology/W19-61> (maí 2020).
- Steinþór Steingrímsson. 2019. Risamálheildin. *Orð og Tunga* 21:159–168. <https://doi.org/10.33112/ordogtunga.21.9>.

Trausti Dagsson og Steinþór Steingrímsson
Stofnun Árna Magnússonar í íslenskum fræðum

