

Orð og tunga

16

Ritstjóri

Ásta Svavarsdóttir
Stofnun Árna Magnússonar í íslenskum fræðum
Neshaga 16
107 Reykjavík
sími 525 4437
asta@hi.is

Ráðgefandi ritnefnd

Anna Helga Hannesdóttir, Háskólanum í Gautaborg
Ari Páll Kristinsson, Stofnun Árna Magnússonar í íslenskum fræðum
Camilla Wide, Háskólanum í Turku
Erla Erlendsdóttir, Háskóla Íslands
Finnur Friðriksson, Háskólanum á Akureyri
Guðrún Kvaran, Stofnun Árna Magnússonar í íslenskum fræðum
Hrafn Loftsson, Háskólanum í Reykjavík
Kendra Willson, Háskólanum í Helsinki
Kristján Árnason, Háskóla Íslands
Lars Vikør, Háskólanum í Osló
Liisa Theilgaard, Háskólanum í Kaupmannahöfn
Margrét Jónsdóttir, Háskóla Íslands
Silvie Cinková, Univerzita Karlova, Prag
Veturliði Óskarsson, Háskólanum í Uppsölum
Zakaris Svabo Hansen, Fróðskaparsetri Færeyja, Þórshöfn
Þorsteinn G. Indriðason, Háskólanum í Bergen

Orð og tunga birtir fræðigreinar um mál og málfræði með áherslu á greinar sem snerta orðaforðann, þ. á m. nöfn og íðorð, og greinar um orðabókafræði og orðabókagerð. Í hverju hefti eru annars vegar greinar um fyrirfram ákveðið þema og hins vegar aðrar greinar á efnissviði tímaritsins almennt. Auk þess eru birtir ritdómar, bókafrægnir og frásagnir af rannsóknarverkefnum og ráðstefnum. Greinar skulu að jafnaði ritaðar á íslensku með útdrætti á ensku en einnig eru birtar greinar á norðurlandamálum og ensku með íslenskum útdrætti. Allar greinar í tímaritinu eru ritrýndar þannig að auk ritstjóra lesa a.m.k. tveir ónafngreindir sérfræðingar hverja grein. Frekari upplýsingar um ritstjórnarstefnu, skil og frágang greina má finna á vefsíðu tímaritsins: http://www.amastofnun.is/page/timarit_ot.

Orð og tunga

16

Ritstjóri
Ásta Svavarsdóttir



Stofnun Árna Magnússonar í íslenskum fræðum
Reykjavík 2014

© Stofnun Árna Magnússonar í íslenskum fræðum, 2014.

Öll réttindi áskilin.

Bók þessa má eigi afrita með neinum hætti, svo sem ljósmyndun, prentun, hljóðritun eða á annan sambærilegan hátt, að hluta eða í heild, án skriflegs leyfis höfunda og útgefanda.

ISSN 1022-4610

Umbrot: Bessi Aðalsteinsson.

Hönnun kápu: Björg Vilhjálmsdóttir.

Prentun og bókband: Prentmet ehf.

Efnisyfirlit

Formáli ritstjóra.	vii
Þemagreinar:	
Heimir Freyr van der Feest Viðarsson: Þáttur málstöðlunar í afstöðu sagnar til neitunar í 19. aldar íslensku — „málgreinir, sem mjer fannst eitthvert danskt óbragð að“	1
Magnus Breder Birkenes & Jürg Fleischer: Intetkjønn med referanse til personer. To ulike normeringstradisjoner i nynorsk og tysk	25
Stephan Elspaß & Konstantin Niehaus: The standardization of a modern pluriareal language. Concepts and corpus designs for German and beyond	47
Vanessa Isenmann: Insight into computer-mediated communication as a new variety of written Icelandic	69
Aðrar greinar:	
Ari Páll Kristinsson & Haraldur Bernharðsson: Íslenska eða enska í íslensku háskólastarfi	93
Kristín Bjarnadóttir: Beygingarlýsing íslensks nútímamáls: Regluverk eða beygingardæmi	123
Jónína Hafsteinsdóttir: Þveit	141
Ritfregnir	149
Ráðstefnur	159

Formáli ritstjóra

Orð og tunga kemur nú út í sextánda sinn. Þema þessa heftis er „Málstöðlun: Tilurð, viðhald og endurnýjun staðalmálsins“ (e. *Language Standardization: The origin, maintenance and renewal of the standard language*). Mörg undanfarin ár hefur verið efnt til málþings um þema hvers árs til undirbúnings heftisins en í þetta sinn var brugðið út af vananum og kallað eftir greinum með boðsbréfi sem sent var fjölmörgum sérfræðingum innan lands og utan. Þetta var viðleitni til þess að stækka og breikka þann hóp sem boðið er að senda greinar í þemahluta heftisins. Í boðsbréfinu, sem sent var bæði á íslensku og ensku, sagði m.a.:

Málstöðlun og staðalmál eru áhugaverð frá málfræðilegu og félagsmálfræðilegu sjónarhorni en tilurð, viðhald og endurnýjun málstaðalsins tengist m.a. málrækt og málstýringu, viðhorfum til tungumálsins, mállýsingum og gerð orðabóka og annarra uppfléttrita um mál og málnotkun. Málstöðlun felur í sér samræmingu og þar með val á milli afbrigða, hvort sem það snýr að mismunandi málum og málbrigðum eða tilbrigðum í einstökum atriðum. Hlutverk málstöðlunar er yfirleitt hvort tveggja í senn, hagnýtt og hugmyndafræðilegt. Hið hagnýta gildi felst í því að greiða fyrir samskiptum innan ríkisins þvert á landfræðileg og félagsleg skil og auðvelða samræmingu í menntakerfinu, opinberri stjórnsýslu, útgáfumálum o.fl. Hinn pólitíski eða hugmyndafræðilegi þáttur snýr að því að þjappa þjóðinni saman um eitt tungumál sem er sameiginlegt öllum sem henni tilheyrir og aðgreinir hana um leið frá öðrum þjóðum.

Meðal þeirra sem fengu boðsbréfið voru þátttakendur í málstofunni „Foundations of Language Standardization“ (Grundvöllur málstöðl-

unar) sem var hluti 25. norrænu málvísindaráðstefnunnar, sem haldin var í Reykjavík vorið 2013. Málin æxluðust þannig að greinarnar fjórar í þemahlutanum eiga allar rætur að rekja til þeirrar málstofu og byggja á fyrirlestrum sem þar voru haldnir. Fyrir bragðið er alþjóðlegri blær yfir heftinu en oft áður og greinarnar sýna íslensku í samhengi við önnur tungumál, sem er eitt af því sem að var stefnt með breyttu fyrirkomulagi.

Tvær greinanna í þemahlutanum fjalla um íslensku frá ólíkum sjónarhornum og á mismunandi tímasteiðum, hinar tvær um önnur evrópsk málsamfélög. Heimir Freyr van der Feest Viðarsson ræðir í grein sinni, *Þáttur málstöðlunar í afstöðu sagnar til neitunar í 19. aldar íslensku — “málsgreinir, sem mjer fannst eitthvert danskt óbragð að”,* um tvenns konar orðaröð í aukasetningum. Hann gerir grein fyrir rannsóknnum sínum á útbreiðslu og dreifingu afbrigðanna á 19. öld og tengir þróunina við hugmyndafræðilegar hræringar á tímabilinu, sérstaklega viðhorf til slíkra tilbrigða í málinu og tilburði til málstöðlunar. Grein Vanessu Isenmann, *Insight into computer-mediated communication as a new variety of written Icelandic*, fjallar aftur á móti um málþróun í íslensku á allra síðustu árum. Viðfangsefni hennar er málnotkun í netsamskiptum og ýmis frávík frá ríkjandi ritmálsstaðli sem þar má sjá, bæði í rithætti og orðaforða. Hún gerir grein fyrir sérkennum þeirrar málnotkunar sem sjá má í samskiptum á Facebook og skoðar þau bæði í ljósi þeirra viðmiða sem almennt gilda í íslensku ritmáli, t.d. um stafsetningu, og einkenna slíkra samskipta í öðrum málsamfélögum. Þau Heimir og Vanessa eru bæði doktorsnemar við Háskóla Íslands og tengjast greinarnar viðfangsefnum ritgerðanna sem þau eru með í smíðum. Þeir Magnus Breder Birkenes og Jürg Fleischer, frá háskólanum í Marburg í Þýskalandi, skrifa saman greinina *Intetkjønn med referanse til personer. To unlike normeringstradisjoner i nynorsk og tysk*. Þar fjalla þeir um notkun hvorugkyns til þess að vísa til fólks af báðum kynjum eins og gert er í íslensku, t.d. í setningum eins og *Þau komu bæði of seint*. Þetta er hvorki gert í þýsku né nýnorsku en nútímastaðallinn í þeim málsamfélögum á sér ólíka sögu með tilliti til þessa einkennis. Í eldri nýnorsku, sem á rætur í norskum mállýskum, var notað hvorugkyn eins og í íslensku en það hefur horfið vegna breytinga í beygingarkerfi mállýsknanna. Hvorugkyn af þessu tagi varð hins vegar aldrei hluti af þýsku staðalmáli jafnvel þótt það hafi verið notað í þýskum mállýskum því málstöðlunin byggðist í upphafi á öðrum sjónarmiðum en í nýnorsku. Fjórða greinin í þemahlutanum, *The standardization of a modern pluriareal*

language. Concepts and corpus designs for German and beyond, er eftir þá Stephan Elspaß og Konstantin Niehaus frá háskólanum í Salzburg. Þar fjalla þeir um málstöðlun á víðáttumiklum málsvæðum með fleiri en eitt kjarnasvæði eins og t.d. því þýska sem nær til fleiri en eins ríkis og margra málsvæða. Þeir gera grein fyrir gerð málheilda sem endurspeгла slíka svæðisskiptingu og byggjast á fjölbreytilegum textum. Dæmi þeirra sýna að í samanburði við málsöfn sem einungis fela í sér formlega ritmálstexta, ekki síst bókmenntaverk, birta þessar málheildir nokkuð aðra mynd af málstöðlun.

Auk þemagreinnanna eru í heftinu þrjú greinar um ýmis efni. Sú fyrsta þeirra, *Íslenska eða enska í íslensku háskólastarfi* eftir þá Ara Pál Kristinnsson og Harald Bernharðsson, er reyndar ekki alls ótengd greininum í fyrri hlutanum því hún fjallar líka um málstefnu. Munurinn er sá að hér er hugað að stöðu málsins í samfélaginu og gagnvart öðrum málum sem þar eru notuð fremur en búningi þess. Höfundarnir gera grein fyrir athugun sinni á málstefnu íslenskra háskóla með sérstöku tilliti til sambúðar íslensku og ensku og á því hvaða mál eru notuð í raun á ýmsum sviðum háskólastarfsins, einkum sem kennslumál á ólíkum stigum námsins, þ.á m. í doktorsritgerðum, og í skrifum háskólakennara. Í grein sinni, *Beygingarlýsing íslensks nútímamáls: Regluverk eða beygingardæmi*, fjallar Kristín Bjarnadóttir um uppbyggingu og einkenni beygingarlýsingarinnar og gerir grein fyrir margvíslegum vandamálum samfara gerð slíks málsafns. Um leið er vikið að ýmsum einkennum beygingarkerfisins í íslensku nútímamáli og möguleikum sem gagnasafnið býr yfir til rannsókna á því. Síðasta greinin í heftinu fjallar um örnefnið *Þveit* og er eftir Jónínu Hafsteinsdóttur. *Þveit* er sem kunnugt er stöðuvatn í Nesjum í Hornafirði og ekki er vitað til þess að örnefnið finnist annars staðar á Íslandi. Samsvarandi heiti er aftur á móti algengt víða annars staðar á Norðurlöndum. Í greininni rekur Jónína uppruna, aldur og útbreiðslu þessa nafnliðar í Norður- og Vestur-Evrópu og setur fram tilgátu um tilurð nafnsins á Íslandi.

Aftan við greinarnar eru ritfregnir. Þar er fjallað í stuttu máli um ýmis áhugaverð rit á sviði tímaritsins, einkum rit sem komu út á síðasta ári eða skömmu fyrr. Flestar lýsingarnar eru skrifaðar af ritstjóra en Hallgrímur J. Ámundason gerir grein fyrir nýlegum ritum um nafnfræði. Loks eru sagðar fréttir af ráðstefnum. Þar eru bæði lýsingar á nýliðnum ráðstefnum og tilkynningar um ráðstefnur sem haldnar verða á þessu ári og því næsta.

Greinar sem birtast í tímaritinu eru ritrýndar sem þýðir að auk

ritstjóra hafa a.m.k. tveir sérfræðingar lesið hverja grein og lagt mat á hana. Ritrynar njóta nafnleyndar og verða því ekki taldir upp hér en þeim er öllum þakkað fyrir vandaðan yfirlestur og umsagnir. Öðrum sem hafa lagt ritstjóra lið við frágang heftisins er einnig þakkað fyrir sitt framlag. Þar má einkum nefna meðlimi ritnefndarinnar og samstarfsfólk á Stofnun Árna Magnússonar í íslenskum fræðum, ekki síst Bessa Aðalsteinsson sem annast hefur umbrot tímaritsins eins og áður.

Að lokum er vert að minnast Christophers Sanders, ritstjóra við fornorrænu orðabókina í Kaupmannahöfn, sem lést á síðasta ári. Hann átti sæti í ráðgefandi ritnefnd *Orðs og tungu* frá því henni var komið á fót fyrir tveimur árum og hafði áður tekið þátt í málþingum tímaritsins og ritað greinar í það. Blessuð sé minning hans.

Heimir Freyr van der Feest Viðarsson

Þáttur málstöðlunar í afstöðu sagnar til neitunar í 19. aldar íslensku

„málgreinir, sem mjer fannst eitthvert
danskt óbragð að“

1 Inngangur

Á nítjándu öld reið yfir Evrópu málstöðlunarhrina sem venjulega er rakin til pólitísku umróts í álfunni og er talin nátengd þjóðernisstefnu sem átti vaxandi fylgi að fagna (sjá t.d. Kjartan G. Ottósson 1990, Leerssen 1999).¹ Í íslensku samhengi hefur fyrri hluta aldarinnar verið lýst sem „gullöld málhreinsunar“ með aukinni áherslu á formgerðarlega þætti þegar líða tók á öldina (sbr. Kjartan G. Ottósson 1990:50).

¹ Greinin er byggð á erindi á *25th Scandinavian Conference of Linguistics*, 13.–15. maí 2013, fluttu á málstofunni *Foundations of Language Standardization*. Ég þakka þátttakendum í málstofunni og ónafngreindum ritrýnendum gagnlegar athugasemdir. Rannsóknin er hluti af doktorsverkefni mínu innan verkefnisins *Málbreytingar og tilbrigði í íslensku máli á 19. öld: tilurð opinbers málstaðals* sem unnið er í samstarfi fræðimanna við Stofnun Árna Magnússonar í íslenskum fræðum, Háskóla Íslands, Uppsalaháskóla og Vrije Universiteit Brussel. Verkefnið hefur hlotið styrk úr Rannsóknarsjóði og Háskólasjóði Eimskipafélags Íslands.

Thomason (2001:9) telur jafnvel þátt málstöðlunar í því að íslenska hafi breyst fremur lítið ekki síður veigamikinn en meinta einangrun málsins, sem hún dregur í efa. Sjaldnast búa þó ítarlegar rannsóknir á málnotkun að baki slíkum staðhæfingum.

Í þessari grein verður sjónum beint að setningafræðilegu atriði sem fyrst er nefnt í ritdómi Konráðs Gíslasonar í 7. árg. Fjöltnis og felst í afstöðu sagnar í persónuhætti til neitunar í aukasetningum. Þar amast Konráð við neitun á undan persónubeygðri sögn (í aukasetningu með sýnilegu frumlagi) og telur það til danskra áhrifa (1844:85, lítillaga aðlagð):

- (1) a. allt það, er vér **áttum ekki** von á (málviðmið Fjöltnis)
 b. allt það, er vér **ekki áttum** von á (dönskuleg orðaskipun)

Venja er að tala um afstöðu sagnar til neitunar í (1a) sem S2 og í (1b) sem S3 til marks um að sögnin sé í öðru/þriðja sæti setningar (sjá t.d. Þórhall Eyþórsson 1997–98 og Höskuld Þráinsson 2010 um sagnfærslu). Þetta verður einnig gert hér þrátt fyrir ágreining um hvort hliðstæð tilbrigði í nútímamáli felist í breytilegri stöðu sagnar eða neitunar (sbr. Höskuld Þráinsson 2010 og tilv. þar).

Fyrri rannsóknir á S3 í íslensku hafa bent til þess að sú orðaröð færist í aukana eftir 1600 en að hún virðist víkja á síðari helmingi 19. aldar og finnst vart í útgefnum ritum á 20. öld (t.d. Heycock & Wallenberg 2013). Tilgáta flestra sem hafa skrifað um orðaröðina er sú að S3 hafi verið stílfræðilegt ritmálsfyrirbrigði sniðið eftir erlendum málum, einkum dönsku, og hafi því aldrei verið hluti af daglegu máli. Þetta er þó sérkennilegt í ljósi þess að máltengsl hafa almennt verið takmörkuð og ekki síður af því að tilbrigðin finnast (takmarkað) í nútímamáli (sbr. Ásgrím Angantýsson 2011) þar sem S3 er talmálsleg ef eitthvað er.

Fyrir liggur að Fjöltnismenn gerðu tilraun til þess að brennimerkja S3 fyrir miðja öldina og ljóst er að breyting varð á útbreiðslu orðaraðarinnar í útgefnum textum. Eftir er hins vegar að gera fullnægjandi grein fyrir breytingunni og að leggja mat á það hvort meðvitund um þetta brennimerkta afbrigði í orðaröð hafi verið nægjanleg til þess að ný málviðmið næðu almennt eyrum málnotenda. Inn í þessa umræðu fléttast einnig meintur munur á lærdum stíl og alþýðustíl, þar sem mál „alþýðu“ á að hafa verið svo að segja ósnortið af erlendum áhrifum á mál og stíl (sjá t.d. Jakob Jóh. Smára 1920:9–15). Slíkar hugmyndir er nauðsynlegt að setja í sögulegt samhengi.

Í greininni verður leitað svara við eftirfarandi spurningum:

- Hversu útbreidd var notkun S3 í íslensku máli á 19. öld?
- Hvenær varð annað afbrigðið (þ.e. S2) hluti nýs málstaðals og hver urðu áhrif þess?

Rannsóknin hefur almennt fræðilegt gildi að því er varðar möguleika á að snúa við málbreytingum, sem oft er talið illmögulegt (t.d. Labov 2001), og sértækt setningafræðilegt gildi í tengslum við hugmyndir um að ríkulega sagnbeygð tungumál eins og íslenska leyfi ekki S3 (sjá t.d. Höskuld Þráinsson 2010, Ásgrím Angantýsson 2011, Heycock & Wallenberg 2013). Því miður er ekki rúm til að gera setningafræðilegu hliðinni fullnægjandi skil hér og fær sú umfjöllun því að bíða betri tíma.

Bygging greinarinnar er sem hér segir: Í öðrum kafla verður stiklað á stóru í umfjöllun um meðvitund málnotenda almennt um tilbrigði í máli og möguleikum á „handvirkum“ breytingum í tengslum við málstöðlun, sem svo er sett í sögulegt samhengi. Í þriðja kafla verður stuttlega fjallað um fyrri athuganir, þeim textasöfnum lýst sem notuð verða í nýrri athugun og málbreytan afmörkuð. Í fjórða kafla verða fyrstu niðurstöður textaathugunar birtar, annars vegar á útgefnum blaða- og tímaritsgreinum og hins vegar persónulegum sendibréfum frá 19. öld, þar sem kemur í ljós að meiri breytileika er að finna en áður var talið. Greininni lýkur á umræðu um þátt málstöðlunar í breytingum á notkun S3 en nýr vitnisburður úr ólíkum áttum bendir til að meðvitund um umrædd tilbrigði í orðaröð takmarkist við síðasta fjórðung 19. aldar, seinna en við var að búast.

2 Fræðilegur grundvöllur

2.1 Meðvitund um mál og tilbrigði

Meðvitund um tilbrigði í máli er einn af þeim grundvallarþáttum sem taldir eru hafa áhrif á þróun málbreytinga innan þeirrar félagsmálfræðilegu rannsóknarhefðar sem kennd er við bandaríska fræðimanninn William Labov. Hann gerir greinarmun á breytingum sem breiðast út án þess að það sé almennt vitund um þær í samfélaginu (breytingar „neðan frá“) og breytingum sem almennir málnotendur vita af í þeim skilningi að þeir bera kennsl á þau tilbrigði sem af þeim spretta (breytingar „ofan frá“) (Labov 1972, 2001). Breytingar ofan

frá eiga það sameiginlegt að gjarnan er reynt að sporna gegn þeim, t.d. af hálfu opinberra stofnana, á meðan breytingar neðan frá koma samfélaginu að mestu að óvörum.²

Málstöðlun felur í sér ákvarðanir opinberra aðila eða áhrifamanna í málsamfélaginu um val á viðmiðum. Eigi hún að skila árangri þurfa málnotendur almennt að þekkja viðmiðin og vera færir um að fylgja þeim. Þótt einstaklingar hafi yfirleitt takmörkuð áhrif á aðra er talið að svonefndir ofurpaurar (e. *mega-actors*) geti haft bein áhrif á grundvallarþætti félagslífs (sjá Deumert 2003:53 o.áfr.). Því kemur ekki á óvart að í umræðu um áhrif málstöðlunar sé lögð áhersla á hlutverk menntastofnana, einkum eftir 1800 í tengslum við þjóðernisstefnu í Evrópu (sjá hér á eftir), og gert er ráð fyrir að áhrif slíkra stofnana hafi verið býsna mikil, a.m.k. til lengri tíma lítið (sbr. Deumert 2003:66–67).

Mikilvæg spurning sem þetta vekur er hvort takmarkanir eru á hve víðtæk bein „inngríp“ geti verið. Það er ekki síst áhugavert í ljósi þess að málnotendur eru ekki endilega meðvitaðir um eigin málnotkun (sjá t.d. Höskuld Þráinsson o.fl. 2010:37, nmgr. 12) og gera sér ekki alltaf fullkomlega ljóst í hverju brennimerkt fyrirbæri, t.d. „þágufalls-sýki“ (sbr. tilv. rit:59–60), raunverulega felast. Ólík svið málfræðinnar eru ennfremur talin ólík að þessu leyti; setningafræðileg atriði lenda t.a.m. neðst á skala Laycocks & Mühlhäuslers (1990:849) yfir möguleika á „handvirkum“ breytingum á tungumálum. Einnig er oft bent á að málfarsleg tilmæli beri jafnan lítinn árangur ef marka má raunverulega málnotkun og þau þjóni því litlum tilgangi (t.d. Labov 2001:4, Poplack & Dion 2009). Þrátt fyrir það hafa fræðimenn talið að málstöðlun eigi sinn þátt í að íslenska hafi breyst minna en mörg önnur mál:

The archaic features of Standard Icelandic are no doubt due in part to the relatively great degree of isolation from other

² Líkan Labovs (1972:237–251) gerir ráð fyrir að félagsmálfræðilegar breytur séu af þrennu tagi: vísar (e. *indicators*), merki (e. *markers*) og staðalímyndir (e. *stereotypes*). Vísar eru mállýskubundin tilbrigði sem málnotendur eru sér ekki meðvitaðir um. Þeir eru því breytingar að neðan. Sé nægileg vitund um tilbrigði í samfélaginu til þess að framkalla stílbundinn breytileika teljast þau merki eða staðalímyndir. Munurinn á þessu tvennu felst í því að þótt málnotendur séu að einhverju leyti meðvitaðir um marka geta þeir ekki lýst fyrirbærinu, en staðalímyndir eru fullkomlega meðvitaðar og hægt er að skírskota til þeirra í umræðu um mál og stíl. Staðalímyndir eru skýr dæmi um breytingar ofan frá.

languages, but they are also in part a result of the process by which the language was standardized: the creators of Standard Icelandic deliberately archaized the language's structure, making it look older so as to bring it closer to the language of the Eddas (Thomason 2001:9)

Sú staðhæfing að stöðlun íslenskunnar hafi falið í sér meðvitaða fyrningu á formgerð málsins er fremur djörf. Þekkt íslenskt dæmi um meinta málstýringu sem tengist setningafræðinni er t.d. að notkun lausa greinisins *hinn* í stað *sá* sé til marks um áhrif (staðlaðs/fyrnds) ritmáls á talmál, því að lausi greinirinn *hinn* muni hafa verið horfinn úr málinu, a.m.k. með nafnorðum.³

Rétt er að hafa í huga að ekki er sjálfgefið að sögulegar breytingar sem birtast í varðveittum textum endurspegli beinlínis málfræðilegar breytingar innan málsamfélagsins. Þær geta einnig komið til af stílfræðilegum breytingum, t.d. ef ritmálsleg viðmið færast nær mæltu máli (sjá t.d. Mair 2009:88–89). Sumt af því sem sagt hefur verið um íslenska málhreinsun ber einmitt frekar keim af stílfræðilegum breytingum en eiginlegum málbreytingum, sbr. orð Jakobs Jóh. Smára (1920:14) að „endurreisn“ málsins hafi byggt að verulegu leyti á alþýðumáli. Hér er þó að ýmsu að hyggja og ljóst að umræða um fornlegt alþýðumál 19. aldar var hápólitísk, eins og nú verður stuttlega vikið að.

2.2 Málhreinsun í sögulegu samhengi

Málhreinsun 19. aldar tengist svonefndum þjóðlegum þankagangi (e. *national thought*), þjóðarvitund skilgreindri út frá uppruna, landmærum, menningu, tungumáli o.s.frv., sem um 1800 varð að pólitískri hugmyndafræði (sbr. Leerssen 1999:9–13). Af aukinni miðstýringu einveldanna, þéttriðnara neti ríkja, vaxandi áhuga á hefðum og svæðisbundnu sjálfræði leiddi sá hugsunarháttur að þjóðir hefðu sitt eigið, einstaka eðli og ættu því tilvistarrétt sem sjálfstæð ríki (sjá einnig t.d. Guðmund Hálfðanarson 2007:192–193, 197–198).

Eins og ítarlega er fjallað um í ofangreindum ritum léku tungumál mikilvægt hlutverk innan þessarar hugmyndafræði því að tilkall til sjálfstæðis byggðist m.a. á því að sýna mætti fram á tilvist tungumáls

³ Sjá t.d. Kjartan Ottosson (2003:128–130), þar sem kemur fram að *hinn* taki við af *sá* á 19. öld beinlínis fyrir tilstilli málhreinsunar, svo að *sá* einskorðast fyrst við talmálið og hörfar þar svo smám saman líka. Óljóst er þó á hvaða vitnisburði sú staðhæfing er byggð.

sem átti sér merka sögu. Upp hófst leit að fornum rithefðum sem víða gat af sér hreintungustefnu. Skýr greinarmunur var gerður á tungumálum og mállýskum og þurfti hvert tungumál að hafa ritmálsstaðal.⁴ Einnig var mikilvægt að geta sýnt fram á að ritmálið ætti sér einhverja stoð í samtímanum á meðal alþýðu manna og litið var á þjóð og tungumál sem einingu, þar sem tungumálið væri ein helsta birtingarmynd þjóðareðlisins (sbr. Leerssen 1999:103–107). Órofin tengsl forníslensku og alþýðumáls samtímans fólu því í sér beinan pólitískan ávinning fyrir land og þjóð og ummælum um slíkt ber að taka með miklum fyrirvara.

3 Afmörkun og aðferðafræði

3.1 Fyrri athuganir

Í sögulegri umfjöllun um S3-orðaröð í íslensku er jafnan gert ráð fyrir að tímabilið 1600–1850 skeri sig úr vegna aukinna erlendra áhrifa (sbr. t.d. Jakob Jóh. Smára 1920:13, Heycock & Wallenberg 2013). Heycock & Wallenberg (2013) athuguðu stöðu persónubeygðrar sagnar í auka-setningum með neitun og öðrum miðlægum⁵ atviksorðum í íslenskum frásagnartextum frá 12. öld til 21. aldar. Þau komust að því að á umræddu tímabili hafi S3 tíðkast með allt að 40% hlutfallstíðni í ákveðnum gerðum setninga, en síðan horfið nær algerlega. Þau vekja athygli á að á 17. og 18. öld komi S3 í íslensku einmitt helst fyrir í

⁴ Aðgreining milli tungumáls og mállýsku kemur skýrt fram hjá Rask (1818:278–279) í umfjöllun hans um íslensku andspænis færeysku. Eins og skandinavísku meginlandsmálin er íslenska að hans mati tungumál („språket på sjelfva Island“) en aftur á móti talar hann um færeysku sem mállýsku („en folkdialekt“).

⁵ Mismunandi ao. hegða sér á ólíka vegu (sjá t.d. Maling 1980, Höskuld Práinsson 2010, Ásgrím Angantýsson 2011 og tilv. í þessum ritum). Staða sumra ao. er mjög frjálst en annarra fastskorðaðri, t.d. *náttúrulega* andspænis *alltaf* í aðalsetningum (< > táknar breytilega stöðu, * að tiltekin staða sé ótæk):

(i) Jón <náttúrulega> les <náttúrulega> fyrir svefninn <náttúrulega>

(ii) Jón *<alltaf> les <alltaf> fyrir svefninn *<alltaf>

Í (ii) sést að *alltaf* verður að standa strax á eftir persónubeygðri sögn, ólíkt *náttúrulega* í (i). Atviksorð eins og *alltaf* eru þess vegna kölluð miðlæg og í þann flokk falla m.a. tíðaratviksorð og neitun. Vegna fastrar stöðu sinnar veita þau traustari vísbendingar um setningafræðilega stöðu annarra hluta setningarinnar.

umhverfi þar sem sú orðaröð er ráðandi í skandinavísku meginlandsmálunum, t.d. í tilvísunarsetningum eins og í (2), en síst í setningum þar sem S2 gengur hvað best í þessum málum, t.d. í skýringarsetningum eins og (3).⁶

- (2) og frá þeim tíma komst eg í vinfengi hans er hann **ei sleit** meðan hann lifði (18. öld, dæmigert S3-umhverfi)
- (3) Ekki er það nóg, að vér **ei drygjum** þá glæpi, sem vinna til útsetningar af kirkjunni (18. öld, dæmigert S2-umhverfi)

Dreifing eftir setningafræðilegu umhverfi líkist því málkunnáttu málhafa sem hefði dönsku að móðurmáli. Heycock & Wallenberg (2013) bjóða ekki upp á neina endanlega túlkun á niðurstöðum sínum en telja sennilegt að S3 hafi verið ritmálsfyrirbæri meðal menntamanna, e.t.v. sniðið meðvitað að dönsku, en hafi aldrei haft áhrif í töluðu máli.⁷ Eftir 1850 dragi verulega úr notkun S3 og á 20. og 21. öld eru nær engin dæmi um orðaröðina í þeim textum sem voru athugaðir. Eldri athugun á dagblaðatextum frá lokum 20. aldar sýndi einnig að S2-orðaröð, hliðstæð (1a), væri einráð (sjá Höskuld Þráinsson 2010:1075, nmgr. 19). Þá hefur jafnvel verið gert ráð fyrir að S3 með neitun komi alls ekki fyrir í nútímaíslensku (t.d. Maling 1980:177, nmgr. 4).

Ekki er ljóst af fyrri umræðu hvort S3 með miðlægum atviksorðum í aukasetningum hafi fyrst komið upp um 1600, t.d. fyrir erlend áhrif, eða hvort sú orðaröð hafi verið möguleg þegar í fornmáli, síðar orðið algengari um tíma – komist í tísku, ef svo má segja – þar til S3 þótti af einhverjum ástæðum ekki lengur eftirsóknarverð eftir 1850. Athugun Heycock & Wallenbergs (2013) gefur til kynna að í elsta máli finnist örfá dæmi um S3 í umhverfi hliðstæðu (2) að ofan og fleiri dæmi af sama tagi á 15. öld. S3 virðist því ekki beinlínis vera nýjung í íslensku á 17. öld; möguleikinn hafi verið fyrir hendi í eldra máli þó að hann hafi e.t.v. verið sjaldnar nýttur.⁸

⁶ Heycock & Wallenberg (2013) sýna ekki dæmi en (2) og (3) eru úr sömu málheild og þau notuðu; fyrra dæmið úr ævisögu Jóns Steingrímssonar, setn. nr. „138.1081“, það síðara úr Vídalínspostillu, setn. nr. „1001“ (sjá Wallenberg o.fl. 2013 og tilv. þar).

⁷ Í dönsku er S3 almennt talin sjálfgefin orðaröð í aukasetningum. Sú röð tók jafnt og þétt við af S2 en S3 varð ekki ríkjandi fyrir en undir lok 17. aldar (sbr. Sundquist 2003:240).

⁸ Einmitt slík léttvæg formgerðarleg áhrif (breytingar á mynstrum/tíðni) telur Thomason (2001:70-71) leiða af vægum erlendum áhrifum, sem fela einnig í sér sérhæfð inntaksorð og kerfisorð, þ.e. annað stig af fjórum á

Færð hafa verið rök fyrir því að í nútímaíslensku sé S3 á einhvern hátt „afbrigðileg“ orðaröð og hana sé helst að finna í setningum þar sem dregnar eru upp andstæður til áherslu, einkum í tilvísunarsetningum, sbr. (4). Annars sé notkun S3 yfirleitt mjög takmörkuð, t.d. í skýringarsetningum, sbr. (5) (sjá ítarlega umfjöllun hjá Ásgrími Angantýssyni 2001, 2011:64–72).

- (4) Jón sagði margt áhugavert en áhugaverðast var það sem hann **ekki sagði**
- (5) *María hélt að hann **ekki segði** sannleikann

Eins og rakið er í mörgum þeim ritum sem vitnað er til hér að ofan hefur sú tilgáta verið sett fram að „afbrigðileiki“ S3 í íslensku sé engin tilviljun og að ríkuleg sagnbeyging í íslensku, ólíkt t.d. dönsku, leiði til þess að S3 gæti aldrei orðið hin ómarkaða orðaröð. Þetta er þó ekki óumdeilt og samræmist fremur illa háu hlutfalli S3 í sumum textum. Slík tengsl beygingarfræðilegra og setningafræðilegra atriða kunna þó að skýra það að S3 hafi ekki steipt S2 af stóli (sbr. Heycock & Wallenberg 2013) en því fer fjarri að fullreynt sé að það leiði ekki af almennari þáttum í máltöku.

3.2 Textasöfn

Fyrri rannsóknir hafa einkum miðast við formlega texta sem margir standa býsna fjarri daglegu máli (sjá t.d. Elspaß 2012). Nýrri athugun á safni sendibréfa og úrvali blaða- og tímaritstexta er ætlað að bæta hér úr að nokkru með samanburði á ólíkum textagerðum, sbr. (6) og (7).

- (6) Blöð og tímarit (1800–1850, 1875, 1900); um 450 þúsund lesmálsorð
- (7) Safn sendibréfa frá 19. öld; um 900 þúsund lesmálsorð

Vert er að hafa í huga að athugunin er enn á frumstigum, m.a. vegna þess að textasafnið er aðeins tilbúið að hluta og enn vantar bakgrunnsupplýsingar um marga málhafa. Bládatextarnir eru fengnir úr safni Landsbókasafns–Háskólabókasafns á *Tímarit.is*.⁹ Í þessari fyrstu

styrkleikaskala hennar. Ætla má að þetta eigi við um mál menntamanna á umræddu tímabili, 1600–1850, þó að máltengsl hafi e.t.v. verið minni almennt séð. Ekki er þó hægt að útiloka sterkari áhrif máltengsla á alþýðu í formi keðjuverkunar í gegnum menntamenn og ritað mál.

⁹ Ég þakka Ástu Svavarsdóttur fyrir aðgang að þessum textum á rafrænu formi. Um er að ræða leiðréttan texta tímarita úr stafrænu safni Landsbókasafns Íslands – Háskólabókasafns (sjá töflu 7 í viðauka). Einnig lagði

atrennu er einungis notast við afmarkað úrval frá ólíkum tímum en ætlunin er að síðar verði fleiri textar athugaðir yfir samfellt tímabil. Málnotkun í blöðum og tímaritum er fremur formleg og textarnir eru einkum skrifaðir af menntamönnum. Því er líklegt að þeir endurspegli að verulegu leyti opinber (eða hálf-opinber) málviðmið á 19. öldinni.

Bréfin eru flest skrifuð á 19. öld af körlum og konum, yfirleitt óskólagengnum, frá ólíkum landshlutum (sjá Harald Bernharðsson & Jóhannes Gísla Jónsson 2012). Enn er verið að bæta við safnið en þar eru nú alls 1639 bréf, þar af 1416 bréf frá 19. öld. Frá ofanverðri 19. öld og upphafi þeirrar 20. er einnig að finna bréf Vestur-Íslendinga.

3.3 Afmörkun málbreytunnar

Fyrri rannsóknir á S3 í íslensku hafa tekið mið af umhverfi þar sem sýnilegt frumlag stendur fremst í aukasetningu og neitun eða miðlægt atviksorð fer á undan persónubeygðri sögn. Til þess að einfalda sam-
anburð verður hér aðeins miðað við afstöðu persónubeygðrar sagnar og neitunar, sbr. (8).

- (8) a. [_{Aðalsetn.} [_{Aukasetn.} frumlag – sögn_{persb.} – neitun]] S2
 b. [_{Aðalsetn.} [_{Aukasetn.} frumlag – neitun – sögn_{persb.}]] S3

Notað var forritið AntConc (Anthony 2012) til þess að sækja setningar með neitun (*ekki, ei, eigi* o.s.frv.) og dæmin voru síðan greind handvirk. Til þess að finna allar ritmyndir í textasöfnunum (t.d. *eckji, ekge, eii, ej*) var AntConc látið útbúa vélrænan orðalista sem farið var yfir handvirk.

Ofangreind afmörkun er venjulega réttlætt með því að breytileiki milli S2 og S3 komi aðeins fram ef frumlag er fremst. Nauðsynlegt er hins vegar að réttlæta þá afmörkun á forsendum 19. aldar. Það kemur nefnilega á daginn að afmörkun við frumlag (í aukasetningu) er e.t.v. ekki jafn sjálfgefin og ætla mætti að óathuguðu máli. Í (9)–(10) eru sýnd frávík frá umræddri reglu en þau eru fátíð:¹⁰

ég til inngangsorð allra árganga Ármanns á Alþingi.

¹⁰ Vitnað er til sendibréfa með upplýsingum um ritunartíma, kyn og aldur bréfitara. Við tilvitnun í tímarit er aðeins skráð útgáfuár.

- (9) a. og ætlar því annar þeirra að ríða norður að Miklabæ til að semja við Arna til næsta árs og menn **ekki hika** við að bjóða honum 4–500 kr. í kaup (1879, kk., 36)
- b. [...] er jeg orðin mjög efnalítill, og vegna elli lasleika **ekki get** lifað án þess að þiggja styrk af öðrum (1881, kvk., 70)
- (10) a. mjer finnst hreint að hjer **ekki hafi** verið neinn vetur [...] (1887, kvk., 24)
- b. það er huggun og von mín að sá sem í öðrum heymi, **ekki lætur** einn vatns drók, ólaunaðann. verða, hann [...] (1885, kvk., ?)
- c. Það er merkilegt, að Hertoginn af Würtemberg var sá einasti Fursti í öllum heimi, sem þann 10da Maí 1802 **ecki hafði** samid fullkominn frid vid Bounaparte. (1803)

Í (9) fer neitun á undan sögn í aðalsetningu þrátt fyrir að fyrsta sæti sé fyllt, af frumlagi í (9a) en forsetningarlið í (9b). Í (10) eru hliðstæð frávik í aukasetningu þar sem liðir aðrir en frumlög fylla fyrsta sætið. Það hefði mikil áhrif á heildarniðurstöðurnar ef breytan yrði skilgreind óháð fremsta lið vegna þess að yfirgnæfandi meirihluti þessara dæma er með S2-röð, eins og vænta má samkvæmt nútímamáli.

Alþekkt er í sögulegri málfræði að stundum verður að líta framhjá dæmum sem finnast og gera ráð fyrir að þau séu beinlínis röng, þ.e. ótækar „raunverulegar“ setningar (sbr. t.d. Eirík Rögnvaldsson 1998). Fáein dæmi um óvenjulega orðaröð eru því ekki endilega nægjanleg ástæða til þess að varpa rótgrónum greiningum fyrir róða. Annar möguleiki er að líta á dæmin í (9)–(10) sem vitnisburð um frjálsari stöðu neitunar hærra í formgerðinni en oft er talið. Þar með veittu þau nýja röksemd fyrir þeirri greiningu Höskuldar Þráinssonar (2010), meðal annarra, að S3 í íslensku felist ekki í breytilegri stöðu sagnar heldur atviksorðs.

4 Niðurstöður athugunar

4.1 Blöð og tímarit (1800–1850, 1875, 1900)

Það úrval blaða- og tímarita sem notast var við býður upp á þann möguleika að skoða þróun í opinberri/formlegri málnotkun á 19. öld. Í efni af þessu tagi er helst von á að finna breytingar sem mætti túlka sem áhrif breyttra málviðmiða eða nýs málstaðals. Með hliðsjón af athugun Heycock & Wallenbergs (2013) á frásagnartextum er þess að

vænta að hlutfall S3-orðaraðar sé hærra á fyrri hluta aldarinnar en eftir 1850. Þessi spá gengur að miklu leyti eftir, sbr. töflu 1.

Tímabil	S2	S3	Hlutfall S3
f.hl. 19. aldar	225	178	44,2%
1875	161	28	14,8%
1900	350	41	10,5%

Tafla 1. Breytan S2/S3 í blöðum og tímaritum, 19. öld

Fyrir 1850 er hlutfall S3 hátt og í sumum textum svo hátt (hæst tæp 78%) að ekki er augljóst hvort bæri að telja S2 eða S3 hina ómörkuðu orðaröð (sjá viðauka, töflu 7). Árin 1875 og 1900 hefur hlutfall S2 aukist mjög á kostnað S3, þó að í sumum tímaritum sé hlutfall S3 enn á bilinu 20–35%. Munurinn á 1800–1850 og 1875 er tölfræðilega marktækur en ekki munurinn á 1875 og 1900.¹¹ Þessar niðurstöður þarf auðvitað ekki að túlka svo að eiginleg málbreyting hafi orðið, t.d. ef stílfræðileg viðmið ritmáls hafa færst nær mæltu máli. Þá er ekki óhugsandi að málstýring ritstjóra á einstökum blöðum skipti máli; sumir ritstjórar á síðari hluta 19. aldar munu hafa verið miklir málhreinsumarmenn (sbr. kafla 5). Þetta þarf hvort tveggja að kanna nánar og á stærra safni en hér liggur til grundvallar.

Til þess að fræðast nánar um eðli breytileikans er forvitnilegt að skoða hlutfall S3 í ólíku setningafræðilegu umhverfi. Tafla 2 sýnir dreifingu eftir tegund aukasetningar og sýnd eru dæmi um hverja gerð í (11) og (12), annars vegar S2 og hins vegar S3.

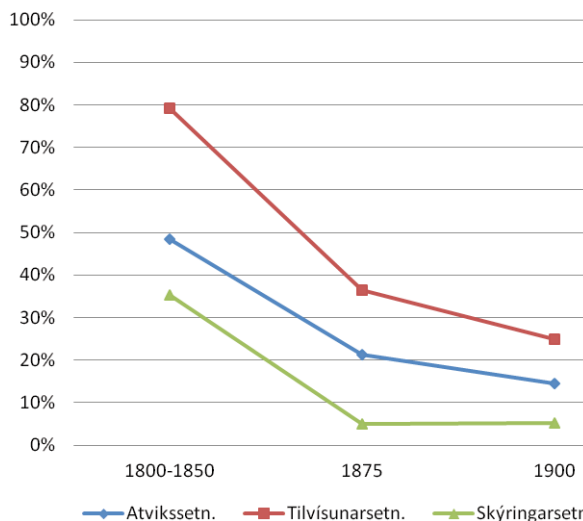
Tegund aukasetningar	1800–1850	1875	1900
Atvikssetningar	48,5% (94/194)	21,3% (20/94)	14,6% (23/158)
Tilvísunarsetningar	79,3% (23/29)	36,4% (4/11)	25,0% (6/24)
Skýringarsetningar	35,4% (58/164)	4,9% (4/81)	5,3% (11/207)
Spurnarsetningar	50,0% (3/6)	0,0% (0/3)	50,0% (1/2)

Tafla 2. Hlutfall S3 af heildardæmafjölda eftir tegund aukasetningar.

¹¹ Kí-kvaðratspróf Pearsons sýnir að tölfræðilega marktækur munur er á f.hl. 19. aldar og 1875, p-gildi = 0,000000000005291. Á 1875 og 1900 er munurinn ekki tölfræðilega marktækur, p-gildi = 0,1699.

- (11) a. [...] reyna síðan af alhuga að bæta hag manna hér á landi, svo að hann **verði ekki** lakari en í öðrum síðuðum löndum (1900)
 b. Margir dattu um grjótid, sem vid **sáum ekki** í náttmyrkrinu (1835)
 c. blöð þeirra hrópa hástöfum til kjósendanna að þeir **skuli ekki** þola þá fulltrúa, sem ekki fylgi flokk þeirra blindandi (1875)
 d. það er eptir að vita hvort þær hinar fáu skepnur er menn setja á vetur í haust, **drepast ekki** úr beinasjúkdómum eða öðrum kvillum (1875)
- (12) a. Borgfirðingar hafa í vor verið að bæta þeim ær, er skáru þær í vetur, svo þeir **eigi yrðu** málnytlaus í sumar (1875)
 b. Ýmsar aðrar ræður voru þar fluttar, sem vjer **eigi finnum** sjerstaklega ástæðu til að minnst á (1875)
 c. þegar þjóðin sjálf er farin að virða og elska módurmál sitt, þá má fullyrða að það **ekki muni** verða án vitnisburðar (1846)
 d. gétur verid að Nordurland hafi haft gott af þessu, hvört **sóttin ekki** var komin á undan þessari ritgjörð (1835)

1800–1850 er hlutfall S3 áberandi hátt í tilvísunarsetningum og í spurnarsetningum (sem aðeins örfá dæmi fundust um) en þar leyfist S3 einmitt helst í nútímamáli. Þó að dragi úr tíðninni er hlutfall S3 áfram hæst í tilvísunarsetningum 1875 og 1900. Myndræn framsetning á þessum upplýsingum sýnir hve stöðugt lækkandi hlutfall S3 er innan hverrar tegundar aukasetningar yfir þessi þrjú tímabil, sbr. mynd 1.



Mynd 1. Hlutfall S3 af heildardæmafjölda eftir tegund aukasetningar (sbr. töflu 2).

Til þess að kanna þá staðhæfingu að skilyrði fyrir notkun S3 fram til 1850 séu svipuð og hjá málhafa sem hefði dönsku að móðurmáli (sbr. kafla 3.1) er forvitnilegt að líta á dreifingu S3 eftir því umhverfi þar sem S3 er algengust í skandinavísku meginlandsmálunum og umhverfi sem leyfir S2 greiðlega.¹² Þetta er sýnt í töflu 3.

Tegund umhverfis	1800–1850	1875	1900
Dæmigert S2	35,3% (65/184)	5,1% (5/98)	5,2% (12/232)
Dæmigert S3	57,8% (26/45)	28,6% (4/14)	26,9% (7/26)
Blandað umhverfi	50,0% (87/174)	24,7% (19/77)	16,5% (22/133)

Tafla 3. Hlutfall S3 af heildardæmafjölda eftir kjörumhverfi í skandinavísku meginlandsmálunum.

Mínar niðurstöður eru hliðstæðar niðurstöðum Heycock & Wallenbergs (2013), þ.e. hlutfall S3 er hæst í dæmigerðu S3-umhverfi, lægst í dæmigerðu S2-umhverfi og blandað umhverfi lendir þarna á milli (líkast dæmigerðu S3-umhverfi). Á fyrri hluta aldarinnar er munurinn tölfræðilega marktækur.¹³ Þetta afstaða hlutfall helst á milli tímabila þrátt fyrir að mjög dragi úr heildartíðni S3; m.ö.o. virðist ekki verða eðlisbreyting á notkuninni.¹⁴

Stóra spurningin sem niðurstöðurnar vekja er hvort lág tíðni S3 á síðari hluta aldarinnar endurspeglir daglegt mál þorra landsmanna eða hvort tíðnin sé „óeðlilega“ lág í þeim tímaritum þar sem hún er lægst, e.t.v. til marks um að málstöðlun sé þar að verki. Til þess að varpa ljósi á þessa spurningu skal nú gerð tilraun til þess að komast nær notkun S3 í textum sem ætla má að standi nær mæltu máli.

¹² Ég fylgi skilgreiningum hjá Heycock & Wallenberg (2013) og kalla umhverfi blandað sem ekki er greinilegt kjörumhverfi S2 eða S3. Þau telja dæmigert S3-umhverfi vera samanburðarsetningar, tilvísunarsetningar og óbeinar spurnarsetningar, dæmigert S2-umhverfi hins vegar skýringarsetningar og stigsetningar (*svo X að ...*, þar sem X er oft lo.) og blandað umhverfi allar aðrar atvikssetningar. Blandað umhverfi flokka þau þó með S3-umhverfi í greiningu á skandinavísku málunum.

¹³ Kí-kvaðratspróf Pearsons sýnir að á f.hl. 19. aldar er p -gildið = 0,003008. Væntingartíðni S3 í dæmigerðu S3-umhverfi árin 1875 og 1900 er (mun) lægri en 5 og heimilar því ekki fyrrnefnt marktækniþróf.

¹⁴ Umdeilt er hvað megi ráða af breytingum á tíðni, ekki síst hvort túlka beri það sem breytingar á málkunnáttu ef notkun setningafræðilegrar formgerðar eykst eða minnkar eða hvort það sýni e.t.v. fremur sveigjanleika í málbeitingu, ef fyrirbærið kemur fyrir á annað borð (sjá t.d. Margréti Guðmundsdóttur 2008:30–32).

4.2 Sendibréf á 19. öld

Ólíkt opinberu og tiltölulega formlegu máli blaða- og tímaritstexta er þess að vænta að almennt sýni (persónuleg) sendibréf ýmis málfarsleg einkenni sem staðlaðir og ritstýrðir textar þegja um (sbr. Elspaß 2012). Til þess að meta áhrif málstöðlunar er því nauðsynlegt að bera slíka texta saman við útgefin rit sem líklegri eru til að fylgja ákveðnum viðmiðum.

Sláandi munur kemur fram á notkun S3 þegar sendibréf eru borin saman við textana sem lýst var hér á undan. Í bréfunum verður lítil breyting á hlutfallslegri tíðni S3 alla öldina: hlutfallið er um 10%, sbr. töflu 4.¹⁵ Mikilvægt er þó að hafa í huga að bréfasafnið inniheldur enn sem komið er mjög fá bréf frá fyrri hluta aldarinnar.

Tímabil	S2	S3	Hlutfall S3
1800–1850	199	18	8,3%
1850–1875	1118	130	10,4%
1875–1900	964	126	11,6%

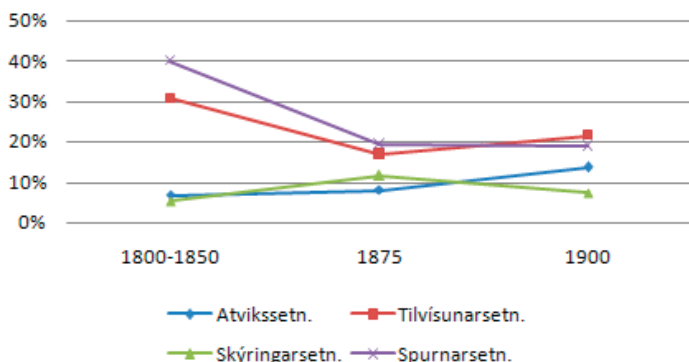
Tafla 4. Breytan S2/S3 í sendibréfum á 19. öld

Dreifing eftir setningafræðilegu umhverfi, sbr. töflu 5 og mynd 2, sýnir að skilyrði fyrir S3 eru talsvert ólík því sem gerist í nútímamáli. Eins og í blöðum og tímaritum er S3 tíðust í tilvísunar- og spurnarsetningum, líkt og nú er. S3 er hins vegar síst að finna í atviks- og skýringarsetningum, þó að hlutfall S3 í skýringarsetningum sé býsna hátt á tímabilinu 1850–1875, eða tæp 12%, í ljósi þess hve illa S3 gengur þar í nútímamáli. Dæmi eru sýnd í (13)–(14).

Tegund aukasetningar	1800–1850	1850–1875	1875–1900
Atvikssetningar	6,7% (7/104)	8,0% (51/641)	13,6% (71/523)
Tilvísunarsetningar	30,8% (4/13)	16,9% (12/71)	21,5 (14/65)
Skýringarsetningar	5,3% (5/95)	11,6% (55/474)	7,3% (34/465)
Spurnarsetningar	40,0% (2/5)	19,4% (12/62)	18,9% (7/37)

Tafla 5. Hlutfall S3 af heildardæmafjölda eftir tegund aukasetningar.

¹⁵ Munurinn á tímabilunum þremur er tölfræðilega ómarktækur, þ.e. dreifingin bendir ekki til þess að breyting hafi orðið; kí-kvaðratspróf Pearsons á töflu 4 sýnir að p -gildi = 0,3238. Munur innan hvers tímabils er einnig fjarri því að vera tölfræðilega marktækur.



Mynd 2. Hlutfall S3 af heildardæmafjölda eftir tegund aukasetningar (sbr. töflu 5).

- (13) a. Brjefið til Kristjáns er þjer sent svo hann **geti ekki** þrætt fyrir að hafa fengið það (1890, kk., 35)
- b. það má þó margt læra hér sem maður **þekkir ekki** heima (1899, kvk., 20)
- c. Jeg bid þig bezti pabbi að fyrirgefa mér að eg **skrifadi** þér **ekki** í vetur (1889, kvk., 19)
- d. Þú spyrð mig í brjefinu þínu hvert mig **langi ekki** heim aptur ef þú gætir útvegað mjer atvinnu sem eg gæti lifað af (1875, kk., 19)
- (14) a. en stódu utan um og forsigladu hana svo kallinn **ekki gėti** skèmt hana (1857, kk., 19)
- b. og var sagt að posturin hefði tekið hana sjálfur og sent með einhverjum sem hann **ekki þegti** (1868, kvk., ?)
- c. hún biður að fyrir gefa að hún **ekki skrifi** (1887, kvk., 25)
- d. þú biður mig að seigja þjer í einlægni hvert mig **ekki lángi** heim aptur (1876, kk., 36)

Notkun S3 í sendibréfum í kjörumhverfi S2/S3 í skandinavísku meginlandsmálunum sýnir nú gerólíka dreifingu, sbr. töflu 6 (í samanburði við töflu 3 hér á undan). Hlutfall S3 er hæst í blönduðu umhverfi og á tímabilinu 1875–1900 er hlutfallið einmitt **lægst** í dæmigerðu S3-umhverfi. Þessi munur er tölfræðilega marktækur á síðari hluta aldarinnar.¹⁶ Því virðast aðrar hömlur vera að verki; líkindin eru yfirborðsleg.

¹⁶ Á f.hl. 19. aldar er væntingartíðni S3 í blönduðu umhverfi of lág fyrir kí-kvaðratspróf Pearsons en dreifingin er marktæk 1850–1875 (p-gildi = 0,00166) og 1875–1900 (p-gildi = 0,0001294).

Tegund umhverfis	1800–1850	1850–1875	1875–1900
Dæmigert S2	6,2% (6/97)	7,4% (41/553)	14,1% (65/461)
Dæmigert S3	5,9% (6/102)	11,8% (65/553)	7,6% (39/516)
Blandað umhverfi	33,3% (6/18)	16,9% (24/142)	19,5% (22/113)

Tafla 6. Hlutfall S3 af heildardæmafjölda eftir kjörumhverfi í skandinavísku meginlandsmálunum.

Ekki hefur tekist að tengja notkun S3 við neina tiltekna félagslega breytnu, svo sem kyn, landfræðilega staðsetningu eða þjóðfélagshóp. Þvert á móti fannst S3 í máli bæði karla og kvenna: prestfrúa, vinnukvenna, vinnumanna, smiða, bænda og bændakvenna svo fátt eitt sé nefnt. Þá fundust dæmi um S3-orðaröð hjá langflestum bréfríturum. Undantekningar voru einkum bréfrítarar sem höfðu fá dæmi um málbreytuna S2/S3 yfirleitt, sem bendir til þess að notkun S3 hafi verið fremur útbreidd þó að tíðnin sé lægri en í blöðum og tímaritum.

5 Umræða um þátt málstöðlunar

Ef meta á áhrif málstöðlunar á málfar fólks er mikilvægt að glöggva sig á hvenær meðvitund um tilbrigði í máli er orðin slík að líta megi á málbreytuna sem breytingu ofan frá (sbr. kafla 3.1 hér á undan). Athuganir á S3 benda til að í útgefnum textum dragi úr tíðni S3 á 19. öld og við það verður málnotkun í þeim líkari því sem gerist í sendibréfum. Þegar nær dregur aldamótunum 1900, og ekki síst á 20. öld, dregur enn frekar úr tíðninni og það er því á þessu skeiði sem áhrif málstöðlunar koma helst til álita.

Hefð var fyrir að vísa til dönskulegrar eða óíslenskulegrar orðaraðar á 19. öld líkt og það væri á allra vitorði í hverju hún fælist. Þannig talar Baldvin Einarsson (1830:vi) til dæmis um málnotkun hins tilbúna Seltirnings, Öundur, sem orðinn er vanur við „blendíng í málinu“, í tímaritinu Ármann á Alþingi:

Tilgángerinn er sá, að ef einhvörjir kynnu að vera orðnir svo vanir við blendíng í málinu, að þeir eigi aðvældiga vissu hvört það væri íslenska eða danska sem þeir mæltu, þá skyldu þeir þó verða að sjá hvørsu hrakliga ljótt það er að aflaka [svo] málid með útlendum ordum, og skræma það með með [svo] útlendri orðaröð (stilhætti) sem á svo illa við íslensku.

Athygli vekur að Baldvin notar S3 í þessari málsgrein: *að þeir eigi aðvældlega vissu*. Því verður að teljast ólíklegt að orðalagið „útlendur

stílsáttur“ vísi til þessa atriðis eða að Baldvin sé meðvitaður um það. Mál Önundar einkennist fremur af dönskuslettum og föstum samböndum en beinlínis af útlendri orðaröð, auk þess sem S3 kemur ekki síður fram í máli þeirra tilbúnu persóna á landsbyggðinni sem eiga að standa Önundi frammar í málfarslegum efnum.

Mikilvægir áhrifavaldar („ofurpaurar“) sem nefndir hafa verið í tengslum við upphafningu alþýðlegs stíls eru Fjölnismenn. Þar vega þyngst ritdómar Konráðs Gíslasonar (auk annarra) í Fjölni: „Áhrif Fjölnis voru geysimikil og ekki síst á stíl þar sem segja má að hann hafi á skömmum tíma útrýmt útlenskulegum lærdómsstíl úr bókum og sett alþýðlegan stíl í öndvegi“ (Kjartan G. Ottósson 1990:72). Ritstjórar íslenskra tímarita munu einnig upp til hópa hafa verið miklir málhreinsunarmenn, einkum á síðari hluta 19. aldar (tilv. rit:78). Þá má leiða líkur að því að áhrif íslenskukennslu við Lærða skólann með Halldór Kr. Friðriksson í broddi fylkingar hafi ekki síst verið mikilvæg (sjá nánar hér á eftir).

Ritdómar Fjölnis mörkuðu ákveðin tímamót að því leyti að ekki er aðeins vikið með óbeinum hætti að málnotkun heldur eru tekin raunveruleg dæmi úr nýlega útgefnum bókum með tilvísunum og hispurslausri umsögn. Gagnrýnin beindist oft að ætluðum erlendum tökuorðum, yfirleitt dönskum. Nýlundan var ekki síst fólgin í þeirri auknu athygli sem setningafræðileg atriði fengu í Fjölni og þar er minnst á um þrjátíu slík atriði sem betur mættu fara. Eitt þeirra er títtnefnd S3-orðaröð, þar sem röðin neitun–sögn í aukasetningum er sögð „dönskuleg“ og bent á að neitun eigi að fara á eftir sögninni (sbr. Konráð Gíslason 1844:85). Athugasemd Konráðs eru fyrstu beinu merki um að amast hafi verið við S3.

Meint tengsl við dönsku vekja þá spurningu hvort S3 hafi verið talin bundin við (rit)mál menntamanna. Vísbendingu í þá átt er að finna í ummælum Sigurðar Gunnarssonar (1878) um orðaröðina í máli „Styrbjarnar á Nesi“.¹⁷ Kveðst Sigurður (1878:10) vilja „bera í tal við þig [þ.e. Styrbjörn] dönskuslettunar þínar í greininni, sem eg nefndi, eða þær málsgreinir, sem mjer fannst eitthvert danskt óbragð að“; flest dæma hans snúa að S3:

Sjaldan er orðum svo skipað í daglegu tali, og í ágætum fornbókum

¹⁷ Um var að ræða ritdeilu vegna nýútkomins smásagnasafns Péturs Péturssonar biskups. „Styrbjörn á Nesi“ hafði ýmislegt að athuga við málið á safninu en Sigurður var í hópi þeirra sem þótti hann ósanngjarn og „stranda á sömu skerjum, er þú vítir aðra fyrir“ (tilv. rit:9).

hefi eg óvíða fundið líka orðaskipun. Þó því bregði fyrir á einstöku stað, eins og jeg man eptir t. a. m. á einum stað í Eglu, þá tel eg líklegra að ritarar hafi víxlað þar orðunum, en að þau standi svo skipuð í frumritinu. (Sigurður Gunnarsson 1878:10)

Dæmi Sigurðar eru á ýmsa lund og með þeim slæðast dæmi sem að betur athuguðu máli tilheyra ekki málbreytunni.¹⁸

Færa má sterk rök fyrir því að um þetta leyti, á síðasta fjórðungi 19. aldar, sé raunverulega fyrst „tekin ákvörðun“ um að gera S3 brottræka úr staðalmálinu. Um þetta vitnar ákaflega mikilvæg en vannýtt heimild: verkefni skólapilta við Lærða skólann í Reykjavík (sjá Braga Þorgrím Ólafsson 2004). Á Þjóðskjalasafni Íslands eru varðveitt prófverkefni úr ýmsum greinum frá tímabilinu 1846–1904 með leiðréttingum þess sem fór yfir, bæði á málfari og stíl. Það er merkilegt við þessi gögn að yfirleitt kemur fram hver leiðrétti og þennan einstæða brunn mætti nýta til þess að kafa dýpra ofan í mótun málstefnu á síðari hluta 19. aldar og byrjun þeirrar tuttugustu.

Ég hef kannað efni í þessu safni, m.a. verkefni frá ólíkum tímum sem nefnast íslenskur stíll, sem (yfirleitt) eru þýðingar úr dönsku á íslensku, og íslenskar ritgerðir, þ.e. frjáls ritun um ákveðið efni. Ég athugaði skipulega þessar tvær tegundir verkefna frá árunum 1847–48, 1852, 1855, 1860, 1875 og 1890 og skráði hjá mér allar leiðréttingar á málfræðilegum atriðum. Niðurstaða mín af þessari athugun er sú að a.m.k. fram til 1875 hafi ekki verið amast við notkun S3, enda þótt slík orðaröð væri býsna algeng frá upphafi bæði í íslenskum stílum og ritgerðum nemenda.

Fyrstu merki sem ég hef fundið um að litið sé á S3-orðaröð sem málvillu eru frá 1882 og árið 1890 er orðaröðin líka leiðrétt, eins og sjá má á myndum 3 og 4. Á báðum myndum sjást leiðréttingar Halldórs Kr. Friðrikssonar. Í þessu sambandi er rétt að nefna að Halldór Kr. er sagður hafa haft gríðarmikil áhrif og „brýndi svo rækilega fyrir nemendum að varast tiltekna málvillur að þær sáust vart í ritum

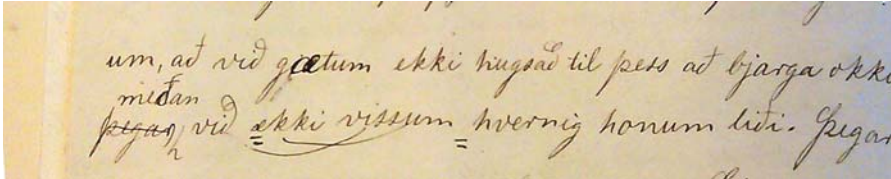
¹⁸ Í fyrsta lagi er um að ræða stílfærslu (sjá t.d. Maling 1980) sem Jakob Jóh. Smári (1920) telur einnig til S3:

(i) líka fyrir manninn, sem í **raun og veru hefur** unnið verkið (Jakob Jóh. Smári 1920:258)

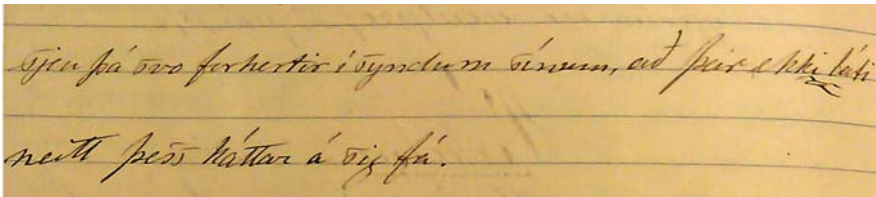
(ii) sem **svo víða hefir** vel vandað (Sigurður Gunnarsson 1878:10)

Greiningin byggist sennilega á því að *sem* sé tilvísunarfornafn og gegni hlutverki frumlags í setningunni. Sigurður (1878:10) gengur lengra og gagnrýnir orðaröð þar sem atviksliðir fara á undan sögn óháð því hvort um auka- eða aðalsetningu er að ræða, sbr. og *maður svo hins vegar væri viss um; en aptur var vel kunnugur þeim og hann að maklegleikum á.*

þeirra sem hann hafði kennt“ (Kjartan G. Ottósson 1990:95). Kjartan bendir einnig á að kennsluferill Halldórs hafi spannað nær allan síðari helming 19. aldar.



Mynd 3. Leiðrétting á sögn í þriðja sæti í íslenskum stíl við árspróf í 3. bekk 1882: „meðan við ekki vissum hvernig honum líði.“



Mynd 4. Leiðrétting á sögn í þriðja sæti í íslenskri ritgerð við árspróf í 4. bekk 1890: „svo forhertir í syndum sínum, að þeir ekki láti neitt þess háttar á sig fá.“

Samkvæmt þessum vitnisburði mætti fyrst eiga von á áhrifum at-
hugasemda við S3-orðaröð undir lok aldarinnar, samfara vaxandi
meðvitund um breytuna, þó að ekki sé hægt að útiloka að ritstjórar
einstakra tímarita hafi tekið strax af skarið í kjölfar ritdóms Konráðs.
Ef breytingin tengdist meðvitaðri ákvörðun ritstjóra eða yfirlesara
mætti búast við að hlutfall S3 væri lægra en í sendibréfunum en sú
er hins vegar ekki raunin, a.m.k. ekki ef á heildina er lítið. Frekari
rannsókn er þörf til þess að ákvarða nánar tímasetningu breytinga á
málviðmiðum og áhrifum þeirra.

6 Lokaorð

Gagnstætt því sem áður hefur verið talið virðist S3 hafa náð töluverðri
útbreiðslu meðal þorra landsmanna og ekki einskorðast við mál
menntamanna. Þetta sést hvað skýrast á því að lítið er um „hreina“
S2-málhafa meðal bréfitara. Þó kom fram eðlismunur á notkun
málbreytunnar í blöðum og tímaritum annars vegar og sendibréf-
um hins vegar miðað við kjörumhverfi S2/S3 í skandinavísku megin-
landsmálunum. Þetta er hliðstætt niðurstöðum Heycock & Wallen-

bergs (2013) og mætti túlka, eins og þau gera, á þann hátt að opinbert mál á fyrri hluta 19. aldar beri keim af dönsku málkerfi en það er forvitnilegt að í sendibréfum eru skilyrðin einmitt ólík þeim dönsku. Eðli þessa munar þarf að kanna nánar.

Dreifing raðanna í ólíku setningafræðilegu umhverfi sýndi að notkunin er talsvert frábrugðin kjörumhverfi S3 í nútímamáli. Meirihluti dæmanna teldust frávik ef miðað er við lýsingar á nútímamáli og því virðist ekki með sömu rökum vera hægt að líta á S3 sem jaðarfyrirbæri. Færð voru ný rök fyrir þeirri tilgátu að breytileiki milli S2 og S3 í (19. aldar) íslensku tengist e.t.v. sveigjanlegri stöðu neitunar. Þó að þetta mætti nota sem röksemd með þeirri umdeildu tilgátu að ríkulega sagnbeygt tungumál eins og íslenska gæti ekki þróað með sér ómarkaða S3-orðaröð er hátt hlutfall þessarar orðraðar í fjölbreyttu setningafræðilegu umhverfi á 19. öld mjög óvænt ef S3 væri „afbrigðileg“ orðaröð.

Ný rök voru færð fyrir því að meðvitund um málbreytuna hafi ekki verið til staðar fyrr en á síðasta fjórðungi 19. aldar, þrátt fyrir gagnrýni Fjölnismanna á S3 þegar á fimmta áratug aldarinnar. Þetta mælir gegn því að breytingar á tíðni S3 í útgefnum ritum þegar eftir miðja öldina tengist málstöðlun (eða málhreinsun). Þá er tíðni S3 í útgefnum ritum á þessu tímabili ekki lægri en í sendibréfunum. Leiðréttingar í verkefnum nemenda í Lærða skólanum í Reykjavík benda til þess að síðustu tvo áratugi aldarinnar sjáist fyrst merki um að farið sé að líta á S3 sem málvillu.

Frekari rannsókna er þörf til þess að kanna dreifinguna nánar í þeim textategundum sem hér hafa verið til umræðu, bæði hvað varðar félagsmálfræðileg og „málfræðileg“ skilyrði, ekki síður merkingafræðileg og pragmatísk en setningafræðileg. Einnig er margt á huldu um stöðu S3 í eldra máli, t.d. í forníslensku, sem þyrfti að kanna nánar.

Tilvísanir

- Anthony, Laurence. 2012. AntConc (útg. 3.2.2). Tokyo: Waseda University. [http://www.antlab.sci.waseda.ac.jp]
- Ásgrímur Angantýsson. 2001. Skandinavísk orðaröð í íslenskum aukasetningum. *Íslenskt mál og almenn málfræði* 23:95–122.
- Ásgrímur Angantýsson. 2011. *The Syntax of Embedded Clauses in Icelandic and Related Languages*. Reykjavík: Hugvísindastofnun.
- Baldvin Einarsson. 1830. [Formáli]. *Armann á Alþingi eða almennur Fundur Íslendinga: Arsrít fyrir búhólda og bændafolk á Islandi* 2:iii–viii.

- Bragi Þorgrímur Ólafsson. 2004. *Landsins útvöldu synir: Ritgerðir skólapilta Lærða skólans í íslenskum stíl 1846–1904*. Sýnisbók íslenskrar alþýðumenn- ingar 7. Reykjavík: Háskólaútgáfan.
- Deumert, Ana. 2003. Bringing speakers back in? Epistemological reflections on speaker-oriented explanations of language change. *Language Sciences* 25:15–76.
- Eiríkur Rögnvaldsson. 1998. Heimildatúlkun í sögulegri setningafræði. Í: Baldur Sigurðsson, Sigurður Konráðsson & Örnólfur Thorsson (ritstj.), *Greinar af sama meidi*, bls. 317–334. Reykjavík: Rannsóknastofnun Kenn- araháskóla Íslands.
- Elsaß, Stephan. 2012. The Use of Private Letters and Diaries in Sociolin- guistic Investigation. Í: Juan Manuel Hernández-Campoy & Juan Camilo Conde-Silvestre (ritstj.), *The Handbook of Historical Sociolinguistics*, bls. 156–169. Blackwell Handbooks in Linguistics. Oxford: Wiley-Blackwell.
- Guðmundur Hálfðanarson. 2007. *Íslenska þjóðríkið: uppruni og endimörk. Íslensk menning*. Reykjavík: Hið íslenska bókmenntafélag/Reykjavík- Akademían. [2. prentun]
- Haraldur Bernharðsson & Jóhannes Gísli Jónsson. 2012. Íslenskt mál á nítjándu öld — bréfasafn. Ópr. handrit.
- Heycock, Caroline & Joel Wallenberg. 2013. How variational acquisition drives syntactic change: The loss of verb movement in Scandinavian. *Journal of Comparative Germanic Linguistics*. [doi:10.1007/s10828-013-9056-0]
- Höskuldur Thráinsson [=Þráinsson]. 2010. Predictable and unpredictable sources of variable verb and adverb placement in Scandinavian. *Lingua* 120: 1062–1088.
- Höskuldur Þráinsson, Ásgrímur Angantýsson, Einar Freyr Sigurðsson, Sig- rún Steingrímsdóttir & Þórhallur Eypórsson. 2013. Efnissöfnun og að- ferðafræði. Í: Höskuldur Þráinsson, Ásgrímur Angantýsson og Einar Freyr Sigurðsson (ritstj.), *Tilbrigði í íslenskri setningagerð. I. Markmið, aðferð- ir og efniviður*, bls. 19–68. Reyjavík: Málvísindastofnun Háskóla Íslands.
- Jakob Jóh. Smári. 1920. *Íslensk setningafræði*. Reykjavík: Bókaverzlun Ársæls Árnasonar.
- Kjartan G. Ottósson. 1990. *Íslensk mállhreinsun: sögulegt yfirlit*. Rit Íslenskrar málnefndar 6. Reykjavík: Íslensk málnefnd.
- Kjartan Ottosson [=Kjartan G. Ottósson]. 2003. Heimenorsk innverknad på islandsk språk i mellomalderen, særleg i morfologien. Í: Kristján Árnason (ritstj.), *Útnorður: West Nordic Standardisation and Variation*, bls. 111–152. Papers from a Symposium in Stockholm October 7th 2001. Reykjavík: Institute of Linguistics/University of Iceland Press.
- Konráð Gíslason. 1844. „Bókafregn“. *Fjölur: Árrit handa Íslendingum* 7:71–104.
- Laycock, Donald C. & Peter Mühlhäusler. 1990. Language engineering: special languages. Í: N. E. Collinge (ritstj.), *An encyclopaedia of language*, bls. 843–875. London/New York: Routledge.

- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Labov, William. 2001. *Principles of Linguistic change. Volume II: Social Factors*. Oxford: Blackwell.
- Leerssen, Joep. 1999. *Nationaal denken in Europa: Een cultuurhistorische schets*. Amsterdam: Amsterdam University Press.
- Maling, Joan. 1980. Inversion in Embedded Clauses in Modern Icelandic. *Íslenskt mál og almenn málfræði* 2:175–193.
- Mair, Christian. 2009. *Twentieth-Century English: History, Variation, and Standardization*. Studies in English Language. Cambridge: Cambridge University Press.
- Margrét Guðmundsdóttir. 2009. Málbreytingar í ljósi málkunnáttufræði. *Íslenskt mál og almenn málfræði* 30:7–52
- Poplack, Shana & Nathalie Dion. 2009. Prescription vs. praxis: the evolution of future temporal reference in French. *Language* 85.3:557–587.
- Rask, Erasmus Christian. 1818. *Anvisning till Isländskan eller Nordiska Fornspråket: Från Danskan öfversatt och omarbetad af Författeren*. Stockholm: A. Wiborgs förlag.
- Sigurður Gunnarsson. 1878. „Styrbjörn á Nesi“. *Ísafold* 5.3: 9–10.
- Sundquist, John D. 2003. The Rich Agreement Hypothesis and Early Modern Danish embedded-clause word order. *Nordic Journal of Linguistics* 26.2: 233–258.
- Thomason, Sarah G. 2001. *Language Contact*. Washington, D.C.: Georgetown University Press.
- Wallenberg, Joel, Anton Karl Ingason, Einar Freyr Sigurðsson & Eiríkur Rögnvaldsson. 2011. *Icelandic Parsed Historical Corpus*. http://www.linguist.is/icelandic_treebank. [Útgáfa 0.9]
- Þórhallur Eyþórsson. 1997–1998. Uppruni sagnfærslu í germönskum málum. *Íslenskt mál og almenn málfræði* 19–20:133–180.

Viðauki

1800–1850			
Minnisverð tíðindi	77,8% (21/27)	Ármann á Alþingi	40,7% (50/123)
Reykjavíkurbósturinn	56% (56/100)	Sunnanpósturinn	32,5% (13/40)
Búnaðarrit suður- amtsins	46,3% (31/67)	Margvíslegt gaman og alvara	15,2% (7/46)
1875			
Íslendingur	32,5% (13/40)	Norðlingur	6,3% (3/48)
Þjóðólfur	23,8% (5/21)	Ísafold	5,9% (2/34)
Norðanfari	10,9% (5/46)		
1900			
Reykvíkingur	35% (7/20)	Fjallkonan	4,8% (2/42)
Reykjavík	25% (1/4)	Ísafold	3,2% (2/63)
Austri	23,9% (11/46)	Þjóðólfur	1,9% (1/53)
Stefnir	22,7% (10/44)	Kvennablaðið	0% (0/20)
Þjóðviljinn	9,3% (4/43)	Bjarki	0% (0/22)
Framsókn	8,8% (3/34)		

Tafla 7. Hlutfall S3 af heildardæmafjölda innan hvers blaðs/tímarits.

Lykilorð

málstöðlun, sagnfærsla, orðaröð, tilbrigði, málbreytingar

Keywords

standardization, verb movement, word order, variation, language change

Abstract

The position of the finite verb with regard to e.g. negation, so-called verb movement, has figured prominently in the literature over the past decades. It has been argued that a morphologically rich language such as Icelandic requires verb movement (V2) and that the verb follows negation (V3) only as an exception. However, linguists have pointed out that during the period 1600–1850, such exceptions were more common than today and that the language of certain texts bear a resemblance to the language of a Danish speaker (Heycock & Wallenberg 2013).

A new study of two 19th-century corpora, newspapers/periodicals and private letters of people with various social backgrounds, supports this result in the public texts. However, V3 is more frequent in private letters than expected if characteristic of Danish influence on learned individuals, as often thought. After 1850 the frequen-

cy of V3 in newspapers/periodicals drops from around 45% to 10–15%, whereas in private letters V3 is stable around 10%. The question is posed whether the change is due to standardisation but it seems that the drop in frequency precedes efforts to oust V3, e.g. in Lærði skólinn college in Reykjavík. This casts doubt on such a hypothesis but more research is needed.

*Heimir Freyr van der Feest Viðarsson
Háskóla Íslands / Stofnun Árna Magnússonar í íslenskum fræðum
Neshaga 16, 107 Reykjavík
hfv3@hi.is*

Magnus Breder Birkenes og Jürg Fleischer

Intetkjønn med referanse til personer

To ulike normeringstradisjoner i nynorsk og tysk¹

1 Innledning

I germanske språk kan et pronomen i intetkjønn referere til to eller flere personer av forskjellig kjønn. Det dreier seg her om en kongruensstrategi som av Corbett (2006:243–253) omtales som "gender resolution",² dvs. at en regel eller et sett regler avgjør hvilket kjønn en bøyingsform som viser tilbake på to eller flere entiteter av ulikt kjønn skal stå i. Formålet med en slik regel er å løse en potensiell kongruenskonflikt. "Gender resolution" i intetkjønn er i dag særlig utbredt i islandsk og færøysk. I det følgende eksempelet hentet fra

¹ Bakgrunnen for artikkelen er et forskningsprosjekt om utviklingen av kongruenssystemer i fire indoeuropeiske språk, ved Philipps-Universität Marburg i Tyskland (prosjekttittel: Diachrone Entwicklung von Kongruenzsystemen in vier indogermanischen Sprachen). Forskningsprosjektet ledes av Jürg Fleischer, Elisabeth Rieken og Paul Widmer og er finansiert av Deutsche Forschungsgemeinschaft DFG (RI 1730/4-1). Vi vil gjerne takke medarbeiderne på prosjektet og dessuten Kristine Breder, Hans-Olav Enger, Helge Sandøy, Elisabeth Solberg og Arne Torp for nyttige innspill til artikkelen.

² I mangel av et bedre avløsningsord på norsk vil vi holde oss til Corbetts begrep "gender resolution" gjennom artikkelen (jf. Corbett 1991; 2006).

Halldór Laxness' roman *Atómstöðin* viser formene *þau* (N.PL.), *bæði* (N.PL.), *hvort* (N.SG.) og *annað* (N.SG.) enten tilbake på begge eller én av nominalfrasene *stúlkan* (F.SG.) 'jenta' og *pilturinn* (M.SG.) 'gutten':

- (1) 1 a) Brauðsalan var opin, stúlkan fyrir innan disk, pilturinn fyrir framan, **þau** (n.pl.) [...] tóku ekki eftir þegar ég bauð gott kvöld (Laxness 2012:49)
- b) Brødbutikken var åpen, piken bak disken, den unge mannen foran, **de** (m./f./n.pl.) la ikke merke til at jeg sa god kveld (Laxness 1966:35)
- 2 a) Híngaðtil höfðu **bæði** (n.pl.) verið alvarleg, nú var **þeim** (n.pl.) altíeinu skemt, horfðu **hvort** (n.sg.) framaní **annað** (n.sg.) og hlógu. (Laxness 2012:51)
- b) Hittil hadde de vært alvorlige **begge** (m./f./n.pl.), men nå moret **de** (m./f./n.pl.) seg plutselig, så på **hverandre** (recp.) og lo. (Laxness 1966:37)

Hvilken kongruensform man skal velge i et tilfelle som (1), er et spørsmål som særlig stiller seg i språk med rik bøyingsmorfologi. I det konkrete eksempelet tilsvarende *þau* som intetkjønnsform ingen av de to kjønnene i nominalfrasene den viser tilbake på (*stúlkan* eller *pilturinn*). Dette er derfor et eksempel på det Corbett kaller "syntactic gender resolution", i motsetning til "semantic gender resolution" der for eksempel semantiske faktorer som animathet kan overstyre morfologisk genus. Corbett (1991:283) setter opp følgende regler for islandsk:

1. hvis alle ledd er hankjønn, brukes hankjønn
2. hvis alle ledd er hunkjønn, brukes hunkjønn
3. ellers brukes intetkjønn

I den norske oversettelsen av Laxness-romanen gjengis nøytrumsformene med former som ikke lenger bøyes i kjønn (heretter kun glossert som RECP. for "reciprocal"). Her gjør altså en mangel på morfologiske distinksjoner "gender resolution"-reglene ovenfor overflødige. Et liknende tilfelle som i norsk finner vi interessant nok også i et språk som tysk, som etter islandsk og færøysk er det germanske språket med mest bevart bøyingsmorfologi. Her oversettes de islandske formene med hhv. *sie* 'de', *beide* 'begge' og *einander* 'hverandre', alle uten noen form for kjønnsbøyning (jf. Laxness 2011:40–41).

Samtidig finnes det klare paralleller til de islandske formene i norske og tyske dialekter og på eldre språktrinn, selv om bruken rett nok

er mer innskrenket enn i islandsk grunnet tap av genusdistinksjoner i flertall. Islandsk på den ene siden og norsk og tysk på den andre har altså åpenbart gått forskjellige veier hva gjelder disse kongruensreglene. I islandsk er kongruens i intetkjønn en del av standardspråket, i norsk (nynorsk og bokmål) og tysk (høytysk) er dette ikke tilfellet. I denne artikkelen ønsker vi å vise at dette ikke bare skyldes en ren intern morfologisk utvikling, men at også eksterne årsaker, som preskriptivisme og språkkontakt, spiller en rolle. Et sentralt spørsmål vil dermed være: hva skjedde på veien til de moderne standardspråkene? Normeringen av tysk og norsk på 16-, 17- og 1800-tallet forløp særdeles ulikt, men resultatet ser ut til å ha vært likt hva angår temaet for denne artikkelen, nemlig at intetkjønnsregelen ble ekskludert eller i det minste forsvant fra standardspråket.

Artikkelen er bygd opp på følgende måte: Først tar vi for oss intetkjønnsregelen slik vi kjenner den fra eldre germanske språk og fra eldre språktrinn av tysk og norsk. I hoveddelen går vi så gjennom grammatikker fra det 17. til det 20. århundre på jakt etter stemplende karakteristikk eller beskrivelser som bringer oss nærmere en forklaring på hvorfor konstruksjonen ikke ble en del av de moderne standardspråkene tysk og nynorsk. Til slutt forsøker vi å se resultatene i lys av de ulike normeringstradisjonene i Norge og i de tyskspråklige landene.

2 "Gender resolution" i germanske språk

"Gender resolution" i intetkjønn ser ut til å ha vært en innovasjon i germansk: I andre indoeuropeiske språk vil hankjønn være vanlig i en tilsvarende kontekst (jf. Corbett 1991:279–280). De første eksemplene på kongruens i intetkjønn flertall finnes allerede i gotisk og med Askedal (1973) foreligger en utførlig behandling av fenomenet i germansk og spesielt tysk. At det dreier seg om et gammelt, germansk fenomen, fremgår av de følgende eksemplene fra gotisk, gammelhøytysk og nyislandsk.³ For å sørge for bedre sammenliknbarhet mellom

³ Den eldste norrøne bibeloversettelsen, *Stjórn*, omfatter kun første del av det gamle testamentet (fra 1. Mosebok til 2. Kongebok). Da det ikke finnes noen oversettelse til mellomnorsk, så vi oss nødt til å la den eldste nytestamentlige bibeloversettelsen til islandsk av Oddur Gottskálksson fra 1540 representere nynorsk snarere enn den danske bibeloversettelsen fra 1550 (Christian III's bibel). Vi er oss selvfølgelig bevisst problematikken i dette.

språkene, har vi valgt oversettelser av det samme bibelverset (Luk 1,5–6) i den eldste tilgjengelige versjonen:

- (2) 1 was in dagam Herodes þiudanis Iudaias gudja namin Zakarias, usafar<am> Abijins, jah qeins is us dauhtrum Aharons, jah namo izos Aileisabaiþ. wesunuh þan **garaihta** (n.pl.) **ba** (n.pl.) in andwairþja gudis [...] (Gotisk Bibel, 85)
- 2 uuas in tagun herodes thes cuninges Iudeno sumer biscof namen zacharias fon themo uuehsale abiasen inti quena imo fon aarones tohterun inti ira namo uuas elisabeth. **siu** (n.pl.) uuaron **rehtiu** (n.pl.) **beidu** (n.pl.) fora gote [...] (Tatian 25,29–26,3)
- 3 A dogū Herodis konungs i Judea / var sa keñimadr er Zacharias er nefndr af stiettū Abia / & hñs husfru var af Dætrū Aarons / su er hiet Elizabeth / eñ **þau** (n.pl.) voru **bædi** (n.pl.) **riettlat** (n.pl.) fyrer gudi (Nýja testamenti Odds Gottskálkssonar, 149 blaðsíða [143])

‘I Herodes’ dager, konge av Judea, var det en prest som het Sakarias av slekten Abia. Og hans kone var av Arons døtre og het Elisabet. Og de var begge rettferdige for Gud [...]’

I (2)1–3 står både personlig pronomen, ubestemt pronomen og predikativt adjektiv i intetkjønn flertall. Siden disse viser tilbake både på Sakarias og Elisabet, er kun regel 3 ovenfor mulig.

Som eksempel (1) i innledningen viste, finnes det “gender resolution”-former både i intetkjønn flertall og intetkjønn entall i moderne islandsk. I gotisk og gammelhøytysk er riktignok bare flertallsformene kjent. Derimot er entallsformene en del av norrønt fra første stund, men her må det tilføyes at den norrøne overleveringen begynner mye senere enn den gotiske og gammelhøytyske; dessuten er det norrøne materialet både mer rikholdig og preget av en selvstendig literær tradisjon.⁴ I tysk finnes det belegg for entallsformene først fra og

⁴ Det at vi i gotisk og gammelhøytysk (som begge er relativt dårlig overlevert) bare finner flertallsformer, kan etter vår mening henge sammen med at flertallsformene er vanligere enn entallsformene i en “gender resolution”-sammenheng. Generelt er det selvfølgelig slik at den kategorielle tokenfrekvensen i entall er betydelig høyere enn i flertall (jf. Corbett 2000:281–286), men i denne spesielle konstruksjonen tyder mye på at det forholder seg omvendt: Her har vi jo å gjøre med en koordinasjon av to nominalfraser og en henvisning til kun én av referentene må kunne betraktes som mer markert enn en henvisning til begge. Av denne grunn går vi ut fra at både entalls- og flertallsformene står i en nær sammenheng med

med middelhøytysk tid (jf. Behaghel 1928:33, Dal 1966:161, Askedal 1973: 87; 135–139, Harnisch 2009: 78, Fleischer og Schallert 2011: 119–120). De følgende eksemplene, (3) og (4), fra en tidlig middelhøytysk tekst, Wiener Genesis fra 1100-tallet, og en nyhøytysk tekst, en roman fra slutten av 1700-tallet, viser hver for seg intetkjønn entall:

- (3) Er unt sîn wîb chêrten fon einanderen ir lîp, daz ir **newederez** (n.sg.) nesâhe wie **daz andere** (n.sg.) getân wâre. (Wiener Genesis 724–727)

‘Han og hans kone vendte kroppene sine fra hverandre, slik at ingen av dem kunne se hvordan den andre var skapt’

- (4) so wechselte die Gräfin mit Wilhelm bedeutende Blicke über die ungeheure Kluft der Geburt und des Standes hinüber, und **jedes** (n.sg.) glaubte an **seiner** (m/n.sg.) Seite, sicher **seinen** (m/n.sg.) Empfindungen nachhängen zu dürfen. (Goethe 1898:285)

‘slik vekslet grevinnen uttrykksfulle blikk med Wilhelm over fødselens og rangens veldige kløft, og hver (av dem) trodde på sin side å kunne hengi seg til disse følelsene’

Også i norrønt og mellomnorsk lar eksempler på ”gender resolution”-former i intetkjønn seg lett finne, noe (5) fra Heimskringla (1200-tallet) og (6) fra et østnorsk diplom fra 1400-tallet viser:

- (5) skildu **þau** (n.pl.) svá sína rœðu, ok gekk sína leið **hvárt** (n.sg.) þeira (Heimskringla, sitert etter Nygaard 1905, § 81)

‘avsluttet de (kongen og Ingjerd) samtalen sin, og hver av dem gikk sin vei’

- (6) Oc loffuade **hvert adro** (n.sg.) Ceciliae oc Odder at **hwort** (n.sg.) skuldæ **annet** (n.sg.) styrkie j allo godo (DiplNor, Nr. 674, 1421)

‘Og Cecile og Odd lovet hverandre at de skulle styrke hverandre i alle gode dager’

I de fleste moderne germanske standardspråk er genusdistinksjonene i flertall (i pronomener, adjektiver, tallord osv.) morfologisk utjevnet, med unntak av islandsk og færøysk. I tysk har pronomener og adjektiver fra og med tidlig nyhøytysk (1350–1650) stort sett en flertallsform som ikke skiller mellom kjønn (se for eks. Ebert et al. 1993:214

hverandre, på tross av at førstnevnte ikke er belagt i gotisk og i tysk sammenheng først dukker opp på 1000-tallet.

om personlig pronomen; kjønnsbøyningen holdt seg lengst ved tallordet '2', jf. Ebert et al. 1993:206). I norsk går de samme distinksjonene i det store og det hele tapt i løpet av den mellomnorske perioden (jf. Indrebø 1951:255). Siden normeringen av tysk og norsk tok til på et tidspunkt da kjønnsbøyningen i flertall for det meste var gått tapt, vil entallsformene stå i sentrum for denne artikkelen.

I alle eksemplene vi har diskutert hittil, har det dreid seg om to mennesker av forskjellig kjønn og hvor man kan vise anaforisk tilbake på de to nominalfrasene. I enkelte germanske språk finnes det utover dette også eksempler på intetkjønnsformer i ren generisk bruk, altså hvor en ikke nærmere definert mengde mennesker gjenopptas med en intetkjønnsform (likevel er det grunn til å tro at hankjønnsformer er vanligere i slike sammenhenger). Et mulig eksempel på generisk bruk av intetkjønn flertall finnes i færøysk, noe som fremgår av (7).⁵ Samtidig ser vi at islandsk, som ellers har utstrakt bruk av "gender resolution" i intetkjønn, tyr til hankjønn i et slikt tilfelle.

- (7) Færøysk: *Summi siga* ... 'noen (n.pl.) sier'
 Islandsk: *Sumir segja* ... 'noen (m.pl.) sier'

(se Thráinsson et al 2012:132)

I tysk, men ikke i norsk, er intetkjønn i generisk bruk også kjent. De følgende eksemplene, det ene fra en litteraturspråklig tekst av Goethe, det andre fra en moderne bayersk dialekt, viser dette:

- (8) Die gewöhnliche Gesellschaft war beisammen, und **jedes** (n.sg.) hatte etwas zu erzählen, zu sagen, zu bemerken [...]
 (Goethe 1889:295)

'Det vanlige selskapet var samlet, og enhver hadde noe å fortelle, å si, å kommentere'

- (9) Då kon **oàns** (n.sg.) gråd machà, wås **s** (n.sg.) wui (Merkle 1975: 152)

'Da kan man gjøre akkurat hva man vil'

I resten av artikkelen skal vi gå nærmere inn på spørsmålet om hvordan man forholdt seg til intetkjønn som "gender resolution"-form under standardiseringen av tysk og norsk. Det vil vise seg at det er visse

⁵ En anonym fagfelle har gjort oss oppmerksom på at man på færøysk bruker *fólk* i entall på en annen måte enn i de andre nordiske språkene. Her betyr *eitt fólk* 'en person', mens det på skandinavisk betyr 'nasjon'. *Summi* kan altså også ganske enkelt svare til flertallsformen *summi fólk*. Likevel er denne analysen ikke uten problemer fordi leksetet *fólk* ikke er språklig uttrykt.

forskjeller mellom de to språkene og at disse forskjellene muligens henger sammen med ulike normeringsprosesser (avsnitt 5).

3 Tysk standardisering

I moderne standardtysk brukes det stort sett hankjønnsformer både ved henvisning til et par og ved generisk referanse.⁶ Eksempel (10) illustrerer dette: I a) har vi ført opp et utdrag fra en tysk roman fra 1670 sett fra perspektivet til en kvinnelig *jeg*-forteller: Her finner vi tre eksempler på den omtalte intetkjønnskonstruksjonen (*das eine* 'det ene', *es* 'det', *das andere* 'det andre'); i b) har vi gjengitt det samme utdraget i en nyoversettelse fra 2010 hvor intetkjønnsformene er blitt erstattet med hankjønnsformer (*jeder* 'hver', *er* 'han', *den anderen* 'den andre':

(10) 1 Jch und mein Mann bekamen einander je länger je lieber / und schetzte sich als **das eine** (n.sg.) glücklich / weil **es** (n.sg.) **das andere** (n.sg.) zum Ehe gemacht hatte [...]
(Grimmelshausen 1992:65)

2 Ich und mein Mann, wir liebten uns, je länger wir zusammen waren, desto mehr. **Jeder** (m.sg.) schätzte sich glücklich, dass **er** (m.sg.) **den anderen** (m.sg.) geheiratet hatte [...]
(Grimmelshausen 2010:59)

'Jeg og mannen min likte hverandre bedre jo mer vi var sammen, og den ene så på seg som lykkelig siden han/hun hadde ektet den andre'

Med tanke på at eksempler på bruk av intetkjønnsformer i liknende sammenhenger både finnes i eldre kilder og i moderne dialekter, er det verdt å spørre seg hvordan det kan ha seg at intetkjønn som "gender resolution"-form praktisk talt er fullstendig forsvunnet i moderne standardtysk.

I nyere forskning på normeringen av tysk er det ved flere anledninger blitt vist til betydningen av såkalte "Stigmatisierungen", dvs.

⁶ Først i og med den feministiske språkkritikken fra 1970-tallet av er det i økende grad blitt stilt krav om alternativer til hankjønnsformene i generisk bruk, for eksempel den såkalte parformen: i forbindelse med pronomenet *jed-*, for eksempel, skal det etter noen språkkritikers ønske hete både *jede* og *jeder* '*eihver, enhver' istedenfor kun *jeder* '(en)hver'. I skriftlige tekster blir dette delvis også forkortet til *jederR*, selv om denne skrivemåten ikke er offentlig godkjent (jf. Duden Grammatik 2009: 257).

fordømmelser av bestemte språklige uttrykk (se for eks. Davies & Langer 2006; Elspaß 2005). Allerede siden begynnelsen av den tyske grammatikktradisjonen, spesielt siden 1600- og 1700-tallet, er bestemte konstruksjoner blitt ekskludert fra det som av mange grammatikere ofte ble omtalt som den "høyere", "edlere" og "mer kultiverte" språkformen, nemlig standardspråket. Dette gjelder særlig for konstruksjoner som dobbelt negasjon (se for eks. Elspaß 2005:275–283, Davies & Langer 2006:246, Elspaß & Langer 2012) eller omskrivning med *tun* (jf. Elspaß 2005:254–268, Davies/Langer 2006:211–224), som ikke er ulikt engelsk omskrivning med *to do*, men som i motsetning til denne ikke er blitt en del av standardspråket.

Når vi i det følgende skal gi en redegjørelse for grammatikernes behandling av "gender resolution", er det viktig å understreke at den omtalte konstruksjonen svært sjelden blir tematisert i grammatikker fra det aktuelle tidsrommet. Som den første, ifølge vår undersøkelse, behandler Carl Friedrich Aichinger (1717–1782) kongruensformene i forbindelse med tallordet '2': Aichinger, som stammer fra det sørtyske dialektområdet (i tysk dialektologi kalt *Oberdeutsch*), oppgir her distinksjonen hankjønn: *zween* – hunkjønn: *zwo* – intetkjønn: *zwey* (jf. Aichinger 1754:236) og legger til:⁷

Zu zwey Dingen ungleiches Geschlechtes braucht man allezeit das neutrum zwey. Z.B. so ein Mann und Weib redet, so sprechen sie: wir zwey. (Aichinger 1754:415–416)

'Om to ting av ulikt kjønn brukes alltid intetkjønnsformen *zwey*. For eks. når mann og kone snakker, sier de: *wir zwey*'

Her blir altså intetkjønnsformen omtalt som den vanlige formen for "gender resolution". Derimot bemerker den klart mest innflytelsesrike grammatikeren på 1700-tallet, Johann Christoph Adelung, i sin *Grammatisch-kritisches Wörterbuch*:

Im gemeinen Leben wird das Neutrum *jedes* auch häufig für jedermann, jede Person gebraucht. (Adelung 1796:1429)

'I dagligtalen brukes intetkjønnsformen *jedes* også ofte for enhver person'

Henvisningen til "das gemeine Leben" kan tolkes dithen at konstruksjonen av Adelung blir sett på som upassende "i det høyere språk". Selv om dette ikke er en eksplisitt fordømmelse av formen, er sitatet

⁷ Også i moderne sørtyske dialekter er bruken av intetkjønnsformen i tallordet '2' som "gender resolution"-form kjent (jf. Kollmer 1987:406 for en bayersk dialekt).

relevant: Adelong representerte, i motsetning til Aichinger, det øst-middeltyske språkområdet som var spesielt viktig for standardspråket høytysk: På et annet sted i ordboka (under oppslagsordet *ein*) finnes det en kort opplysning om intetkjønnsformen uten noen form for negativ konnotasjon.

Das einer behält hier sein männliches Geschlecht, wenn gleich die Person, die es gebraucht, und von sich verstanden wissen will, weiblichen Geschlechtes ist. Sonst ist auch das Neutrum für alle Geschlechter üblich. (Adelong 1796:1685)

'*einer* beholder her hankjønnsformen, selv om personen som bruker den er kvinne og refererer til seg selv. Utover det er også intetkjønn for alle kjønn vanlig'

Utover disse eksemplene fant vi ingen videre kommentarer til intetkjønn som "gender resolution"-form. Mange grammatikker går ikke inn på konstruksjonen i det hele tatt eller fører kun opp hankjønnsformer uten nærmere spesifisering. Først i det 21. århundret dukker "gender resolution" i intetkjønn opp i en preskriptiv grammatikk igjen, nemlig i Duden-grammatikken.⁸ Eksempel (11) representerer den sveitsiske varianten av moderne standardtysk og er hentet fra internett (jf. Duden Grammatik 2009: 257). Det er verdt å legge til at konstruksjonen er svært utbredt i moderne sveitsertyske dialekter.

(11) Den Lunch bringt **jedes** (n.sg.) selber mit. (kilde: <http://www.cevi-alpin.ch/touren/infos/>; dato: 23.06.2013)

'Lunsjen tar man selv med'

Alt i alt må utbyttet fra tysk normeringshistorie sies å være heller magert, men én interessant observasjon har vi gjort: Vi har sett at intetkjønn som "gender resolution"-form er kjent fra eldre språktrinn og moderne dialekter og til og med det siste opplaget av Duden-grammatikken viser til et moderne internetteksempel. Samtidig er konstruksjonen nærmest fullstendig fraværende i det moderne tyske standardspråket, og dette ser ut til å ha skjedd uten noen sterk preskriptiv mobilisering mot intetkjønnsformen.

⁸ Likevel er det viktig å understreke at Duden-grammatikken fra og med den 7. utgaven (2005) er blitt mer deskriptiv (det ser man for eksempel på den utstrakte bruken av internetteksempler). Dette kan altså være grunnen til at eksemplene på intetkjønn dukker opp igjen i en preskriptiv grammatikk i det 21. århundret.

4 Nynorsk standardisering

Etter Norges frigjøring fra Danmark i 1814 begynte man å stille seg spørsmålet hva som skulle være skriftspråket i et ikke lenger danskstyrt Norge. Ivar Aasen, som først gjorde seg bemerket som forfatter av en språkpolitisk artikkel i 1830-årene, ønsket som kjent å bygge opp et norsk standardspråk på bakgrunn av dialektene, ikke på dansken eller på den dannede dagligtale i storbyene. Et viktig første steg på veien til et slikt standardspråk var utgivelsen av *Det norske Folkesprogs Grammatik* i 1848, hvor Aasen systematisk kartla de norske dialektene og dermed innledet norsk dialektologisk forskning. Materialet samlet han selv under reiser i hele Norge i perioden 1843–46 (jf. Torp og Vikør 2003:152–153). I motsetning til den senere *Norsk Grammatik* blir *Det norske Folkesprogs Grammatik* tradisjonelt regnet som deskriptiv (jf. Walton 1996:396). Aasen selv så på den som en materialsamling til en fremtidig norm. I 1848-grammatikken behandler Aasen "gender resolution" i intetkjønn flertall og entall utførlig.⁹ I ordlæren kan vi lese:

307. Et Adjektiv, Pronomen eller Talord maa være af samme Kjøen, som det Substantiv, hvortil det hører, eller i hvis Sted det træder. Naar et saadant Ord svarer til to eller flere Subst. som ere af eet Kjøen, da sættes det i Fleertal af dette Kjøen. Naar det derimod svarer til to eller flere Subst. af forskjelligt Kjøen, sættes det altid i Neutrum; dette skeer ogsaa oftest, naar der tales om den ene af et Par. (Aasen 1848: § 307)

Under denne regelen nevner Aasen tre eksempler på konstruksjonen:

- (12) anten baae **tvau** (n.pl.) eller **eit** (n.sg.) av dei
- (13) enten beggie eller **kvart** (n.sg.) fyre seg
- (14) **da eina** (n.sg.) æ **stort** (n.sg.), aa **hitt** (n.sg.) æ større
(Aasen 1848: § 307)

(12) viser kongruens i intetkjønn flertall og entall i samme setning, i (13) kongruerer flertallsformen *beggie* ikke i kjønn, men entallsformen *kvart* står entydig i intetkjønn. I (14) viser både påpekende pronomen og predikativt adjektiv regelen. I alle tre tilfeller refereres det til en kvinne og en mann. I en anmerkning under dette avsnittet legger Aa-

⁹ I *Søndmørsk Grammatik* fra 1851 spiller kongruens derimot en mindre prominent rolle og de nevnte formene tas ikke med der (man kan kun lese at det heter "tvau" i intetkjønn flertall).

sen til at intetkjønnsregelen også er kjent fra norrønt ("i det gamle Sprog"), men at det fenomenet var mer utpreget der siden kjønnsbøyning særlig i flertall er mer begrenset i moderne dialekter og at det i dag primært er tallordene som bevarer dette bøyningsmønsteret.¹⁰ I *Ordbog over det norske Folkesprog* fra 1850, som Aasen selv regnet som sitt hovedverk, finnes liknende eksempler under de respektive oppslagsordene.

Samlingen *Prøver af Landsmaalet i Norge* fra 1853 inneholder målprøver fra hele Norge og normaliserte tekster av Aasen (*Prøver af et almindeligt Landsmaal*) som peker frem mot normen, "Grunndialekten", som presenteres i 1864-grammatikken. Interessant nok finnes det ingen eksempler på "gender resolution" i intetkjønn i målprøvene. Snarere tvert i mot finner vi et moteksempel, som illustrert i (15) fra Seljord i Telemark:

- (15) Der æ ein Gar, som heiter Juvik, aa der va ei Gjente som heitte Margjit; denna Gjenta ha'en huglagt, aa ho vilde au gjønne hava'n; men dæ va inkje Raa te, at dei kunne faa **einann'** (recp.) (Aasen 1853:54)

I (15) har vi en et resiprokpronomen som viser tilbake på et (hypotetisk) par, men det dreier seg om en ubøyd form eller en historisk hankjønnsform. Etter Aasen (1848: § 307) ville man ha forventet en entydig intetkjønnsform i denne konteksten. Aasen velger da også å kommentere formen i noteapparatet: "Egentlig: ein-annan (hinanden). Her hvor der tales om En af hvert Kjøen, skulde det egentlig hedde: eit-anna, ligesom 'kvart-anna' i Bergens Stift" (Aasen 1853:55). Bruken av modalverbet "skulde" er tvetydig: Etter all sannsynlighet jevnfører Aasen denne dialekten med systemet i norrønt, men samtidig kunne en se for seg at han også ønsket at intetkjønnsformene, som var vanlige i norrønt, skulle inngå i den nye nynorsknormalen. At dette kan ha vært et ønske fra Aasens side, får vi en viss bekreftelse på i de normaliserte prøvene i verkets annen del, som fremgår av (16):

- (16) Men Skiperen og Kono hans fekk atter Garden sin og stod seg godt og livde so væl saman, at dei visste ikje kor snilde dei skulde vera mot **kvart-annat** (n.sg.) (Aasen 1853:81)

¹⁰ En kan legge til at sterke adjektiv i eldre mål viste et skille mellom hankjønn/intetkjønn på den ene siden og hunkjønn på den andre: Dermed heter det *Konurna var rika*, men *Kono og Mannen var rikje*. Vi takker en anonym fagfelle for denne opplysningen (se også Beito 1970:215). Men da hankjønn/intetkjønn her begge ender på -e, er det uklart om vi har "gender resolution" etter hankjønn eller intetkjønn her.

I (16) ser vi tydelig at resiprokpronomenet står i intetkjønn, men det er viktig å fremheve at Aasen ikke uttaler seg eksplisitt om konstruksjonen her. Aasens mulige preferanse for intetkjønnsformene går også frem av en annen kilde fra samme tid, den mye omdiskuterte oversettelsen av sagaen om Fridtjof den Frøkne fra 1858. Det dreier seg her om en oversettelse av en eldre islandsk tekst. Ved å parallellisere oversettelsen med en tidlig tekstutgave Aasen kan ha brukt, blir også det språkhistoriske forbildet for konstruksjonen klart. Resultatet kan man se i (17). Teksten i a) baserer seg på det yngre håndskriftet (B) som inneholder en lengre tekst enn den eldre håndskriftgruppen (A¹ og A²). Den islandske teksten i håndskrift B stammer muligvis fra 1400-tallet (jf. Larsson 1901); b) representerer Aasen (1858).

- (17) 1 a) gefa mún ek þèr hringinn, ef þú heitir, at lóga honum eigi, ok senda mèr, ef þú villt eigi ega, ok hèrmeð skulu við játa **hvort** (n.sg.) **öðru** (n.sg.) trú sinni (Rafn 1829:69)
- b) Denne Ringen skal eg giva deg, um du lovar, at du inkje vil lata honom fraa deg, og at du vil senda honom til meg, naar du inkje vil hava honom; og hermed skulo mid lova **kvartannat** (n.sg.) Truskap. (Aasen 1997:III, 338)
- 2 a) með þessari trúlofan skipta **þau** (n.pl.) hringnum (Rafn 1829:69)
- b) Med denne Trulovingi skifte **dei** (m./f./n.pl.) Ringarne sine (Aasen 1997:III, 338)

I (17)1 viser ikke resiprokpronomenet *kvartannat* tilbake på to koordinerte nominalfraser, men til det deiktiske pronomenet *við* 'vi' som her refererer til en mann og en kvinne. I likhet med *Prøver af et almindeligt Landsmaal* ser vi at entallsformene holder seg, mens flertallsformene er ubøyde. Det er i disse skriftene, *Prøver af Landsmaalet i Norge* og *Fridtjofs saga* at Aasens landsmålsnormal kommer tydeligere og tydeligere frem. Man kunne være fristet til å si at den er sterkt inspirert av islandsk, noe mange også kritiserte ham for, særlig etter nevnte oversettelse, men samtidig ligger den tett opp til vestnorske dialekter som han anså for å ha de mest "fuldkomne" formene.

Men Aasen baserte seg ikke bare på eldre vestnorske målføre i arbeidet med den nye normalen. I og med 1864-grammatikken, hvor Aasen-normalen foreligger i ferdig form, ser vi at han velger å gi regelen om "gender resolution" en mindre prominent plass i *Landsmaalet* enn beskrivelsen fra 1848 legger opp til. Fra å være en del av hoved-

teksten i 1848-grammatikken, er den delen som omfatter intetkjønn entall i *Norsk Grammatik* flyttet til anmerkingsapparatet (jf. Aasen 1864: § 309). Det er en kjent sak at anmerkningene tiltar i forhold til 1848-grammatikken, hvor først og fremst mer perifere former som finnes i eldre mål føres opp (jf. Walton 1996:504). Aasen selv skriver i fortalen til *Norsk Grammatik*:

at kun Sprogets almindelige Stof og de vigtigste dertil hørende Regler ere afhandlede i Bogens egentlige Text, medens derimod Dialekternes afvigende Former ere omtalte i de tilføjede Anmærkninger (Aasen 1864:XXI)

Aasen later med dette til å forkaste regelen fra 1848-grammatikken som en eksplisitt norm for landsmålet. En bekreftelse på dette synet mener vi også å finne i en av Aasens senere og mer ukjente tekster, *Heimsyn*, som ble utgitt anonymt i 1875. Formålet med denne teksten var å vise at landsmålet kunne brukes på alle fagområder, i dette tilfellet dreier det seg om "ei snøgg Umsjaaing yver Skapningen og Menneskja, tilmaatad fyre Ungdomen". I skapelsesberetningen heter det:

(18) I beste Aldren heve baade Mann og Kvinna ein medskapad Hug til aaliva saman med ein Make, og den rette Samlivnaden er daa, at **kvar** (m./f.sg) skal hava sin Make fyre Livstid (Aasen 1997:II, 267)

Etter eksempel (13) fra 1848-grammatikken ville man her forventet *kvart* i det ubestemte pronomenet. Aasens nedtoning av disse formene kan ha vært et resultat av mottakelsen av Fridtjofs saga, idéen om et enhetlig mål og reaksjoner fra målfolk ellers. En rask gjennomgang av nynorskklitteraturen på 1800- og 1900-tallet, ved hjelp av Nasjonalbibliotekets tjeneste Bokylla.no som per dags dato omfatter om lag 100.000 trykte verk i Norge frem til år 2000, viser at intetkjønnsformene må kunne betraktes som relativt marginale; disse hadde tydeligvis lite fotfeste blant nynorskfolk. Faktisk ser det kun ut til å være Kristoffer Janson som har brukt disse formene noenlunde konsekvent. Så heter det i romanen *Han og ho* fra 1876: "Dei kjende inkje meir til kvartannat desse tvau" (Janson 1876:7). Konstruksjonen finnes også i romanen *Marit Skjølte* fra 1905. Men ellers er det lite belegg å finne i det undersøkte materialet: I et anonymt skrift om det nye landsmålet fra 1862 finnes eksempelet "dei toko kvartannat i henderna og gingo saman" (Anonym 1862:33). Ikke bare intetkjønnsformen, men også flertallsbøyningen i sterke verb må her betraktes som konservative trekk.

I senere nynorsk grammatikografi forblir "gender resolution" i intetkjønn fotnotestoff, for siden å fjernes helt. I Marius Hægstads *Norsk maallæra eller grammatik i landsmaalet* fra 1879 trekkes det opp et skille mellom resiprokpronomenene *kvarannan* og *kvartannat*, hvor førstnevnte form brukes ved samme kjønn, sistnevnte ved ulikt kjønn (jf. Hægstad 1879:22): "*dei gjekk ofte til kvarannan (tvo Menner), dei var forlitet kjende med kvartannat (Mann og Kvinna)*", men vel og merke skjer dette i en anmerkning, i likhet med Aasen (1864). I Leiv Heggstads nynorskgrammatikker fra 1914 og 1931 karakteriseres intetkjønnsformene som avleggs: "I gamaldags mål brukar dei inkjekynsformi um parfolk: dei tala om kvartanna (no vanleg kvarandre)" (Heggstad 1914 og 1931:116). En parallellisering av Aasens første skuespill *Ervingen* fra 1855 og en modernisert utgave fra 1949 oppsummerer utviklingen i nynorsk godt:

- (19) 1 a) Soleids kann me stevjast og kivast ofta, naar me er **tvau** (n.pl.) eine [...] (Aasen 1997:II, 54)
- b) Såleis kan me stevjast og kivast ofte, når me er **to** (m./f./n.pl.) eine [...] (Aasen 1949:40)
- 2 a) Her standa no desse tvo stridande Partarne so meinlause midt imot kvar-annan, og ingen av dei veit, at dan eine skal jaga dan andre ut. Faa dei no vita aldri so litet, so verda dei rædde **kvartannat** (n.sg.), og so faa me inkje meir Moro. (Aasen 1997:II, 55)
- b) Her står nå desse to stridande partane så meinlause midt imot einannan, og ingen av dei veit at den eine skal jaga den andre ut. Får dei no vita aldri så lite, så vert dei redde **einannan** (recp.), og så får me ikkje meir moro. (Aasen 1949:41)

Mens intetkjønnsformene i originalutgaven, som ser ut til å ha vært mest fremtredende i vestnorske dialekter (se under) opprinnelig var en del av Aasens normalformer, lempes det på reglene utover på 1800-tallet, noe som blir synlig allerede i Aasens senere arbeider. I den moderne utgaven av *Ervingen* fra 1949 er intetkjønnsformene både i flertall og entall erstattet med ubøyde former. I innledningen til denne utgaven skriver Per Thorsen at teksten er "haldi så nær Aasens som det er høve til – med bruk av nokre tillatne sideformer". Alt i alt mener vi å kunne sette opp følgende utvikling i nynorsk: regel (Aasen 1848; 1853) > dialektalt avvik fra normen (Aasen 1864 og Hægstad 1879) > begrenset til konservative dialekter (Heggstad 1914; 1931).

Den viktigste nynorskgrammatikken fra 1900-tallet, Beito (1970:247), påpeker at former med "gender resolution" i intetkjønn kun hører hjemme i eldre mål, i overenstemmelse med Heggstad (1931). En gjennomgang av materialet i Norsk Ordbok (oppslagsord: *annan*, *ein*, *einkvan*, *kvarannan*) viser at intetkjønnsformer i denne bruken primært har vært å finne i Hordaland, Sogn, på Sunnmøre og Nordmøre og i Trøndelag. I Norsk Ordboks seddelarkiv på kortet for Verdal finner vi følgende bemerkning om pronomenet *kvarannen*:

kvaranna blir brukt og er visst likså rett Verdalsmål som kvaranner. Dette siste blir mest alltid brukt no. Eg trur det første høyrer til ei litt eldre tid, det er ikkje sjeldan å høyra det no heller. Og det er visst rett at det må vera om kar og kvinne. (Setelarkivet – Norsk Ordbok: 289 Verdal, understreking i original)

Et utvalg dialektgrammatikker fra de aktuelle områdene i Sør-Norge støtter dette utsagnet. Søreide (1930:60) sier om Nordfjordmålet at resiprokpronomenet normalt er ubøyd, men at „ei inkjekynsform høyrer ein endá“. Jenstad (1985:175) sier om tallordet *to* i Sunndalsmålet på indre Nordmøre at "[i] dette ordet finst det no så å seie ikkje att restar av gammal kjønnsbøying [...]". Vår hypotese er at formene forsvant fra standardspråket fordi de var døende i dialektene allerede på Aasens tid, og dette synes ikke usannsynlig i lys av disse arbeidene.

Én ting er at formene forsvant fra standardspråket fordi de ikke lenger var levende i dialektene, en annen er spørsmålet om hvorfor formene forsvant fra dialektene i utgangspunktet. Dette er ikke tema for denne artikkelen, men noen få tanker rundt dette er på sin plass. Vi mener det er både interne og eksterne grunner til at denne formen for "gender resolution" ikke har overlevd i norsk: Én intern grunn kan være, som Beito (1970:247) fremhever, at skillet mellom resiprokpronomen som viser til to personer og slike som viser til flere, aldri har vært strengt gjennomført i praksis, og må ha vært sett på som uviktig (se også Knudsen 1949:49–52). Resiprokpronomenet *kvarannan* utviklet seg fra å være et syntagme til å bli ett ord i en uni-verberingsprosess; det siste stadiet i grammatikaliseringen av resiprokpronomenet var at genus- og numeruskongruensen fra tidligere språktrinn forsvant.

Denne forklaringen gjelder naturligvis ikke for de tilfellene hvor entallsformene fremdeles kongruerer i kjønn, som i de ubestemte pronomenene *kvar* og *ein*. Her blir forklaringen mer tentativ: Alle dia-

lekter som kjenner "gender resolution" i entall her, ser også ut til å ha opprettholdt den i flertall i tallordet *tvau*. Tap av denne distinksjonen i flertall, kan ha ført til bortfallet av entallsformene, på tross av at morfologien tillater en distinksjon her. Dette passer også godt til språkhistorien: flertall er vanligst og først dokumentert, entall er sjeldnere og dukker senere opp i kildene. En ekstern grunn til at formene forsvant, kan ha vært at dansk, som stod i nær kontakt med norsk fra 1300-tallet, mistet denne formen for "gender resolution" tidlig (se Falk og Torp 1900:57).

I denne artikkelen har vi kun sett på nynorsk, og utelatt bokmålet, fordi de aktuelle formene ikke er belagt der (jf. Western 1921; Næs 1972; Faarlund et al. 1997). En mulig rest av fenomenet i moderne bokmål kan riktignok finnes i uttrykket "å gå hvert til sitt" som vanligvis brukes om et ektepar som skiller lag, eller om en sammensveiset gruppe, som et band. Et moderne internetteksempel illustrerer dette:

(20) Kutcher og Moore skilles. Går **hvert** (N.SG.) til sitt etter seks år som ektefolk (<http://www.dagbladet.no/>, 17.11.2011)

I (20) dreier det seg trolig om en leksikalsk rest som ikke er en del av det grammatiske systemet. Det blir derfor problematisk å snakke om "gender resolution" her: Intetkjønnsformen har også en synkron fonologisk motivasjon siden det neste ordet begynner på /t/. Disse formene er heller ikke akseptert av alle (Hans-Olav Enger, personlig kommunikasjon). Men uttrykket bærer noe likhet med eksempel (5) fra Heimskringla.

5 Sluttord

I denne artikkelen har vi sett på en spesiell type kongruensstrategi som inntreffer når det skal vises tilbake på referenter av forskjellig kjønn. Bruk av intetkjønn i en slik kontekst var en innovasjon i germansk og finnes blant annet i islandsk, norsk og tysk. Av de moderne germaniske standardspråkene finnes regelen kun i islandsk og færøysk, mens den i norsk og tysk kun inngår i substandardvarieteter. I denne artikkelen har vi forsøkt å kaste lys over hva som skjedde på vei til de moderne standardspråkene nynorsk og høytysk.

Mens intetkjønnsformene ble sett på som folkelige og derfor uegnede for et dannet tysk høyspråk, fremheves de av Ivar Aasen som bevaringsverdige former med klar referanse til systemet i norrønt og

i islandsk, og underforstått, i motsetning til dansk. Situasjonen i tyske og norske dialekter var derimot stikk motsatt det grammatikerne ønsket: På bakgrunn av eldre og nyere tyske dialektstudier er det god grunn til å anta at intetkjønnsformene var svært vanlige i tyske dialekter frem til 1800-tallet. I norsk tyder samtidig mye på at formene kun fantes i eldre mål allerede da Ivar Aasen samlet språkmaterialet til sitt *Landsmaal*: Dette viser seg ikke bare i hans egne arbeider (jf. *Prøver af Landsmaalet i Norge*), men gjenspeiler seg også i senere nynorskgrammatikker og dialektgrammatikker hvor formene karakteriseres som arkaiske og ikke lenger er påkrevd. Resultatet er altså i begge språk det samme: I motsetning til islandsk hvor disse formene er grammatikaliserte, er de forsvunnet ut av tysk og norsk standardspråk.

Ett paradoks er det verdt å dvele litt ved: Selv om intetkjønnsformene må kunne sees på som relativt marginale i en norsk sammenheng, er det ikke vanskelig å finne referanse til dem i grammatikkene. I høytyske grammatikker er derimot henvisninger til dialektformer svært uvanlig. Dette er tilfellet fordi grammatikkene springer ut av to høyst forskjellige normeringstradisjoner: I det språklig sett svært heterogene tyske språkområdet gikk bestrebelsene fra 1600-tallet ut på å skape et enhetlig, samlende høyspråk, på tross av dialektene, og med utgangspunkt i et rasjonalistisk språkideal (for eksempel gjennom eksklusjon av dobbelt negasjon, perifrastiske former, *constructio ad sensum* osv.), mens nynorskens oppgave var å gjenreise et norsk språk på bakgrunn av likhetstrekk nettopp i dialektene. Det er selvsagt en klar sosial dimensjon her: Mens tysk normering hadde et klart elitepreg hvor det ikke var vanskelig å ta avstand fra bestemte former i dagligtalen, måtte talsmennene for landsmålet spille på bondestandens språk. Når de fleste dialektene ikke lenger opprettholdt kjønnsbøyning i flertall og i bestemte entallsformer, var det vanskelig å få gjennomslag for den gamle intetkjønnsregelen.

Bibliografi

- Adelung, Johann Christoph. 1796. *Grammatisch-kritisches Wörterbuch der Hochdeutschen Mundart, mit beständiger Vergleichung der übrigen Mundarten, besonders aber der Oberdeutschen*. Zweyte vermehrte und verbesserte Ausgabe. Zweyter Theil, von F–L. Leipzig: Härtel.
- Aichinger, Carl Friedrich. 1754. *Versuch einer teutschen Sprachlehre*. Wienn: Johann Paul Kraus. [Nachdruck Hildesheim/New York 1972: Georg Olms.]

- Anonym. 1862. *Skrift og Umskrift i Landsmaalet*. Kristiania: Werner.
- Askedal, John Ole. 1973. *Neutrum Plural mit persönlichem Bezug im Deutschen unter Berücksichtigung des germanischen Ursprungs*. Oslo: Universitetsforlaget.
- Behaghel, Otto. 1928. *Deutsche Syntax: eine geschichtliche Darstellung. Band III: die Satzgebilde*. Heidelberg: Winter.
- Beito, Olav T. 1970. *Nynorsk grammatikk*. Oslo: Samlaget.
- Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Dal, Ingerid. 1966. *Kurze deutsche Syntax auf historischer Grundlage*. 3., verbesserte Auflage. Tübingen: Niemeyer.
- Davies, Winifred V., & Nils Langer. 2006. *The making of bad language: lay linguistic stigmatisations in German: past and present*. (VarioLingua 28.) Frankfurt a. M. u.a.: Lang.
- DiplNor = *Diplomatarium Norvegicum*. Digitalisert av Dokumentasjonsprosjektet ved Universitetet i Oslo. www.dokpro.uio.no/dipl_norv/diplom_felt.html
- Duden Grammatik. 2009. *Duden: die Grammatik. Unentbehrlich für richtiges Deutsch*. Herausgegeben von der Dudenredaktion. 8., überarbeitete Auflage. (Duden 4.) Mannheim u.a.: Dudenverlag.
- Ebert, Robert Peter, Oskar Reichmann, Hans-Joachim Solms & Klaus-Peter Wegera. 1993. *Frühneuhochdeutsche Grammatik*. (Sammlung kurzer Grammatiken deutscher Dialekte 12.) Tübingen: Niemeyer.
- Elspaß, Stephan. 2005. *Sprachgeschichte von unten: Untersuchungen zum geschriebenen Alltagsdeutsch im 19. Jahrhundert*. (Reihe Germanistische Linguistik 263.) Tübingen: Niemeyer.
- Elspaß, Stephan, & Nils Langer. 2012. Jespersen's cycle and the history of German negation – challenges from a sociolinguistic perspective. *Neuphilologische Mitteilungen* 113:275–292.
- Falk, Hjalmar, & Alf Torp. 1900. *Dansk-norskens syntax i historisk fremstilling*. Kristiania: Aschehoug.
- Fleischer, Jürg, & Oliver Schallert. 2011. *Historische Syntax des Deutschen: eine Einführung*. Tübingen: Narr.
- Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Givón, Talmy. 1970. The resolution of gender conflicts in Bantu conjunction: When syntax and semantics clash. In: *Papers from the Sixth Regional Meeting, Chicago Linguistic Society*, s. 250–261. Chicago: Chicago Linguistic Society.
- Goethe, Johann Wolfgang. 1889. *Aus meinem Leben. Dichtung und Wahrheit*. In: Goethes Werke. Werke im engeren Sinne. Abteilung 1. Band 26. Weimar: Böhlau.

- Goethe, Johann Wolfgang. 1898. *Wilhelm Meisters Lehrjahre*. In: Goethes Werke. Werke im engeren Sinne. Abteilung 1. Band 21. Weimar: Böhlau.
- Gotisk Bibel = *Die Gotische Bibel*, herausgegeben von Wilhelm Streitberg. Band 1: Der gotische Text und seine griechische Vorlage. Mit Einleitung, Lesarten und Quellennachweisen sowie den kleineren Denkmälern als Anhang. Mit einem Nachtrag von Piergiuseppe Scardigli. 7. Auflage. 2000. Heidelberg: Winter.
- Grimmelshausen, Hans Jacob Christoffel von. 1992. *Courasche*. In: Hans Jacob Christoffel von Grimmelshausen: Werke. Herausgegeben von Dieter Breuer I/2: 9–151.) Frankfurt am Main: Deutscher Klassiker Verlag.
- Grimmelshausen, Hans Jacob Christoffel von. 2010. *Lebensbeschreibung der Erzbetrügerin und Landstörzerin Courage / Der seltsame Springinsfeld: zwei simplicianische Romane*. Aus dem Deutschen des 17. Jahrhunderts und mit einem Nachwort von Reinhard Kaiser. Frankfurt am Main: Eichborn.
- Harnisch, Rüdiger. 2009. Genericity as a principle of paradigmatic and pragmatic economy: the case of German *wer* 'who'. In: Steinkrüger, Patrick O., & Manfred Krifka (red.), *On inflection*, s. 69–88. (Trends in Linguistics Studies and Monographs 184.) Berlin/New York: de Gruyter.
- Heggstad, Leiv. 1914. *Norsk grammatik for skuleungdom*. Kristiania: Norli.
- Heggstad, Leiv. 1931. *Norsk grammatikk. Større utgåve. Andre utgåva av Norsk Grammatikk for skuleungdom*. Oslo: Norli.
- Hægstad, Marius. 1879. *Norsk maallæra eller grammatik i landsmaalet*. Bergen: Nygaard.
- Indrebø, Gustav. 1951. *Norsk målsoga*. Bergen: John Grieg.
- Janson, Kristofer. 1876. *Han og ho*. Bergen: Giertsen.
- Janson, Kristofer. 1905. *Marit Skjølte*. Kristiania: Omtvedt.
- Jenstad, Tor Erik. 1985. *Sunndalsmålet. Eit grensemål på Indre Nordmøre*. Norske Studiar IV. Utgjevne av Norsk Målførearkiv, Universitetet i Oslo. Oslo: Novus.
- Kaiserchronik = *Die Kaiserchronik eines Regensburger Geistlichen*. Herausgegeben von Edward Schröder. (Deutsche Chroniken und andere Geschichtsbücher des Mittelalters 1, 1.). 1892. Hannover: Hahn.
- Knudsen, Trygve. 1949. *Pronomener*. Oslo: Universitetets studentkontor.
- Kollmer, Michael. 1987. *Die schöne Waldlersprach: von Wegscheid bis Waldmünchen, von Passau bis Regensburg. 1. Band: Lautliche und grammatische Beschreibung der Waldlersprache (mit Berücksichtigung der Sprache des Flachlandes rechts der Donau)*. Prackebach: Eget forlag.
- Larsson, Ludvig. 1901. *Friðþjófs Saga ins frækna*. Halle: Niemeyer.
- Laxness, Halldór. 1966. *Land til salgs*. Oversatt fra islandsk av Ivar Eskeland. Oslo: Tiden.
- Laxness, Halldór. 2011. *Atomstation*. Aus dem Isländischen von Hubert Seelow. Göttingen: Steidl.
- Laxness, Halldór. 2012. *Atómstöðin*. 5. útgáfa. Reykjavík: Vaka-Helgafell.

- Merkle, Ludwig. 1975. *Bairische Grammatik*. München: Heimeran. [Nachdruck 1990]
- Norsk Ordbok = *Norsk ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. 1966–. Oslo: Samlaget.
- Nygaard, Marius. 1905. *Norrøn syntax*. Kristiania: Aschehoug.
- Nýja testamenti Odds Gottskálkssonar* = Þetta er hid nya Testament, Jesu Christi eigenlig ord oc Euangelia huer hann sialfr predikadi oc kendi, hier i heime, Sem hans postular oc Gudz spialla menn sidan skrifudu. Roskilde 1540. (*Bækur.is*: http://baekur.is/bok/000037812/THetta_er_hid_nya_Testament (10.10.2013).)
- Næs, Olav. 1972. *Norsk grammatikk. Elementære strukturer og syntaks*. Oslo: Fabritius.
- Rafn, Carl Christian. 1829. *Fornaldar sögur Nordrlanda eptir gömlum handritum útgefnar af C.C. Rafn*. Annat bindi. København.
- Setelarkivet - Norsk ordbok*. Digitalisert av Dokumentasjonsprosjektet ved Universitetet i Oslo. <http://www.edd.uio.no/perl/search/search.cgi> (10.10.2013).
- Søreide, Lars E. 1930. *Nordfjordmålet*. (Norske maalføre XII.) Oslo: Norli.
- Tatian = Die lateinisch-althochdeutsche Tatianbilingue Stiftsbibliothek St. Gallen Cod. 56. Unter Mitarbeit von Elisabeth De Felip-Jaud herausgegeben von Achim Masser. (Studien zum Althochdeutschen 25.) 1994. Göttingen: Vandenhoeck & Ruprecht.
- Thráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen & Zakaris Svabo Hansen. 2012. *Faroese. An overview and reference grammar*. Second edition. Tórshavn/ Reykjavík: Faroese University Press/Institute of Linguistics, University of Iceland.
- Torp, Arne, & Lars S. Vikør. 2003. *Hovuddrag i norsk språkhistorie*. 3. utgåve. Oslo: Gyldendal.
- Walton, Stephen J. 1996. *Ivar Aasens kropp*. Oslo: Samlaget.
- Western, August. 1921. *Norsk riksmålsgrammatikk. For studerende og lærere*. Kristiania: Aschehoug.
- Wiener Genesis = *Die frühmittelhochdeutsche Wiener Genesis*. Kritische Ausgabe mit einem einleitenden Kommentar zur Überlieferung von Kathryn Smits. (Philologische Studien und Quellen 59.) 1972. Berlin: Erich Schmidt.
- Aasen, Ivar. 1848. *Det norske Folkesprogs Gammatik*. Kristiania.
- Aasen, Ivar. 1850. *Ordbog over det norske Folkesprog*. Kristiania.
- Aasen, Ivar. 1851. *Søndmørsk Grammatik*. Kristiania.
- Aasen, Ivar. 1853. *Prøver af Landsamaalet i Norge*. Kristiania.
- Aasen, Ivar. 1858. *Fridtjofs Saga: i Omskrift i det nyere Landsmaal*. Kristiania: Selskabet.
- Aasen, Ivar. 1864. *Norsk Grammatik*. Kristiania.
- Aasen, Ivar. 1873. *Norsk Ordbog*. Kristiania.

Aasen, Ivar. 1949. *Ervingen. Songspel i ei vending. Innleiing og ord-forklåringar ved Per Thorson*. Oslo: Gyldendal.

Aasen, Ivar. 1997. *Skrifter i samling. Trykt og utrykt. I-III*. Kristiania og København 1911–12. Elektronisk utgave. Oslo: Det norske Samlaget.

Lykilorð

kynjablöndun, kynjasamræmi, hvorugkyn, málstöðlun

Keywords

gender resolution, agreement, neuter, standardization

Útdráttur

Í greininni er fjallað um sérstaka leið eða reglu sem varðar kynjablöndun í sumum germönskum málum, þar sem hvorugkyn, í fleirtölu eða eintölu, getur vísað til hóps af fólki af ólíku kyni. Slík regla er t.d. til staðar í íslensku staðalmáli. Hana má líka finna á eldri stigum þýsku og norsku, svo og í nútímamállýskum, en ekki í staðalmálunum nú á tímum. Spurningin er hvernig þessi staða kom upp. Jafnvel þótt saga málstöðlunar í þýsku og nýnorsku sé mjög ólík er niðurstaðan að því leyti sú sama að reglan um að við kynjablöndun sé hvorugkyn haft um fólk af ólíku kyni hefur horfið í staðalmálunum. Að okkar mati orsakast þetta af ólíkum málfræðihefðum: Í Þýskalandi leiddi skynsemishyggja 17. og 18. aldar til þess að litið var á mállyskurnar sem ógn við staðalmálið og hvorugkynsmyndunum var hafnað með tilvísun til notkunar þeirra í “hversdagslífinu”. Þegar nýnorska mótaðist í andrúmslofti þjóðernisrómantíkur á 19. öld urðu mállýskurnar aftur á móti grundvöllur hins nýja staðalmáls og hvorugkynsmyndirnar þóttu þess vegna æskilegar í upphafi. En þar sem hvorugkynið varð æ sjaldgæfara vegna útjöfnunar í beygingakerfinu hvarf það síðan einnig úr staðalmálinu.

Abstract

In this paper we investigate a specific gender resolution rule: in some Germanic languages neuter plural and singular may refer to persons of different sex. For example, this rule is part of the Icelandic standard language. It can also be found in older stages of German and Norwegian, as well as in modern dialects, but not in their present-day standard languages, which raises the question how this situation came about. Although the standardization histories of German and Norwegian (Nynorsk) are very different, the result is equal insofar as the neuter resolution rule was not maintained in the standard languages. In our opinion this results from two different grammatical traditions: In the German 17th/18th century rational tradition, dialects were seen as a threat to the standard language, and the neuter forms were excluded with reference to their use in “common life”. In the more national-romanticist 19th century tradition of Nynorsk, however, the dialects were the fundament of the new

standard language, and the neuter forms were therefore favored in the beginning. But since the neuter construction became increasingly rare due to morphological leveling, it did not survive in the standard language.

Magnus Breder Birkenes / Jürg Fleischer
Institut für Germanistische Sprachwissenschaft
Philipps-Universität Marburg
Deutschhausstraße 3
D-35037 Marburg
birkenem@staff.uni-marburg.de / fleischj@staff.uni-marburg.de

Stephan Elspaß and Konstantin Niehaus

The standardization of a modern pluriareal language

Concepts and corpus designs for German and beyond

1 Overview

In this paper, we address the topic “Foundations of Language Standardization” from a theoretical and a methodological point of view, aiming at a more detailed account of standardization processes of a modern pluriareal language like German in the last two hundred years. Part 2 sets off with a short reflection on dominating notions and ideologies about language standardization, which will be contrasted with alternative views. Part 3 introduces a practical corpus design that follows from these considerations. In part 4, we will present three case studies from the standardization history of German which are based on our corpus design. We will compare results from studies on both formal and informal registers from historical (mainly nineteenth century) and present-day German with particular regard to the degree of ‘monocentricity’ or ‘pluriareality’ displayed in this corpora. Part 5 has a short conclusion.

2 Conflicting concepts on the standardization and the variation of modern standard languages

Traditionally, textbooks on the histories of ('big') modern standard languages such as English or German present rather monolithic standard varieties as results, closing stages, and sometimes even as objectives of standardization processes. The most powerful linguistic ideologies behind such conceptualisations of standard language and standardization can be identified as 'homogeneity' and 'standard language ideology'. The "dogma of homogeneity" consists of "a view of society in which differences are seen as dangerous and centrifugal and in which the 'best' society is suggested to be one without intergroup differences" (Blommaert & Verschueren 1998:194–5). Although Blommaert & Verschueren have multilingual societies in mind, the basic concept also applies to pluricentric and pluriareal settings. Closely related to homogeneity is the 'standard language ideology', i. e. "a bias toward an abstract, idealized homogeneous language, which is imposed and maintained by dominant bloc institutions and which names as its model the written language" (Lippi-Green 2012:67, mainly based on Milroy & Milroy 1985).

In the following, we will present two fairly new alternative concepts (or ideologies) to these two ideologies. The first one, 'language history from below', is an approach which in our view can be beneficial for standardization studies of *all* modern standard languages. The pluriareal concept will be less helpful for small languages like Icelandic, but, as we think, essential for the investigation of standardization processes of big(ger) languages. Such alternative approaches could, or rather should have consequences for the design of corpora, as we will demonstrate in Chapter 3, where German will serve as a case for illustration.

2.1 Standardization in a view 'from below'

In textbooks on language histories, standardization processes are often depicted in a teleological way, as long marches toward uniform standard varieties. Moreover, for the modern period such standardization (hi)stories are mostly accounts of the development of *printed* language and thus of conceptually written texts which were written

– or rather: composed – by professional writers. Such histories may be termed “language histories from above”. Opposed to such a dominant ‘bird’s eyes view’, there has been a strong plea for ‘language histories from below’ in the study of standardization (cf. Elspaß 2005, Elspaß, Langer, Scharloth & Vandebussche 2007). The idea is, firstly, to account for the language use of the vast majority of speakers of modern language communities and its role in standardization processes. Secondly, a view ‘from below’ advocates a radical shift from conceptually written texts to texts representing the “informal/oral type of linguistic conception”, in short “language of immediacy” (Oesterreicher 1997:193–4), as the starting point of standardization studies. “Language of immediacy” differs from “language of distance” with respect to various communicative parameters. Prototypically, such as in an intimate face-to-face conversation, it is characterized by a familiarity between communication partners in private settings and in spatio-temporal proximity, by a free development of topics, spontaneity and a maximum of cooperation between communication partners, etc. (as opposed to unfamiliarity of communication partners in public settings and in spatio-temporal distance, by fixed topics, reflection and a minimum of cooperation between communication partners, etc.). In the context of standardization studies, a shift to a view ‘from below’ therefore requires a principal focus on historical texts which are as close to ‘language of immediacy’ and their reflections in grammar and lexis as possible. Private letters, in particular, fulfil this requirement (cf. Elspaß 2012, Elspaß in print). Elspaß (2005) and Nobels (2013) are the first two monograph-size studies applying this approach to periods in the standardization phases of two different Germanic languages, namely nineteenth century German and seventeenth century Dutch respectively.¹

The second aspect, the inclusion of conceptually oral language (and its inherent variation), is at the heart of the ‘principle of viability’ which Vilmos Ágel has formulated as a fundamental requirement for the analysis of linguistic phenomena.

Jede linguistische Beschreibung (bzw. Erklärung) muss mit der Beschreibung (bzw. Erklärung) der Geschichte des zu beschreibenden (bzw. zu erklärenden) Phänomens konform sein.

¹ The approach has also proved fruitful for studies on older language periods. Such studies include Schulte (2009) on the younger Fupark, Ernst & Elspaß (2011) on Old High German glosses and Graser & Tlusty (2012) on Early New High German street songs.

'Every linguistic description (or explanation) has to fit into the description (or explanation) of the history of the phenomenon under description (or explanation).'

(Ágel 2001:319, 2003:2)

2.2 Standardization of a modern pluriareal language: the case of German

For 'big' languages such as English or German, an alternative view on standardization would also account for the co-existence of different standard language varieties (such as 'the Englishes') and for the idea of an internal variability of these varieties.

A first step away from the notion of a 'uniform' standard language was to conceptualize German as a pluricentric language. Up until the 1980s, many authors saw the standard variety in West Germany as the only legitimate German standard language. According to the widespread notion of pluricentricity, going back mainly to Clyne (1992), (pluricentric) German consists of the three 'national varieties' in Germany, Austria and Switzerland and other varieties of standard German in smaller countries and regions (Liechtenstein, Luxemburg, East Belgium and South Tyrol, cf. Ammon, Bickel, Ebner et al. 2004). This concept, however, has been criticized for different reasons. Firstly, it is an entirely political concept, based on the notion of *Überdachung* of the language area by a political state. As for the recent history of German, this would have had the somewhat odd consequence that on 3 October 1990, the German language has lost an entire national variety, namely GDR German, literally overnight. Secondly, from a linguistic perspective, one may ask whether 'national varieties of German' are really varieties? One can argue that less than two per cent of variation in standard German lexis and pronunciation and even less variation in grammar does hardly make a 'variety'. A third problem from the linguistic side can be raised in view of empirical evidence. While some national variants do exist, in a lot of cases the diatopic extension of Standard German variants is not limited to national borders. They are not absolute variants, as terms such as 'Germanism', 'Helvetism' or 'Austriazism' may suggest. Rather, in many cases they are relative variants, i.e. they are employed in a part of the respective country or by a part of the members of the speech community only. Moreover, diatopic standard variants do cross borders, i.e. their distribution is

not restricted to one country alone. ‘Pure’ national variants may even be the exceptional case in the German speaking countries. Figure 1 (adopted from Scherr & Niehaus 2013:78) may serve to illustrate a more realistic picture of the areal variation in the German-speaking countries.

Because of such empirical evidence, some researchers would prefer to talk of Standard German as a pluriareal language (e.g. Wolf 1994, Scheuringer 1997). This is the approach that we adopt, hence step 2 on the way to our corpus design.

The competing concepts of ‘pluricentric German’ vs. ‘pluriareal German’ have sparked a controversy (mainly between German and Austrian scholars) about the historical adequacy of the two models (cf. von Polenz 1999, Reiffenstein 2001). In our view, the debate would profit from empirical evidence from present-day as well as historical data, as we will try to demonstrate in section 4 of this paper.

<u>area A</u> variant x: 70% variant y: 30%	<u>area C</u> variant x: 20% variant y: 20% variant z: 60%
<u>area B</u> variant x: 30% variant y: 70%	

Fig. 1. Relative areal variation in a pluriareal language community (example)

3 A corpus design for a modern pluriareal language

What follows from such considerations and alternative concepts for the design of corpora to allow for painting a more realistic (or rather adequate) picture of standardization processes? The requirement for such a corpus is that it contains historical as well as present-day data which can help us to explain the areal variation and also the on-going change of a modern language in a long-term perspective. As for German, we introduce a corpus design which both accounts for the pluriareality of German and considers historical data ‘from above’ as well as ‘from below’.

In a ‘view from above’, we look at written (standard) language

in regional newspapers, thus printed texts. What is new about our project on modern German is that we work with regionally balanced corpora. Surprisingly, whereas corpora on earlier periods of German up until the eighteenth century are all regionally balanced,² there are no such corpora for the nineteenth and the twenty-first century. So for present-day German, we use the corpus of our *Variantengrammatik* project which is a joint project of a German, an Austrian and a Swiss team, which are now based at the universities of Zurich, Salzburg and Graz.³ For the nineteenth and the beginning of the twentieth century, the Salzburg team compiled a small corpus of newspaper texts which we refer to as the *Kaiserreich* ('Empire') corpus. The two corpora are, of course, very different in size: The corpus of the *Variantengrammatik* project comprises more than 600 million words, the *Kaiserreich* corpus so far only 100,000 tokens.

In a view 'from below', we focus on spoken language data or written data which are as close to speech as possible. For present-day German, we used the so-called Pfeffer corpus with interview data mostly from the 1960s (Pfeffer & Lohnes 1984, with a total of 670,000 words) and we also looked at maps from our *Atlas of Colloquial German* (AdA) (cf. Elspaß & Möller 2003ff.). The historical data 'from below' consist of a 880,000 words corpus of nineteenth century emigrant letters (mainly based on the corpus of Elspaß 2005). Again, both corpora are (more or less) regionally balanced.

4 Case studies

Based on the principle that variation is inherent to a modern standard language and with regard to our corpora, our case studies focus on the following research questions:

- How much variation did printed German allow in the nineteenth century?
- Is (and was) the variation in non-standard German similar to or rather different from printed standard German? Which tenden-

² Such as the corpora of the Middle High German (1050–1350) and the Early New High German (1350–1650) grammar or the German Manchester Corpus (1650–1800), cf. Paul (2007), Reichmann & Wegera (1993) and Scheible et al. (2011).

³ The project is funded by the major research grant organizations of the three countries: the *Schweizerischer Nationalfonds* (SNF) [100015L-134895], the *Deutsche Forschungsgemeinschaft* (DFG) [EL 500/3-1] and the *Austrian Science Fund* (FWF) [I 716-G18].

cies in standard and non-standard German are discernible in the last 150 years?

- Does printed (standard) German change from a stage of ‘monocentricity’ to ‘pluricentricity’, as von Polenz (1989:15) claimed?

Von Polenz’ hypothesis needs some explanation here. He assumed that socio-pragmatic criteria played a much bigger role for New High German than for any other period in the history of German. Thus, in his periodization model for New High German, the time between 1770 to 1870 is labelled ‘Establishment of an educated bourgeois variety’ (“Etablierung der bürgerlichen Bildungssprache”), whereas the ‘Establishment of the modern German standard language’ took place between 1870 and 1950⁴ and was characterized by ‘monocentric tendencies’ (“Etablierung der modernen Standardsprache, monozentrische Tendenzen”). In contrast, the time after 1950, which he simply called ‘Contemporary German’ (“Gegenwartsdeutsch”), could be portrayed as the time of ‘a polycentric drift’ (“polyzentrische Weiterentwicklungen”). Whereas the first phase saw rapidly increasing literacy rates, the second phase can be identified as the crucial period of time of the emergence of a national literature and the standard language ideology (cf. von Polenz 1989:16–27).

To illustrate our approach, we present three case studies, one involving a lexical variable, and two focussing on grammatical variables in standard German. For grammatical variation, we will investigate the serialization patterns in three-verb clusters (4.1) and discontinuous pronominal adverbs (4.2). For lexical variation, we will look at the distribution of variants for ‘Saturday’ (4.3).

4.1 Grammatical variation I: serialization in three-verb clusters

The serialization of verb clusters in German has received much attention in the research literature.⁵ In our case study, we will focus on

⁴ This is noteworthy, because traditional German textbooks on the history of German would have it that the German standard language had been established by the end of the 18th century.

⁵ To name but a few: Bech (1955/57), Van de Velde (1981) and Askedal (1986) discussed the most important theoretical aspects in detail. The benchmark in historical research on verb clusters is still the study by Hård (1981), which was taken up by Ágel (2001) to exemplify some ‘historical principles in today’s grammar’.

a particular instance of word order variation in verb clusters. Variation may occur in different respects. We will restrict ourselves to the variation of the finite auxiliary verbs *haben* and *werden* in three-verb clusters, all in subordinate clauses with so-called “Ersatzinfinitiv” (‘substitutive infinitive’) constructions.

A few notes are necessary on some basic theoretical aspects and the ‘rules’ as codified by today’s grammars. In three-verb clusters containing a past participle, the finite verb in subordinate clauses would be put at the end of the clause:

- (1) ... *damit er gesehen³ werden² kann¹*
 ... so that he seen^{3-inf} become^{2-inf} can^{1-fin}
 ‘so that he can be seen’

The dependencies within the verbal cluster therefore usually run from right to left. This serialization pattern has proved to be relatively constant in the history of New High German (cf. Härd 1981:13, 167; Ágel 2001:322), and it is also given as a codified present-day ‘basic rule’ by the most authoritative present-day grammar, the Duden grammar (cf. Duden grammar 2009:474). However, this rule does not apply for cases in which the second dependency is filled by a modal verb. In such clusters, modal verbs do not have past participle forms, but an infinitive. This is considered to be a substitution of the past participle, hence the name “Ersatzinfinitiv” (‘substitutive infinitive’). This grammatical feature also alters the word order; here, the finite verb occurs in the initial position:

- (2) ... *dass er das Spiel hat¹ sehen³ können²*
 ... that he the match has^{1-fin} watch^{3-inf} can^{2-inf}
 ‘that he has been able to watch the match’

The Duden grammar (2009:474) pronounces this a ‘special rule’ in present-day standard German, which is compulsory in cases where the finite verb is *haben*. If the finite verb is *werden*, both positions (initial and final) are possible (cf. Duden grammar 2009:474–475). In contrast to this variation of the finite verbs, the order of the infinite parts of the cluster is restricted to 3–2.

It is noteworthy, that a third possible pattern is not mentioned in the Duden grammar, namely 3–1–2. Here, the finite verb is put in between the infinite parts (so-called “Zwischenstellung”). According to Härd (1981:117), this variant gradually fell into disuse as from the seventeenth century and was virtually extinct in the nineteenth century. Although it is not codified in grammars of standard German,

the 3–1–2 order is considered to be a feature of Austrian standard German (cf. Härd 1981:178; Zeman 1988:78 and 2002:51–53; Askedal 1989:25; Bader & Schmid 2009:219).

So what findings do the corpus studies provide? Which serialization patterns occur in standard language usage? The major caveat which concerns the validity of our results is that all corpora used for this study except for the corpus *Variantengrammatik* are rather small, compared to today’s standards in corpus-linguistics, and that, in general, three-verb clusters are a rather rare phenomenon, let alone variation in these clusters. This means that we have to deal with small numbers of variants, particularly in our historical corpora.

The *Variantengrammatik* corpus, however, with ca. 640 million words compiled from some 60 regional newspapers in the German-speaking countries, clearly shows pluriareal variation (also cf. Scherr & Niehaus 2013), cf. figure 2.

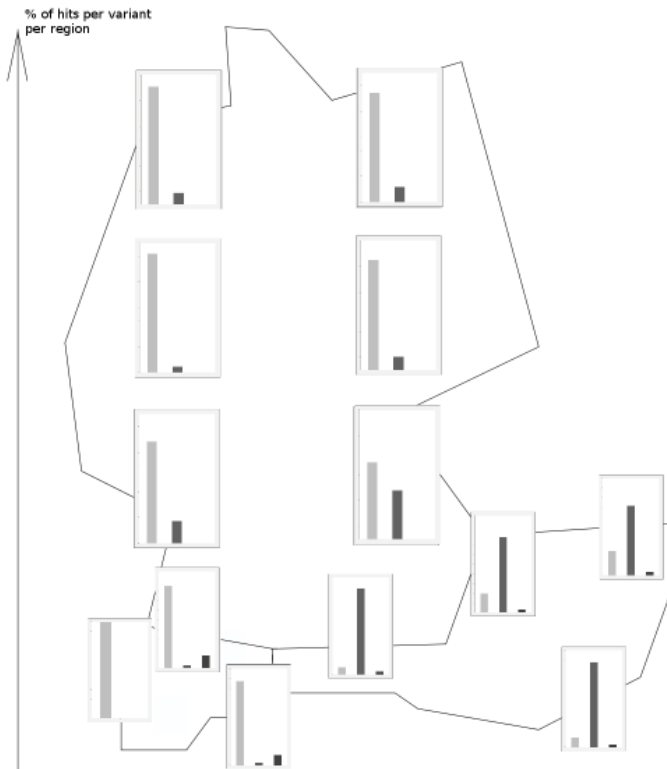


Fig. 2. Serialization in verb clusters in the corpus *Variantengrammatik* (main German-speaking countries) (light grey: 1-3-2; dark grey: 3-1-2; black: 3-2-1)

These proportions (and also a look at true positives) show that in this corpus all three variants are well documented. The 3–1–2 order is very noticeable in the south-eastern regions of the German-speaking language area. In Austria, it is even the dominant (but not absolute) variant. As leading grammar books such as the Duden grammar consider themselves *grammars of usage*, it is surprising that the 3–1–2 variant has not been marked as a standard variant yet, even more so, as it is also highly common in the adjoining areas of Liechtenstein, South Tyrol and the south-east of Germany (i.e. large parts of Bavaria).

Is this variation and this particular distribution a result of recent developments in standard German? As mentioned above, three-verb clusters are relatively rare, and we cannot expect high numbers of instances in the three smaller corpora. The *Kaiserreichkorpus* gives no hints as to regional variation in nineteenth century printed standard German, there is evidence for the 1–3–2 order only. This may support the monocentric tendencies after 1870, which von Polenz mentioned. But note that the corpus is very small, so that the lack of data for variants may simply be due to limitations of the corpus size.

For the view ‘from below’, the Pfeffer Corpus and the Emigrant letter corpus, render a different picture. In spite of the low numbers, it quickly becomes clear that the areal distribution of the three-verb cluster variants resembles the spread of the individual variants in present-day standard German. Moreover, there has been a continuity of usage in serialization patterns since the nineteenth century. In southern Germany and Austria the 3–1–2 order is the only variant that occurs in the Emigrant letter corpus (5 matches), while in northern Germany also the initial position of the finite verb (1–3–2) shows up and is highly dominant (> 70% within the variable). The distribution in the Pfeffer corpus is very similar, with the only difference that the 1–3–2 order seems to have become stronger in the southern areas. With all due caution, we may conclude then, firstly, that the areal distribution of the three variants in the *Variantengrammatik* corpus is matched by the pattern that we find in the non-standard corpora and, secondly, that there has been a stable pattern in non-standard German over the last 150 years.

From the results of the corpus study, it becomes evident that an explanation for the variation in verb clusters must be grammatically as well as historically and sociolinguistically founded. Grammatically speaking, prosodic-induced focusshifts (esp. focus on the lexical verb) might have an effect on word order in verbal clusters, according to

dialectal and diachronic studies (cf. Dubenion-Smith 2010, 2011; Sapp 2011:204–205). Sociolinguistically speaking, Bader & Schmid believe that the dominance of the 3–1–2 order in Austrian standard German could be caused by a more ‘casual’ handling of standard-norms in Austria (cf. Bader & Schmid 2009:219), but this seems disputable as perhaps also the opposite seems likely: the 3–1–2 order might be considered a ‘national’ norm by Austrian speakers (cf. Patocka 1997:281). This does not change the empirical fact that this word order variant is commonly used in standard context in other German-speaking countries as well and stands as a general feature of south-eastern standard German.

4.2 Grammatical variation II: discontinuous pronominal adverbs

Another grammatical construction which variationist studies have paid a lot of attention to in recent years – also for other Germanic languages such as Dutch, Frisian and Afrikaans⁶ – is the pronominal adverb and its discontinuous variants. Example (3) presents a pronominal adverb in its continuous variant, examples (4) and (5) illustrate the two most common discontinuous forms⁷ of the German pronominal adverb *damit* ‘with this’.

- (3) *damit habe ich nichts zu tun*
 this-(PRO)-with have I nothing to do
 ‘I have nothing to do with this’
- (4) *da habe ich nichts mit zu tun*
 this-(PRO) have I nothing with to do
- (5) *da habe ich nichts damit zu tun*
 this-(PRO) have I nothing this-(PRO)-with to do

In the discontinuous variant (4), the pronominal (usually *da-*, but also *wo-* and *hier-*) and the prepositional element of the adverb are separated in such a way that at least one other part of the sentence is moved in between. In most of such cases, the pro element is moved to the be-

⁶ Cf. Fleischer (2002a, b) and Negele (2012) for an overview.

⁷ Sometimes the first pronominal element is immediately followed by a ‘full’ pronominal adverb, resulting in a ‘contact’ position: *da damit habe ich nichts zu tun* ‘this (PRO) this-(PRO)-with have I nothing to do’. In another variant, the pronominal element is dropped: *__habe ich nichts mit zu tun* ‘Ø (PRO) ‘have I nothing with to do’.

ginning of the sentence.⁸ A similar variant is a construction in which the pro-element seems to be ‘doubled’ (5). Except for some dialects, this ‘split construction’ occurs only when the preposition begins with a consonant (e.g. *mit*, *zu*, *von*). In pronominal adverbs connecting *da-* or *wo-* and a preposition with an initial vowel, an *r* is inserted (e.g. *darauf*, *worüber*). Most grammars of standard German, such as the Duden grammar (2009:581), do not consider discontinuous or other variants of pronominal adverbs as standard. Moreover, according to some authors, e.g. Eisenberg (2004:198), the ‘double PRO construction’ is restricted to cases in which the preposition begins with a vowel; as a consequence the vowel in the second *da-* is usually dropped, e.g. example (6).

- (6) da habe ich nicht d[a]ran gedacht
 this-(PRO) have I not this-(PRO)-*r*-of thought
 ‘I have not thought of this’



Fig. 3. Areal distribution of discontinuous pronominal adverbs in the corpus *Varianten-grammatik* (left column: split construction, right column: double pro construction)

⁸ The prepositional element, however, does not always move to the last position. Therefore, the term ‘preposition stranding’, which is sometimes used in grammars (cf. Eisenberg 2004:198) or the research literature (Fleischer 2002a, b), can be misleading.



Fig. 4. Areal distribution of split construction in 19th and 20th c. colloquial German (dark grey area; symbols indicate southernmost documented places of use in the respective century), adapted from Negele (2012:121)

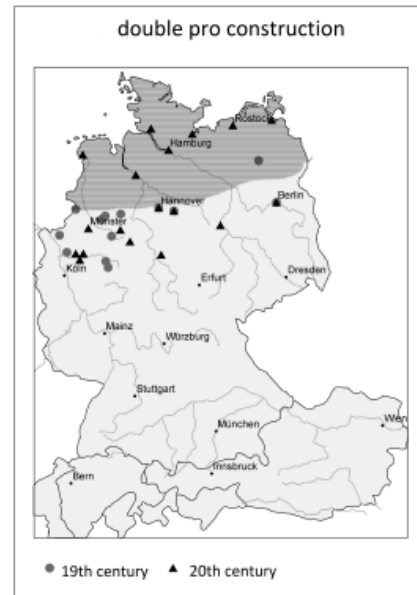


Fig. 5. Areal distribution of double pro constructions in 19th and 20th c. colloquial German (light grey area; symbols indicate northernmost documented places of use in the respective century), adapted from Negele (2012:122)

‘Double pro constructions’ in which the preposition begins with a consonant, however, are also widespread. Negele (2012:242–44) argues that such ‘double pro constructions’ as in example (5), as well as ‘split construction’, are standard German, and she gives evidence from standard German texts which hint at a clear areal distribution. ‘Double pro constructions’ are employed in the north, ‘split construction’ in the south of the German speaking countries. In spite of the relatively few instances in the *Varietengrammatik* corpus, Negele’s findings can to a certain extent be corroborated by the results of our corpus study. Figure 3 shows a preference for the ‘split construction’ in the north and west of Germany, whereas the ‘double pro constructions’ appear mainly in the southern parts of the German language area. The (small) *Kaiserreichkorpus* gives no hints as to a regional variation in nineteenth century printed standard German. Not a single instance of a discontinuous construction could be found.

Again, it will be safe to assume that discontinuous constructions have not fallen from heaven into present-day German, nor will the

regional distribution be coincidental: In fact, Negele's data from her corpus studies on nineteenth and twentieth century colloquial German, mainly based on the emigrant letter and the Pfeffer corpus (cf. fig. 4 and 5, from Negele 2012:121–2, maps 21 and 22), confirm that the present-day north-south division in the preferred use of discontinuous pronominal constructions in standard German echoes the distribution in nineteenth as well as twentieth century colloquial German.

4.3 Lexical variation: *Sonnabend* and *Samstag* 'Saturday'

Our last case study looks at a prominent instance of lexical variation in standard German, the different names for the last day of the week. The largest dictionary concerned with regional lexical variation in standard German is the *Variantenwörterbuch* ('dictionary of [standard German] variants', Ammon, Bickel & Ebner et al. 2004). For the variable 'Saturday', it states that *Sonnabend* is the common variant in northern Germany (ibid.:653, 724), while *Samstag* is used in central and southern Germany, Austria and Switzerland, with the latter variant expanding more and more into the centre and to the north (ibid.:653).

Again, our data from the *Variantengrammatik* corpus draw a somewhat different picture (cf. Fig. 6). Although the *Variantenwörterbuch's* statement (and codification) can be confirmed in general, the corpus study renders one notable difference. In central Germany the usage of *Sonnabend* is very much restricted to the middle-east. (There are a few instances of *Sonnabend* in newspapers from Luxemburg, central-west Germany and even southern Germany, but the respective absolute numbers are so low that they would not show up in the columns of our map, cf. fig. 6.)

This areal distribution in figure 6 is repeated in the other three corpora, the Pfeffer Corpus and the two historical corpora for (standard and non-standard) 19th century German. Thus, the overall picture can be interpreted as another case of 'variational continuity'. (This time, however, both variants were and are marked as standard German by the codices.)

The areal (and non-pluricentric) continuity probably reaches back to the Early Modern period. The structure of this variable might be explained by usage conventionalized in times when denomination was a strong factor in the division of regions, as the *Sonnabend*-area

is very similar to the post-1648 regions of Lutheran denomination (cf. Putzger 2006:107). Thus, one could speculate that the two variants originally represented social in-group markers, an effect which then could have been enhanced by restrictions for Lutheran and catholic printers during the counter-reformation, as testified for other variables (cf. Rössler 2005:364–367).

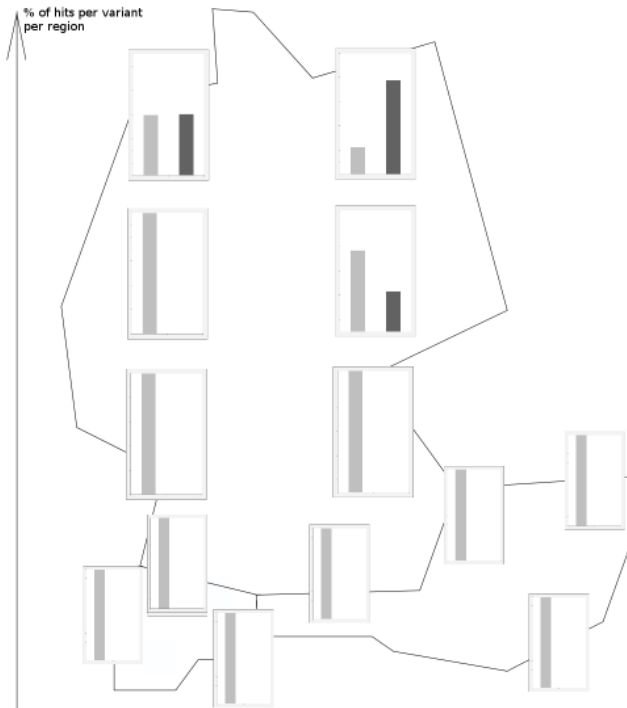


Fig. 6. Lexical variation for 'Saturday' in the corpus *Varietengrammatik* (light: *Samstag*; dark: *Sonnabend*)

5 Conclusion

In this paper, we argued for a consideration of standard as well as non-standard language data from present-day and historical corpora in the study of language standardization. With reference to data from German, we adopted the pluriareal rather than the pluricentric model of 'big' standard languages, as the pluriareal concept seems to depict the factual language variation more adequately and realistically.

As to our research questions, we can now answer them as follows:

- How much variation did printed standard German allow in the nineteenth century? –
The *Kaiserreich* corpus does not indicate any relevant areal variation. The almost complete lack of variation might, however, be due to the small size of the nineteenth century standard German corpus.
- Does printed standard German change from ‘monocentricity’ to ‘pluricentricity’ (cf. von Polenz 1989:15)? –
For the period from 1870 to the First World War, von Polenz’ notion of ‘monocentricity’, which corresponds to a low degree of areal variation, can be supported – at least on the basis of the small corpus that we have so far. Present-day standard German, by contrast, allows for more areal variation. Corresponding findings reported in the research literature (cf. Schmidlin 2011:65; Klein 2013:26) can be supported by the results of our corpus studies. The rise in variation could be attributed to the increasing informality in newspaper style which has become noticeable particularly after the mid-twentieth century (cf. Betz 2006:183–185), as informality is closely associated with the degree of areal variation in language. However, as we can find patterns in the regional distribution of linguistic variants which in many cases are not only determined by political borders, a change ‘from monocentricity to pluriareality’ would describe the linguistic situation more adequately.
- Is (and was) the variation in non-standard German similar to or rather different from printed standard German? Which tendencies in standard and non-standard German are discernible in the last 150 years? –
As for the last 150 years, variation in non-standard German and printed standard German can differ substantially in a historical perspective, as our case studies indicate. In two of our three case studies (4.1 and 4.2), the variational patterns seem to be more similar to each other in the present, not least due to a tendency towards informal styles in standard German (as we just argued). A major finding of the case studies on the serialisation of verb-clusters and the discontinuous pronominal adverbs is that pluriareal patterns which have emerged in today’s standard German can be traced back to present-day as well as earlier non-standard usage. Further corpus studies are needed, of course, to establish whether these are rather singular cases or whether they represent a more general pattern in the recent standardization history.

We hope that our case studies could illustrate one important methodological point and its consequences for research. Studies on standardization and of standard as well as non-standard variation require more and bigger corpora, particularly from the ‘age of standardization’, i.e. the nineteenth century, onwards. In addition, for big languages such as German we also need more regionally balanced text corpora. As the two case studies on grammatical phenomena clearly demonstrated, one major problem can be a ‘descriptive gap’ in standardization studies as well as in grammaticography, when variants are overlooked or ignored because of a lack of adequate corpora. This, in turn, limits the validity of such studies, particularly with respect to an adequate description and explanation of linguistic phenomena in terms of the ‘principle of viability’ (cf. section 2.1). The bigger present-day as well as historic corpora are, and the more they reflect the areal division of a speech community, the better they can help to generate formulations of empirically based ‘diachronic principles of grammar’. The comparatively large *Variantengrammatik* corpus gives us a slight hint at what treasures may be hidden in big historical corpora – even if they turn out to be small in terms of absolute numbers.

References

- AdA = Elspaß, Stephan & Robert Möller. 2003–2013. *Atlas zur deutschen Alltagssprache*. <http://www.atlas-alltagssprache.de/>.
- Ágel, Vilmos. 2001. Gegenwartsgrammatik und Sprachgeschichte. Methodologische Überlegungen am Beispiel der Serialisierung im Verbalkomplex. *Zeitschrift für Germanistische Linguistik* 29:293–318.
- Ágel, Vilmos. 2003. Prinzipien der Grammatik. In: Lobenstein-Reichmann & Oskar Reichmann (eds.), *Neue historische Grammatiken. Zum Stand der Grammatikschreibung historischer Sprachstufen des Deutschen und anderer Sprachen*, pp. 1–46. Tübingen: Niemeyer.
- Ammon, Ulrich, Hans Bickel, Jakob Ebner, et al. 2004. *Variantenwörterbuch des Deutschen. Die deutsche Standardsprache in Österreich, der Schweiz und Deutschland sowie in Liechtenstein, Luxemburg, Ostbelgien und Südtirol*. Berlin/New York: de Gruyter.
- Askedal, John Ole. 1986. Zur vergleichenden Stellungsfelderanalyse von Verbalsätzen und nicht-verbale Satzgliedern im Deutschen. *Deutsch als Fremdsprache* 23:269–273 and 342–348.
- Bader, Markus & Tanja Schmid. 2009. Verb clusters in colloquial German. *Journal of comparative Germanic Linguistics* 12:175–228.
- Bech, Gunnar. 1955/57[1983]. *Studien über das deutsche Verbum Infinitum*. 2nd vol. Tübingen: Niemeyer.

- Betz, Ruth. 2006. *Gesprochensprachliche Elemente in deutschen Zeitungen*. Raddolfzell: Verlag für Gesprächsforschung. <http://www.verlag-gespraechsforschung.de/2006/betz.htm> [accessed 12.07.2013].
- Blommaert, Jan & Jef Verschueren. 1998. The role of language in European nationalist ideologies. In: Schieffelin et al. (eds.), pp. 189–210.
- Clyne, Michael. 1995. *The German language in a changing Europe*. Cambridge: Cambridge University Press.
- Dubenion-Smith, Shannon A. 2010. Verbal Complex Phenomena in West Central German: Empirical domain and multi-causal account. *Journal of Germanic Linguistics* 22,2:99–191.
- Dubenion-Smith, Shannon A. 2011. Der Verbalkomplex im Schlesischen. *Zeitschrift für Dialektologie und Linguistik* 78:281–320.
- Duden grammar (2009) = *Duden. Die Grammatik. Unentbehrlich für richtiges Deutsch. Herausgegeben von der Dudenredaktion*. 8th, rev. ed. Mannheim/Zürich: Dudenverlag.
- Dürscheid, Christa, Stephan Elspaß & Arne Ziegler. 2011. Grammatische Variabilität im Gebrauchsstandard: das Projekt 'Variantengrammatik des Standarddeutschen'. In: Konopka et al. (eds.), pp. 123–140.
- Eisenberg, Peter. 2004. *Grundriss der deutschen Grammatik*. 2nd, rev. ed. Vol. 2: Der Satz. Stuttgart: Metzler.
- Elspaß, Stephan. 2005. *Sprachgeschichte von unten. Untersuchungen zum geschriebenen Alltagsdeutsch im 19. Jahrhundert*. Tübingen: Niemeyer.
- Elspaß, Stephan. 2012. The Use of Private Letters and Diaries in Sociolinguistic Investigation. In: Hernández-Campoy, Juan Manuel & Juan Camilo Conde-Silvestre (eds.), *The Handbook of Historical Sociolinguistics*, pp. 156–169 Oxford: Wiley-Blackwell.
- Elspaß, Stephan (in print). Private letters as a source for an alternative history of Late Modern German. In: Auer, Anita, Daniel Schreier & Richard J. Watts (eds.), *Letter Writing and Language Change*. Cambridge: Cambridge University Press.
- Elspaß, Stephan, Nils Langer, Joachim Scharloth & Wim Vandenbussche (eds.). 2007. *Germanic language histories 'from below' (1700–2000)*. Berlin/New York: de Gruyter.
- Ernst, Oliver & Stephan Elspaß. 2011. Althochdeutsche Glossen als Quellen einer Sprachgeschichte ‚von unten‘. *NOWELE* 62/63:249–283.
- Fleischer, Jürg. 2002a. *Die Syntax von Pronominaladverbien in den Dialekten des Deutschen. Eine Untersuchung zu Preposition Stranding und verwandten Phänomenen*. Stuttgart: Steiner.
- Fleischer, Jürg. 2002b. Preposition Stranding in German Dialects. In: Barbiers, Sjef, Leonie Cornips & Susanne van der Kleij (eds.), *Syntactic Microvariation*. (= Meertens Institute Electronic Publications in Linguistics, 2), pp. 116–151. Amsterdam: Meertens Institute. <http://www.meertens.knaw.nl/books/synmic/pdf/fleischer.pdf> [accessed 12.07.2013].
- Graser, Helmut & B. Ann Tlustý. 2012. Sixteenth Century Street Songs and

- Language History 'from below'. In: Langer, Nils, Steffan Davies & Wim Vandenbussche (eds.): *Language and History, Linguistics and Historiography. Interdisciplinary Approaches*, pp. 362–388. Oxford [etc.]: Lang.
- Hagemann, Jörg, Wolf Peter Klein & Sven Staffeldt (eds.). 2013. *Pragmatischer Standard*. Tübingen: Stauffenburg.
- Härd, John Evert. 1981. *Studien zur Struktur mehrgliedriger deutscher Nebensatzprädikate. Diachronie und Synchronie*. Göteborg: Acta Universitatis Gothoburgensis.
- Klein, Wolf Peter. 2013. Warum brauchen wir einen klaren Begriff von Standardsprachlichkeit und wie könnte er gefasst werden? In: Hagemann, Jörg et al. (eds.), pp. 15–33.
- Konopka, Marek, Jacqueline Kubczak, Christian Mair, František Štícha & Ulrich H. Waßner (eds.). 2011. *Grammar and Corpora 2009. Third International Conference, Mannheim 22.–24.09.2009*. Tübingen: Narr.
- Lippi-Green, Rosina. 2012. *English with an Accent. Language, ideology, and discrimination in the United States*. 2nd edition. London/New York: Routledge.
- Milroy, James & Lesley Milroy. 1985. *Authority in Language. Investigating Language Prescription and Standardisation*. London/New York: Routledge and Kegan Paul.
- Negele, Michaela. 2012. *Varianten der Pronominaladverbien im Neuhochdeutschen. Grammatische und soziolinguistische Untersuchungen*. Berlin/New York: de Gruyter.
- Nobels, Judith. 2013. *(Extra)Ordinary Letters. A view from below on seventeenth-century Dutch*. Utrecht: LOT.
- Oesterreicher, Wulf. 1997. Types of orality in text. In: Bakker, Egbert & Ahuvia Kahane (eds.), *Written voices, spoken signs*, pp. 190–214. Cambridge, MA/London: Harvard University Press.
- Patocka, Franz. 1997. *Satzgliedstellung in den bairischen Dialekten Österreichs*. Frankfurt a. Main [u.a.]: Peter Lang.
- Paul, Hermann. 2007. *Mittelhochdeutsche Grammatik*. Rev. by Thomas Klein et al. Tübingen: Niemeyer.
- Pfeffer, J. Alan & Walter F. W. Lohnes. 1984. *Grunddeutsch. Texte zur gesprochenen deutschen Gegenwartssprache*. Tübingen: Niemeyer.
- von Polenz, Peter. 1989. Das 19. Jahrhundert als sprachgeschichtliches Periodisierungsproblem. In: Cherubim, Dieter & Klaus J. Mattheier (eds.), *Voraussetzungen und Grundlagen der Gegenwartssprache*, pp. 11–30. Berlin/New York: de Gruyter.
- von Polenz, Peter. 1999. Deutsch als plurinationale Sprache im postnationalistischen Zeitalter. In: Gardt, Andrea, UlrikeHaß-Zumkehr & Thorsten Roelcke (eds.): *Sprachgeschichte als Kulturgeschichte*, pp. 115–132. Berlin / New York: de Gruyter.
- Putzger. 2006. *Historischer Weltatlas. Kartenausgabe*. Berlin: Cornelsen.

- Reichmann, Oskar & Klaus-Peter Wegera (eds.). 1993. *Frühneuhochdeutsche Grammatik*. Von Robert Peter Ebert, Oskar Reichmann, Hans-Joachim Solm und Klaus-Peter Wegera. Tübingen: Niemeyer.
- Reiffenstein, Ingo. 2001. Das Problem der nationalen Varietäten. Rezensionssatz zu Ulrich Ammon: Die deutsche Sprache in Deutschland, Österreich und der Schweiz. Das Problem der nationalen Varietäten. Berlin/New York 1995. *Zeitschrift für deutsche Philologie* 120:78–89.
- Rössler, Paul. 2005. *Schreibvariation – Sprachregion – Konfession. Graphematik und Morphologie in österreichischen und bayerischen Drucken vom 16. bis ins 18. Jahrhundert*. Frankfurt a. Main [u.a.]: Peter Lang.
- Sapp, Christopher D. 2011. *The Verbal Complex in Subordinate Clauses from Medieval to Modern German*. Amsterdam/Philadelphia: Johns Benjamins.
- Scheible, Silke, Richard Jason Whitt, Martin Durrell & Paul Bennett. 2011. Investigating diachronic grammatical variation in Early Modern German. Evidence from the GerManC corpus. In: Konopka et al. (2011), pp. 539–548.
- Scherr, Elisabeth & Konstantin Niehaus. 2013. ... weil man den Gebrauchsstandard erheben wird wollen. Variabilität und funktionale Äquivalenz in der Standardsyntax am Beispiel der 'Zwischenstellung' in Verbalkomplexen. In: Hagemann, Jörg et al. (eds.), pp. 75–84.
- Scheuringer, Hermann. 1997. Sprachvarietäten in Österreich. In: Stickel, Gerhard (ed.), *Varietäten des Deutschen. Regional- und Umgangssprachen*, pp. 332–345. Berlin/New York: de Gruyter.
- Schieffelin, Bambi, Kathryn Woolard & Paul Kroskrity (eds.). 1998. *Language Ideologies: Practice and theory*. New York: Oxford University Press.
- Schmidlin, Regula. 2011. *Die Vielfalt des Deutschen: Standard und Variation. Gebrauch, Einschätzung und Kodifizierung einer plurizentrischen Sprache*. Berlin/New York: de Gruyter.
- Schulte, Michael. 2009. Neue Überlegungen zum Aufkommen des Jüngerer Fußparks. Ein Beitrag zur Schriftgeschichte „von unten“. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 131:229–251.
- Van de Velde, Marc. 1981. Zur Reihenfolge von verbalen Elementen. In: Kohrt, Manfred & Jürgen Lenerz (eds.), *Sprache: Formen und Strukturen. Akten des 15. Linguistischen Kolloquiums Münster 1980 (Vol. 1)*, pp. 223–235. Tübingen: Niemeyer.
- Wolf, Norbert Richard. 1994. Österreichisches zum österreichischen Deutsch. Aus Anlaß des Erscheinens von: Wolfgang Pollak: Was halten die Österreicher von ihrem Deutsch? Eine sprachpolitische und soziosemiotische Analyse der sprachlichen Identität der Österreicher. *Zeitschrift für Dialektologie und Linguistik* 61:66–76.
- Zeman, Jaromír. 1988. Zu Stellungsvarianten des Verbs im Österreichischen: „daß nur ich es hören habe können“. *Brünner Beiträge zur Germanistik und Nordistik* 6:71–82.
- Zeman, Jaromír. 2002. *Die deutsche Wortstellung*. Wien: Edition Praesens.

Keywords

language history, variationist linguistics, grammar, lexical variation, grammatical variation (verbal clusters, discontinuous pronominal adverbs)

Lykilorð

málsaga, tilbrigðarannsóknir, málfræði, tilbrigði í orðaforða, málfræðileg tilbrigði

Útdráttur

Lengi vel voru hugtök eins og “einsleitni” og “staðalmálshugmyndafræði” ríkjandi í tengslum við málstöðlun og það leiddi til þess að gögnin sem rannsóknir voru byggðar á einskorðuðust gjarnan við formlegt mál og bókmenntatexta. Á síðustu árum hafa tvær hugmyndir tekið að hreyfa við hefðbundnum viðhorfum til málstöðlunar. Í fyrsta lagi hefur nálgun sem kennd er við málsögu “að neðan” dregið til sín athygli sem nýtt sjónarhorn, ekki bara gagnvart þýsku heldur einnig öðrum germönskum málum (cf. Elspaß, Langer, Scharloth & Vandenbussche 2007). Í rannsóknum innan þessa fræðilega ramma er áhersla lögð á söguleg gögn sem endurspeglu talmálsleg einkenni, einkum texta sem skrifaðir voru af fólki úr lægri lögum þjóðfélagsins, í þeim tilgangi að leiða í ljós breytingar og/eða samhengi í málsögunni sem áður höfðu ekki komið fram. Í öðru lagi hefur líkan sem gerir ráð fyrir mörgum svæðum (fremur en einum kjarna) fengið útbreiðslu, einkum með tilliti til “stórra” mála eins og þýsku. Í greininni er gerð grein fyrir skipulagi málheilda sem tekur tillit til beggja þessara hugmynda við athuganir á stöðlunarferlinu og leitast er við að sýna að rannsóknir sem byggjast á stærri og jafnvægari málheildum geta styrkt grundvöll athugana á málstöðlun.

*Stephan Elspaß / Konstantin Niehaus
Universität Salzburg
Fachbereich Germanistik
Erzabt-Klotz-Straße 1
A-5020 Salzburg
stephan.elspass@sbg.ac.at / konstantin.niehaus@sbg.ac.at*

Vanessa Isenmann

Insight into computer-mediated communication

as a new variety of written Icelandic

1 Introduction

The Internet and constant development of digital communication technology has a tremendous impact in Iceland. According to the *Global Information Technology Report 2013*, 92.6% of Icelandic households have Internet access and based on the estimates of Internet world statistics, 67.2% of Iceland's population are registered on Facebook.

In Iceland, the Internet is used for entertainment, information gathering and communication, just as it is in other parts of the world. Technological innovations now allow communication through a range of electronic devices and an array of platforms including for example blogs, email, (video) chat and (video) telephony. Nevertheless, the digital discourse is still mainly represented as written communication. However, it has been shown that online communication displays structures that are deviant from standard writing norms such as the neglect of spelling rules and punctuation or compensation strategies for non-verbal communication features (cf. Androutsopoulos 2011). Consequently, patterns and style used on a blog or on Facebook are different from those occurring in previous (personal) writing that is written communication outside the digital medium. The scope of rea-

sons ranges from various purposes for online communication (some being formal, others informal) to constantly changing communicative settings. Because of the latter, language has to transform appropriately to comply with new requirements and possibilities (Crystal 2006, Schlobinski 2006) presented by time, space and technological limitations that arise. The result is a set of verbal and non-verbal strategies that vary across national languages, although cross-linguistic similarities are possible (cf. Sveningsson 2003:139).

However, research on online communication, often referred to as *computer-mediated communication* (CMC) (Herring 1996, Thurlow, Lengel & Tomic 2004, Wright & Webb 2011), is still mostly focused on the English language (Baron 2008, Greiffenstern 2010, Herring 2004)¹. Especially smaller languages have hardly been taken into account. This paper aims therefore to broaden the picture and to give an initial insight into Icelandic CMC. Regarding the often claimed “purist” language policy in Iceland, the examination of CMC becomes especially relevant, for it contains features that contrast with what one may call “good Icelandic”. English borrowings (*like, nice*) for example, are frequently used in Icelandic CMC. Moreover, strategies to compensate for essential elements of oral communication, such as tone and body language, can be observed. Examples of this are the use of emoticons and verbalized laughter. Finally, spelling, such as the use of diacritics, and punctuation are not prioritized in the digital discourse. Instead, the creative use of spelling and punctuation in CMC may serve individual and stylistic purposes.

With regard to the analysis of CMC as a new variety of Icelandic, we must first understand its nature. Accordingly, this paper will address the following questions: What linguistic features are significant? How can they be categorized? And what functional purpose do they serve? Answering these questions will aid to assess whether CMC can be characterized as a new variety of written Icelandic.

In order to answer these questions, section 2 will first give an introduction to CMC as a new way for communication and will then

¹ Introduced in the 1980s, the term *CMC* was first used to encompass (interactive) communication through different computer related platforms such as instant messaging, chat and email (cf. Baron 2008: 11). With the coming of new digital communication technologies that are not computers in the narrow sense, such as mobile phones, the term was expanded to communicative interaction through electronic devices (e.g. Thurlow, Lengel & Tomic 2004). This essay will thus use the term in the latter sense and understand CMC as written (interactive) communication carried out on digital electronic devices.

specifically focus on the communicative options of Facebook as it provides the source for the corpus that will be introduced in this paper. In section 3, a case study is discussed. The results are then presented in section 4, with due regard to their communicative function and meaning for CMC as a variety of written Icelandic. Finally, section 5 serves to summarize the most relevant findings.

2 Computer-mediated communication

2.1 A new written variety

In linguistics, scholars have approached CMC from various perspectives, using an array of methods and focusing on a variety of phenomena. Some have tried to give a general overview of CMC specific forms (Runkehl, Schlobinski & Siever 1998, Storrer 2000). Others have surveyed CMC with regard to discourse analytic questions (Schönfeldt 2001, Beißwenger 2003). Yet others discuss the role of CMC for language (Bittner 2003, Crystal 2006) and the impact of English on individual languages, refuting anxieties that it might dominate the digital discourse (Sveningsson 2003, Thurlow, Lengel & Tonic 2004). Language change in and through CMC, and the categorization of CMC as written or spoken language has also been a topic of linguistic investigation (Haase et al. 1997, Smyk-Bhattacharjee 2006).

Nonetheless, CMC is neither written nor simply typed spoken language. Users have to develop strategies to meet communicative challenges involved in interaction without visual contact (Storrer 2001). Moreover, some essential features of oral communication are missing or become functionally reinterpreted. Since the utterance is conveyed as a whole, planning units such as hesitation sounds are not needed but may be used to fulfill stylistic purposes. Furthermore, CMC contains elements of writing, for example punctuation and spelling features such as upper and lower cases. Baron (2000:248) therefore describes CMC as an “emerging language centaur – part speech, part writing”.

CMC also reveals communication codes that go beyond traditional elements of interaction. Emoticons are neither associated with standard writing nor speaking, as are hashtags². Their use however

² Hashtags are character strings with an initial hash that were primarily used as metatags on different platforms to technically link different posts together in or-

is common practice in online communication. Thus the written and oral strategies are complemented by non-verbal symbols, which Crystal (2006:51) describes as “electronically mediated properties”. In his work he argues that CMC is something completely different from previous writing and speaking due to the inclusion of spoken, written and sign language features. At the same time it incorporates structures and strategies that none of the named mediums utilizes. He therefore proposes to see CMC, in contrast to speech, writing and signing, as a “fourth medium” for language (Crystal 2006:272).

In public discourse, CMC has often been perceived as “homogenous language”. Yet it encompasses distinct varieties. Different “Internet situations” imply different varieties (Bittner 2003, Crystal 2006). That is to say, the language used in blogs varies from the language in emails and, furthermore, both differ from the language used in chat groups. According to Androutsopoulos (2003:174), two criteria are hence relevant for the distinction of Internet-situational varieties: the distance of time between production and perception, and the communicative direction. This means that it is important whether the communication is (near-) synchronous (e.g. chat) or asynchronous (e.g. email) and whether it is interactive (e.g. chat) or not (e.g. blog).

For the categorization of CMC components Androutsopoulos (2007:81–3) suggests four themes that are applicable across languages and Internet situations: *conceptual orality*, *semiotics of compensation*, *linguistic economy* and *graphostylistics*. In this regard *conceptual orality* refers to aspects associated with (informal) spoken language, such as interjections. This categorization is based on a model of speaking and writing dimensions proposed by Koch & Oesterreicher (1994) who allocate the terms *spoken* and *written* language not only to the medium in which language is used but also with linguistic style.³

der to facilitate the search for relevant topics. In addition to that however, they have now developed a semantic function, giving indications on what broader topic the sender is referring to.

³ Koch & Oesterreicher (1994) differentiate between the termini *spoken* and *written* language both by the linguistic medium, that is either spoken = phonic or written = graphic, and the linguistic style, the “conception” as they call it. Concerning the medium, the terms *spoken* and *written* language are clearly separated: Language is verbalized either in a spoken or written manner although a change of medium is possible (as in dictation writing). Regarding conception however *spoken* and *written* language are two opposite poles of a continuum which includes numerous graduations. Critical for the location of an utterance or text on the continuum are categories, such as grade of closeness or grade of emotionality, that determine its style.

Semiotics of compensation includes aspects that compensate for non-verbal strategies of communication, such as facial or prosodic features. Prominent elements of this theme are emoticons and verbalized laughter. Furthermore, *linguistic economy* refers to strategies to shorten a message in order to comply with space, time or other limitations, as for example acronyms and abbreviations. Finally, *graphostylistics* describe alternatives of writing, based on the phonetic representation of a word or a word combination. A distinct realization for this is the replacement of lexemes and words by homophonic graphemes or grapheme combinations (see *you* → *cu*).

2.2 Facebook discourse as a test case

Representing the communication trends on the Internet on a micro level, Facebook has developed various ways of (self-) expression and communication. There are synchronous and asynchronous as well as interactive and non-interactive options to communicate. They range from written private messages to posts shared with a broader audience and a video-call application.

This paper however focuses on the *group* utility, which enables its members to engage with others, share items and discuss specific topics. Facebook groups refer to a wide spectrum of purposes as for example work and project related themes or topics of individual interests (sports, music). They are created by an individual and maintained by the group members, who post, update or upload files, comment on updates, poll the group or chat with everyone at once. Furthermore, members are able to remove their updates and comments again.

Facebook groups may be public, private or secret. Whereas any Facebook user may find and join a public group, see the members and their posts, private groups are visible albeit not accessible to the public. In other words, one must be added to the group by another member to be able to participate in the communicative interaction. Finally, secret groups are invisible to the public. Participants are selected and added by other group members and only they may see and edit posts.

As updates are published and commented on at any time, or not at all, communication in groups is asynchronous at large. The tone is prevalently informal and the communicative direction may be both monological and dialogical.

Because the group utility allows a rather easy collection of data, and considering that the language can hardly be described as dif-

ferent from other Facebook communication, it constitutes a suitable corpus for linguistic observation. Because groups consist of a limited number of members that interact with each other in various communication “pairs”, it provides a well-defined and yet multifaceted source for the collection of data. Private groups moreover allow for data collection that complies with ethical standards. Their members are generally more active and their number lower so that it becomes easier to obtain the permission of all group members to use the data for research purposes.

3 Data and Methodology

The rest of this paper focuses on a case study concerning characteristic properties of Icelandic CMC in order to depict and analyze CMC-inherent properties and to define it as a potential new variety of written Icelandic.

3.1 Corpus

The corpus introduced here is an active private Facebook group⁴ with 24 participants aging between 21 and 31. Three members are non-native speakers of Icelandic, and have therefore been excluded from the analysis. This leaves 562 posts and comments, published between June 19th, 2012 and April 25th, 2013. The total number of analyzable tokens is 5,310. It is important to note, that the group members know each other personally which might influence the interaction.

Although the corpus is rather small, it allows an initial overview of the diversity and the potential of Icelandic CMC.

3.2 Methodology

In order to facilitate the analysis, the corpus has been downloaded and sorted into utterances and further into lists of tokens. The lists have been analyzed according to the themes proposed by Androutsopoulos (2007): *conceptual orality*, *semiotics of compensation*, *linguistic economy* and *graphostylistics*. In so doing, properties corresponding

⁴ The members have been asked for permission to use the data for research purposes. They have agreed without exception.

with Androutsopoulos' (2007) categorization model and deviating from previous writing norms have been extracted and counted.

In this analysis, borrowings adapted to Icelandic phoneme-grapheme correspondence, such as *næs* 'nice', which could be categorized as *graphostylistics* due to their alternative writing, are counted as instances of the theme of *conceptual orality*. This coincides with the frequent use of (English) borrowings in spoken Icelandic. Foreignisms are often anxiously observed and contrast with Icelandic purist language policy that aims to keep the Icelandic lexicon "pure" (cf. Hanna Óladóttir 2009, Kristján Árnason 2006, 2009). Full integration of borrowings into Icelandic requires adaption on different levels such as accent, phonology, morphology and grammar (cf. Baldur Jónsson 2002, Guðrún Kvaran 2004). In addition, neologisms are created in order to counteract the distribution of borrowings. Nevertheless, (English) borrowings are one of the most distinctive features in the corpus.

For a deeper analysis the themes have been expanded into sub-categories as shown in the following.

4 Results

The findings, introduced here, do not serve as quantitative or qualitative measures, but rather illustrate features and patterns present in Icelandic CMC. The phenomena observed in the corpus are deviant from (previous) writing outside the digital medium and may hence serve as an indication of the development of a new variety of written Icelandic.

4.1 Conceptual Orality

The following examples, (1)–(4)⁵, illustrate the two most frequent phenomena that were classified as realizations of *conceptual orality* according to the model of Koch & Oesterreicher (1994), English borrowings and interjections. (The content of each example is described in a footnote.)

- (1) **A: Jæja krakkar, boys are back in town. Ætla að henda inn grófri dagskrá fyrir löngu dagana. (...)**
A: róleg á likeinu

⁵ In the examples, each participant is represented by a letter of the alphabet to ensure anonymity. The letters do not refer to the same participant across examples.

B: Langir dagar eru svo skemmtilegir, ég bara varð!
C: Æj!! Mig langar í fjöruferð!
D: næs!
E: Djöfull væri ég til í fjöruferð.
F: ÞESSIR DAGAR VERÐA FRÁBÆRIR!!! OH MY GOD!⁶

- (2) **A:** Langur dagur á morgun. **Jíííhaaaaaaaaa**
B: **Vá**, vandræðalegt, takk fyrir að minna mig á þetta.⁷
- (3) **A:** Er **séns** ad fa far med einhverjum úr midbænum i kvöld?
B: haha það er svo erfitt að horfa á þig **ströggl**a svona
A: Hjááálp!!!⁸
- (4) **A:** Takk fyrir **awesome** kvöld í gær!!!⁹

For a deeper analysis, Table 1 presents all categories of *conceptual orality* as well as their type and token distribution.

	English expressions				Other foreign expressions				Interjections		Lexical Creations	
	Un-adapted		Adapted		Un-adapted		Adapted		Types	Tokens	Types	Tokens
	Types	Tokens	Types	Tokens	Types	Tokens	Types	Tokens				
	55	91	25	51	3	3	1	1	31	76	2	2
Total number of types/tokens	80/138				4/4				31/76		2/2	

Table 1. Conceptual Orality: Categories are English expressions, foreign expressions (non-English), interjections and lexical creations. The first two categories are subdivided into unadapted and adapted expressions, and they include both single words and phrases. Phrases are counted as one expression. The last row shows the total number of types and tokens for each category.

⁶ In (1), the participants converse about the schedule for some long working days.

⁷ In (2), A reminds the audience of a long working day. B thanks for the reminder.

⁸ In (3), A asks the audience to get a ride. B replies that it is hard to see A struggling. A then calls for help.

⁹ In (4), A thanks the audience for a great evening.

Interjections (*jæja*, *æj*, *jííííhaaaaaaaaa*, *vá*) are a frequent phenomenon in the corpus as are English words and phrases¹⁰, both with English spelling (*oh my god*, *awesome*) and adapted to Icelandic orthography and/or grammar (*likeinu*, *næs*, *ströggla*).

Unadapted English borrowings include English single words and phrases that are not adapted according to Icelandic phoneme-grapheme correspondence but are rather used with English orthography (*boys are back in town*, *awesome*). *Unadapted English borrowings* include all terms that follow this pattern even if their form coincides with Icelandic orthography. The most frequent unadapted English borrowing is *kids* and its derivation *kiddos* with a total number of incidences of 13. The word was always used to address the audience either in the expression *hey kids* (5) or as a single word (6).

- (5) **A: Hey kids.** Það er hádegisfæðsla í Halastjörninni núna á föstudaginn frá kl 11.30. Frír matur og mega fjör. (...)
B: YEEES¹¹

- (6) **A: hvað er planið kids??** Er þetta off eða?¹²

Mega is the second most common English word in the corpus and accounts for 11 tokens. It is used both as an adjective (5) and an adverb (7).

- (7) **A: Takk fyrir mega næs fund krakkar (...)**¹³

It can be observed that phrases (*boys are back in town*, *oh my god*) are always unadapted (cf. (1), relevant part repeated as (8)). In the category of *unadapted English borrowings*, 19 out of 91 tokens are phrases. However, the category of *adapted English borrowings* contains no phrases. It is likely that phrases often function as citations and therefore they are not adapted to Icelandic orthography.

¹⁰ With regard to phrases a distinction between *borrowing* and *code switching* could be made but is of no further use for the purposes of this paper. Hence, both single words and phrases are called *borrowings* in this paper. Furthermore, there is dissension in linguistics about how to define and differentiate the terms *borrowing* and *code switching*. For further reading about the distinction of *borrowing* and *code switching* see Thomason 2001 and Winford 2001.

¹¹ In (5), A announces a free lunch.

¹² In (6), A asks what the plan is and whether a plan was cancelled.

¹³ In (7), A thanks for a meeting.

- (8) **A:** Jæja krakkar, **boys are back in town.** (...)
 (...)
 F: ÞESSIR DAGAR VERÐA FRÁBÆRIR!!! **OH MY GOD!**

Adapted English borrowings include lexical items that have been adapted to Icelandic grammar and/or spelling by adjusting the orthography to standard rules of Icelandic phoneme-grapheme correspondence and to rules of Icelandic grammar (e.g. *næs* 'nice', *ströggla* 'to struggle'). With a total of six incidents *næs* is the most frequent adapted English borrowing. It is interesting to observe that the word occurs only once with English spelling but six times with Icelandic orthography. Even though the corpus can hardly be considered representative this may indicate the advanced integration of *næs* into Icelandic. However, relative to English expressions with original spelling (91 tokens), the adaptation to Icelandic is less common in the corpus (51 tokens). Regarding the aforementioned requirements for the full integration of borrowings into Icelandic this becomes especially interesting.

As seen in the word *næs/nice*, some terms may occur in both adapted and unadapted form. The word *shit* occurs three times with English and twice with Icelandic spelling (*sjitt*).

- (9) **A:** **Sjitt** krakkar. Þessi er sko GEIÐVEIKT NÆS.¹⁴
- (10) **A:** **Kids**, hverjir eru með I föstudagsbjör, eg þarf að kynnast ykkur!
A: nei **shit**, kemst ekki, vandræðanlegt.¹⁵

Here, the use of either English or Icelandic spelling seems to be of individual choice. As the examples in (8) and (9) indicate, the same user tends to either adapt all borrowings in their utterance (*sjitt*, *næs*), or none (*kids*, *shit*).

A different strategy may however be observed regarding the expression *like* which occurs several times in the corpus. In these occurrences, *like* has a Facebook-related meaning of pressing the *like*-button in order to express one's appreciation of a post or comment. In three of the six examples that occur in the corpus it is unadapted.

- (11) **A:** mikið elska ég starfsmannafélag
 reykvjavíkurborgar. (...)
B: fánlegt **like** á það!¹⁶

¹⁴ In (9), A finds something very good.

¹⁵ In (10), A proposes to meet for a beer but has to take back the proposal later.

¹⁶ In (11), A expresses his/her love of the union. B approves.

- (12) **A:** Heyheyhey! Við ætlum nokkur á Vegamót á morgun að borða! (...)
B: Matur og Fylleri **LIKE**¹⁷
- (13) **A:** Krakkar, hverjir eru til í eitthvað skemmtistuð um helgina, fös eða lau??
A: **like** = já?¹⁸

Yet, in three examples *like* is adapted to Icelandic grammar though not to Icelandic spelling rules, e.g. *likeinu* (cf. (1), relevant part repeated as (14)).

- (14) (...) **A:** róleg á **likeinu**

In (14), *like* is used as a noun. The stem retains its English form but the word is adjusted to Icelandic grammar by attaching the dative article (*-inu*) in accordance with the preceding preposition *á*.

In the other two cases, *like* is used as a verb, as for example in (15). The spelling is English but the Icelandic infinitive ending *-a* is added, with a dash.

- (15) **A:** Kæra samstarfsfólk. Vilduð þið vera svo góð að kíkja á krotið mitt, hugsanlega **like-a** og í mesta lagi deila, ég verð ykkur ævinlega þakklát.¹⁹

Concerning the verb form, the mixing of English spelling and Icelandic inflection (as in *like-a*) displays an exception since other verbs in the corpus are adapted in both orthography and grammar (see for example *ströggla*). Furthermore, the use of a dash to adhere the inflectional ending can only be observed for *like*. The English orthography becomes especially interesting when taking into account that the button to press says “líkar þetta” in the Icelandic Facebook interface. However, Facebook was initially English speaking and the term *like* as a Facebook-inherent act is entrenched across languages. The preference of *like* over *líkar* in the corpus may thus derive from the endeavor to highlight the Facebook-related meaning in contrast to the intrinsic semantics of *líkar*.²⁰ Since the verb form of *like* is treated

¹⁷ In (12), A proposes dinner at a restaurant. B approves the idea.

¹⁸ In (13), A proposes to meet at the weekend and explains that pressing the *like* button means *yes*.

¹⁹ In (15), A asks the audience to have a look at something, press the *like* button and repost it.

²⁰ This coincides with a tendency in colloquial Icelandic in general to use the bor-

differently from other adapted English verbs (as for example *ströggla*), the dash is most likely a consequence of the orthographically unadapted verb stem and may serve as a clearer visualization of the inflectional ending.

The mixing of an English spelling of the stem and an Icelandic inflectional ending like in the form *likeinu* is not exceptional since the same partial adaptation can be observed in other Facebook and Internet related terms:

(16) (...)

A: minnir að það sé askja mannaðakerfi, vertu viss um að **browserinn** leyfi **popups** því annars opnast glugginn ekki²¹

(17) (...)

A: hver komst inn á **facebookið** þitt?²²

Similar to *like* it can be argued that the terms in (16)–(17) are not adjusted to Icelandic orthography because of their Facebook (*facebookið*) or Internet (*browserinn*, *popups*) related meaning.

Another such example is the term *facereip* in (18):

(18) (...)

A: haha þetta var **facereip**, samt til við tækifæri²³

The term derives from the English neologism *facerape*, which according to the online *Urban Dictionary* refers to an unauthorized use of someone's Facebook account resulting in humorous changes of status updates or posts to friends. Only the second part of the compound is adapted according to Icelandic phoneme-grapheme correspondence, possibly due to the fact that the latter part of a compound dominates the inflection.²⁴

The category *Other foreign borrowings* includes lexical items from foreign languages other than English.

rowed verb *læka* for the meaning 'to put a Like on' rather than applying the psych verb *líka* 'to like' in this context.

²¹ In (16), A explains how to open a website.

²² In (17), A asks another participant who entered his/her Facebook account.

²³ In (18), A explains that someone announced something through his/her Facebook account without permission.

²⁴ This pattern is common in so-called hybrid words where the first part of the compound is foreign based and the second part Icelandic. In this way the hybrid word fits well into the Icelandic inflectional system. Guðrún Kvaran (2004) presents several examples of hybrid words in Icelandic.

- (19) (...)
A: hvað kostar þetta?
B: nada²⁵

The expressions observed include French (*gúrm* 'gourmet') with adapted orthography, as well as Swedish (*älsklings* 'darlings') and Spanish (*nada* 'nothing', *pronto* 'soon'). All these words keep their original spelling. As shown in Table 1, such expressions occur rarely (altogether 4 tokens), indicating the domination of English as the main source of foreign influence in Icelandic digital discourse.

In accordance with earlier CMC research (cf. Androutsopoulos 2011, Storrer 2000) and the model of Koch & Oesterreicher (1994), *interjections* (as for example *jæja*, *æj*, *vá*, *jíííhaaaaaaaaa*) are also classified as instances of *conceptual orality*. Interjections are common in the corpus (altogether 76 tokens), with *hey* being the most frequent one (11 tokens). Other common interjections are *jæja* (nine tokens), *vá* (eight tokens) and *oh* (seven tokens).

Finally, *lexical creations* comprise formations that are created for stylistic reasons:

- (20) **A:** Djöfull sem ég sakna ykkar. kv. Frá **babbadíbubbi-landi**²⁶

Lexical creations are, however, hardly used in the corpus. The only example of an Icelandic creation is the one shown in (20), *babbadíbubbi-landi*. In addition the English formation *alrightyright* occurs once.

In short, interjections and English borrowings are the most frequent features of the *conceptual orality* theme in the corpus. As for English borrowings, single words may occur both adapted and unadapted to Icelandic. Borrowed phrases are always unadapted. Internet and Facebook related terms can have Icelandic inflectional endings but they are not adapted to Icelandic orthography. It has been pointed out that English is the dominant language in the field of computers in Iceland (e.g. Hanna Óladóttir 2009:72). Although many programs and Internet platforms such as Facebook are offered in Icelandic, Icelanders are used to the English terms in the realm of computers. It is therefore not surprising that English computer-related terms are not adapted to Icelandic spelling in the corpus.

²⁵ In (19), the costs of something are the topic of the conversation.

²⁶ In (20), A says that he/she misses the other group members.

4.2 Semiotics of Compensation

The following examples introduce phenomena categorized as *semiotics of compensation*.

- (21) **A:** Jæja krakkar, hverjir eru til í spurningarbombuna á þriðjudaginn, með **L-L-L-Loga** bergmann. (...).
B: Haaaaaa?
C: Takk en nei, takk.
A: Þar að segja að koma sem áhorfendur.
D: hahahaha... fáum við peninga..??
A: já og utanlandsferð
D: hells to the yeahh
C: Ég vil fara til Fijii...²⁷
- (22) **A:** prufan gekk sjúklega vel og ég er komin með vinnu á Vegó! woop woop!!
B: Til hamingju!!!
C: NICE! Til hamingju!
D: Æði, til hamingju! Hlakka til að koma í kokteilA til þín ;-)
A: takk öll :-D hlakka til að fá þig í kokteilA til mín! :-)
E: geggið :-D²⁸

Semiotics of compensation are mainly introduced through laughter and emoticons as well as expressive punctuation and forms of written prosody, for example through the capitalization of words.

Table 2 shows the distribution of types and tokens of the strategies used for the compensation of non-verbal features of interaction.

	Expressive punctuation		Emoticons		Written prosody				Verbalized laughter	
					Foreign based		Icelandic			
	Types	Tokens	Types	Tokens	Types	Tokens	Types	Tokens	Types	Tokens
	6	116	12	86	11	12	28	33	3	31
Total number of types/tokens	6/116		12/86		39/45				3/31	

Table 2. Semiotics of Compensation: Categories are expressive punctuation, emoticons, written prosody and verbalized laughter. Written prosody is subdivided into foreign based words and Icelandic words. The last row shows the total number of types and tokens.

²⁷ In (21), the attendance of a quiz show is topic of the conversation.

²⁸ In (22), the participants congratulate A on a successful job interview.

Expressive punctuation is represented through six different types, the most frequent ones being the multiple representation of dots (50 tokens), multiple exclamation marks (44 tokens) and multiple question marks (14 tokens) as shown in (23) and (24) (repeated sections of (21)–(22)):

(23) (...)

D: hahahaha... fáum við peninga..??

(24) **A:** prufan gekk sjúklega vel og ég er komin með vinnu á Vegó! woop woop!!

B: Til hamingju!!!

Concerning *emoticons*, the regular smiley is the most frequent type (37 tokens). Beyond that, the winking smiley (21 tokens) and the laughing smiley (12 tokens) are commonly represented.

Written prosody is expressed through the capitalization of words and word combinations (27 tokens) and by multiple letters (20 tokens):

(25) **A:** ERUÐ ÞIÐ TILBÚIN?????????

B: þeeeegiðu...(...) ²⁹

Capital letters are in general interpreted as screaming or shouting, whereas multiple letters are perceived as sound lengthening. Both strategies serve to emphasize the word or word combination in question. Even a combination of both strategies occurs as shown in (26) (repeated from (5) above):

(26) (...)

B: YEEES

An interesting observation is the realization of emphasis in emoticons:

(27) **A:** Rétt upp hönd sem ætla á tónleika á lau?! o/

B: 0//³⁰

In (27) the emoticon *o/* (zero or *o* followed by a slash) represents the raising of a hand, with zero respectively *o* depicting the head and the slash illustrating the raised arm. In the response the emoticon is taken over but a second slash is added. The representation through grapheme multiplication is thus used to underline the response.

²⁹ In (25), A asks the audience if they are ready. B replies with “Shut up!”.

³⁰ In (27), A asks the audience to raise their hand if they will be attending a concert.

Verbalized laughter is represented mainly through *haha* (29 tokens). The length however is variable, i.e. *hahaha* or longer manifestations may appear. Apart from that *hehe* and *híhí* (one token each) are realized.

In sum, the above findings highlight the presence of semiotics of compensation in Icelandic CMC as CMC characteristic features. The representations found in the corpus solely contain properties of CMC that are used across languages. However individual strategies appear, such as the emphasizing of emoticons (*o//*), presumably for creative stylistic reasons.

4.3 Linguistic economy

Besides *conceptual orality* and *semiotics of compensation*, examples of *linguistic economy* occur in the corpus. In this analysis the term encompasses both linguistic shortenings (e.g. acronyms) and orthographic economy, such as the non-occurrence of diacritical marks (cf. (28)–(30)), as both result in the simplification of orthographic form.

(28) A: Vitiði hvort það **se** ekki þannig að maður er bara með 3 **manaða** uppsagnarfrest ef maður er **fastraðinn**, er að spyrja fyrir vinkonu mína sem vinnur í **Kringlumyri** og langar að losna³¹

(29) A: NEIOH, það gleymdist að **lata** mig vita að **eg** ætti ekki að vinna **i** dag (...)
B: **lol** á þig (...)³²

(30) A: **eigum** við að panta borð **kl 8**?³³

In Table 3 the categories of *linguistic economy* are listed with their types and tokens.

³¹ In (28), A asks for a period of notice.

³² In (29), A complains about not having been informed about a day off from work.

³³ In (30), A asks to book a table.

	Acronyms		Abbreviations		Shortenings				Spelling			
					Foreign based		Icelandic		Foreign based		Icelandic	
	Types	Tokens	Types	Tokens	Types	Tokens	Types	Tokens	Types	Tokens	Types	Tokens
	3	6	5	15	1	1	5	8	4	4	54	452
Total number of types/tokens	3/6		5/15		6/9				58/456			

Table 3. Linguistic economy: Categories are acronyms, abbreviations, shortenings and spelling deviations. Shortenings and realizations of spelling deviations are subdivided into foreign based words and Icelandic words.³⁴ The last row shows the total number of types and tokens.

As shown in Table 3, variation in spelling (*se* instead of *sé*, *manaða* for *mánaða*, *fastraðinn* for *fastráðinn*, *Kringlumyri* instead of *Kringlumýri*, *lata* for *láta*, *eg* for *ég'*, *i* instead of *í*) occurs frequently in the corpus (altogether 456 tokens). Spelling is facilitated in different ways. The strategies follow traditional techniques of typewriting, which are dominated by a sign-repertoire that does neither consider diacritics nor language-specific characters, due to its endeavor of general validity. Spelling consistency is not required. Special characters may be used but diacritics left out, as in (28) and (29) (*manaða*, *fastraðinn* instead of *mánaða*, *fastráðinn*). In other examples diacritics are used but special characters are left out as in (31) (repeated from (3) above) (*ad* for *að*, *med* for *með*, *midbænum* instead of *miðbænum*):

- (31) **A:** Er séns **ad** fa far **med** einhverjum úr **midbænum** i kvöld?

Also, the correct use of upper and lower case according to official spelling rules becomes irrelevant as in (32) (cf. (30) above).

- (32) **A:** **eigum** við að panta borð kl 8?

Correct spelling does not seem to be required. Only in one example a user refers to a spelling deviation in the response as shown in (33) (repeated from (22) above):

³⁴ The same foreign based terms were also counted in the theme of *conceptual orality*.

- (33) (...)

D: Æði, til hamingju! Hlakka til að koma í **kokteilA** til þín ;-)
- A: takk öll :-D hlakka til að fá þig í **kokteilA** til mín!
- :-)

By repeating the deviation in the response, the spelling is only indirectly referred to. Other than that, spelling deviations seem commonly accepted.

Acronyms have been mentioned as a characteristic feature of CMC (cf. Crystal 2006, Baron 2008). Specifically, *lol* ‘laughing out loud’ has obtained transnational fame. In this corpus *lol* occurs three times. Other acronyms are hardly used in the corpus. The total number of tokens is six and all acronyms that occur are of English origin (*lol*, *asap* ‘as soon as possible’, and *btw* ‘by the way’).³⁵

Abbreviations are more common than acronyms (altogether 15 tokens). They are established Icelandic abbreviations, with *kv* ‘kveðja’ and *kl* ‘klukkan’ being the most frequently used (six tokens each).

Shortenings have a total number of nine tokens. Five out of six types are Icelandic with two being common shortenings for the days of the week (*fös* ‘föstudagur’, *lau* ‘laugadagur’). The others follow a common strategy, the shortening with an *ó*-ending (*vandró* for “vandræði” (‘problems’), *brennó* for “brennibolti” (name of a ball game), *abbó* for “afbrýðisamur” (‘jealous’).

- (34) A: hey **btw** hver er síðan fyrir til að stilla tímana? (...)
- B: ég las þetta hrikalega vitlaust hjá þér, sé það núna, **vandró**.³⁶

The only foreign based shortening is *diff* (‘difference’) which is used in the informal expression *meikar ekki diff*:

- (35) A: (...) akkurat mánuður í að ég sjái ykkur og blessuð börnin! kv. spennt og sakna
- A: eða aðeins svona rúmlega, meikar ekki **diff**..³⁷

To summarize, the strategies to facilitate and shorten spelling are not

³⁵ Due to their English origin they are also counted as occurrences of the theme of *conceptual orality*.

³⁶ In (34), A asks for the website to correct clocked working hours. B then explains that he/she had at first misunderstood the post.

³⁷ In (35), A says that it is exactly a month until he/she will see the other person again but then revises the comment with respect to the exact time even if it does not make a difference.

developed through CMC, even if they are frequently used for economic reasons. The neglect of standard spelling is encouraged by the fact that apparently it does not lead to misunderstandings or rebukes. Other time saving strategies such as acronyms and abbreviations are hardly used. Space-saving techniques are not needed as space limitations are hardly ever reached. As communication in Facebook groups is asynchronous or near-synchronous at most, time pressure is negligible. However, Facebook is increasingly used with smart-phones which often do not provide the Icelandic keyboard. Spelling inconsistency and the lack of diacritics may therefore stem from technical limitations that arise as a result of the device used. Nevertheless, as CMC in Facebook is informal it can be stated that standard writing rules are obsolete in this Internet situation.

As spelling and grammar adaptations of foreign-based expressions towards Icelandic rules were analyzed under the heading of conceptual orality, *graphostylistics* are hardly represented in the corpus. The only representation in this context is @myplace 'at my place':

(36) A: (...) smá reunion í kvöld @myplace (...)³⁸

5 Conclusion

Although the case study introduced in this paper is rather small, its findings illustrate significant characteristics of Icelandic CMC. Some of the features can be compared to previous research on modern Icelandic language use, which have discussed the recent influence of globalization on Icelandic (e.g. Hilmarsson-Dunn & Ari Páll Kristinsson 2010). For example, Leonard & Kristján Árnason (2011:96) state: "From both the speaker's and the language planners' perspective, the enthusiasm for insisting on the ideology of linguistic purism appears to have begun to wane over the last 20 years."

English words and phrases are not a characteristic of CMC in particular as lexical borrowing is a general tendency in informal Icelandic. Ásta Svavardóttir for example argues that English borrowings are more frequent in informal language than in formal language (cf. Ásta Svavardóttir 2004:175). The use of English expressions and of interjections in the corpus thus reflects informal language use in Icelandic CMC.

³⁸ In (36), A proposes to meet at his/her place.

The English borrowings reveal various patterns concerning their adaption to Icelandic grammar and orthography. Some borrowings are orthographically adapted with few or no exceptions (*næs*), whereas in other cases adaption seems to be a matter of personal choice (*shit* or *sjitt*). Computer and Internet related terms are adapted grammatically, by adding inflectional endings, but do not follow the rules of Icelandic phoneme-grapheme correspondence (*like-a*, *likinu*, *browserinn*). It has been argued that English is the dominant language of computers and of the Internet and that borrowings in these settings are often positively perceived (cf. Hanna Óladóttir 2009). Hence, the preference for English spelling of such words is possibly connected to their semantic relation to the Internet or to Facebook.

Apart from English borrowings and interjections, strategies for the compensation of non-verbal communication strategies can be observed in the material, including emoticons, verbalized laughter, written prosody and expressive punctuation. These strategies are cross-linguistic phenomena of CMC that have sometimes been considered to be the only “real” linguistic innovation of the Internet (Androutsopoulos 2007:82). Finally, spelling deviations occur frequently in the corpus in the form of neglecting diacritics, upper case or punctuation. Two reasons may cause spelling deviations: The first reason may be technical limitations that arise when Facebook is used with smartphones that do not provide the Icelandic keyboard. The second reason may be that spelling rules are not prioritized in Facebook as the communication is informal and correct spelling is therefore considered to be less important.

As a fairly new phenomenon it cannot be foreseen yet where CMC is heading in Icelandic and what standards may evolve. Broader and deeper studies are needed to be able to make conclusions about patterns and generalizations. However, the elements and strategies introduced in this paper display significant deviations from other written varieties. Although CMC does not necessarily develop new linguistic features, it incorporates properties that in their combination may form a new variety of written Icelandic with regard to (informal) language use and new means of expression. Beyond that, CMC enables users to unfold their linguistic freedom and creativity in both verbal and non-verbal ways (cf. Smyck-Bhattacharjee 2006). The Internet therefore provides a fruitful new testing ground for the study of (informal) language use in modern Icelandic.

References

- Androutsopoulos, Jannis. 2003. Online-Gemeinschaften und Sprachvariation. Soziolinguistische Perspektiven auf Sprache im Internet. *Zeitschrift für germanistische Linguistik* 31/2:173–197.
- Androutsopoulos, Jannis. 2007. Neue Medien – neue Schriftlichkeit? *Mitteilungen des Deutschen Germanistenverbandes* 1/07:72–97.
- Androutsopoulos, Jannis. 2011. Language change and digital media: A review of conceptions and evidence. In: Kristiansen, Tore & Nikolas Coupland (eds.). *Standard Languages and Language Standards in a Changing Europe*, pp. 145–159. Oslo: Novus Press.
- Ásta Svavardóttir. 2004. English borrowings in spoken and written Icelandic. In: Anna Duszak & Urszula Okulska (eds.). *Speaking from the Margin. Global English from a European Perspective*, pp. 167–176. Frankfurt a.M.: Peter Lang.
- Baldur Jónsson. 2002. Aðlögun tökuorða í íslensku. *Málsgreinar. Afmælisrit Baldurs Jónssonar*, pp. 219–233. Reykjavík: Íslensk málnefnd.
- Baron, Naomi. 2000. *Alphabet to Email*. London: Routledge.
- Baron, Naomi. 2008. *Always on. Language in a Mobile World*. Oxford: Oxford University Press.
- Beißwenger, Michael. 2003. Sprachhandlungskoordination im Chat. *Zeitschrift für germanistische Linguistik* 31/2:198–231.
- Bittner, Johannes. 2003. *Digitalität, Sprache, Kommunikation. Eine Untersuchung zur Medialität von digitalen Kommunikationsformen und Textsorten und deren varietätenlinguistischer Modellierung*. Berlin: Erich Schmidt.
- Crystal, David. 2006. *Language and the Internet*. 2nd ed. Cambridge: Cambridge University Press.
- Global Information Technology Report 2013*. <http://www.weforum.org/reports/global-information-technology-report-2013>. (25.06.2013)
- Greiffenstern, Sandra. 2010. *The Influence of Computers, the Internet and Computer-Mediated Communication on Everyday English*. Berlin: Logos
- Guðrún Kvaran. 2004. The adaption of loanwords in Icelandic. In: *Aspects of Foreign Words. Loanwords in the World's Languages*, pp. 262–271. Tokyo: National Institute for Japanese Language.
- Haase, Martin et al. 1997. Internetkommunikation und Sprachwandel. In: Weingarten, Rüdiger (ed.). *Sprachwandel durch Computer*, pp. 51–85. Opladen: Westdeutscher Verlag.
- Hanna Óladóttir. 2009. *Shake, sjeik eller mjólkurhristingur? Islandske holdningar til engelsk språkpåvirkning*. Oslo: Novus Forlag.
- Herring, Susan (ed.). 1996. *Computer-Mediated Communication. Linguistic, Social and Cross-Cultural Perspectives*. Amsterdam: John Benjamins.
- Herring, Susan. 2004. Slouching toward the Ordinary. Current trends in computer-mediated communication. *New Media & Society* 6/1:26–36.

- Hilmarsson-Dunn, Amanda & Ari Páll Kristinsson. 2010. The language situation in Iceland. *Current Issues in Language Planning* 1/3:207–276.
Internet World Stats. <http://www.internetworldstats.com>. (15.06.2013)
- Kristján Árnason. 2006. Island. In: Kristiansen, Tore & Lars Vikør (eds.). *Nordiske språkhaldningar. Ei meiningsmáling*, pp. 17–34. Oslo: Novus Forlag.
- Kristján Árnason. 2009. Að bera sér orð í munn: Hvenær verður orðið íslenskt? *Orð og tunga* 11:75–100.
- Koch, Peter & Wulf Oesterreicher. 1994. Schriftlichkeit und Sprache. In: Günther, Hartmut & Otto Ludwig (eds.). *Schrift und Schriftlichkeit. Writing and Its Use. Ein interdisziplinäres Handbuch internationaler Forschung. An Interdisciplinary Handbook of International Research*, pp. 587–604. Berlin: de Gruyter.
- Leonard, Stephen & Kristján Árnason. 2011. Language ideology and standardisation in Iceland. In: Kristiansen, Tore & Nikolas Coupland (eds.). *Standard Languages and Language Standards in a Changing Europe*, pp. 91–96. Oslo: Novus Press.
- Runkehl, Jens, Peter Schlobinski & Torsten Siever. 1998. *Sprache und Kommunikation im Internet*. Wiesbaden: Verlag für Sozialwissenschaften.
- Schlobinski, Peter (ed.). 2006. *Von *hdl* bis *cul8r**. *Sprache und Kommunikation in den neuen Medien*. Mannheim: Dudenverlag.
- Schönfeldt, Juliane. 2001. Die Gesprächsorganisation in der Chat-Kommunikation. In: Beißwenger, Michael (ed.). *Chat-Kommunikation. Sprache, Interaktion, Sozialität und Identität in synchroner computervermittelter Kommunikation*, pp. 25–53. Stuttgart: Ibidem.
- Smyk-Bhattacharjee, Dorota. 2006. „A “lyricist” by night, teenager by day and sleepaholic everytime in between”. Language change in/through CMC. *Germanistische Linguistik* 186–187: 68–85.
- Storrer, Angelika. 2000. Schriftverkehr auf der Datenautobahn. Besonderheiten der schriftlichen Kommunikation im Internet. In: Voß, Günther, Werner Holly & Klaus Boehnke (eds.). *Neue Medien im Alltag. Begriffsbestimmungen eines interdisziplinären Forschungsfeldes*, pp. 153–177. Wiesbaden: Verlag für Sozialwissenschaften.
- Storrer, Angelika. 2001. Sprachliche Besonderheiten getippter Gespräche: Sprecherwechsel und sprachliches Zeigen in der Chat-Kommunikation. In: Beißwenger, Michael (ed.). *Chat-Kommunikation. Sprache, Interaktion, Sozialität und Identität in synchroner computervermittelter Kommunikation*, pp. 3–24. Stuttgart: Ibidem.
- Sveningsson, Malin. 2003. The Significance of English Content in a Swedish Web Chat. In: Pettersson, John Sören (eds.). *HumanIT 2003*, pp. 138–158. Karlstad: Karlstad University Studies.
- Thomason, Sarah. 2001. *Language Contact. An Introduction*. Edinburgh: Edinburgh University Press.
- Thurlow, Crispin, Laura Tengel & Alice Tomic. 2004. *Computer Mediated Communication. Social Interaction and the Internet*. London: Sage Publications.

Urban dictionary. <http://www.urbandictionary.com>. (02.11.2013)

Winford, Donald. 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell.

Wright, Kevin & Lynn Webb. 2011. *Computer-Mediated Communication in Personal Relationships*. New York: Peter Lang.

Lykilorð

netsamskipti, íslenska, málbrigði, Internet

Keywords

computer-mediated communication, Icelandic, language variety, Internet

Útdráttur

Á undanförunum árum hefur þróast á Íslandi eins og víða annars staðar sýndarveruleiki til hliðar við raunveruleikann og í gagnkvæmum tengslum við hann. Internetið hefur vaxandi áhrif á líf Íslendinga með þeim samskiptamöguleikum sem það býður upp á. Aukið mikilvægi samskipta gegnum rafræna miðla hvetur til rannsókna á málnotkun við nýjar aðstæður. Því hefur verið haldið fram að netsamskipti (e. *computer-mediated communication*, CMC), þ.e.a.s. (gagnvirk) samskipti með stafrænum rafeindatekjum eins og tölvum og símum, samræmist ekki venjulegum viðmiðum ritmálsins þar sem málnotkunin lagi sig að nýjum möguleikum og takmörkunum. Markmið greinarinnar er að gefa innsýn í íslensk netsamskipti. Athugunin byggist á safni færslna úr virkum Facebook-hópi og valin dæmi úr efninu varpa ljósi á ýmis einkenni slíkra samskipta á íslensku. Meðal þeirra eru einkenni sem gjarnan eru tengd talmáli svo sem ensk lán, bæði stök orð og frasar, og upphrópanir. Einnig birtast einkenni sem endurspeglar aðferðir til þess að koma til skila í riti hljómrænum og myndrænum þáttum munnlegra tjáskipta, t.d. tónfalli, áherslum og svipbrigðum, auk atriða sem spretta af þörf á að einfalda og flýta fyrir ritun. Þar sem margir þessara þátta víkja frá því sem hefur tíðkast í ritmáli almennt er því haldið fram að netsamskipti leiði til nýs afbrigðis af ritaðri íslensku.

Vanessa Monika Isenmann

Íslensku- og menningardeild

Háskóla Íslands

Árnagarði v/Suðurgötu, 101 Reykjavík

vmi@hi.is

Ari Páll Kristinsson og Haraldur Bernharðsson

Íslenska og enska í íslensku háskólastarfi

1 Inngangur

Þessi grein er ávöxtur af samnorrænu verkefni. Árið 2011 styrkti Norræna ráðherranefndin samstarfsverkefni 18 norrænna fræðimanna um að safna saman og vinna úr upplýsingum um notkun ensku annars vegar og viðkomandi þjóðtungu hins vegar í norrænum háskólum.¹ Markmið samstarfsverkefnisins var m.a. að finna og benda háskólayfirvöldum á gagnlegar leiðir til að ná hinu tvíþætta markmiði að styðja rannsóknir og þekkingarmiðlun heima fyrir, á þjóðtungunni, jafnframt því sem háskólarnir laða til sín stúdenta og fræðimenn hvaðanæva, veita framúrskarandi menntun og birta rannsóknarniðurstöður sem eiga erindi við alþjóðlegt rannsóknarsamfélag. Leiðarljós í samstarfsverkefninu er sameiginleg yfirlýsing um markmið á sviði tungumála sem menntamálaráðherrar Norðurlandanna undirrituðu 2006 (*Yfirlýsing um málstefnu Norðurlanda*). Yfirlýsing þessi er þó ekki lagalega bindandi fyrir aðildarþjóðirnar. Þótt framkvæmd meginmarkmiða yfirlýsingarinnar geti verið breytileg frá einu landi til annars ber samþykktin með sér að fulltrúar allra Norðurlandþjóðanna eru uggandi um stöðu þjóðtungnanna gagnvart ensku í háskólakennslu og rannsóknum. Í yfirlýsingunni er mælt með því að háskólarnir móti sér stefnu um samhliða notkun

¹ Netværk for parallelsproglige mål på Nordens internationaliserede universiteter, <http://nordiskparallelsprogsnet.blogs.ku.dk/>. Auk höfunda þessarar greinar tók Ástráður Eysteinnsson prófessor þátt í starfinu af Íslands hálfu, sbr. Ástráð Eysteinnsson (2012).

viðkomandi þjóðtungu og ensku. Markmiðið er að hvorugt málanna ryðji hinu úr vegi í háskólastarfseminni.

Engum blöðum er um það að fletta að enska er ráðandi tungumál um þessar mundir í hinu alþjóðlega menntunar- og vísindasamfélagi sem íslensku háskólarnir eru hluti af (sjá t.d. Ammon 2010:113–116). Háskólayfirvöld standa frammi fyrir því að móta og framfylgja stefnu um sambúð íslensku og ensku í stjórnun, rannsóknum, kennslu og samskiptum almennt í íslenska háskólasamfélaginu; sambúð sem tryggir vöxt og viðgang íslensku sem kennslu- og vísindamáls jafnframt því sem kostir ensku og enskukunnáttu eru nýttir eins og við á í alþjóðlegu menntunar- og fræðaumhverfi. Hvað sem líður stærð og námsframboði ber öllum íslensku háskólunum að rækja það tvíþætta hlutverk að sinna jafnt innlendum þörfum sem alþjóðlegu fræðasamfélagi. Eftirfarandi kaflar fjalla á einn eða annan hátt um það vandasama verkefni.

Í 2. kafla hér á eftir, *Háskólar og skjalfest málstefna*, er greint frá íslenskri málstefnu, lögum um íslenskt mál og skjalfestri málstefnu íslenskra háskóla að svo miklu leyti sem hún liggur fyrir. Í 3. kafla, sem nefnist *Útgáfumál*, er að finna úttekt á því á hvaða tungumáli íslenskir háskólastarfsmenn birta einkum rannsóknir sínar og á hvaða máli doktorsritgerðir eru skrifaðar við íslenska háskóla. 4. kafli nefnist *Kennslumál* og þar má sjá hver hlutur ensku og íslensku er í grunnnámi og í framhaldsnámi í íslensku háskólunum. Í 5. kafla, *Háskólastarfsemi þvert á landamæri og þjóðerni*, er m.a. greint frá samsetningu nemenda- og kennarahópsins í íslensku skólunum m.t.t. ríkisfangs. Í 6. kafla, *Niðurlagi*, er að finna stutta samantekt og lokaorð.

2 Háskólar og skjalfest málstefna

2.1 Opinber málstefna og háskólar á Íslandi

Alþingi samþykkti árið 2011 *Lög um stöðu íslenskrar tungu og íslensks táknumáls*, að tillögu sérfræðinganevndar mennta- og menningarmálaráðherra. Löggjöfin markar ákveðin tímamót í sögu íslenskrar málræktar og málstefnu. Í lögnum er kveðið á um að íslenska sé þjóðtunga Íslendinga og opinbert mál á Íslandi (1. gr.) og að íslenska sé opinbert mál Íslands á alþjóðavettvangi (12. gr.). Lögin mæla svo fyrir að ríki og sveitarfélög beri ábyrgð á að varðveita og efla íslenska tungu og skuli sjá til þess að hún sé notuð (5. gr.); þjóðtungan sé sam-

eiginlegt mál landsmanna og stjórnvöld skuli tryggja að unnt verði að nota hana á öllum sviðum íslensks þjóðlífs (2. gr.). Síðastnefnda ákvæðið felur m.a. augljóslega í sér það markmið að íslenskt mál sé nothæft í öllum greinum vísinda sem stundaðar eru á Íslandi.

Lögin um stöðu íslenskrar tungu og íslensks táknmáls geyma ákvæði um íslensku sem „mál Alþingis, dómstóla, stjórnvalda, jafnt ríkis sem sveitarfélaga, skóla á öllum skólastigum og annarra stofnana sem hafa með höndum framkvæmdir og veita almannþjónustu“ (8. gr.). Ákvæðið um íslensku sem mál „skóla á öllum skólastigum“ varðar vitaskuld íslensku háskólana og verður nánar vikið að því hér á eftir.

Í lögunum er einnig sérstök grein um íslenskan fræðiorðaforða. „Stjórnvöld skulu stuðla að því að íslenskur fræðiorðaforði á ólíkum sviðum eflist jafnt og þétt, sé öllum aðgengilegur og notaður sem víðast“ (11. gr.). Þetta ákvæði sýnir ótvírætt að löggjafinn ætlast til þess að háskólakennarar og aðrir íslenskir fræðimenn í stofnunum íslenska ríkisins hafi vöxt og viðgang íslenskra fræðiorða í huga í störfum sínum og skrifum.

Af öðrum ákvæðum laganna, sem varða íslensku, má nefna að 10. gr. þeirra mælir svo fyrir að „[m]ál það sem notað er í starfsemi ríkis og sveitarfélaga eða á vegum þeirra [skuli] vera vandað, einfalt og skýrt“ og að ríki og sveitarfélög skuli leita samvinnu við Íslenska málnefnd um málstefnu og stöðu íslenskrar tungu (5. gr.).

Samkvæmt háskólalögum (*Lög um háskóla, Lög um opinbera háskóla*) gegna íslenskir háskólar tvíþættu hlutverki eins og vikið var að í inngangi. Annars vegar eru þeir hluti af alþjóðlegu mennta- og vísindasamfélagi. Hins vegar hafa þeir sem íslenskar vísinda- og fræðslustofnanir sérstökum skyldum að gegna við íslenskt samfélag og þeim ber að styrkja innviði þess. Háskólalögin kveða ekki á um mál kennslu og rannsókna. Þar koma aftur á móti til skjalanna lögin um stöðu íslenskrar tungu og íslensks táknmáls sem fjallað var um hér á undan.

Í íslenskri málstefnu (*Íslenska til alls*), sem Alþingi samþykkti í mars 2009² (*Þingsályktun um íslenska málstefnu*), segir m.a. að þótt í *Lögum um háskóla* sé hvergi „minnst á íslenskt mál sem megingtungumál

² Með þingsályktun 12. mars 2009 féllst Alþingi á tillögu mennta- og menningarmálaráðherra um að taka upp, sem íslenska málstefnu, tillögur Íslenskrar málnefndar í ritinu *Íslenska til alls*. Í þingsályktuninni segir: „Alþingi lýsir yfir stuðningi við það meginmarkmið tillagna Íslenskrar málnefndar að íslenska verði notuð á öllum sviðum íslensks samfélags“ (*Þingsályktun um íslenska málstefnu*).

íslenskra háskóla beinum orðum“ hljóti það þó að vera „frumskilyrði þess að háskólar geti styrkt innviði íslensks samfélags“. Þar lýsir Íslensk málnefnd viðhorfi sínu til stöðu íslensku í háskólastarfi þannig:

Án traustrar stöðu tungumálsins stendur samfélagið völtum fótum sem íslenskt samfélag og innviðirnir fúna. Hlutverk háskóla samkvæmt [lögum um háskóla] er meðal annars að stuðla að miðlun þekkingar og færni til samfélagsins alls. Þar öðlast nemendur menntun á flestum sviðum vísinda og tækni og dreifast að námi loknu um þjóðfélagið allt til þess að vinna að sérgreinum sínum og miðla vitneskju sinni til samfélagsins. En þekkingu og færni verður ekki miðlað til íslensks samfélags á öðru máli en íslensku ef sem flestir eiga að geta nýtt sér það sem í boði er. Íslenska þarf þess vegna að vera nothæf í vísindum og fræðum og búa yfir öllum þeim orðaforða sem nauðsynlegur er til þess að eðlileg samskipti geti orðið á milli þess sem þekkinguna hefur numið og þess sem tekur við henni. (*Íslenska til alls* 2009:42)

Eins og fram hefur komið kveður 8. gr. *Laga um stöðu íslenskrar tungu og íslensks táknumáls* á um að íslenska sé mál „skóla á öllum skólastigum“. Það er engum vafa undirorpið að orðalagið „öll skólastig“ merkir að háskólarnir eru hér engan veginn undanskildir.

Eigi að síður hljóta háskólarnir sökum sérstöðu sinnar að hafa ákveðið svigrúm, og þá einkum í meistara- og doktorsnámi, til að kenna einstök námskeið eða jafnvel heilar námsbrautir á öðru tungumáli. Í íslenskri málstefnu (*Íslenska til alls* 2009:48) er tilgreint sem sérstakt markmið að „[í]slenska sé opinbert mál allra háskóla á Íslandi og kennsla fari þar að jafnaði fram á íslensku“. Um grunnnám segir: „Kennsla í grunnnámi ætti öll að fara fram á íslensku, nema um sé að ræða nám í erlendum tungumálum, og lokaritgerðir í grunnnámi ættu einnig að vera skrifaðar á íslensku“ og hvað varðar framhaldsnám er stefnan sú að: „[f]ramhaldsnám í háskólum á Íslandi verði að jafnaði á íslensku“.

Af framansögðu er ljóst að Alþingi hefur á allra síðustu árum, þ.e. í þingsályktun um íslenska málstefnu 2009 og í löggjöf um stöðu íslenskrar tungu 2011, látið í ljósi skýran vilja og fyrirmæli um forgang íslenskrar tungu í íslensku háskólastarfi — jafnframt því sem viðurkenndur er sá veruleiki að íslenska er ekki eina tungumálið sem nota þarf og nota má í vísindum og háskólakennslu á Íslandi.

2.2 Skjalfest málstefna íslenskra háskóla

Sjö háskólar starfa á Íslandi: Háskóli Íslands (HÍ), Háskólinn á Akureyri (HA), Háskólinn á Bifröst (HB), Háskólinn í Reykjavík (HR), Hólaskóli — Háskólinn á Hólum (HH), Landbúnaðarháskóli Íslands (LBHÍ) og Listaháskóli Íslands (LHÍ). *Tafla 1* sýnir nemendafjölda haustið 2011, hverjir þeirra eru hreinir ríkisskólar og hver viðurkennd fræðasvið háskólanna eru (*Ný sýn. Breytingar á vísindagöngum nýsköpunarkerfinu*). Í *töflu 1* eru dregnar helstu línur um námsframboðið en myndin er flóknari. Ef t.d. er litið til lista er ritlist einnig kennd innan Hugvísindasviðs HÍ og nokkrar listgreinar eru kenndar í tengslum við kennslufræði á Menntavísindasviði HÍ.

	Ríkisskóli	Nemar 2011	Viðurkennd fræðasvið							
			Fél. vís.	Heilb. vís.	Hugvís.	Menntavís.	Nátt. vís.	Verk+ tæ.vís.	Auðl.+ búvís.	Listir
HÍ	x	13.919	x	x	x	x	x	x		
HA	x	1.493	x	x					x	
HB		431	x							
HR		2.468	x					x		
HH	x	172							x	
LBHÍ	x	238					x		x	
LHÍ		414								x

Tafla 1. Yfirlit yfir rekstrarform, stærð og vísindasvið íslenskra háskóla.

Mennta- og menningarmálaráðherra beindi því til allra íslensku háskólanna árið 2010 að taka saman og birta málstefnu viðkomandi skóla.

Á heimasíðum sex íslenskra háskóla af sjö er að finna skjalfesta málstefnu sem samþykkt hefur verið af skólayfirvöldum. Hér er um að ræða Háskóla Íslands síðan 2004, Háskólann á Akureyri síðan 2008, Landbúnaðarháskóla Íslands síðan 2011 og Hólaskóla — Háskólann á Hólum, Háskólann á Bifröst og Listaháskóla Íslands síðan 2012.

Háskólinn í Reykjavík sker sig úr. Í *Stefnu Háskólans í Reykjavík* er ekki vikið beinlínis að tungumálum í kennslu og rannsóknum. Af stefnunni almennt má þó ráða að starfsmenn og stúdentar reiði sig töluvert á erlendum tungumálum auk íslensku:

Háskólinn í Reykjavík er alþjóðlegt háskólasamfélag sem notar alþjóðleg viðmið í rannsóknum, kennslu og stjórnun. HR starfar markvisst með leiðandi erlendum háskólum og rannsóknastofnunum og leggur áherslu á að þjálfa nemendur í alþjóðlegum samskiptum.

Hinir háskólarnir sex hafa, eins og fram hefur komið, samþykkt og birt málstefnu sem er öllum aðgengileg á heimasíðu viðkomandi skóla: *Málstefna* [Listaháskóla Íslands], *Málstefna Háskóla Íslands*, *Málstefna Háskólans á Akureyri*, *Málstefna Háskólans á Bifröst*, *Málstefna Hólaskóla – Háskólans á Hólum*, *Málstefna Landbúnaðarháskóla Íslands (LbhÍ)*.

Við töldum áhugavert að fara í gegnum málstefnuskjölin sex og leita eftir meginatriðum þeirra, hvar þeim bæri saman og hvar munur væri á þeim. Helstu niðurstöður þessarar greiningar koma fram í *töflu 2*. Eins og fram hefur komið hefur HR ekki birt sérstaka skjalfesta málstefnu og er skólinn því ekki með hinum sex skólunum í *töflu 2*.

Áhersluatriði í skráðri málstefnu	HÍ	HA	HB	HH	LBHÍ	LHÍ
Íslenska aðal tungumál skólans	x	x	x	x	x	x
Erlendir starfsmenn læri íslensku	x	x	x	x	x	x
Greinarmunur gerður á kennslumáli í grunn- og framhaldsnámi	x	x		x	x	x
Áhersla á íslenskan íðorðaförða, íslenska málrækt	x	x	x	x	x	x
Kennarar og nemendur séu góðir í ensku og öðrum erl. málum		x	x	x		
Tryggja skal erlendum skiptistúdentum næg námskeið á ensku	x	x				
Námskeið haldin um ísl. tungu og menningu fyrir erl. skiptistúd.		x				
Skólinn þarf að geta boðið námskeið/námsleiðir á erl. málum			x			
Ætlast til að erl. nem. geti stundað nám sitt á íslensku				x		
Fleiri erlend mál en enska mikilvæg				x		
Koma skal til móts við þarfir nem. með erl. móðurmál í námsmati						x

Tafla 2. Áhersluatriði í málstefnuskjölum sex íslenskra háskóla.

Málstefnuskjöl skólanna sex tiltaka öll að íslenska sé aðal tungumálið í viðkomandi skóla:

Háskóli Íslands: Málstefnan hefur að leiðarljósi að talmál og ritmál Háskólans er íslenska, jafnt í kennslu, rannsóknnum sem stjórnýslu.

Háskólinn á Akureyri: Háskólinn á Akureyri starfar innan íslensks samfélags, á íslensku málsvæði, og íslenska er opinbert tungumál skólans.

Háskólinn á Bifröst: Íslenska er opinbert tungumál Háskólans á Bifröst og íslenska er almennt tal- og ritmál á öllum sviðum skólans í kennslu, rannsóknnum og stjórnýslu.

Hólaskóli — Háskólinn á Hólum: Íslenska er tungumál skólans en vald á öðrum málum verður einnig að vera til staðar [...] Íslenska er [...] opinbert tungumál Hólaskóla [...] skulu öll skjöl sem skólinn lætur frá sér fara teljast vera í frumriti á íslensku. [...] Fundir innan stjórnýslunnar skulu að jafnaði fara fram á íslensku en fari þeir fram á ensku eða öðru erlendu tungumáli skal þess getið sérstaklega í fundarboði.

Landbúnaðarháskóli Íslands: Íslenska er almennt tal- og ritmál á öllum sviðum skólans, í kennslu, rannsóknnum og stjórnýslu.

Listaháskóli Íslands: Íslenska er helsta samskiptamál í skólanum, bæði sem talmál og ritað mál, hvort sem er í kennslu, rannsóknnum eða stjórnýslu.

Allir hvetja skólarnir erlenda starfsmenn sína til íslenskunáms. Fastast er kveðið að orði í málstefnu HA þar sem erlendum starfsmönnum er settur tiltekinn tímarammi til að ná árangri í íslenskunáminu:

Ætlast er til að fastráðnir, erlendir kennarar leggi sig fram um að læra íslensku enda veitir Háskólinn þeim aðstoð í þeim efnum. Fastráðnir kennarar sem ekki hafa íslensku að móðurmáli skulu innan fimm ára hafa öðlast nægilega færni í tungumálinu til að eiga almenn samskipti, geta setið fundi og geta hagnýtt sér upplýsingar frá stjórnýslu háskólans [...] Háskólinn býður upp á námskeið í íslensku og íslenskri menningu fyrir fastráðna kennara erlendis frá.

Að því er varðar val á kennslutungumáli gera Háskóli Íslands, Háskólinn á Akureyri, Hólaskóli — Háskólinn á Hólum, Landbúnaðarháskóli Íslands og Listaháskóli Íslands með einhverjum hætti greinarmun á grunnnámi og framhaldsnámi, þ.e. lögð er enn ríkari áhersla á

íslenskuna sem kennslumál í grunnnáminu en í framhaldsnámi. Sem dæmi má taka umfjöllun um þetta atriði í málstefnu HÍ en hliðstæð umfjöllun er hjá hinum skólunum:

[...] kennsla og próf til fyrstu háskólagráðu fara að mestu fram á íslensku. Víkja má frá þessari meginreglu ef sérstök ástæða er til, svo sem við kennslu í erlendum málum, ef kennari er erlendur eða kennsla er einnig ætluð útlendingum. Rannsóknnum og framhaldsnámi fylgja erlend samskipti og fleiri mál en íslenska eru notuð í því starfi, einkum enska. Megin kennslumál í framhaldsnámi er þó íslenska, eftir því sem við verður komið [...] Deildir Háskólans geta boðið upp á heilar námsleiðir í framhaldsnámi eða einstök námskeið á erlendum málum, eftir því sem aðstæður leyfa.

Skólarnir sex leggja ríka áherslu á málrækt, íslenskan íðorðaförða og nauðsyn þess að nemendur og kennarar geti fjallað um sérgreinar sínar á íslensku. Orðalag um þessi atriði er svolítið mismunandi en textinn er alls staðar fremur ýtarlegur. Hér skal látið nægja að birta klausu úr málstefnu LBHÍ sem dæmi:

Háskólinn skal með ákveðnum hætti hvetja starfsmenn til að sinna málrækt, t.d. með því taka saman fagorðasöfn og birta rannsóknir á íslensku jafnt sem ensku eða öðrum erlendum málum. Kennarar eru sérstaklega hvattir til að leiðbeina nemendum um málnotkun eftir því sem við verður komið, auk venjubundinna faglegra leiðbeininga. Landbúnaðarháskólinn leggur áherslu á að vinna að varðveislu orðaförða sem og ný-orðasmíð á þeim fræðasviðum sem skólinn sinnir. Þannig tekur skólinn virkan þátt í nýsköpun, viðhaldi og þróun íslenskrar tungu.

Þessu til viðbótar leggja Háskólinn á Akureyri, Háskólinn á Bifröst og Hólaskóli — Háskólinn á Hólum áherslu á nauðsyn þess að kennarar og nemendur séu færir í ensku og öðrum erlendum málum og er hvatt til vandaðrar málnotkunar hvert svo sem tungumálið er. Sem dæmi má taka málstefnu HB en þar segir:

Gerðar eru ríkar kröfur um vandað málfar í kennslu á erlendum málum, rétt eins og í íslensku.

Háskóli Íslands og Háskólinn á Akureyri tilgreina sérstaklega framboð á námskeiðum á ensku handa erlendum skiptinimum sem skilja

ekki íslensku. Í málstefnu HÍ segist skólinn m.a. „leitast við að tryggja erlendum skiptistúdentum nægilegt framboð af sérhæfðum námskeiðum á ensku“. Hjá HA segir að tryggja verði „nægt framboð námsleiða [...] á ensku til að sinna þörfum skiptinema“. Jafnframt er tekið fram að HA bjóði upp á „námskeið í íslenskri tungu og menningu fyrir erlenda skiptinema“.

Háskólinn á Bifröst tiltekur ekki erlenda skiptinema sem afmarkaðan hóp sérstaklega né heldur erlenda tungumálið ensku umfram önnur erlend mál:

[...] eitt af markmiðum [HB er] að byggja upp öflugan alþjóðlegan háskóla. Til þess að svo megi verða þarf skólinn að geta boðið upp á heilar námsleiðir eða einstök námskeið á erlendum málum eftir því sem aðstæður og eftirspurn leyfa.

Hólaskóli — Háskólinn á Hólum tiltekur ekki heldur skiptinema sérstaklega heldur er vísað til erlendra nemenda yfirleitt. Jafnframt er lögð áhersla á að enska sé ekki eina erlenda tungumálið í alþjóðlegu samstarfi. Tekið er fram að erlendir nemendur eigi að geta numið á íslensku.

Listaháskóli Íslands er eini skólinn sem tilgreinir í málstefnu sinni ákveðnar leiðir í námsmati til að koma til móts við nemendur með annað móðurmál en íslensku.

Í inngangi var greint frá sameiginlegri yfirlýsingu norrænu menntamálaráðherranna, undirritaðri 2006, um stefnu um samhliða notkun viðkomandi þjóðtungu og ensku í háskólastarfinu (*Yfirlýsing um málstefnu Norðurlanda*). Þar er markmiðið að hvorugt málanna ryðji hinu úr vegi. Í málstefnu Háskólans á Akureyri er vikið að slíku samspili málanna tveggja: „Sú meginregla gildir um samspil íslensku og ensku innan háskólans að starfsmenn noti ensku ásamt með íslensku, en ekki í stað hennar.“ Í málstefnu Listaháskóla Íslands segir að auk þess sem skólinn geri kröfu um að akademískir starfsmenn miðli þekkingu til nærsamfélagsins séu „starfsmenn og nemendur skólans virkir þátttakendur í alþjóðlegri fagumræðu á öðrum tungumálum“. Þá segir um meistaranám í hönnunar- og arkitektúrdeild, myndlistardeild og tónlistardeild að enska sé helsta samskiptamálið í deildunum en eigi að síður sé „lögð áhersla á að nemendur séu virkir í samfélaginu í umræðu um listir og að þeir miðli þangað þekkingu sinni og færni“.

Sameiginleg meginstef í skráðri málstefnu íslensku háskólanna

sex, sem hér hefur verið rýnt í, eru einkum þau að íslenska sé aðal-tungumálið – og sé vikið frá meginreglunni um að kenna á íslensku eigi það fremur við á framhaldsstigi en í grunnnámi – og að lögd sé rík áhersla á íslenskan íðorðaförða og vandaða íslenska málnotkun. Það skal áréttað að Háskólinn í Reykjavík er ekki með í greiningunni hér á undan enda hefur skólinn ekki birt skjalfesta málstefnu.

3 Útgáfumál

3.1 Fræðileg skrif háskólamanna

Enda þótt það sé yfirlýst stefna að íslenska sé aðalmál háskóla á Íslandi er hún fjarri því að vera aðalmálið í fræðilegum skrifum. Tölur byggðar á framtali starfa akademískra starfsmanna við tvo stærstu háskólana á Íslandi, Háskóla Íslands og Háskólann í Reykjavík, sýna að árið 2011 var enska aðal-tungumál í fræðilegum skrifum þeirra eins og sýnt er á mynd 1 hér á eftir. Þessar tölur og aðrar í þessum kafla eru byggðar á gögnum frá Vísinda- og nýsköpunarsviði Háskóla Íslands og Rannsóknarþjónustu Háskólans í Reykjavík.³

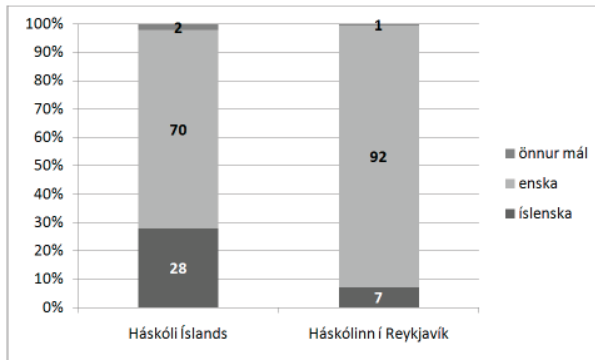
Árið 2011 voru 27,9% framtalinna skrifa akademískra starfsmanna við Háskóla Íslands á íslensku, 69,8% á ensku og aðeins 2,3% á öðrum tungumálum. Af skrifum akademískra starfsmanna Háskólans í Reykjavík voru 92,3% á ensku, 7,1% á íslensku og aðeins 0,6% á öðrum tungumálum. Skyld er að hafa í huga að þessi gögn eru ekki fyllilega samanburðarhæf. Í fyrsta lagi er nokkur munur á matskerfum skólanna.⁴ Í annan stað eru skólarnir sjálfir æði ólíkir. Við Háskólann

³ Gögn úr Háskóla Íslands eru byggð á árlegu framtali starfa akademískra starfsmanna skólans sem varðveitt eru í gagnagrunni Vísinda- og nýsköpunarsviðs. Farið var yfir gögn fyrir árin 2005–2011 í þessum gagnagrunni, alls ríflega 13.200 færslur. Þar eru gögnin flokkuð eftir *Matskerfi opinberra háskóla* en Haraldur Bernharðsson ber ábyrgð á greiningu tungumáls ritverkanna. Leitast var við að útrýma tvískráningum eða fjölskráningum fjölhöfundaverka en slíkt er reyndar afar erfitt vegna misræmis í skráningu verkanna. Þá verður að hafa í huga að matskerfið hefur tekið nokkrum breytingum á þessum tíma. Þrátt fyrir þessa ágalla ættu gögnin að gefa nokkuð góða mynd af stöðu mála. Höfundar þakka starfsmönnum Vísinda- og nýsköpunarsviðs Háskóla Íslands, einkum Reyni Erni Jóhannssyni, fyrir aðgang að gögnunum og margháttða greiðvikni við úrvinnslu þeirra.

Gögn úr Háskólanum í Reykjavík eru byggð á skýrslunni *Styrkur Háskólans í Reykjavík í rannsóknnum* og viðbótargögnum frá Rannsóknarþjónustu Háskólans í Reykjavík. Höfundar þakka Kristjáni Kristjánssyni, forstöðumanni Rannsóknarþjónustu Háskólans í Reykjavík, aðstoð og veittar upplýsingar.

⁴ Torvelt er að fá fyllilega samanburðarhæf gögn. Í tölunum hér að ofan var ritdóm-

í Reykjavík eru fjórar deildir, Viðskiptadeild, Tölvunarfræðideild, Lagadeild og Tækni- og verkfræðideild en við Háskóla Íslands eru fimm fræðasvið, Félagsvísindasvið, Heilbrigðisvísindasvið, Hugvísindasvið, Menntavísindasvið og Verkfræði- og náttúruvísindasvið. Fræðileg skrif á íslensku eru mest í hug-, félags- og menntavísindum eins og fram kemur hér að neðan en engin hugvísindadeild er við Háskólann í Reykjavík og lagadeild Háskólans í Reykjavík samsvarar aðeins litlum hluta Félagsvísindasviðs Háskóla Íslands. Því kemur ekki á óvart að fræðileg skrif á íslensku skuli reynast miklu minni við Háskólann í Reykjavík en Háskóla Íslands. Ef til vill er eðlilegast að bera Háskólann í Reykjavík saman við Verkfræði- og náttúruvísindasvið Háskóla Íslands (sjá hér á eftir).



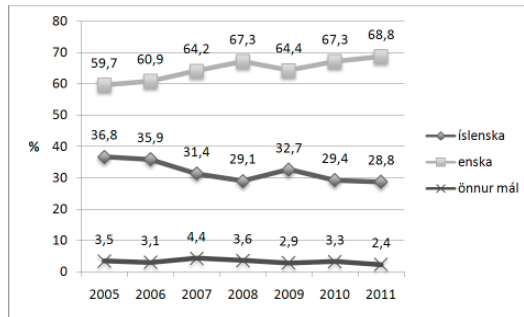
Mynd 1. Tungumál fræðilegra skrifa akademískra starfsmanna við Háskóla Íslands og Háskólann í Reykjavík árið 2011. Íslenska er neðst á myndinni, þá enska og efst eru önnur mál.

Við bæði Háskóla Íslands og Háskólann í Reykjavík hafa skrif á ensku aukist nokkuð undanfarin ár á kostnað annarra tungumála þótt breytingarnar séu reyndar ekki miklar eins og sjá má á *mynd 2* og *mynd 3* hér á eftir.⁵

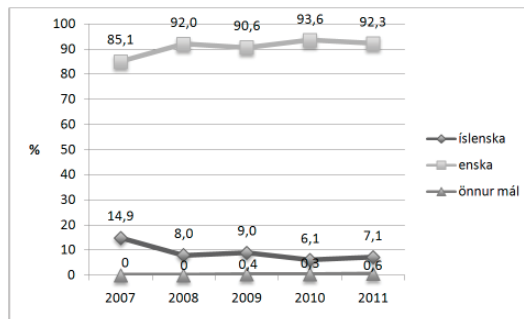
Tungumál fræðilegra skrifa er æði breytilegt eftir fræðasviðum eins og þegar hefur verið vikið að. Þetta sést glögglega þegar gögn um skrif akademískra starfsmanna við Háskóla Íslands árið 2011 eru flokkuð eftir sviðum eins og gert er á *mynd 4*.

um sleppt í skrifum við Háskóla Íslands þar sem ritdómar eru ekki taldir með í gögnunum frá Háskólanum í Reykjavík. Að ritdómum meðtöldum í Háskóla Íslands væru skrif á íslensku 28,8%, á ensku 68,8% og á öðrum málum 2,4%.

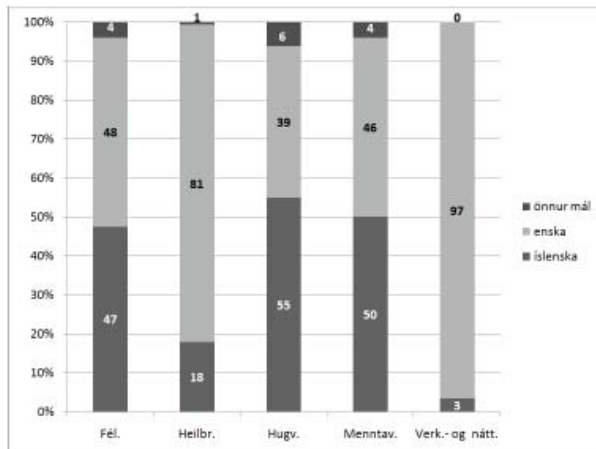
⁵ Tölur um skrif á norrænu málunum færeysku, norsku, sænsku og dönsku eru lágar, samtals að meðaltali 2,2% í Háskóla Íslands 2005–2011 og enn lægri við Háskólann í Reykjavík.



Mynd 2. Tungumál fræðilegra skrifta í Háskóla Íslands 2005–2011.



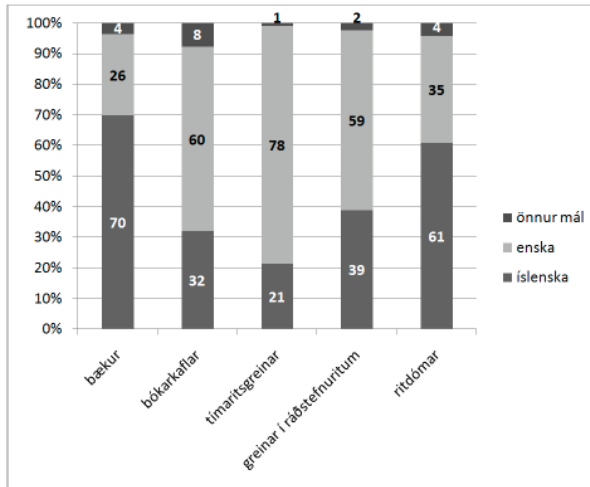
Mynd 3. Tungumál fræðilegra skrifta í Háskólanum í Reykjavík 2007–2011.



Mynd 4. Fræðileg skrif í Háskóla Íslands 2011 eftir fræðasviðum skólans: Félagsvísindasviði (Féll.), Heilbrigðisvísindasviði (Heilbr.), Hugvísindasviði (Hugv.), Menntavísindasviði (Menntav.) og Verkfræði- og náttúruvísindasviði (Verk- og nátt.). Íslenska er neðst á myndinni, þá enska og efst eru önnur mál.

Skrif á íslensku eru mest á Hugvísindasviði (55%), Menntavísindasviði (50%) og Félagsvísindasviði (47,5%) og þar er líka skrifað svolítið á öðrum tungumálum en íslensku og ensku þótt í litlum mæli sé. Á Verkfræði- og náttúruvísindasviði og Heilbrigðisvísindasviði eru skrif akademískra starfsmanna aftur á móti að langmestu leyti á ensku (96,7% og 81,4%).

Í matskerfi opinberra háskóla er akademískum verkum skipað niður í flokka. Á mynd 5 er sýnd greining á tungumáli framtalinna skrifa við Háskóla Íslands árið 2011 í fimm flokkum: bókum, bókar-köflum, tímaritsgreinum, greinum í ráðstefnuritum og ritdómum. Þar eru skrif á íslensku mest í flokki bóka og ritdóma en skrif í þessum tveimur flokkum eru nær eingöngu á Hugvísindasviði, Menntavísindasviði og Félagsvísindasviði. Staða íslensku er langveikust í flokki tímaritsgreina.

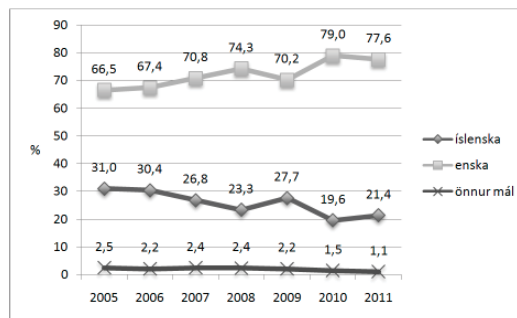


Mynd 5. Fræðileg skrif í Háskóla Íslands 2011. Hlutfallsleg skipting eftir tungumáli í mismunandi flokkum. Íslenska er neðst á myndinni, þá enska og efst eru önnur mál.

Í flokki tímaritsgreina við Háskóla Íslands hefur orðið hægfara fjölgun greina á ensku undanfarin ár á kostnað greina á öðrum málum. Á mynd 6 hér á eftir er sýnt yfirlit yfir framtaldar tímaritsgreinar akademískra starfsmanna árin 2005–2011 eftir tungumálum. Á þessu tímabili eru skrif langmest á ensku, að meðaltali 72%, en skrif á íslensku eru að meðaltali 26%. Tímaritsgreinar á Norðurlandamálunum færeysku, norsku, dönsku og sænsku og öðrum málum (þar á meðal þýsku, frönsku og spænsku) eru samtals að meðaltali 2%.

Í íslensku háskólasamfélagi, eins og við háskóla víðast hvar annars

staðar, er lögð rík áhersla á ritrýni allra fræðilegra skrifa; skrif á öðrum vettvangi eru lítils metin. Auk þess eru skrif á alþjóðlegum vettvangi að jafnaði meira metin en önnur skrif. Verk akademískra starfsmanna við Háskóla Íslands eru metin eftir *Matskerfi opinberra háskóla*. Þar eru, svo dæmi sé tekið, veitt 15–20 stig fyrir grein í tímariti á alþjóðlegum lista *Thomson Reuters Web of Knowledge (ISI Web of Knowledge)* eða í A- eða B-flokki á lista *European Research Index for the Humanities (ERIH)*. Ritryndar greinar birtar annars staðar eru metnar til 5–10 stiga. Hlutfall (12,5%) af launum allra akademískra starfsmanna rennur í sérstakan sjóð (Vinnumatssjóð eða Ritlauna- og rannsóknasjóð prófessora) sem greitt er úr einu sinni á ári í samræmi við rannsóknavirkni sem mæld er í rannsóknastigum. Þessar greiðslur geta verið umtalsverðar, oft jafngildi einna til tvennra mánaðarlauna. Tilgangur þessa fyrirkomulags er vitanlega að örva rannsóknavirkni en um leið ræður það miklu um val birtingarvettvangs og þar með einnig tungumáls því að afar erfitt er að fá skrif á íslensku metin í efsta flokki. Bókaforlög í efsta matsflokknum gefa ekki út bækur á íslensku og nauðafá tímarit í efstu matsflokkunum birta greinar á íslensku.



Mynd 6. Tungumál tímaritsgreina í Háskóla Íslands 2005–2011.

3.2 Doktorsritgerðir

Háskóli Íslands var stofnaður árið 1911 og er elsti og jafnframt langstærsti háskólinn á Íslandi. Fyrsta doktorsritgerðin var varin 1919 en doktorsvarnir voru fáar lengi framan af og þegar Háskóli Íslands fagnaði hálfri aldar afmæli sínu 17. júní 1961 höfðu aðeins verið varðar þar 25 doktorsritgerðir (Guðni Jónsson 1961:129). Doktorsritgerðum við Háskóla Íslands tók þó að fjölga verulega eftir aldamótin 2000 eins og nánar verður greint frá hér á eftir.

Aðeins þrjár doktorsritgerðir voru varðar við Kennaraháskóla Ís-

lands áður en hann sameinaðist Háskóla Íslands 1. júlí 2008. Þær voru allar varðar síðasta starfsár skólans og eru tvær þeirra á íslensku en ein á ensku (*Doktorsritgerðaskrá Landsbókasafns Íslands – Háskólabókasafns*; einnig: *Háskóli Íslands – Doktorsvarnir*).

Við Háskólann í Reykjavík er í boði doktorsnám við fjórar deildir skólans, Tölvunarfræðideild, Tækni- og verkfræðideild, Viðskipta-deild og Lagadeild. Fyrsta doktorsvörnin við Háskólann í Reykjavík fór fram í árslok 2009 í tölvunarfræðum og var ritgerðin á ensku (*Fréttasafn á vef Háskólans í Reykjavík og Háskólinn í Reykjavík – Árskýrsla 2010*, bls. 10).

Doktorsritgerðir hafa ekki verið varðar við aðra háskóla á Íslandi. Umfjöllun um tungumál doktorsritgerða við íslenska háskóla hlýtur því fyrst og fremst að beinast að doktorsritgerðum við Háskóla Íslands.

Varðar voru 75 doktorsritgerðir við Háskóla Íslands frá stofnun hans 1911 til ársins 1999. Nokkuð vantar því upp á að þar hafi verið varin ein ritgerð á ári á þessu tímabili. Doktorsvörnum fjölgaði þó jafnt og þétt eftir því sem leið á tímabilið, eins og sýnt er á mynd 7.

Doktorsritgerðir varðar við Háskóla Íslands á þessu tímabili voru á þremur tungumálum. Nokkurt jafnræði er með íslensku og ensku, eins og sjá má á töflu 3 hér á eftir, en enskan hefur þó yfirhöndina. Helmingur ritgerðanna (50%) var á ensku, hátt í helmingur, eða 44,7%, á íslensku en fjórar ritgerðir eða 5,3% voru á þýsku. Þrír Íslendingar skrifuðu ritgerðir sínar á þýsku, allar í læknisfræði (árin 1933, 1935 og 1942) en höfundur fjórðu ritgerðarinnar, sem fjallaði um tónlist og var varin 1959, var þýskur að uppruna (*Doktorsritgerðaskrá Landsbókasafns Íslands – Háskólabókasafns*).

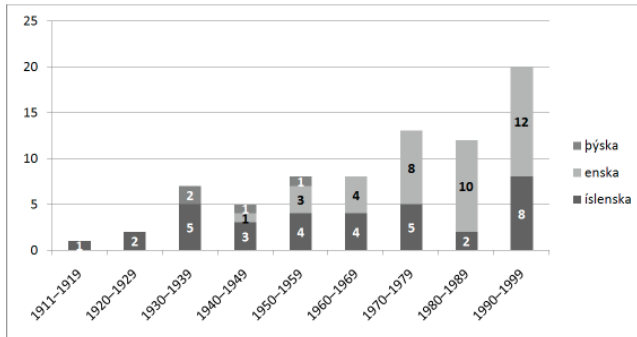
Tungumál	Fjöldi	Hlutfall
þýska	4	5,3
íslenska	34	44,7
enska	38	50,0
alls	76	100

Tafla 3. Tungumál doktorsritgerða við Háskóla Íslands 1911–1999.

Hlutdeild hvers tungumálanna þriggja á tímabilinu 1911–1999 er sýnd á mynd 7. Athygli vekur hve mjög hlutur ensku eykst síðustu þrjá áratugi tuttugustu aldarinnar en þá hefur notkun þýsku alveg lagst af fyrir nokkru.

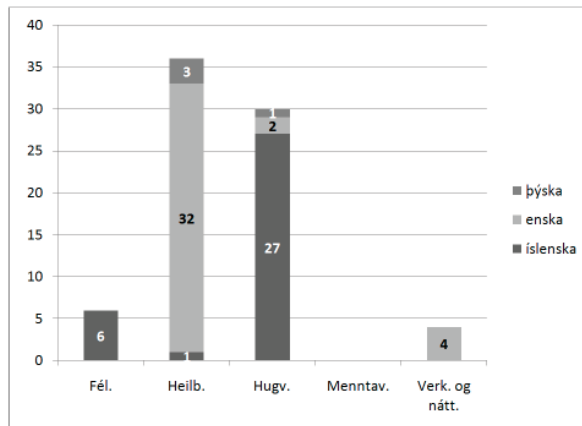
Ritgerðir á öðrum málum en íslensku eru langflestar í læknisfræði

og öðrum heilbrigðisvísindum (geðlæknisfræði, tannlækningum og lyfjafræði), fáeinar í jarðfræði og jarðeðlisfræði en jafnframt er ein á ensku í guðfræði, ein á þýsku í tónlist og ein á ensku í íslenskum bókmenntum. Á mynd 8 er gerð tilraun til að skipa doktorsritgerðum á tímabilinu 1911–1999 undir núverandi fræðasvið Háskóla Íslands.



Mynd 7. Tungumál doktorsritgerða við Háskóla Íslands 1911–1999.

Ritgerðirnar, sem þarna eru taldar til Félagsvísindasviðs, eru allar í lögfræði og þær sem taldar eru til Verkfræði- og náttúruvísindasviðs eru í jarðfræði, jarðeðlisfræði og verkfræði. Við Heilbrigðisvísindasvið eru fyrst og fremst ritgerðir í læknisfræði en ritgerðir við Hugvísindasvið eru flestar í íslenskum fræðum (íslenskum bókmenntum, málfræði og sagnfræði) en auk þess í guðfræði, listfræði og tónlist.



Mynd 8. Tungumál doktorsritgerða við Háskóla Íslands 1911–1999 miðað við núverandi skiptingu skólans í fræðasvið.

Á tíunda áratug tuttugustu aldar hófst skipulagt doktorsnám við Háskóla Íslands en fram að því höfðu doktorsefni við skólann ekki

hlotið þar neina formlega leiðbeiningu við doktorsrannsóknir sínar (Gunnar Karlsson o.fl. 2011:671–675). Í kringum aldamótin 2000 fer hið nýja doktorsnám að bera ávöxt og þá verða ákveðin tímamót því að doktorsritgerðum tekur að fjölga mikið og langmestur hefur vöxturinn verið frá 2008 eins og sjá má á mynd 9 (*Háskóli Íslands – Doktorsvarnir*).

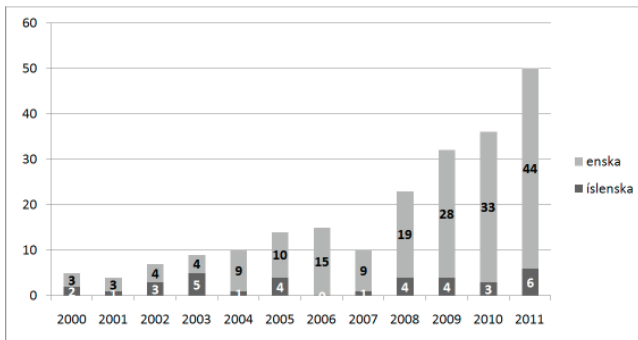
Á tímabilinu 2000–2011 voru varðar alls 215 ritgerðir við Háskóla Íslands. Mikill meirihluti þeirra er á ensku, alls 181 ritgerð eða 84,2%, eins og sýnt er á töflu 4 en aðeins 34 eru á íslensku, 15,8%. Ekki er um önnur tungumál að ræða (*Háskóli Íslands – Doktorsvarnir*).

Tungumál	Fjöldi	Hlutfall
íslenska	34	15,8
enska	181	84,2
alls	215	100

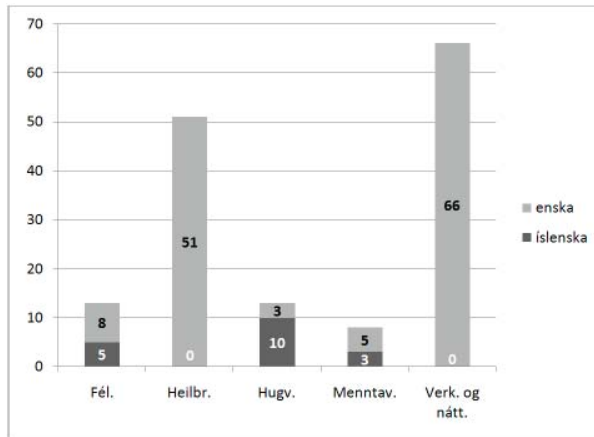
Tafla 4. Tungumál doktorsritgerða við Háskóla Íslands 2000–2011.

Sá mikli vöxtur, sem færast í doktorsritgerðaskrif á þessu tímabili, er að heita má allur á ensku eins og sýnt er á mynd 9.

Ef doktorsritgerðir áráanna 2007–2011, alls 151 talsins, eru flokkaðar eftir fræðasviðum skólans kemur glögglega í ljós að langflestar doktorsritgerðir á þessu tímabili eru á Heilbrigðisvísindasviði, 51 ritgerð eða 33,8%, og Verkfræði- og náttúruvísindasviði, 66 ritgerðir eða 43,7%. Yfirgnæfandi meirihluti doktorsritgerða við Háskóla Íslands á þessu tímabili, 77,5%, er við þessi tvö fræðasvið en þær eru allar á ensku eins og sjá má á mynd 10. Meirihluti ritgerða við Félagsvísindasvið og Menntavísindasvið er á ensku en Hugvísindasvið sker sig úr því að þar er meirihluti ritgerðanna á íslensku.



Mynd 9. Tungumál doktorsritgerða við Háskóla Íslands 2000–2011.



Mynd 10. Tungumál doktorsritgerða við Háskóla Íslands árin 2007–2011 eftir sviðum.

Nýtt stjórnskipulag Háskóla Íslands og skipting í fimm fræðasvið tók gildi 1. júlí 2008 og í kjölfar þess hefur átt sér stað ýmiss konar endurskipulagning og endurnýjun á reglum, þar á meðal reglum um doktorsnám (*Háskóli Íslands — Yfirlit um reglur HÍ*).

Í reglum um doktorsnám og doktorspróf við Heilbrigðisvísindasvið (nr. 1252/2011) er mælt fyrir um að doktorsritgerð skuli skrifa á ensku en henni skuli fylgja útdráttur bæði á íslensku og ensku (18. gr.). Ekki er getið um mögulegar undanþágur frá þessu skilyrði.

Í reglum um doktorsnám við Hugvísindasvið (nr. 154/2011), Menntavísindasvið (nr. 641/2011) og Félagsvísindasvið (nr. 500/2011) segir að ritgerð skuli að jafnaði vera á íslensku eða ensku og á öllum þremur sviðum er jafnframt ætlast til þess að útdráttur á báðum tungumálum fylgi. Í reglum Hugvísindasviðs og Menntavísindasviðs er jafnframt kveðið á um að doktorsritgerð megi vera á öðrum erlendum tungumálum ef doktorsnefnd eða doktorsráð samþykkir og við Hugvísindasvið getur deild veitt doktorsefni leyfi til að tala erlent tungumál við doktorsvörn.

Í reglum um doktorsnám við Verkfræði- og náttúruvísindasvið (nr. 642/2011) er ekki kveðið sérstaklega á um tungumál doktorsritgerðar eða útdrátt.

Á öllum fræðasviðum Háskóla Íslands er lögð áhersla á að andmælendur við doktorsvarnir séu fengnir utan deildar eða sviðs, annar eða jafnvel báðir. Öll svið nema Menntavísindasvið leggja á það áherslu að annar andmælandi að minnsta kosti sé fenginn utan Háskóla Íslands, sé þess kostur. Það þýðir í reynd að oft þarf að leita

til erlendra sérfræðinga og í reglum bæði Verkfræði- og náttúruvísindasviðs og Hugvísindasviðs er beinlínis ætlast til þess að fenginn sé sérfræðingur frá erlendri rannsóknastofnun sé þess kostur. Þessi ákvæði um andmælendur setja tungumáli ritgerðar og doktorsvarnar nokkrar skorður þar sem á mörgum sviðum getur reynst torvelt að fá utanaðkomandi andmælendur sem læsir eru á íslensku.

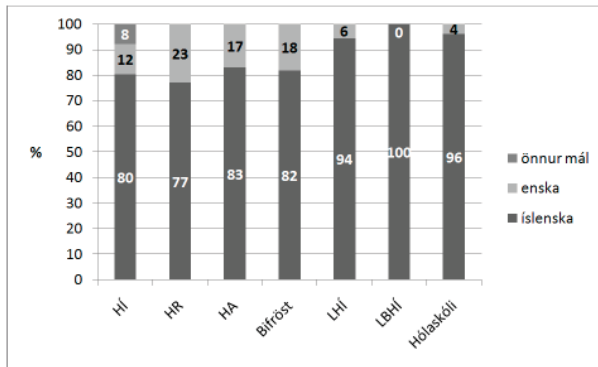
4 Kennslumál

4.1 Yfirlit

Íslenska er aðalkennslumál allra íslenskra háskóla. Við Háskóla Íslands er mikil kennsla í erlendum tungumálum sem fer eðlilega að miklu leyti fram á viðkomandi erlendu máli. Kennsla í erlendum málum er óveruleg eða engin við aðra háskóla. Við flesta háskólana fer kennsla að einhverju leyti fram á ensku. Á mynd 11 eru sýndar hlutfallstölur byggðar á fjölda námskeiða skólaárið 2011–2012 (en reyndar 2010–2011 fyrir Háskóla Íslands og Háskólann á Bifröst).⁶

Við Landbúnaðarháskóla Íslands fór öll kennsla fram á íslensku skólaárið 2011–2012 en við Hólaskóla eru 3,8% námskeiðanna á ensku og við Listaháskóla Íslands 5,6%. Kennsla á ensku er meiri við hina skólana, frá 11,7% við Háskóla Íslands til 23% við Háskólann í Reykjavík.

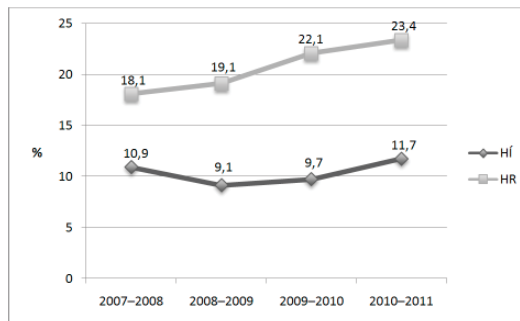
⁶ Upplýsingar um námskeið í Háskóla Íslands frá Sveini Klausen (í tölvupósti í mars 2012), námskeið í Háskólanum í Reykjavík frá Steini Jóhannssyni (í tölvupósti í apríl 2012), námskeið í Háskólanum á Akureyri frá Steinunni Aðalbarnardóttur og Stefáni Jóhannssyni (í tölvupósti í apríl 2012), námskeið í Listaháskóla Íslands frá Björgu Jónu Birgisdóttur (í tölvupósti í apríl 2012), námskeið í Landbúnaðarháskóla Íslands frá Þórunni Eddu Bjarnadóttur (í tölvupósti í apríl 2012), í Hólaskóla frá Hjördísi Gísladóttur (í tölvupósti í apríl 2012). Upplýsingar um námskeið við Háskólann á Bifröst eru fengnar úr *Kennsluskrá Háskólans á Bifröst skólaárið 2010–2011. Grunnám í öllum deildum* og *Kennsluskrá Háskólans á Bifröst skólaárið 2010–2011. Meistaránám í öllum deildum*.



Mynd 11. Tungumál í kennslu við íslenska háskóla skólaárið 2011–2012 (eða 2010–2011). Hlutfallstölur allra námskeiða í grunn- og framhaldsnámi.

4.2 Þróunin 2007–2011

Torvelt er að fá tölur langt aftur í tímann um fjölda námskeiða á ensku við íslenska háskóla en óhætt er að fullyrða að kennsla á ensku (ef frá eru talin námskeið sem beinlínis eru ætluð til að kenna tungumálið) hafi verið næsta fátíð fyrr en undir lok tuttugustu aldar. Fyrst fer verulega að kveða að notkun ensku sem vinnumáls í kennslustofu á fyrstu árum þessarar aldar.



Mynd 12. Hlutfall námskeiða á ensku af öllum námskeiðum við Háskóla Íslands og Háskólann í Reykjavík 2007–2011.

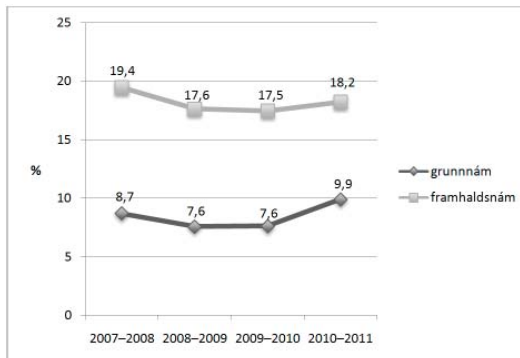
Á mynd 12 eru bornar saman tölur frá tveimur stærstu háskólunum, Háskóla Íslands og Háskólanum í Reykjavík, um fjölda allra námskeiða á ensku sem hlutfall af öllum kenndum námskeiðum skólaárin 2007–2011. Í Háskóla Íslands er hlutfallið að meðaltali 10,4% og í Háskólanum í Reykjavík 20,6%. Í Háskóla Íslands er hlutfallið á bilinu

10,9–11,7% á þessu tímabili og jókst heldur 2009–2011 eftir nokkra lækkun frá skólaárinu 2007–2008 til 2008–2009 eins og sjá má á mynd 12. Í Háskólanum í Reykjavík hefur hlutfallið aftur á móti aukist frá ári til árs öll þessi fjögur skólaár og fer hæst í 23,4% skólaárið 2010–2011 en reyndar má bæta því við að hlutfallið var 23% skólaárið 2011–2012 í Háskólanum í Reykjavík (sbr. mynd 11 að framan) svo að aukningin hefur stöðvast þar.

4.3 Grunnám og framhaldsnám

Ef greint er á milli grunnnáms og framhaldsnáms við Háskóla Íslands kemur í ljós að hlutfall námskeiða á ensku er drjúgum hærra í framhaldsnámi en í grunnnámi. Skólaárið 2007–2011 var hlutfall námskeiða á ensku að meðaltali 8% í grunnnámi en 18% í framhaldsnámi en tölur fyrir þessi ár eru sýndar á mynd 13.

Í Háskólanum á Akureyri er kennsla á ensku einnig mun meiri í framhaldsnámi en í grunnnámi. Skólaárið 2011–2012 voru 8,6% námskeiða í grunnnámi kennd á ensku en 36% í framhaldsnámi.⁷ Í Árskýrslu skólans um árið 2010 kemur fram að kennt er á ensku í allmörgum námskeiðum í félagsvísindadeild og lagadeild (bls. 20). Við lagadeild er BA-nám í lögfræði að hluta kennt á ensku og enn fremur er boðið upp á alþjóðlegt meistaranám í heimskautarétti sem alfarið er kennt á ensku (bls. 23–24).



Mynd 13. Hlutfall námskeiða á ensku við Háskóla Íslands eftir námsstigi skólaárin 2007–2011.

Athygli vekur að þessu er öfugt farið við Háskólann á Bifröst; þar er kennsla á ensku fyrst og fremst í grunnnámi. Í *Kennsluskrá* skólans fyr-

⁷ Skv. upplýsingum frá Steinunni Aðalbjarnardóttur og Stefáni Jóhannssyni í tölvu-pósti í apríl 2012.

ir grunnnám (staðnám) í öllum deildum skólaárið 2010–2011 voru 53 námskeið. Þar af voru 36 eða 68% sögð kennd á íslensku en 17 á ensku, eða 32%. Námskeið kennd á ensku voru einkum í viðskiptafræði. Í *Kennsluskrá* fyrir meistaranám í öllum deildum skólaárið 2010–2011 voru 46 námskeið. Af þeim voru 44 sögð kennd á íslensku.

4.4 Áhrif á náms- og kennsluaðferðir

Athuganir benda til þess að það geti haft merkjanleg áhrif á náms- og kennsluaðferðir ef enska leysir þjóðtunguna af hólmi sem kennslumál. Flestar rannsóknir benda til þess að það hafi fremur neikvæð heildaráhrif á námsárangur (sjá Salö 2010:17–19). Hér er þó enn margt ókannað. Sænskar rannsóknir sýna að nemendur spyrja síður og svara síður spurningum ef kennt er á ensku í stað sænsku, þeir eiga erfiðara með að fylgja röksemdafærslu fyrirlesara og halda þræði í fyrirlestrum, skrifa minna hjá sér af því sem kennarinn segir en einbeita sér frekar að því að hlusta. Aðrar breytingar eru m.a. að þeir lesa námsefnið fremur áður en þeir sækja fyrirlestrana og tala minna í kennslustundum (Salö 2010:18). Samkvæmt hollenskri rannsókn leiðir annað kennslutungumál en þjóðtungan til lakari námsárangurs fyrsta kastið en nemendur virðast ná sér á strik eftir að þeir hafa vanist kennslumálinu (s.st.).

Athuganir á stöðu ensku á Íslandi hafa m.a. leitt í ljós að um 90% af námsefni á háskólastigi eru á ensku og að íslenskir háskólakennarar virðist almennt þeirrar skoðunar að því fylgi lítil eða engin vandamál (Hafdís Ingvarsdóttir og Birna Arnbjörnsdóttir 2010). Raunar kemur fram í sömu könnun að langflestir hjálpa þeir íslenskum nemendum sínum á ýmsan hátt með enska orðaforðann. Þá sýna rannsóknir á viðhorfum stúdenta (Birna Arnbjörnsdóttir og Hafdís Ingvarsdóttir 2010) að a.m.k. þriðjungur háskólanema á Íslandi á í einhverjum erfiðleikum með að skilja fræðitexta á ensku en það er breytilegt eftir námsgreinum. Almennt telja íslenskir stúdentar sig góða í ensku þegar þeir koma í háskóla en samt sem áður telja þeir að það auki vinnuálág þeirra að nota ensku í stað íslensku í náminu og að þeir þurfi að beita ólíkum aðferðum við að tileinka sér námsefnið (Birna Arnbjörnsdóttir og Hafdís Ingvarsdóttir 2010). Í kjölfar mikillar fjölgunar á námskeiðum á ensku á meginlandi Evrópu, einkum í framhaldsnámi, hafa t.a.m. Þjóðverjar sums staðar verið að prófa sig áfram með að hafa hluta kennslunnar einnig á þjóðtungunni (House 2008:78).

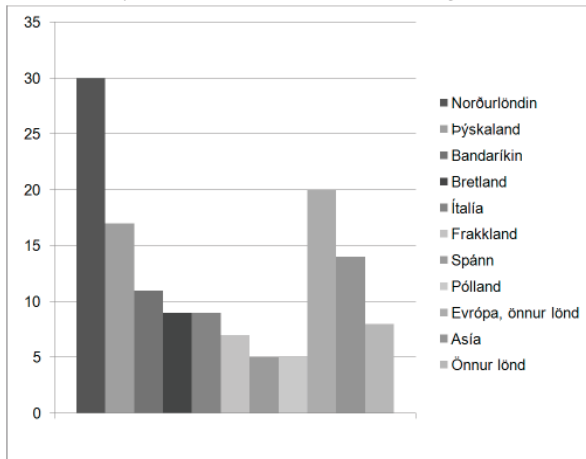
Vert er að ítreka að rannsóknir í þessu efni eru ekki komnar mjög

langt á veg. Einkum er þörf á frekari kennslufræðilegum rannsóknum til að unnt sé að meta tap eða ávinning af ensku sem kennslumáli í stað íslensku í stúdentahópi sem samanstendur af íslenskum og erlendum stúdentum — og veita kennurum íslenskra háskóla leiðbeiningar um hvort og hvernig þeir þurfi að breyta kennsluáðferðum sínum.

5 Háskólastarfsemi þvert á landamæri og þjóðerni

Allir íslensku háskólarnir sækja fram á alþjóðavettvangi. Þeim er í mun að rækja vel það hlutverk sitt að vera hluti alþjóðlegs mennta- og vísindasamfélags. Þeir leggja aukna áherslu á rannsóknir og rannsóknarnám og að laða til sín erlenda stúdentu og fræðimenn.

Aðsókn erlendra stúdentu að háskólanámi á Íslandi hefur vaxið mjög síðan um aldamótin 2000, kennaraskipti við erlenda háskóla eru orðin mun tíðari en áður og sífellt fleiri erlendir fræðimenn sækjast eftir störfum í íslenskum háskólum. Stærsti háskóli landsins, Háskóli Íslands, hefur sett sér sem markmið að verða einn af 100 „bestu háskólum heims“ á skrá *Times Higher Education Supplement* og hefur þegar náð þeim árangri að vera meðal 300 bestu háskólanna á þeim lista. Íslensku háskólarnir leggja stóru áherslu á framhaldsnám, kennara- og stúdentaskipti, alþjóðleg rannsóknarnet og birtingu rannsóknarniðurstaðna hjá virtum erlendum forlögum.



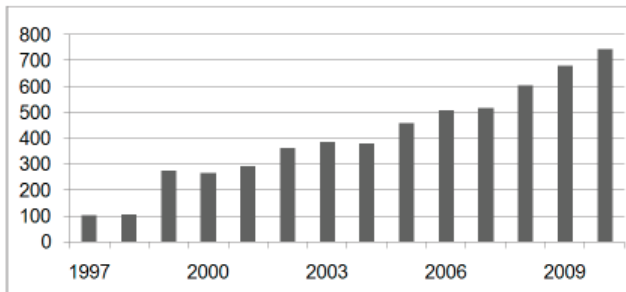
Mynd 14. Erlendir akademískir starfsmenn íslenskra háskóla (135 manns) í nóvember 2010, eftir ríkisfangi. Heimild: Deildarstjóri mennta- og menningarmáladeildar Hagstofu Íslands, janúar 2013.

Samkvæmt tölum frá Hagstofu Íslands⁸ voru akademískir starfsmenn í íslensku háskólunum 2.216 talsins (í 1.455 stöðugildum) í nóvember 2010. Þar af voru starfsmenn með erlent ríkisfang 135 manns (í 90 stöðugildum) eða sem svarar 6,1% (6,2% stöðugilda). Þeir komu víða að eins og mynd 14 sýnir. Því fer fjarri að akademísku starfsmennirnir komi að jafnaði frá hinum hefðbundnu enskumælandi málsvæðum heimsins. Þó má gera ráð fyrir að flestir þeirra búi yfir haldgóðri enskukunnáttu.

Ísland var meðal þeirra ríkja sem stóðu að svonefndri Bologna-yfirlýsingu árið 1999 en markmið hennar er að Evrópa sé eitt samfellt menntasvæði á háskólastigi. Með því að samræma námslýsingar og námskröfur er stúdentum gert kleift að fá nám í ólíkum löndum metið á sambærilegan hátt til háskólagráðu. Auk skiptistúdenta, sem sækja námskeið í íslenskum háskólum sem hluta af námi sínu, hefur þeim útlendingum fjölgað sem koma til lengri dvalar og ljúka námi að fullu á Íslandi.

Hvar sem borið er niður sést að erlendum stúdentum hefur fjölgað til muna í íslenskum háskólum undanfarinn hálfan annan áratug.

Nýnemar með erlent ríkisfang, sem hófu nám íslenskum háskólum haustið 2010, voru 745 talsins en haustið 1997 voru erlendir stúdentar 101. Mynd 15 er byggð á tölum Hagstofu Íslands (*Talnaefni*) sem ná fram til ársins 2010.



Mynd 15. Nýnemar í íslenskum háskólum með erlent ríkisfang 1997–2010. Byggt er á tölum Hagstofu Íslands (*Talnaefni*).

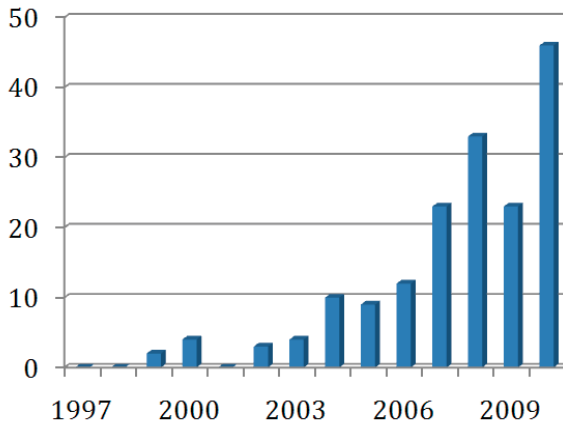
Ef einungis eru taldir þeir útlendingar sem hefja doktorsnám í íslenskum skólum á hausti hverju kemur í ljós að þar hefur aðsóknin vaxið hlutfallslega meira en í öðru námi.

Enginn nýnemi með erlent ríkisfang hóf doktorsnám á Íslandi 1997 og 1998 og árin 1999–2003 aðeins örfáir. Þeir voru hins vegar 10 talsins

⁸ Þakkir til Ástu Urbancic, deildarstjóra mennta- og menningarmáladeildar Hagstofunnar, fyrir að útvega gögnin.

2004 og hefur síðan fjölgað verulega og voru 46 talsins haustið 2010. Þeir leggja einkum stund á náttúruvísindi.

Erlendum skiptistúdentum í íslensku háskólunum fjölgaði um ríflega 500% á tímabilinu 1996–2010, þ.e. úr 126 á ári í 656 á ári, nokkuð jafnt og þétt allan tímann.⁹ Um tveir þriðju hlutar skiptistúdenta á Íslandi koma á vegum Erasmus-áætlunarinnar. Framan af tímabilinu 1996–2010 stunduðu erlendir skiptistúdentar oftast nám í Háskóla Íslands og Kennaraháskóla Íslands (samtals um 83% hópsins) en á síðustu árum hefur hlutur sumra yngri háskólanna vaxið hlutfallslega mjög mikið. Veturinn 2010–2011 sóttu um 16% skiptistúdentanna nám í Háskólanum í Reykjavík, um 8% í Háskólanum á Akureyri og um 7% í Háskólanum á Bifröst. Um 62% skiptistúdentanna voru þá í (sameinuðum) Háskóla Íslands.



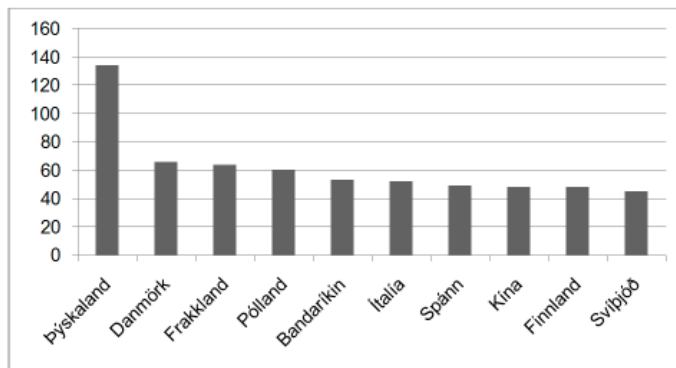
Mynd 16. Nýnemar á doktorsstigi með erlent ríkisfang 1997–2010. Byggt er á tölum Hagstofu Íslands (*Talnaefni*).

Tölur Háskóla Íslands sýna að erlendir stúdentar koma frá sífellt fleiri löndum. Árið 2001 voru upprunalöndin 53 en áratug síðar, 2011, voru þau 86 talsins.

Staða ensku sem háskólakennslu- og vísindatungumáls á alþjóðavísu kemur m.a. fram í því að hún verður fyrirferðarmeiri í kennslu, ritgerðasmíð, rannsóknum og samskiptum í íslenskum og öðrum norrænum háskólum enda þótt erlendir stúdentar, sem þangað leita, komi alls ekki endilega frá „hefðbundnum enskumælandi svæðum“. *Mynd 17*, sem sýnir upprunalönd erlendra stúdenta á Íslandi, sýnir að ýmis lönd þar sem stór hluti landsmanna hefur ensku sem móðurmál

⁹ Þakkir til Óskars Eggerts Óskarssonar fyrir að útvega gögnin.

eða fyrsta mál (s.s. Bretland, Írland, Kanada, Ástralía, Nýja-Sjáland, Suður-Afríka) eru allsekki meðal 10 helstu upprunalanda stúdenta sem leita til Íslands. Bandaríkjamönnum hefur raunar fjölgað í íslenskum háskólum frá 2001 en hlutfallslega þó mun minna en stúdentum frá Þýskalandi, Frakklandi, Póllandi, Kína og fleiri löndum þar sem enska er ekki algengasta fyrsta mál. Hvað varðar háskólakennarana er svipaða sögu að segja. Af erlendum akademískum starfsmönnum (135 manns) við íslensku háskólana haustið 2010 voru t.a.m. aðeins 17% (23 manns) samanlagt frá Bretlandi, Bandaríkjunum og Kanada. Þetta sýnir að aukin enskunotkun í háskólastarfi, m.a. á Íslandi, stafar fyrst og fremst af því að enska er alþjóðlegt samskiptamál og vísindamál sem þjóðir heims nota til viðbótar við eigin tungumál. Það er m.ö.o. daglegt brauð í háskólum eins og í öðrum samfélagsgeirum að enska, töluð og rituð, sé samskiptamál tveggja eða fleiri einstaklinga sem tala hana ekki sem móðurmál eða fyrsta mál.



Mynd 17. Tíu alengustu upprunalönd erlendra stúdenta í Háskóla Íslands árið 2011. Byggt er á tölum á vef Háskóla Íslands (*Erlendir nemendur í Háskóla Íslands*).

6 Niðurlag

Enska hefur orðið sífellt fyrirferðarmeiri í íslenskri háskólakennslu og rannsóknarstarfi undanfarna áratugi og einkum eftir aldamót. Erlendum nemendum og starfsmönnum hefur jafnframt fjölgað stórlega, sbr. yfirlitið í 5. kafla hér á undan. Kennsla fer að nokkru leyti fram á ensku, einkum í framhaldsnámi. Í 4. kafla hér á undan kom t.a.m. fram að skólaárið 2010–2011 voru 11,7% allra námskeiða í Háskóla Íslands og 23,4% allra námskeiða í Háskólanum í Reykjavík kennd á ensku. Lokaritgerðum á ensku fjölgar. Úttekt í 3. kafla sýnir

t.d. að allar doktorsritgerðir á sviðum heilbrigðis- og verkfræði og náttúruvísinda við Háskóla Íslands á árunum 2007–2011 voru á ensku. Meginhluti námsefnis í íslenskum háskólum er á ensku, um 90% samkvæmt rannsóknum Hafðisar Ingvarsdóttur og Birnu Arnbjörnsdóttur (2010). Þá birta íslenskir fræðimenn niðurstöður rannsókna sinna í mjög miklum mæli á ensku. Í 3. kafla kom m.a. fram að árið 2011 voru um 78% allra tímaritsgreina, sem akademískir starfsmenn Háskóla Íslands birtu, á ensku.

Rauði þráðurinn í málstefnuskjölum íslensku háskólanna, sem könnuð voru í 2. kafla, er að íslenska skuli vera aðal tungumál þessara skóla.¹⁰ Þar er jafnframt lögð rík áhersla á íslenskan íðorðaförða og vandaða íslenska málnotkun.

Þegar hugað er að skjalfestri málstefnu háskólanna má hins vegar ekki gleyma því að orð á blaði eru annað en veruleikinn. Þeim verður að fylgja eftir og það kann að hafa í för með sér togstreitu í framkvæmd þegar ólíkum markmiðum er raðað í forgangsröð eftir mikilvægi. Tími háskólamanna er takmarkaður og fjárráð háskólanna sömuleiðis.

Enda þótt frekari rannsóknir þurfi að fara fram á þeim áhrifum sem val á tungumáli hefur á háskólastarfsemina er eigi að síður ljóst og óhætt að fullyrða að það val hefur áhrif á nám og kennslu og á þann vettvang sem valinn er fyrir birtingu rannsóknarniðurstaðna og jafnframt á það hvernig til tekst við miðlun þekkingar til nærsamfélagsins. Jafnframt er augljóst af hinum öru breytingum á undanförunum fáeinum áratugum, einkum eftir árþúsundamótin, að ekki dugir að bíða átekta; háskólar og vísindasamfélagið í heild þarf að bregðast við breyttum aðstæðum með markvissum hætti í nánustu framtíð eigi að lánast að stýra aðstæðum og þróun í sem heppilegastan farveg.

Hér er mikið í húfi: innviðir nútímalegs samfélags á Íslandi, þjóðtunga Íslendinga og notkunarsvið hennar, menntunar- og tækniþróun á Íslandi og samkeppnisskilyrði íslensku háskólanna í alþjóðlegum samanburði.

Heimildir

Ammon, Ulrich. 2010. World Languages: Trends and Futures. Í: N. Coupland (ritstj.) *The Handbook of Language and Globalization*. Bls. 101–122. Malden/Oxford: Blackwell.

¹⁰ Um er að ræða skjöl frá sex háskólum af sjö, þ.e. öllum nema Háskólanum í Reykjavík sem hefur ekki sent frá sér skjalfesta málstefnu.

- Ársskýrsla 2010. Háskólinn á Akureyri. <http://www.unak.is/static/files/Arsskyrsla%20HA%202010-vefutg%203.pdf> (21. febrúar 2014).
- Ástráður Eysteinnsson. 2012. Íslensk sprogpolitik, kultur og akademisk praksis. Domæner, store og små. *Nordand – nordisk tidsskrift for andre-språksforskning* 2012, 7,(2):143–159.
- Birna Arnbjörnsdóttir og Hafdís Ingvarsdóttir. 2010. Coping with English at University. Students' Beliefs. *Ráðstefnurit Netlu – Menntakvika 2010*. Reykjavík: Menntavísindasvið Háskóla Íslands. <http://netla.hi.is/menntakvika2010/008.pdf> (21. febrúar 2014).
- Doktorsritgerðaskrá Landsbókasafns Íslands – Háskólabókasafns*. <http://doktor.landsbokasafn.is/> (21. febrúar 2014).
- Erlendir nemendur í Háskóla Íslands*. http://www.hi.is/adalvefur/erlendir_nemendur_i_haskola_islands (21. febrúar 2014).
- Fréttasafn* á vef Háskólans í Reykjavík. <http://www.ru.is/haskolinn/frettir/nr/24310> (21. febrúar 2014).
- Guðni Jónsson. 1961. *Saga Háskóla Íslands. Yfirlit um hálfra aldar starf*. Reykjavík: Háskóli Íslands.
- Gunnar Karlsson (ritstjóri), Guðmundur Hálfdanarson, Sigríður Matthíasdóttir og Magnús Guðmundsson. 2011. *Aldarsaga Háskóla Íslands 1911–2011*. Reykjavík: Háskólaútgáfan.
- Hafdís Ingvarsdóttir og Birna Arnbjörnsdóttir. 2010. Coping with English at Tertiary Level: Instructors' Views. *Ráðstefnurit Netlu – Menntakvika 2010*. Reykjavík: Menntavísindasvið Háskóla Íslands. <http://netla.hi.is/menntakvika2010/010.pdf> (21. febrúar 2014).
- Háskóli Íslands – Doktorsvarnir*. <http://www.hi.is/adalvefur/doktorsvarnir> (21. febrúar 2014).
- Háskóli Íslands – Yfirlit um reglur HÍ*. http://www.hi.is/is/skolinn/yfirlit_um_reglur_hi (21. febrúar 2014).
- Háskólinn í Reykjavík – Ársskýrsla 2010*. http://www.ru.is/media/baeklingar/HR_Arsskyrsla-2010.pdf (21. febrúar 2014).
- House, Juliane. 2008. English as lingua franca in Europe today. Í: G. Extra og D. Gorter (ritstj.) *Multilingual Europe: Facts and Policies*. Bls. 63–86. Berlín / New York: Mouton de Gruyter.
- Íslenska til alls. Tillögur Íslenskrar málnefndar að íslenskri málstefnu samþykktar á Alþingi 12. mars 2009*. 2009. Reykjavík: Mennta- og menningarmálaráðuneytið. http://www.islenskan.is/islenska_til_all_2009-11-06.pdf (21. febrúar 2014).
- Kennsluskrá Háskólans á Bifröst skólaárið 2010–2011. Grunn nám í öllum deildum*. http://bifrost.is/Files/Skra_0051427.pdf (21. febrúar 2014).
- Kennsluskrá Háskólans á Bifröst skólaárið 2010–2011. Meistaránám í öllum deildum*. http://bifrost.is/Files/Skra_0051428.pdf (21. febrúar 2014).
- Lög um háskóla*, nr. 63 13. júní 2006. <http://www.althingi.is/lagas/140b/2006063.html> (21. febrúar 2014).

- Lög um opinbera háskóla*, nr. 85 12. júní 2008. <http://www.althingi.is/lagas/nuna/2008085.html> (21. febrúar 2014).
- Lög um stöðu íslenskrar tungu og íslensks táknmáls*, nr. 61 7. júní 2011. <http://www.althingi.is/lagas/140a/2011061.html> (21. febrúar 2014).
- Matskerfi opinberra háskóla*. http://www.hi.is/sites/default/files/matskerfi_opinberra_haskola_des_2013.pdf (21. febrúar 2014).
- Málstefna* [Listaháskóla Íslands]. http://lhi.is/media/filer_private/2012/08/14/malstefna_listahaskola_islands_1.pdf (21. febrúar 2014).
- Málstefna Háskóla Íslands*. Samþykkt á 13. háskólafundi 21. maí 2004. http://www.hi.is/adalvefur/malstefna_haskola_islands (21. febrúar 2014).
- Málstefna Háskólans á Akureyri*. Samþykkt í háskólaráði 22. febrúar 2008. http://www.unak.is/static/files/Stjornsysla_stefnumal/Malstefna%20HA_%20samthykt%20220208.pdf (21. febrúar 2014).
- Málstefna Háskólans á Bifröst*. Samþykkt í háskólaráði 29. ágúst 2012. <http://www.bifrost.is/islenska/um-haskolann/stefna-og-hlutverk/malstefna-haskolans-a-bifrost/> (21. febrúar 2014).
- Málstefna Hólaskóla – Háskólans á Hólum*. http://holar.is/almennt_efni/malstefna (21. febrúar 2014).
- Málstefna Landbúnaðarháskóla Íslands (LbhÍ)*. <http://vefur.lbhi.is/forsida/Haskolinn/Malstefna> (21. febrúar 2014).
- Ný sýn. Breytingar á vísinda- og nýsköpunarkerfinu*. 2012. Reykjavík: Forsætisráðuneytið. <http://www.forsaetisraduneyti.is/media/Skyrslur/ny-syndes-2012.pdf> (21. febrúar 2014).
- Salö, Linus. 2010. *Engelska eller svenska? En kartläggning av språksituationen inom högre utbildning och forskning*. Stockholm: Språkrådet, Institutet för språk och folkminnen.
- Stefna Háskólans í Reykjavík*. <http://www.ru.is/haskolinn/stefna/> (21. febrúar 2014).
- Styrkur Háskólans í Reykjavík í rannsóknnum*. 2012. Reykjavík: Rannsóknarþjónusta HR. <http://www.ru.is/media/almennt/Styrkur-Haskolans-i-Reykjavik-i-rannsoknum-2007-2012.pdf> (21. febrúar 2014).
- Talnaefni*. <http://www.hagstofa.is> (21. febrúar 2014).
- Yfirlýsing um málstefnu Norðurlanda*. 2006. 2007. ANP 2007:746. Kaupmannahöfn: Norræna ráðherranefndin. <http://www.norden.org/da/publikationer/publikationer/2007-746> (21. febrúar 2014).
- Dingsályktun um íslenska málstefnu*. Samþykkt á Alþingi 12. mars 2009. <http://www.althingi.is/altext/136/s/0699.html> (21. febrúar 2014).

Lykilorð

háskólar, enska, íslenska, kennslutungumál, tungumál fræðilegra skrifa

Keywords

universities and colleges, English, Icelandic, language of instruction, language of academic publications

Summary

The article describes the relationship between the use of English and Icelandic in Icelandic universities and academia.

While Iceland's seven universities and colleges specialize in different fields of science and technology, they all have to meet both domestic demands for research, teaching, various service and dissemination of knowledge in Iceland, and also, at the same time, to recruit international faculty, participate in international academic cooperation, and provide study offers for non-Icelandic students. According to Icelandic law, Icelandic is the "language of schools at all levels", which is a provision that holds for universities and colleges. Presently, six out of seven Icelandic universities and colleges have issued their language policies. All these policies provide that Icelandic is the main language of the institutions, and the policies all emphasize Icelandic terminologies. Obviously, however, English must play a key role along with Icelandic, due to a number of external factors.

At present, approximately 6% of students and faculty at Icelandic higher education institutions are non-Icelandic citizens, and English is partly the language of instruction. The proportion of individual classes taught in English ranges from 0% to 23%, depending on the nature of the institutions and the studies that they offer. It also depends on levels of study (English is less frequent as language of instruction in undergraduate studies, and more frequent in graduate studies, on the whole). Academic staff are strongly encouraged to publish their research results in international fora. About 80% of academic publications of faculty members at the two largest universities in Iceland (i.e., at the University of Iceland, and Reykjavik University) are in English. However, the language of writing differs considerably from one field to another; about half of academic publications are in English in the social sciences, education sciences, and humaniora, while the figure is about 97% in engineering, and natural sciences. Presently, doctoral dissertations in the medical sciences, engineering, and natural sciences are exclusively written in English.

Ari Páll Kristinsson

Stofnun Árna Magnússonar í íslenskum fræðum, Málræktarsviði / Háskóla Íslands

Neshaga 16, 107 Reykjavík

aripk@hi.is

Haraldur Bernharðsson

Íslensku- og menningardeild

Háskóla Íslands

Árnagarði v/Suðurgötu, 101 Reykjavík

haraldr@hi.is

Kristín Bjarnadóttir

Beygingarlýsing íslensks nútímamáls

Regluverk eða beygingardæmi

1 Inngangur

Beygingarlýsing íslensks nútímamáls (BÍN) er gagnasafn um beygingarkerfið í íslensku nútímamáli. Gagnasafnið hefur nú að geyma ríflega 270 þúsund beygingardæmi með u.þ.b. 5,8 milljón beygingarmyndum, þ.e. orðmyndum með greiningarstrengjum eða mörkum.¹ Beygingarlýsingunni var upprunalega ætlað að þjóna tvíþættum tilgangi, að vera gagnasafn um beygingarmyndir til nota í máltækni-verkefnum og til þess að veita almenningi aðgang að beygingardæmunum.

Vinna við BÍN hófst árið 2002 og var hluti af tungutækniátaki menntamálaráðuneytisins (Eiríkur Rögnvaldsson o.fl. 2009). Fyrsta útgáfa BÍN til máltækninota var gefin út á geisladiski í mars 2004 og vefsetrið var opnað síðar sama ár. Gögn úr BÍN hafa verið notuð í ýmis máltækni-verkefni, t.d. leitarvélar, mörkun og leiðréttingarhugbúnað, og í kennsluefni og orðabókum. Beygingarlýsingin og *Mörkuð íslensk málheild* (MÍM) (Sigrún Helgadóttir o.fl. 2012) eru systurverkefni sem unnin hafa verið samtímis á Stofnun Árna Magnússonar í íslenskum

¹ Mark er strengur af skammstöfunum fyrir málfræðilegar formdeildir, þ.e. orðflokks- og beygingargreiningu.

fræðum. Upprunaleg gögn sem notuð voru í BÍN voru orðabókar-gögn, aðallega úr rafrænni útgáfu af *Íslenskri orðabók* (Mörður Árnason 2000), en þar voru u.þ.b. 135 þúsund flettiorð, og úr gagnasöfnum Orðabókar Háskólans. Leitað hefur verið í ýmsar aðrar heimildir um orðaforða en næsta stig er að bæta orðaforðanum úr MÍM við.

Frá upphafi var ætlunin að sýna íslenska beygingarkerfið eins og það er í raun, með eins góðri lýsingu á beygingarafbrigðum og nokkur kostur er. Vegna þessa var strax ákveðið að búa til safn beygingardæma þar sem hver beygingarmynd er geymd og sýnd á sínum stað í beygingardæminu, í stað þess að setja saman virkt kerfi beygingarreglna þar sem orð eru merkt eftir því hvaða regla á við í hverri formdeild. Meginástæðan fyrir þessu var gagnaskortur en í ljós kom að nauðsynleg gögn til að setja fram slíkt reglukerfi voru einfaldlega ekki til þrátt fyrir langa sögu rannsókna á íslensku máli.

Vandamálið er að reglukerfi sem sett er fram á grunni ónógra gagna um beyginguna er bæði of- og vanvirkt, þ.e. það framleiðir mun fleiri orðmyndir en æskilegt er (t.d. fleirtölumyndina *sykrar* af karlkynsorðinu *sykur*) og sleppir öðrum (t.d. þágufallsmyndinni *hönd* í merkingunni *rithönd*). Til gagnagreiningar og til nota í leitarvélum kemur ofvirkni ekki verulega að sök þar sem tilbúnar orðmyndir koma einfaldlega aldrei fram en vanvirkni verður til þess að tækar orðmyndir verða útundan, t.d. í leit og greiningu, þ.e. þær verða að óþekktum orðum sem eru eitt af helstu vandamálum í textagreiningu. Ofvirkt reglukerfi er líka ótækt til framleiðslu á beygingarmyndum, a.m.k. ef útkoman á að vera rétt beygingardæmi. Þegar slík gögn eru notuð í máltækniþúnaði sem skilar frá sér texta verður niðurstaðan ekki góð, ef tilbúnar orðmyndir koma þar í stað þess sem rétt er. Þetta atriði er mjög mikilvægt, t.d. í ljósi þess að gögn úr BÍN eru notuð til samhengisháðrar leiðréttingar. Þá gera notendur BÍN á vefnum kröfu um að beygingardæmin sem birtast á síðunni séu rétt og skili sér ekki sem afrakstur af ofvirkum reglum.²

2 Flækjustigið í íslenska beygingarkerfinu

Hlutfallið á milli fjölda beygingardæma og beygingarmynda í BÍN

² Notendur BÍN á vefnum eru mjög duglegir við að senda fyrirspurnir um vafaatriði í beygingardæmunum og leiðréttingar, þar sem þess er þörf. Í maí 2013 voru heimsóknir á síðuna 53.907 og flettingar 211.902. Notendur voru 20.301, þar af 79% innanlands.

sýnir hve margar beygingarmyndirnar eru. Beygingarmyndir nafnorðs eru allt að 16, án afbrigða, beygingarmyndir lýsingarorðs eru allt að 120, og sagnorðs allt að 107, ef hvergi eru göt í beygingardæmunum.³ Hverri einstakri beygingarmynd í BÍN fylgir greiningarstrengur eða mark, t.d. NF-ET fyrir nefnifall eintölu af nafnorði og GM-NH fyrir nafnhátt sagnar í germynd. Markamengið fyrir íslensku er stórt og í því eru yfir 700 mörk (Sigrún Helgadóttir 2012), enda eru beygingarmyndirnar margar.

Beygingarkerfið er líka talsvert flókið, þar sem beygingarendingar eru ekki einræðar. Sama beygingarendingin getur verið birtingarvísir fyrir fleiri en eina formdeild og í fleiri en einum orðflokki, og sama formdeildin getur birst sem meira en ein beygingarending. Beygingarendingin *-a* er t.d. birtingarvísir margra beygingarformdeilda í öllum orðflokkum (Eiríkur Rögnvaldsson 1990:23) og endingin í eignarfalli eintölu af sterkum karlkynsorðum með nefnifallsendinguna *-ur* getur verið *-s* eða *-ar*, t.d. *hundur*, *hunds*; *köttur*, *kattar*.

Í umfjöllun um íslenska beygingarkerfið er almennt litið svo á að orð flakki á milli beygingarflokka þegar tilbrigði eru í beygingarmyndum. Sögnin *fela* er þannig ýmist veik (þt. *faldi*) eða sterk (þt. *fól*) og nafnorðið *grein* er ýmist *greinar* eða *greinir* (1. og 2. flokkur kvenkynsnafnorða, skv. Valtý Guðmundssyni 1922:62). Við gerð BÍN var sú leið farin að setja upp beygingardæmi fyrir hvert orð fyrir sig í einu lagi, þannig að í hverju beygingardæmi birtust öll afbrigði í einu sem hluti af heildarmynstri fyrir orðið. Hvert beygingardæmi er því ein heild en ekki tilvísun í margra beygingarflokka eftir því sem afbrigði gefa tilefni til. Beygingardæmin í BÍN byggjast samt á nefnimyndum (eða uppflertimyndum) þannig að afbrigði í nefnimynd skilja á milli beygingardæma. Orðin *sannleikur* og *sannleiki* eru t.d. tvö beygingardæmi, þó að málnotendur blandi beygingunum saman á mjög áhugaverðan hátt.⁴ Í hverjum beygingarflokki í BÍN eru öll afbrigði sem birtast í orði tekin með og því er beygingarflokkur einungis mengið af þeim beygingarreglum sem þarf til þess að búa til beygingardæmi fyrir orð, eitt eða mörg, eftir því sem verkast vill. Þessi reglu knippi sem standa að baki beygingarflokkunum hafa í sjálfu sér ekki fræðilegt gildi (sjá Ásta Svavarsdóttir 1993) en þau eru tól til að halda utan um flókin gögn.

³ Þarna vantar samt spurnarmyndir sagna, t.d. *ferðu*, *fórstu* o.s.frv.

⁴ Tíðnitölur benda til þess að sterka beygingin sé notuð í nefnifalli en veika beygingin í aukaföllunum: *sannleikurinn er sagna bestur*; í *sannleika sagt*.

3 Keyrslan á beygingardæmunum

Upprunalega voru beygingardæmin í BÍN búin til með því að nota tiltölulega einfaldar skriftur þar sem númeraðir orðstofnar í færslum fyrir einstök orð voru keyrðir inn í ramma með beygingarendingum, þ.e. beygingarflokk. Niðurstaðan var safn af XML-skrám sem síðan voru notaðar til að búa til það úttak sem þörf var fyrir, t.d. til birtingar á vefnum eða til nota í máltækniverkefni. Gagnagrunnurinn sem nú hýsir BÍN virkar á svipaðan máta. Færslan um hvert orð geymir öll stofnbrigði og upplýsingar um það hvar í beygingardæminu hvert stofnbrigði á sér stað. Nafnorð eru merkt eftir því hvort þau koma fyrir í eintölu, fleirtölu eða hvort tveggja, sagnir eru merktar eftir því hvort þær koma fram í germynd, miðmynd eða hvort tveggja o.s.frv.⁵ Hver beygingarmynd er síðan geymd í gagnagrunninum með sínu marki, t.d. *akur*<þgf.et.>, sjá Töflu 1.

Án greinis (+)				
	Eintala (+)			Fleirtala (+)
Nf.	1+0	<i>akur</i>	3+ar	<i>akrar</i>
Þf.	1+0	<i>akur</i>	3+a	<i>akra</i>
Þgf.	3+i	<i>akri</i>	2+um	<i>ökrum</i>
Ef.	1+s	<i>akurs</i>	4+a	<i>akra</i>
Með greini (+)				
Nf.	1+inn	<i>akurinn</i>	3+arnir	<i>akrarnir</i>
Þf.	1+inn	<i>akurinn</i>	3+ana	<i>akrana</i>
Þgf.	3+inum	<i>akrinum</i>	2+unum	<i>ökrunum</i>
Ef.	1+sins	<i>akursins</i>	4+anna	<i>akranna</i>

Tafla 1. Ramminn fyrir nafnorðið *akur*.

Í Töflu 1 er ramminn fyrir nafnorðið *akur*, en þar eru engin göt í beygingardæminu og engin afbrigði. Gildin fyrir beygingarformdeildirnar eru því alltaf '+', þ.e. eintala = +et., fleirtala = +ft., og ákveðni = +ángr., +mgr. (Í fleirtöluorði væri gildið fyrir eintöluna -et. og þá verður sá hluti beygingardæmisins tómur.) Stofnbrigðin *akur* (1), *akr* (3 og 4),⁶ *ökr* (2) eru tölusett og eru hluti af inntaki fyrir orðið. Á vefsíðunni birtist beygingardæmið á mjög svipaðan hátt og í Töflu 1 en 2. og 4. dálki er sleppt.

Alengasta úttak á gögnum úr BÍN til máltækninota er einfaldur

⁵ Nákvæmar upplýsingar um rofana sem ráða birtingu beygingarmynda er að finna á vefsíðu BÍN: <http://bin.amastofnun.is>.

⁶ Sama stofnbrigði fær af tæknilegum ástæðum stundum fleiri en eitt númer.

listi með sex sviðum fyrir hverja beygingarmynd, eins og í úttakinu fyrir nafnorðið *akur* í Töflu 2. Sviðin eru uppflöttiorð, auðkennisnúmer, orðflokkur eða kyn nafnorðs, hluti orðaforðans í BÍN⁷ (hér 'alm', fyrir 'almennt mál'), beygingarmynd og málfræðilegt mark eða greiningarstrengur.

akur;472164;kk;alm;akur;NF.ET.
 akur;472164;kk;alm;akurinn;NF.ET.MGR.
 akur;472164;kk;alm;akur;PF.ET.
 akur;472164;kk;alm;akurinn;PF.ET.MGR.
 akur;472164;kk;alm;akri;PGF.ET.
 akur;472164;kk;alm;akrinum;PGF.ET.MGR.
 akur;472164;kk;alm;akurs;EF.ET.
 akur;472164;kk;alm;akursins;EF.ET.MGR.
 akur;472164;kk;alm;akrar;NF.FT.
 akur;472164;kk;alm;akrarnir;NF.FT.MGR.
 akur;472164;kk;alm;akra;PF.FT.
 akur;472164;kk;alm;akrana;PF.FT.MGR.
 akur;472164;kk;alm;ökrum;PGF.FT.
 akur;472164;kk;alm;ökrunum;PGF.FT.MGR.
 akur;472164;kk;alm;akra;EF.FT.
 akur;472164;kk;alm;akranna;EF.FT.MGR.

Tafla 2. Úttak úr BÍN til máltækninota.

Rammarnir (sbr. Töflu 1) eru búnir til eftir því sem þörf krefur, í hvert skipti sem nýfundid afbrigði passar ekki í þá ramma sem þegar eru til, þ.e. þegar knippi af beygingarreglum dugar ekki fyrir öll afbrigði einhvers orðs. Vegna þessa er BÍN byggð á beygingu einstakra orða, ef í það fer, þ.e. kerfið er búið til neðan frá ('bottom-up'). Dæmi um þetta eru samsett orð sem hafa hvorugkynsorðið *vegur* sem síðari hluta. Ósamsetta orðið *vegur* er oftast *vegar* í eignarfalli eintölu, þó að beygingarmyndinni *vegs* bregði einnig fyrir, t.d. í orðasambandinu *komast til vegs* og *virðingar*. Í sumum samsetningum af orðinu er eignarfallsendingin *-s* hins *vegar* ráðandi, t.d. í orðunum *farvegur*, *jarðvegur* og *útvegur*. Orðin *vegur* og *akvegur* annars *vegar* og *farvegur*, *jarðvegur* og *útvegur* skiptast því í tvo beygingarflokka, ef. et: *-ar/-s* og *-sl/-ar*. Þarna er ráðandi ending á undan en röðin á beygingarafbrigðum í BÍN ákvarðast yfirleitt af tíðni þó að þar geti margir aðrir þættir komið við sögu.

Beygingarflokkar í BÍN eru nú yfir 630 og í sumum þeirra eru tugþúsundir orða en í öðrum eru stök orð. Þetta stafar stundum af

⁷ Helstu „hlutar BÍN“ eru almennt mál, mannanöfn og örnefni og sérorðaforði ýmiss konar.

Því að gamlar beygingarmyndir koma enn fyrir í nútímamáli, þó svo að beygingarreglurnar sem að baki liggja séu ekki lengur virkur hluti af beygingarkerfinu. Úreltu beygingarmyndirnar eru þá gjarnan samhengisháðar í nútímamáli. Sem dæmi má nefna beygingarmyndina *refarnir* í orðasambandinu *til þess eru refarnir skornir* í stað *refirnir* sem er nefnifall fleirtölu í nútímamáli og beygingarmyndina *skjöldu* í sambandinu *ganga fram fyrir skjöldu*, í stað *skildi* sem er nútímamálsmynd þolfalls fleirtölu. Orðmyndirnar *refar* og *skjöldu* eru strangt til tekið ekki hluti af beygingarkerfi nútímamáls en þær koma vissulega fyrir í textum og eru því bráðnauðsynlegar í BÍN, bæði til máltækni nota og á vefnum. Afbrigði af þessu tagi fjölga beygingarflokkum í BÍN þar sem hvert einasta afbrigði á sér stað í beygingarflokknum og breytir skilgreiningunni á honum.

Uppflettiorð á bilinu	Beygingarflokkar	Uppflettiorð fleiri en	Beygingarflokkar
<10.000	5		
5.000-9.999	7		
1.000-4.999	29	<1000	41
500-999	21	<500	62
100-499	72	<100	134
50-99	47	<50	181
10-49	126	<10	307
2-9	146	<2	453
1	156		Alls: 609

Tafla 3. Stærð og fjöldi beygingarflokka nafnorða, lýsingarorða og sagna.

Beygingarflokkar í BÍN er misjafnlega stórir, sjá Töflu 3, en athyglisvert er að í u.þ.b. fjórðungi beygingarflokka meginorðflokka, nafnorða, lýsingarorða og sagna, er aðeins eitt orð. Í u.þ.b. 42% þessara beygingarflokka er beygingin alveg „regluleg“, þ.e. þar er aðeins ein regla um hverja formdeild. Þar eru engin afbrigði af beygingarmyndum, engin innri beyging (þ.e. beyging fyrri hluta, t.d. í *litliputti*, þf. *litlaputta*) eða afbrigðileg göt í beygingardæmunum, eins og t.d. í kvenkynsorðinu *kröggur* en þar hafa engin dæmi fundist um eignarfall fleirtölu.⁸ Í BÍN eru 49.296 afbrigðapör (t.d. eignarfallið *þröskuld-ar/þröskulds*) og 262 afbrigðaprenndir (t.d. þolfallið *Berglind/Berglindi/Berglindu*), þ.e. afbrigði með sama greiningarstreng innan beygingardæmis (feb. 2012).

⁸ Þeir sem spurðir eru um eignarfallsmyndina neita yfirleitt að svara, jafnt málfræðingar sem aðrir.

4 Gagnaskortur

Upprunalegu gögnin í BÍN voru orðabókarefni, úr rafrænni útgáfu *Íslenskrar orðabókar* (2000), en hún kom fyrst út 1963, og úr Ritmálsafni Orðabókar Háskólans, með milliskrefi í *Norræna verkefnum* svokallaða (Kristín Bjarnadóttir 1998). Önnur meginheimild í upphafi var bókin *Nöfn Íslendinga* (Guðrún Kvaran og Sigurður Jónsson frá Arnarvatni, 1991), en í henni eru u.þ.b. 4.800 mannanöfn. Í fyrstu útgáfu BÍN (2004) voru u.þ.b. 176 þúsund beygingardæmi og orðaforðinn var aðallega úr þessum þremur heimildum. Viðbótarorðaforði í síðari útgáfum BÍN er úr ýmsum áttum og hann er m.a. afrakstur samvinnu þar sem gögn úr BÍN hafa verið notuð í leitarvélar, t.d. á vef Hins íslenska biblíufélags⁹ og í símaskránni¹⁰, sem er góð uppspretta nafna, örnefna og fyrirtækjaheita.

Upplýsingar um beygingu orða er að finna í *Íslenskri orðabók* og í *Nöfnum Íslendinga* en sú lýsing er ekki alltaf nægileg til þess að setja fram heil beygingardæmi. Í öðrum heimildum eru yfirleitt engar upplýsingar um beygingu. Það var því ljóst frá upphafi að leita þyrfti í málfræðirit til að finna frekari upplýsingar en síðar kom í ljós að meira að segja þar voru gögn brotakennd og ófullnægjandi.

Hefðin í íslenskum orðabókum er að gefa kenniföll nafnorða (eignarfall eintölu og nefnifall fleirtölu) og kennimyndir sagna (þrjár fyrir veikar sagnir og fjórar fyrir sterkar sagnir) (sjá t.d. Eiríkur Rögnvaldsson 1998 og Kristín Bjarnadóttir 2006). Ýmist eru endingarnar einar látnar duga (*bátur*, kk. *-s*, *-ar*) eða beygingarmyndir skrifaðar fullum stöfum þar sem stofnbrigði koma fram (*köttur*, kk. *kattar*, *kettir*). Beygingarmyndir aðrar en kenniföll og kennimyndir eru yfirleitt ekki sýndar nema í undantekningartilfellum, þ.e. þegar beyging þykir óregluleg. Samt er beygingarkerfinu þannig háttáð að aðrar beygingarmyndir eru ekki endilega fyrirsegjanlegar út frá þessum upplýsingum og nægir þar að nefna þágufall eintölu í sterkum karlkynsorðum (Friðrik Magnússon 1984). Þágufall af *bátur* er t.d. ýmist *báti* eða *bát* en með greini er þágufallið *bátnum* en ekki **bátinum*. Ófyrirsegjanleiki af þessu tagi kemur fram í lýsingarorðum og sögnum líka, í mismiklum mæli eftir stofngerð. Lýsingu á beygingu lýsingarorða er þröngur stakkur skorinn í íslenskum orðabókum og henni er ekki sinnt nema í undantekningartilvikum, þ.e. þar sem

⁹ <http://bibliian.is>

¹⁰ <http://ja.is>

óreglan þykir vera veruleg. Niðurstaðan um upplýsingar um beygingu orða í orðabókum er því að hún sé brotakennd en á móti kemur að orðaforðinn er verulegur og að því leyti eru orðabækur góðar heimildir í BÍN.

Í málfræðiritum, sérstaklega í yfirlitsritum, er annar háttur á. Þar er greinargóð lýsing á beygingarkerfinu sem slíku, með beygingardæmum um valin orð. Markmiðið er að lýsa kerfinu sjálfu sem heild og nota til þess viðeigandi dæmi.¹¹ Rannsóknarsagan er orðin býsna löng; fyrsta umfjöllunin er í *Grammaticæ islandicæ rudimenta* eftir Jón Rúgmann sem kom fyrst út í Kaupmannahöfn 1651 (sjá Guðrún Kvaran 2005). Hefðin er líka sterk og segja má að Rask hafi mótað aðferðina í *Vejlæðningun* 1811. Sett er fram kerfi vel afmarkaðra beygingarflokka og tekin dæmi um þau. Sömu dæmin eru gjarnan notuð aftur og aftur þannig að orðaforðinn er mjög lítill, þó svo að allar heimildir séu lagðar saman.¹²

Málfræðiritin eru ekki góð heimild um beygingu einstakra orða, nema þar sem svo vill til að orðið er tekið sem dæmi. Megináhersla er á erfðaorðaforðann og dæmin eru fá. Vegna þessa verða sumir hlutar orðaforðans út undan, t.d. tökuorð, talmálsorð og slangur, jafnvel þó að þar sé sennilega helst von á nýbreytni í beygingum. Til dæmis má nefna beygingarmyndir í þágufalli eintölu í sterkum hvorugkynsorðum. Þar eru málfræðibækur sammála um að endingin *-i* sé altæk, utan þriggja eða fjögurra orða, þ.e. *fé*, *hné* (*kné*) og *tré*.¹³ Í nýlegri handbók um beygingar- og orðmyndunarfræði (Guðrún Kvaran 2005) koma aðeins fram tvær *i*-lausar myndir. Þágufallið *tré* er sýnt í beygingardæmi ((20–43)d, bls. 242) og þar fylgir athugasemd: „Aðeins *hné* beygist eins og *tré*“. Í yfirliti um beygingu sterkra hvorugkynsorða segir svo: „Þágufall eintölu sterkra beygingar endar alltaf á *-i*.“ (bls. 244).

Í vinnunni við BÍN kom fljótlega í ljós að undantekningarnar eru mun fleiri þar sem ósamsett fleirkvæð hvorugkynsorð eru ýmist end-

¹¹ Undantekningin er kandiðatsritgerð Ástu Svavarsdóttur (1993) en þar er beygingu nafnorða lýst, eftir tíðnitölum, og stuðst við tiltekið gagnasafn. Annars er viðfangsefnið oftast beygingarkerfið í heild eða tilteknir flokkar innan þess.

¹² Yfirgrípsmesta og gagnlegasta beygingarlýsingin við vinnuna við BÍN reyndist vera bók Valtýs Guðmundssonar *Íslandsk grammatik* (1922), sérstaklega um beygingu sagna, en verkið hefur samt sem áður sömu einkenni og önnur verk af sama toga; þar er gerð grein fyrir meginflokkum og undantekninga getið í viðbótargreinum.

¹³ Hjá Valtý Guðmundssyni er orðið *hlé* haft þarna með. Í nútímamáli virðist það yfirleitt fá endinguna *-i* í þágufalli.

ingarlaus eða fá *-i* í þágufalli eintölu. E.t.v. ráða málhreinsunarsjónarmið því að slíku efni er sleppt í málfræðiritum en jafnframt verður áherslan á kerfið sem heild til þess að afbrigði og frávik verða út undan. Flækjustigið verður óneitanlega meira þegar í ljós kemur að tökuorð haga sér ekki endilega alveg eins og erfðaorðaforðinn. Þetta gildir líka um rótgróin tökuorð, t.d. orð sem er að finna í textum frá því fyrir tíma Rasks skv. Ritmálssafni Orðabókar Háskólans, t.d. hvorugkynsorðin *alkóhól*, *flannel*, *fennel*, *vitriól* og *indígó* (sjá Kristín Bjarnadóttir 2010). Þágufallsafbrigðin af orðunum *fennel* og *arsenik* eru t.d. tvö í blaðatextum á *Tímarit.is*:

- (1) a Hvönn er svolítið lík fennel<þgf.et.>
b ...ásamt brytjuðu fenneli<þgf.et.>
- (2) a ...byrjað eitur, drepinn með arsenik<þgf.et.>
b Hann var myrtur með arseniki<þgf.et.>

Gagnaskorturinn birtist því með ólíkum hætti í heimildunum. Í orðabókum og orðabókarefni er mikill orðaforði en upplýsingar um beygingu einstakra orða eru ekki tæmandi. Í mállýsingum eru hins vegar ítarlegar upplýsingar um hluta orðaforðans, þ.e. lýsing á kerfinu sem slíku, en eins og dæmið um þágufalls-*i* sýnir er sú lýsing ekki heldur tæmandi. Vegna þessa var óhjákvæmilegt að safna meiri gögnum við gerð BÍN.

5 Rannsóknargildi BÍN

Gagnvart BÍN er einkum þörf á því að leita að ófyrirsegjanlegum beygingarmyndum en talsverð vinna fólst í því að fylla í þau göt sem hér hafa verið nefnd í heimildunum. Til þessa eru öll gögn notuð sem hægt er að komast yfir að skoða og á það bæði við um gögn innan stofnunar og utan. Þar er fyrst að nefna söfn Orðabókar Háskólans, sérstaklega Ritmálssafn, prentaðar orðabækur og rafræn textasöfn (t.d. *Íslenskt textasafn* og safn Landsbókasafns Íslands – Háskólabókasafns af textum úr blöðum og tímaritum).¹⁴ Mat á gögnum styðst við málkennð þess sem vinnur verkið og er hún stundum eina haldreipið þegar engin dæmi finnast. Þá er oft leitað til allra þeirra sem gefa færi á sér og notendur BÍN á vefnum senda líka inn athugasemdir og viðbætur. Í neyð er stuðst við leit á vefnum (með Google), þrátt fyrir alla

¹⁴ *Íslenskt textasafn*: http://www.arnastofnun.is/page/gagnasofn_textasafn; og *Tímarit.is*.

þá annmarka sem því fylgja, bæði við að meta heimildir og einnig vegna þess hve slík leit er tímafrek og ómarkviss. Þrátt fyrir þetta er niðurstaðan sú að sumar beygingarmyndir eru með öllu ófinnanlegar, alveg sama hvaða aðferðum er beitt.

Dæmi um þetta er þágufall eintölu af orðinu *Yggdrasil*. Í fornum textum hefur ekki tekist að finna eitt einasta dæmi um þágufallið sem hlýtur að vera annaðhvort *Yggdrasli* (sbr. reglulega beygingu karlkynsnafnorða sem hafa endinguna *-ill* í nefnifalli, t.d. *bendill*, *brimill*, *depill*, þgf. *bendli*, *brimli*, *depli*) eða *Yggdrasil*, sem væri þá sambærileg undantekning og endingarleysið í þágufalli í *Huginn* (þgf. *Hugin*, ekki **Hugni*) og *Muninn* (þgf. *Munin*, ekki **Munni* (!)). E.t.v. eru það líkindin við hvorugkynsorðið *drasl* sem verða til þess að fólk virðist veigra sér við að nota þágufallsmyndina *Yggdrasli* og finnst hún jafnvel fyndin, en þágufallsmyndin *Yggdrasil* þykir ekki alveg nógu góð heldur.¹⁵ Flestir virðast því reyna að komast hjá því að nefna verslunina *Yggdrasil* í þágufalli og haga máli sínu á annan veg.

Eins og áður sagði er ekki alltaf hægt að finna dæmi um beygingarmyndir og þá er málkenndin ein til leiðsagnar og m.a. vegna þessa verður BÍN seint fullunnið verk. Þar er giskað á myndir þegar gögnin duga ekki til og þar sem heimildir þykja ekki traustar. Reynt er að leiðbeina notendum vefjarins með því að bæta inn athugasemdum um málnotkun ofan við beygingardæmin þannig að fólk geti sjálft myndað sér skoðun á efninu. Í athugasemd við orðið *Yggdrasil* er t.d. tekið fram að heimildir um þágufallið séu ekki áreiðanlegar.

Það er m.ö.o. ekki svo að allar orðmyndir sem finnast eigi greiða leið í BÍN, þó að BÍN sé beygingarlýsing en ekki forskrift. Orðmyndir sem enginn íslenskur málnotandi tæki gildar koma fyrir, sérstaklega á vefnum, svo sem fleirtölumyndin *fótarnir* í stað *fætarnir* af karlkynsnafnorðinu *fótur* og þágufallsmyndin *fótnum* í stað *fætinum*, og þær eiga ekki heima í BÍN. Venjulega er hægt að útiloka slíkar myndir með því að styðjast við tíðnitölur, a.m.k. í algengum orðum, en á endanum er það ákvörðun þess sem vinnur verkið hvaða beygingarmyndir á að sýna.¹⁶ Til máltækninota, sérstaklega til greiningar, er kostur að taka sem allra flestar beygingarmyndir með, þar sem fjöldi óþekktra orðmynda minnkar við það. Það er hins vegar jafnvægislist að meta hvaða beygingarmyndir eru afbrigði og hverjar eru villur, t.d. þegar

¹⁵ Á vefnum finnast dæmi um báðar myndir en dæmi um *Yggdrasli* eru helst í skrifum um þágufallið sjálft.

¹⁶ Mörkin á milli afbrigðis af beygingarmynd og villu eru stundum ljós og stundum ekki en í öllum textum leynast villur.

nota á gögnin í leiðréttingarhugbúnað. Lausnin í BÍN er að styðjast við tíðnitölur og málfenndina. Orðmyndum sem koma sárásjaldan fyrir, jafnvel í ótraustum heimildum, og þykja að auki verulega ankanna-
legar er því sleppt. Þetta er of frjálst fyrir þá sem vilja að verkið sýni málið aðeins eins og það á að vera „að bestu manna vitund“ en of íhaldssamt fyrir þá sem vilja fá allt með, ef það hefur á annað borð sést í málinu. Þar sem BÍN er ætlað að þjóna tveimur herrum, máltækni og almennum notendum, verður að fara bil beggja. Sérstök ástæða er til að fara varlega vegna þess að notendur vefsíðunnar virðast oft gera ráð fyrir því að BÍN sé vottunaraðili um það sem talið er rétt mál, þrátt fyrir að í athugasemdunum með beygingardæmunum sé oft tekið fram að tiltekin afbrigði séu ekki í samræmi við nógildandi stafsetningu eða stefnu stífustu málræktarmanna.¹⁷

Lýsingin á beygingarkerfinu í BÍN takmarkast af því að leitað er að einstökum beygingarmyndum eftir því sem ástæða þykir til en ekki viðhöfð allsherjar orðmyndaleit í ómörkuðum texta, enda er slík aðgerð óframkvæmanleg. Ástæðan er m.a. sú að margræðni beygingarmynda í íslensku er gríðarleg, bæði innan beygingardæmis og á milli þeirra. Sem dæmi má nefna orðmyndina *minni* sem kemur fram í fjórum orðum, alls í 36 beygingarmyndum.¹⁸

Margræðnin er augljós af tölum um beygingarmyndir og orðmyndir í BÍN. Þar eru 5,8 milljón beygingarmyndir (þ.e. orðmyndir með marki eða greiningarstreng) en orðmyndir eru 2,8 milljónir. Af þeim eru 1,8 milljón orðmyndir sem bara koma fyrir á einum stað, þ.e. ótvíræðar eða einræðar orðmyndir. Milljón orðmyndir eru því margræðar. Í Töflu 4 eru tölur um tvíræðni beygingarmynda en þar eru tölurnar hærri þar sem beygingarmynd er orðmynd með marki eða greiningarstreng.

¹⁷ Í BÍN er ætlast til þess að notendur myndi sér skoðun sjálfir. M.ö.o. er spurningunni um það hvað er gott og rétt ekki svarað í BÍN. Til þess skortir einfaldlega rannsóknir.

¹⁸ Orðmyndin *minni*: miðstig af lýsingarorðinu *lítill* (20 beygingarmyndir), sögnin *minna* (10 beygingarmyndir), hvorugkynsnafnorðið *minni* (5 beygingarmyndir) og svo þágufall eintölu, kvenkyn af eignarfornafninu *minn*.

Beygingarmyndir í BÍN	5.881.374	
Einræðar beygingarmyndir	1.850.090	31,5%
Margræðar innan beygingardæmis ¹⁹	3.619.482	61,5%
Margræðar milli beygingardæma ²⁰	63.641	1,1%
Margræðar innan og milli beygingardæma ²¹	348.161	5,9%

Tafla 4. Margræðni beygingarmynda í BÍN.

Vegna tvíræðninnar er markaður texti, eins og *Mörkuð íslensk málheild*, einstakt tæki til þess að finna dæmi um einstakar beygingarmyndir.

6 Samanburður við Markaða íslenska málheild

Í *Markaðri íslenskri málheild* (MÍM) eru u.þ.b. 25 milljón lesmálsorð en samanburðurinn á MÍM og BÍN sem hér er lýst byggist á hluta MÍM, u.þ.b. 17,7 milljónum lesmálsorða. Samanburðurinn var gerður með það í huga að endurbæta BÍN og auka.²² Talnastrengjum, táknum o.þ.h. var sleppt úr MÍM, þannig að lesmálsorðin úr MÍM sem borin voru saman við BÍN voru 16.245.429.

Lesmálsorð	16.245.429
Einræðar markaðar orðmyndir	737.856
Þar af í BÍN	425.238
Þar af ekki í BÍN	312.618

Tafla 5. Einræðar markaðar orðmyndir í samanburði MÍM og BÍN.

Þess ber að geta að efnið úr MÍM er ekki yfirfarið og leiðrétt en samanburðurinn fólst í því að keyra saman lista um beygingarmyndir (þ.e. orðmynd með marki) úr MÍM og sambærilegt efni úr BÍN.²³ Beygingarmynd úr MÍM fylgir nefnimynd (þ.e. uppflettimynd eða lemma) en talsvert er um villur í lemmuninni. Síðan var farið yfir niðurstöðuna á skjá, línu fyrir línu. Í samanburðinum var orðflokksgreining úr MÍM

¹⁹ Dæmi: Orðmyndin *penna* er þolfall, þágufall og eignarfall eintölu af nafnorðinu *penni*.

²⁰ Dæmi: Orðmyndin *burstanna* er eignarfall fleirtölu af nafnorðnum *burst* og *bursti*.

²¹ Dæmi: Orðmyndin *minni* er 34 beygingarmyndir af fjórum orðum, lo. *lítill*, so. *minna*, hk. *minni* og fn. *minn*.

²² Sumt af efninu í þessum hluta MÍM verður í reynd ekki hluti af aðalsafni MÍM en kemur samt að fullum notum við vinnuna við BÍN. Mörkun á MÍM var heldur ekki lokið þannig að ýmislegt af því sem hér er lýst byggir ekki á endanlegri gerð MÍM. Hluti af afrakstri samkeyrslunnar var einmitt hugsaður til þess að koma að gagni við bæði verkin.

²³ Markamengin eru ekki eins en þau voru samræmd fyrir samkeyrsluna.

leiðrétt og lemmunin líka, í því efni sem telst vera íslenskur orðaforði.²⁴ Þess ber að geta að hugbúnaðurinn sem notaður er í lemmunina, svokallað lemmald (Anton Karl Ingason o.fl. 2008), erfir villur í orðflokkgreiningu úr mörkunum í MÍM en textinn er fyrst markaður og síðan lemnaður. Rangur orðflokkur í marki verður þess vegna til þess að sjálf nefnimyndin eða lemman verður óhjákvæmilega röng. Í yfirferðinni var efni sem ekki telst til íslensks orðaforða flokkað sérstaklega og merkt eins og komið verður að hér á eftir.

Alls var farið yfir 312.618 orðmyndir (eða markaða strengi) úr MÍM, þ.e. það efni sem ekki fannst í BÍN. Ríflega þriðjungur (34,5%) þessara orðmynda var rétt orðflokksmerktur og lemnaður í MÍM en hafði einfaldlega ekki ratað inn í BÍN úr öðrum heimildum. Í Töflu 6 er fjöldi beygingarmynda úr meginorðflokkinum þremur úr samanburðarefninu úr MÍM. Í 1. dálki er fjöldi beygingarmynda skv. mörkuninni í MÍM og í 2. dálki er fjöldi rétt markaðra beygingarmynda þar. Mismunurinn á 1. og 2. dálki sýnir ranga greiningu og/eða lemmun í MÍM.²⁵ Í 3. dálki eru tölur um beygingarmyndir í hverjum orðflokki eftir handleiðréttingu á samanburðarefninu. Þar er m.ö.o. niðurstöðutala um beygingarmyndir í hverjum orðflokki í því efni úr MÍM sem ekki finnst í BÍN, eftir leiðréttingu.

	Heildartala í MÍM	Rétt í MÍM	Niðurstaða
hk.	70.770	31.942	54.317
kk.	142.630	42.910	64.709
kvk.	78.908	46.207	64.499
lo.	23.835	11.713	18.345
so.	7.639	1.035	3.210

Tafla 6. Fjöldi beygingarmynda í MÍM, fyrir og eftir leiðréttingu.

Samanburðurinn á MÍM og BÍN var aðallega gerður til þess að finna íslenskar beygingarmyndir sem vantaði í BÍN. U.þ.b. 60% af öllum mörkuðum strengjum úr MÍM sem ekki fundust í BÍN reyndust vera íslenskar beygingarmyndir. Í yfirferðinni var afgangurinn, u.þ.b. 40%, flokkaður lauslega. Þar var erlent efni fyrirferðarmest (24,6%), en einnig má nefna villur (5,8%), skammstafanir (1,6%) og ýmiss konar tölvustrengi, vefföng, slóðir o.þ.h. (0,7%). Þetta efni skiptir ekki bein-

²⁴ Til dæmis má nefna mark og lemmun á beygingarmyndinni *Miklagarði* sem greind er sem þolfall eintölu af kvenkynsnafnorðinu *Miklagarð* í MÍM, í stað þágufalls eintölu af karlkynsnafnorðinu *Mikligarður*.

²⁵ Athugið að villufjöldinn á aðeins við efni sem ekki finnst í BÍN og hlutfallið gefur alls ekki mynd af heildarvillufjölda í greiningu og/eða lemmun í MÍM.

línis máli í vinnunni við BÍN en getur nýst sem ýtarefni í máltækni, t.d. í leiðréttingarbúnað o.þ.h., þar sem eitt af meginvandamálum í máltækni er greining á óþekktum „orðum“ eða stafastrengjum.

Fyrsta niðurstaða úr samanburðinum á MÍM og BÍN er sú að bæta þurfi 125 þúsund beygingardæmum við í BÍN. Meðalfjöldi orðmynda í beygingardæmi í BÍN er rétt yfir 20, þannig að alls gætu viðbótarorðmyndirnar orðið yfir 2,5 milljónir en heildarfjöldi beygingardæma yrði þá hátt í 400 þúsund og beygingarmyndir yfir 8,3 milljónir. Þetta eru bráðabirgðatölur þar sem bera þarf saman uppflettiorð í BÍN og MÍM og gáta lemmunina en því verki verður ekki lokið fyrr en allt efnið er komið inn í BÍN. Í vinnuferlinu verður gengið úr skugga um að allar beygingarmyndir í MÍM séu á sínum stað í BÍN og er það lokastigið í samanburðinum.

Vinna við viðbótarefnið í BÍN er hafin. Beygingarflokkur orða er fundinn með samanburði við efni sem fyrir er í BÍN, m.a. með því að nota greiningartól fyrir samsett orð.²⁶ Síðan er færslan fyrir hvert orð fullunnin, með því að merkja gildi á beygingarformdeildum sem koma eiga fram í hverju beygingardæmi, t.d. +/-et, +/-ft. o.s.frv., sbr. Töflu 1. Beygingarmyndir úr MÍM verða hafðar til hliðsjónar, ásamt efni úr öðrum heimildum eftir þörfum, sbr. það sem sagt er um gagnaoöflun hér að framan. Jafnframt verða beygingardæmi sem þegar eru í BÍN endurskoðuð eftir því sem viðbótarbeygingarmyndir úr MÍM gefa tilefni til.

7 Gagnaskortur

Samanburðarefnið úr MÍM er 17,7 milljónir lesmálsorða, alls 737.856 strengir sem teljast vera orðmyndir. Þar af eru beygingarmyndirnar sem hér eru til athugunar tæplega 623 þúsund (84%). Þetta er lítið brot af beygingarmyndunum sem þegar eru í BÍN, þ.e. 5,8 milljónum.²⁷ Þessar tölur gefa vísbendingu um það hve stór málheild þyrfti að vera til þess að hægt væri að nota hana til þess að setja fram beygingarlýsingu á borð við BÍN eða samsvarandi beygingarreglur um ís-

²⁶ Greiningartólið er verk Jóns Friðriks Daðasonar í samvinnu við Stofnun Árna Magnússonar í íslenskum fræðum og er hluti af meistaraverkefni hans um leiðréttingu á ljóslesnum texta (Jón Friðrik Daðason 2012).

²⁷ Af tæknilegum ástæðum er ekki enn gerlegt að bera saman heildarfjölda uppflettiorða í MÍM og BÍN þar sem gera verður ráð fyrir talsverðum fjölda af villum í lemmun í MÍM á orðmyndum sem einnig eru BÍN. Það efni hefur ekki enn verið athugað.

lensku. Beygingarlýsing sem aðeins byggði á dæmum úr MÍM væri ekki stór enda er beygingarmyndafjöldinn í MÍM innan við tíundi hluti af beygingarmyndafjöldanum í BÍN og orðaforðinn er aðeins lítill hluti af því sem venjulegur Íslendingur hefur á valdi sínu. Þetta vandamál kemur íslenskum orðabókarmönnum ekki á óvart, þar sem sams konar gagnapurrd eða gagnaskortur kemur fram þegar leitað er að beygingarmyndum m.t.t. setningarlegrar stöðu eða samhengis (Jörgen Pind o.fl. 1993).

Eftir fyrirhugaðar viðbætur við BÍN þar sem byggt er á samanburðinum við MÍM verða orðmyndir þar alls 8,3 milljónir. Úr 623 þúsund orðmyndum úr MÍM bætast því við 2,5 milljónir nýrra beygingarmynda í BÍN og beygingardæmin þar verða sennilega yfir 395 þúsund. Til samanburðar má nefna að í Ritmálssafni Orðabókar Háskólans eru yfir 700 þúsund uppflettiorð. Þar er að vísu sögulegt efni, þ.e. orðaforði frá miðri 16. öld til nútímans, en samt sem áður er þetta vísbending um hve gríðarlegur orðaforðinn er í raun og veru. Orðmyndun í íslensku er virk og því þarf að nota allar aðferðir til þess að fá heildarmynd af orðaforðanum og allar heimildir, orðasöfn, málheildir, og hvað eina sem til fellur, auk greiningartóla og reglna.

8 Niðurstaða

Eins og fram kom í inngangi var upprunalegur tilgangur BÍN sá að búa til efni til máltækninota og jafnframt að koma efninu á framfæri til almennra notenda. Við vinnuna hefur gildi BÍN við rannsóknir á beygingarkerfinu smám saman komið betur og betur í ljós. Þrátt fyrir aldalanga hefð í lýsingu á íslenska beygingarkerfinu er enn mikið verk óunnið. Íslenska beygingakerfið er flókið, þar er mikill fjöldi beygingarflokka, afbrigði eru algeng og jafnframt er skrið á orðum og orðmyndum á milli beygingarflokka. Vegna þessa var strax frá upphafi ákveðið að í BÍN yrðu sett upp beygingardæmi fyrir einstök orð og þau geymd sem ein heild í stað þess að búa til kerfi af ofvirkum beygingarreglum. Hverju beygingardæmi fylgja tvenns konar athugasemdir. Í vinnuathugasemdum eru niðurstöður af athugunum á einstökum orðum og annað tilfallandi efni um vinnuna en í athugasemdum til notenda á vefnum eru leiðbeiningar um málnotkun, t.d. um notkun afbrigða, eftir merkingu, stíl o.s.frv.

Enn sem komið er byggist BÍN á því að rannsaka einstök orð og geyma beygingardæmin sem heild. Smám saman verður raunhæft í ljósi þess sem áður er rakið að búa til reglakerfi til þess að lýsa íslenska beygingarkerfinu, eftir því sem gögnin í BÍN aukast. Slíkt tól kæmi áreiðanlega að góðum notum í ýmsum máltækniverkefnum, sérstaklega ef það væri notað ásamt orðhlutagreini, þar sem BÍN er ekki og verður aldrei tæmandi.

Rétt er að minna á að BÍN er enn verk í vinnslu. Gögn úr BÍN eru aðgengileg án endurgjalds á vefsíðu Stofnunar Árna Magnússonar í íslenskum fræðum, með þeim skilmálum sem þar koma fram.

Heimildir

- Anton Karl Ingason, Hrafn Loftsson, Sigrún Helgadóttir og Eiríkur Rögnvaldsson. 2008. A Mixed Method Lemmatization Algorithm Using a Hierarchy of Linguistic Identities (HOLI). Í: Raante, Arne, og Bengt Nordström (ritstj.): *Advances in Natural Language Processing*, s. 205–216. (Lecture Notes in Computer Science, Vol. 5221.) Berlín: Springer.
- Ásta Svavarsdóttir. 1993. *Beygingakerfi nafnorða í nútímaíslensku*. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Eiríkur Rögnvaldsson. 1990. *Íslensk orðhlutafræði*. Reykjavík: Málvísindastofnun Háskóla Íslands.
- Eiríkur Rögnvaldsson 1998. Málfræði í íslenskri orðabók: Hvernig og til hvers? *Orð og tunga* 4:25–32.
- Eiríkur Rögnvaldsson, Hrafn Loftsson, Kristín Bjarnadóttir, Sigrún Helgadóttir, Anna B. Nikulásdóttir, Matthew Whelpton og Anton Karl Ingason. 2009. Icelandic Language Resources and Technology: Status and Prospects. *Proceedings of the NODALIDA 2009 Workshop Nordic Perspectives on the CLARIN Infrastructure of Language Resources*. Odense.
- Friðrik Magnússon. 1984. Ein lítil beygingarending. *Mímir* 32:33–43.
- Guðrún Kvaran. 2005. *Orð*. Handbók um beygingar- og orðmyndunarfræði. Íslensk tunga 2. Reykjavík: Almenna bókafélagið.
- Guðrún Kvaran og Sigurður Jónsson frá Arnarvatni. 1991. *Nöfn Íslendinga* (1. útg.). Reykjavík: Heimskringla.
- Hrafn Loftsson, Jökull H. Yngvason, Sigrún Helgadóttir og Eiríkur Rögnvaldsson. 2010. Developing a PoS-tagged corpus using existing tools. Sarasola, Kepa, Francis M. Tyers og Mikel L. Forcada (ritstj.), *SaLTMiL Workshop on Creation and Use of Basic Lexical Resources for Less-Resourced Languages*, LREC 2010, s. 53-60. Valetta, Möltu. [Rafræn útgáfa 8.7.2013: <https://notendur.hi.is/eirikur/corpustagging.pdf>.]
- Jón Friðrik Daðason. 2012. *The Post-Correction of Icelandic OCR Text*. Skemman.is. [<http://hdl.handle.net/1946/12085>]

- Jörgen Pind, Friðrik Magnússon og Stefán Briem. 1991. *Íslensk orðtíðnibók*. Reykjavík: Orðabók Háskólans.
- Jörgen Pind, Kristín Bjarnadóttir, Jón Hilmar Jónsson, Guðrún Kvaran, Friðrik Magnússon og Ásta Svavarsdóttir. 1993. Using a Computer Corpus to Supplement a Citation Collection for a Historical Dictionary. *International Journal of Lexicography* 6/1:1–18.
- Kristín Bjarnadóttir. 1998. *Norræna verkefnið*. [8.7.2013: <http://lexis.hi.is/kristinb/norr.pdf>.]
- Kristín Bjarnadóttir. 2006. Málfræði í orðabókum. *Orð og tunga* 8:27–43.
- Kristín Bjarnadóttir. 2010. *Algilt -i eða hverfult: Um þágufall eintölu í sterkum hvorugkynsnafnorðum*. 24. Rask-ráðstefnan 30. janúar 2010. [<http://lexis.hi.is/kristinb/KB-Rask2010.pdf>.]
- Kristín Bjarnadóttir. 2012. The Database of Modern Icelandic Inflection. *Proceedings of "Language Technology for Normalization of Less-Resourced Languages"*, SaLTMiL 8 – AfLaT 2012. Istanbul, Tyrklandi. [Rafræn útgáfa 8.7.2013: <http://www.lexis.hi.is/kristinb/lrec2012-dmii.pdf>.]
- Mörður Árnason (ritstj.). 2000. *Íslensk orðabók*. Rafræn útgáfa. Reykjavík: Edda hf.
- Rask, Rasmus Christian. 1811. *Vejledningen til det Islandske eller gamle nordiske Sprog*. København: Schubothes Forlag.
- Sigrún Helgadóttir, Ásta Svavarsdóttir, Eiríkur Rögnvaldsson, Kristín Bjarnadóttir og Hrafn Loftsson. 2012. The Tagged Icelandic Corpus (MÍM). *Proceedings of "Language Technology for Normalization of Less-Resourced Languages"*, SaLTMiL 8 – AfLaT 2012. Istanbul, Tyrklandi. [Rafræn útgáfa 8.7.2013: <http://www.lexis.hi.is/kristinb/lrec2012SH-saltmil.pdf>.]
- Valtýr Guðmundsson. 1922. *Íslensk grammatik*. Kaupmannahöfn: Hagerup.

Lykilorð

máltækni, gagnasöfn, beyging, málheild, gagnaskortur

Keywords

language technology, language resources, inflection, corpora, data scarcity

Abstract

The topic of the article is the *Database of Modern Icelandic Inflection* (DMII), a collection of about 270 thousand paradigms with over 5.8 million inflectional forms. The DMII was created as a multipurpose resource, for use in language technology, lexicography, and as a reference tool for the general public online. In spite of a long tradition of morphological research, it was clear from the beginning that the necessary data for a productive rule system for Icelandic inflection was not available, in a manner that would produce all and only the appropriate inflectional forms. The result is therefore

a database containing the full set of inflectional forms, with grammatical tags. The article contains a description of the creation of the DMII and its sources, initially almost completely lexicographic. The creation of the *Tagged Icelandic Corpus* (MÍM) has, for the first time, made an empirical study of Icelandic inflection possible, but comparison of the data from the DMII and MÍM shows that data scarcity makes it imperative to use both kinds of data for a thorough description of the inflectional system, due to the huge tag set in a very rich inflectional system.

Kristín Bjarnadóttir

Stofnun Árna Magnússonar í íslenskum fræðum / Háskóli Íslands

Neshaga 16

107 Reykjavík

kristinb@hi.is

Jónína Hafsteinsdóttir

Þveit

Suðvestan við Ketillaugarfjall í Nesjum, A-Skaftafellssýslu, er stöðuvatn sem nefnt er *Þveit*, eitt stærsta stöðuvatn sýslunnar. Hálend eyja er í Þveitinni, nefnd *Rauðhamar* (Byggðasaga I 1971:239). Þetta örnefni er hvergi til annars staðar á Íslandi svo vitað sé en það er þekkt í nágrannalöndum.

Merking orðsins *tveit* (*þveit* 'lítill landskiki' í fornu máli) er í nýnorsku 'grasflöt í skógi eða á milli kletta' eða 'slægjustykki í skógi'. *Tveit* (<*Þveit*) er algengt í norskum bæjanöfnum. Það kemur fyrir um 200 sinnum sem ósamsett bæjanafn og tvöfalt fleiri dæmi eru um það sem síðari lið í samsettum nöfnum (*-tveit*, *-tvet* o.fl.). Þá er það notað í meira en 80 seljanöfnum svo kunnugt sé og er ósamsett í fjórðungi þeirra. Hins vegar er þetta sjaldgæft sem fyrri liður samsettra gamalla bæjanafna (*Tveit*-, *Tveitar*-). Flestir þessara bæja munu vera frá víkingatíma, jafnvel eldri. Fyrri hluti samsettra bæjanafna er oftast mannsnafn eða viðurnefni, stundum guðsheiti (*Frøys*-, *Tors*-, *Ull*-), en dæmi eru þó um fyrri liði af öðru tagi. Kristin mannanöfn koma ekki fyrir. Sé útbreiðsla nafnanna í Noregi skoðuð kemur í ljós að þau eru algengust á svæðinu frá Þelamörk til Hörðalands. Í þessum tveimur fylkjum er t.d. meira en helmingur ósamsettu nafnanna. Fá nöfn eru á Vestur-Ögðum sem þó eru ekki fjarri hinum fylkjunum. Ríflega 30 ósamsett nöfn eru á Austlandi (í Akurshúsfylki, á Austfold, Heiðmörk og Upplöndum) en athygli vekur að umræddur nafnliður er nánast óþekktur norðan Sognsfjarðar. Annar algengur liður í norskum bæjanöfnum er *-rud* og merkir svæði sem rutt hefur verið í skógi til ræktunar eða byggingar, í fornu máli *ruð*, af so. *ryðja*. Nöfn með endingunni *-ruð* koma ekki fyrir á Íslandi. Nafnliðirnir *-tveit* og *-rud*

eru því taldir merkja nokkurn veginn það sama en nöfnin í síðari flokknum munu vera yngri en hin (*Norsk stadnamnleksikon* 2007:466–467). Orðið *tveit* hefur m.a. einnig verið notað um skurð eða axarfar í tré sem verið er að höggva (Aasen 1918:850). Bent hefur verið á að býli með *þveitar*-nöfnum hafi yfirleitt ekki talist til betri bæja og nefnir Magnus Olsen **Þrælaþveit* sem dæmi en það (*Threltuedt*) er meðal horfinna nafna sem Olof Rygh tíundar í áttunda bindi ritverks síns *Norske Gaardnavne* (Olsen 1939:22; Rygh 1905:209).

Nafnliðurinn *-tved* er í u.þ.b. 125 dönskum byggðaheitum („Bebyggelsesnavne“) og kemur fyrir í ýmsum samsetningum. 54 þeirra nafna eru á eyjunum og um 70 á Jótlandi þar sem þau eru nær eingöngu bundin við skógi vaxin svæði (Knudsen 1939:107–109). Peder Dam og Johnny Gøgsig Jakobsen hafa kortlagt dreifingu þessara og fleiri nafna af svipuðum toga (þ. á m. *-rød*) í Danmörku eins og hún birtist í jarðabók Kristjáns V. frá 1688 (Dam og Jakobsen 2007:35–59). Athygli hefur vakið að þrátt fyrir verulega útbreiðslu slíkra „ruðningsnafna“ (*-tved* og *-ryd*) í Danmörku eru þau ekki til á Borgundarhólmi (Holmberg 2007:100).

Nöfn mynduð með þessum lið eru fleiri en ætla mætti því að hann hefur tilhneigingu til að verða *-sted* þegar fyrri liður endar á *-s* (*-s-tved* > *-sted*: *Dysted*, *Søndersted*) og hann á líka til að stytast (*Harte*, *Vedde*). Menn hafa deilt um merkingu nafnliðarins í dönskum örnefnum, t.d. talið hann merkja ‘eitthvað sem afmarkast af vatni, spildu umlukta skógi eða ræktaðan skika umluktan óræktuðu landi’. Gunnar Knudsen rekur slóð nafnliðarins í öðrum norrænum málum, nefnir m.a. vesturnorræna orðið *þveit* sem haft er um ‘lítinn jarðarskika’ og samsvarandi orð getur í norskum mállýskum táknað ‘skurð í trjábol, engjablett í skógi’ o.fl. Svipuð merking hefur verið lögð í sænsk staðahæiti af sama uppruna og hún birtist einnig í enskum mállýskum þar sem orðið hefur skotið upp kolli sem norrænt tökuorð (Knudsen 1939:108–109).

Aage Houken segir um nafnið *Tved* að það sé leitt af sögn sem svarar til sagna í öðrum norrænum málum (vnorr. *þveita*, no.máll. *tveita*, sæ.máll. *tveta*) sem allar merki ‘höggva, höggva af’ eða eitthvað í þá veru (Houken 1976:253). Upprunalega merki orðið því ‘eitthvað afhöggið’ og þar af spretti merkingin ‘höggvinn skógur, ruddur blettur’. Aage Houken getur þess í umfjöllun sinni um *rød* (*ryd*) að nöfn með þeim lið séu upp komin eftir lok víkingatíma þar sem þau sé hvorki að finna í Danalögum né Normandí (Houken 1976:190).

Samnafnið *tved* er ekki þekkt ósamsett í dönsku – einungis sem

nafnliður, dreginn af sagnorði sem merkir 'skera'. Orðliðurinn er kvenkyns og merking hans 'staður þar sem tré eða runnar hafa verið felldir; ruddur blettur'. Nöfn, samsett með þessum lið, eru án efa mörg frá tímum víkinga því að hann kemur fyrir í dönskum nöfnum í Danalögum og Normandí. Nöfn með endingunni *-tved* hafa oft tekið miklum breytingum og má bæta ýmsum dæmum við þau sem áður voru nefnd (*Humble, Hoed*) (Jørgensen 2008:307). Dæmi eru þó þess að *-tved*-nöfn eigi sér aðra skýringu: Nafnið *Egtved* er til á tveimur stöðum í Danmörku. Annar er á Sjálandi, þar sem heitir *Egtuid* um 1525 og merkir væntanlega 'rudda svæðið milli eikanna', hinn á Jótlandi og heitir *Ekthyuf* um 1325. Fyrri hluti nafnsins er að vísu „eik“ en seinni hlutann rekja menn til forndanska orðsins **thiuf* sem merkir 'kjarr' – *Eikarkjarr* (Jørgensen 2008:66). Ekki er því allt sem sýnist.

Þorps- eða bæjarnafnið *Tvet* (*Þveit* 1383) (*Tveta, Tvetane*) er þekkt á nokkrum stöðum í Svíþjóð, einkum í Bohusléní. Þar um slóðir leggja menn sömu merkingu í nafnið og annars staðar og rekja allt aftur til víkingatíma þegar orðið (*tveit*) hefur þýtt 'ruddur blettur eða svæði í skógi' (Ortnamnen 1938:49–50; Wadström 1983:79). Til er í sænskum mállýskum nafnorðið *tvet* í merkingunni 'höggspæmir' og er einnig haft um 'rauf eða skurð sem myndast við rót trés þegar það er fellt', rétt eins og sambærilegt orð í norsku (*tveit*). Fleirtöluorðið *tvetar* í sænskum mállýskum getur merkt 'trjábútar, trjabolir, felld tré, höggspæmir' (Ortnamnen 1938:49–50).

Til er engilsaxnesk sögn af sömu rót og þær norrænu sagnir sem nefndar hafa verið, *þwítan* 'skera af, skera í sundur', og raunveruleg merking nafnorðs (d. *tved*) af þeirri rót því 'það sem er skoríð í sundur'. Ekki er nema steinsnar frá merkingunni 'að höggva tré' yfir í 'ryðja' og orðið, sem haft er um staðinn, sprettur af því sem þar var áður og verkinu sem unnið var (Knudsen 1939:109; Houken 1976:253). Nafnliðurinn *-þveit* er algengur í skógi vöxnum sveitum í Normandí og má gera ráð fyrir að mörg þeirra nafna séu upp komin löngu eftir að víkingatíð er um garð gengin og hið sama á að líkindum við um ýmis héruð á Englandi. Áður en það gerðist hafði nafnliðurinn orðið að tökuorði, *þveit*, bæði í ensku og frönsku (Fellows-Jensen 1994:79). Ekki hefur vafist fyrir Engendingum að nota nöfnin sem bástust þeim með víkingum. Sem dæmi má nefna nafnliðina *þorp*, *þveit* og *toft* sem hafa lifað af í máli þeirra með framburði líklega áþekktum þeim sem var á víkingatíma (*Thorpe, Thwaite, Toft*). Öðru máli gegndi um frönsku enda hún fjarlægari norrænni tungu í öllum skilningi. Þar er *þ* borið fram sem *t* í nöfnum eins og *Le Torp* (*þorp*) og *Le Tuit* (*þveit*)

og samhljóðaklasar fara sínar eigin leiðir. Þeir eru oft einfaldaðir með því að bæta við innskotssérhljóði (d. *svarabhaktivokal*) eða fella niður annað samhljóð af tveimur, t.d. í nafninu *Etalondes* (*Stanelonde* 1056 < *steinn-lundr) (Fellows-Jensen 1994:74).

Á Englandi þekkja menn á þriðja hundrað nafna sem enda á *-þwait* og eru höfð um 'nýrækt'. Sambærileg nöfn eru til í Normandí (Houken 1976:253) þar sem liðurinn hefur orðið að *-tuit*. Um fyrri liði slíkra nafna er sama að segja og fyrr var nefnt, þeir eru sóttir í mannanöfn frá víkingatíma en nöfn úr kristnum sið fyrirfinnast ekki (Hald 1950:151).

Notkun og útbreiðsla örnefna af norrænum toga, m.a. nafna með endingunni *-þveit*, hefur orðið fræðimönnum notadrjúg við að kortleggja landvinninga og búsetu víkinga á Englandi og í Normandí (Fellows-Jensen 1994:79).

Ekki er auðséð hvernig koma má vatnsheitinu *Þveit* heim og saman við það sem nú hefur verið rakið og því nauðsynlegt að skyggjast aðeins víðar um. Norska orðið *tveit* hefur fleiri merkingar en áður hafa verið nefndar, t.d. 'renna höggvin í ís' (Aasen 1918:850). Renna – eða vök – höggvin í ís á vatni hlýtur að vera gerð í ákveðnum tilgangi. Víða hafa menn stundað það að höggva vakir til veiða niður um ís á sjó og vötnum. Bændur, a.m.k. á Stapa, nutu fyrrum hlunninda af silungsveiði í Þveitinni þó hún hafi farið þverrandi í seinni tíð (Byggðasaga I 1971:239). Þá vaknar sú spurning hvort Nesjamenn hafi á einhverju skeiði höggvið „þveit“ í ísinn á vatninu til veiða. Guðrún Guðmundsdóttir, sem fædd var á Miðskeri 1863 og átti heima í Nesjum fram á fullorðinsár, talar um silungsveiði í vatninu í endurminningum sínum (Guðrún Guðmundsdóttir 1975:36). Hún nafngreindir mann nokkurn sem var hneigður mjög til veiðiskapar og veiddi silung í Þveitinni sumar og vetur. „Þegar hann gat ekki lagt net sín fyrir ís, hjó hann vakir á ísinn og dorgaði silung upp um vakirnar“ (Guðrún Guðmundsdóttir 1975:91).

Eldri heimildir geta einnig um silungsveiði í Þveitinni. Ritari lýsingar Austur-Skaftafellssýslu um miðja 18. öld sér ástæðu til að láta þess getið að í Þveitinni sé „idkudnockur silungs veide“ (*Sýslulýsingar 1744–1749* 1957:14–15). Í lýsingu Bjarnaness- og Hoffellssókna 1840 segir að í vatninu sé „strax djúpt við löndin so silunganet eru lögð í það frá landi og veiðist þar á ýmsum tímum ársins lítið eitt af hönum“ (Skaftafellssýsla 1997:57). Önnur lýsing frá árinu 1873 lætur þess getið að í vatni þessu sé „silungur oft veiddur haust og vor í lagnetum“ (sama rit:91). Gæti ekki vatnið – e.t.v. nafnlaust áður – hafa

fengið heiti sitt af „þveit“ sem höggvin var í ísinn? Í nágrannalöndum fengu ruddir blettir og skikar í skógum nafnið *Þveit* og voru nýttir til ræktunar og búsetu. Gæti ekki „þveitin“ – rennan eða vökin – í ísilögðu vatninu á sama hátt hafa orðið ígildi nafns og að lokum fest sig í sessi sem slíkt og tekið til alls vatnsins?

Sagnir eru um nykur og fleiri kynjaskepnur í Þveitinni og hafa ýmsir, sem farið hafa yfir hana á ís, einkum síðla dags, heyrt þar drunur og ólæti mikil og hefur ekki þótt ráðlegt að vera þar einn á ferð í myrkri (Guðmundur Jónsson Hoffell 1946:84–85). Svipaðar sagnir skráði Sigfús Sigfússon í þjóðsagnasafni sínu (*Íslenskar þjóðsögur og sagnir* IV 1982:93–94) en nefnir vatnið *Þveiti* (hk.). Freistandi er að álykta að þarna kunni að búa að baki varúðarráðstöfun, jafnvel sumpart vegna „þveitarinnar“ í ísnum. Brestir í ís geta einmitt verið merki þess að hann sé ótraustur og vert að fara þá að öllu með gát. Slysavarnafélagið Landsbjörg hefur séð ástæðu til að brýna fyrir almenningi að fara ekki út á ís sem orðinn sé ótraustur vegna hláku og snúa ævinlega við ef brestir heyrast (*Skessuhorn* 24. janúar 2008).

Heimildaskrá

- Aasen, Ivar. 1918. *Norsk ordbok med dansk forklaring*. 4. útgáfa. Kristiania: Cammermeyer.
- Byggðasaga = *Byggðasaga Austur-Skaftafellssýslu*. I–III. 1971–1976. Reykjavík: Bókautgáfa Guðjóns Ó. Guðjónssonar.
- Dam, Peder og Johnny Gøgsig Jakobsen. 2007. Danske middelalderlige rydningsbebyggelser. Í: Birgit Eggert, Bente Holmberg og Bent Jørgensen (ritstj.). *Nordiske navnes centralitet og regionalitet*. Rapport fra NORNAs 35. symposium på Bornholm 4.–7. maj 2006, bls. 35–59. NORNA-Rapporter 82.
- Fellows-Jensen, Gillian. 1994. Vikinger i England og Normandiet: Hvad stednavnene fortæller. Í: Gillian Fellows-Jensen og Bente Holmberg (ritstj.). *Vikingetidens sted- og personnavne*. Rapport fra NORNAs 22. symposium i København 14.–16. januar 1993, bls. 67–87. NORNA-Rapporter 54.
- Guðmundur Jónsson Hoffell. 1946. *Skaftfellskar þjóðsögur og sagnir*. Akureyri: Þorsteinn M. Jónsson.
- Guðrún Guðmundsdóttir. 1975. *Minningar úr Hornafirði*. Reykjavík: Hið íslenska bókmenntafélag.
- Hald, Kristian. 1950. *Vore stednavne*. Kaupmannahöfn: Udvalget for Folkeoplysnings Fremme.
- Holmberg, Bente. 2007. Centralitet og regionalitet i Bornholms sted- og personnavne. Í: Birgit Eggert, Bente Holmberg og Bent Jørgensen (ritstj.).

- Nordiske navnes centralitet og regionalitet*. Rapport fra NORNAs 35. symposium på Bornholm 4.–7. maj 2006, bls. 91–106. NORNA-Rapporter 82. Houken, Aage. 1976. *Håndbog i danske stednavne*. 2. útgáfa. Kaupmannahöfn: Gyldendal.
- Íslenskar þjóðsögur og sagnir*. IV. 1982. Sigfús Sigfússon safnaði og skráði. Ný útgáfa. Reykjavík: Þjóðsaga.
- Jørgensen, Bent. 2008. *Danske stednavne*. 3. útgáfa. Kaupmannahöfn: Gyldendal.
- Knudsen, Gunnar. 1939. De danske Stednavne. Í: Magnus Olsen (útg.). *Stedsnavn*. Nordisk kultur V, Stockholm, bls. 76–123.
- Norsk stadnamnleksikon*. 2007. 4. útgáfa. Ritstjórar: Jørn Sandnes og Ola Stems-haug. Oslo: Samlaget.
- Olsen, Magnus. 1939. „Norske Gaardnavne.“ – Gård, Bygd og Vei. Í: Magnus Olsen (útg.). *Stedsnavn*. Nordisk kultur V, Stockholm, bls. 5–52.
- Ortnamnen = *Ortnamnen i Göteborgs och Bohus län*. XVIII. 1938. Göteborg: Institutet för ortnamns- och dialektforskning vid Göteborgs Högskola.
- Rygh, Olof. 1905. *Norske Gaardnavne. Oplysninger samlede til Brug ved Matrikelens Revision*. 8. bindi. Nedenes Amt. Kristiania: Fabritius.
- Skaftafellssýsla = *Skaftafellssýsla. Sýslu- og sóknalýsingar Hins íslenska bókmenntafélags 1830–1873*. 1997. Jón Aðalsteinn Jónsson og Svavar Sigmundsson sáu um útgáfuna. Reykjavík: Sögufélag.
- Skessuhorn*. Fréttaveita Vesturlands. 24. janúar 2008. Á vefsíðunni <http://www.skessuhorn.is/frettir/nr/67016/>. (1. okt. 2013).
- Sýslulýsingar 1744–1749*. 1957. Sögurit 28. Reykjavík: Sögufélag.
- Wadström, Roger. 1983. *Ortnamn i Bohuslän*. Stockholm: AWE/Geber.

Lykilord

örnefni, orðsifjar, stöðuvatn

Keywords

place names, etymology, lake

Abstract

The topic of the article is the origin of the toponym *Þveit*, a name of a lake in South-Eastern Iceland but otherwise unknown as a place name in Iceland. The corresponding morpheme is well known in Scandinavia and Denmark, both by itself and as the last part of place names, esp. in the names of (small) farms, and it occurs in place names in earlier viking settlements in Britain and France. The author discusses the meanings of the word *þveit* and its cognates in modern Scandinavian, the age and distribution of place names derived from it, and the possible origin of the Icelandic name. Her conclusion is that the name may originally derive from the habit of cutting

a hole in the ice for fishing, as one meaning of *þveit*, known from Norwegian dialects, is 'a channel cut into ice'.

Jónína Hafsteinsdóttir
Stofnun Árna Magnússonar í íslenskum fræðum
Nafnfræðisvið
ninahafst@vortex.is

Ritfregnir

Rit um málstöðlun

Kristiansen, Tore & Nikolas Coupland (ritstj.). *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus Press. 2011. (239 bls.) ISBN 978-82-7099-659-9.

Kristiansen, Tore & Stefan Grondelaers (ritstj.). *Language (de)standardisation in Late Modern Europe: Experimental Studies*. Oslo: Novus Press. 2013. (404 bls.) ISBN 978-82-7099-741-1.

Árið 2009 var svonefndri SLICE-áætlun hleypt af stokkunum með tveimur vinnustofum í Kaupmannahöfn. SLICE stendur fyrir „Standard Language Ideology in Contemporary Europe“ og felur í sér víðtækt rannsóknarsamstarf milli evrópskra fræðimanna. Sjónum er beint að þróun staðalmála og breytingum í málstöðlun í nútímanum, einkum m.t.t. hugmyndafræðilegra forsendna hennar. Frumkvæðið að áætluninni kemur frá Kaupmannahöfn en fræðimenn og rannsóknarstofnanir í fjölmörgum Evrópulöndum eiga aðild að henni (sjá nánar: <http://lanchart.hum.ku.dk/slice/>). Bækurnar tvær, sem báðar eru greinasöfn, eru fyrsti afrakstur samvinnunnar og þær eru m.a. forvitnilegar í tengslum við þessa þetta heftis af Orði og tungu.

Í fyrri bókinni er lagður fræðilegur grundvöllur að áætlunum rannsóknum innan vébanda SLICE-áætlunarinnar. Hún skiptist í þrjá hluta. Fyrst er inngangur eftir ritstjórana, Tore Kristiansen prófessor í Kaupmannahöfn og Nikolas Coupland prófessor í Cardiff, þar sem bæði er fjallað ítarlega um fræðilegan ramma rannsókna og ýmis grundvallarhugtök sem snerta málstöðlun, hugmyndafræði og málviðhorf og gerð grein fyrir efni og efnisSKIPUN bókarinnar. Því næst eru þrettán greinar sem hver um sig fjallar um stöðu mála í tilteknu málsamfélagi („Community reports“), þ.á m. grein um íslensku, „Language ideology and standardisation in Iceland“, eftir Stephan Pax Leonard og Kristján Árnason; önnur málsamfélög sem fjallað er um eru Austurríki, Danmörk, England og Wales (enska), Finnland (greinar bæði um finnsku og sænsku), Þýskaland, Írland (írsku), Litháen, Holland, Noregur, Svíþjóð og Wales (velska). Í síðasta hluta bókarinnar eru fimm greinar undir yfirskriftinni „Theoretical contributions“ þar sem fjallað er um tiltekin fræðileg viðfangsefni sem snerta rannsóknaráætlunina með dæmum

úr ýmsum tungumálum. Þar er m.a. fjallað um fjölmiðlamál og áhrif þess á þróun staðalmála, bæði mál hefðbundinna fjölmiðla eins og útvarps og sjónvarps og málnotkun í nýjum rafrænum miðlum, og um eðli breytinga á (opinberri) málnotkun á síðustu árum og áratugum, hvort þær felist heldur í því að viðmiðin sem staðalmálið byggir á hafi tekið breytingum (demotisation) eða í því að staðalmál séu á undanhaldi (destandardisation).

Í síðari bókinni eru kynntar fyrstu niðurstöður rannsókna innan ramma SLICE-áætlunarinnar. Eins og í fyrri bókinni skrifa ritstjórnir, sem í þessu tilviki eru Tore Kristiansen og Stefan Grondelaers prófessor í Nijmegen, ítarlegan inngang þar sem fjallað er um rannsóknaráætlunina og fræðilegan ramma hennar og gerð grein fyrir efni ritsins og skipulagi þess. Meginviðfangsefnið eru rannsóknir á málviðhorfum sem byggjast á tilraunum og prófum og sækja m.a. fyrirmyndir til langrar rannsóknarhefðar á því sviði í Danmörku. Í innganginum er því gerð allítarleg grein fyrir dönsku rannsóknunum, aðferðum sem þar hefur verið beitt og helstu niðurstöðum þeirra. Greinunum er skipað í tvo flokka. Í þeim fyrri, sem hefur yfirskriftina „(De)standardisation studies using Speaker Evaluation Experiments“, eru átta greinar sem fjalla um athuganir á máltilbrigðum þar sem athygli þátttakenda er beint að málnotendum fremur en málnotkuninni sem slíkri og áhersla er lögð á talað mál. Í síðari hlutanum eru sjö greinar undir yfirskriftinni „Methodological concerns and alternative approaches“. Þar á meðal er grein um mat á rituðum íslenskum textum, „Evaluation of different registers in Icelandic written media“, eftir Ara Pál Kristinsson og Amöndu Hilmarsson-Dunn. Einnig eru í þessum hluta greinar um aðferðafræði og fræðilegan grundvöll athugana á málviðhorfum og málhugmyndafræði frá ýmsum sjónarhornum.

Rit um málstefnu

Knudsen, Karin Jóhanna L., Hjalmar P. Petersen & Kári á Rógvi (ritstj.). *4 or more languages for all: Language policy challenges of the future*. Oslo: Novus forlag. 2012. (153 bls.) ISBN 978-82-7099-715-2.

Ritið er safn átta greina sem byggjast á úrvali fyrirlestra af þverfaglegri ráðstefnu sem haldin var í Færeyjum sumarið 2011 undir yfirskriftinni „4 or more languages for all: Language policy challenges of the future“. Höfundarnir koma úr ólíkum áttum, bæði landfræðilega og faglega, því meðal þeirra eru sérfræðingar á sviði málfræði, sagnfræði, bókmenntafræði, félagsfræði og lögfræði. Greinarnar spanna því vítt svið og höfundar nálgast viðfangsefnið frá mismunandi sjónarhorni. Sem dæmi um ólíka nálgun má nefna sögulega grein um fjöltyngi í Danaveldi fyrir daga Þjóðríkisins eftir sagnfræðinginn Steen Bo Frandsen, grein Malan Marnersdóttur um þýðingar á verkum Williams Heinesens á færeysku, þau málviðhorf sem þar birtast og áhrif þeirra á textann, og greinina „Who is `one of us'“ eftir Dimitry Kochenov, prófessor

í Evrópurétti, sem fjallar um sambandið milli ríkisborgararéttar, tungumáls og minnihlutasamfélaga víðs vegar í Evrópu. Í tveimur greinum er vísað beint til málstefnu Norðurlanda frá 2006 (http://www.norden.org/is/utgafa/utgefing-efni/2007-746/at_download/publicationfile) þar sem m.a. er kveðið á um samhliða tungumálanotkun, bæði hvað varðar sambúð opinbers máls og minnihlutamála í einstökum ríkjum og sambúð þjóðtungnanna og ensku. Annars vegar fjallar Anne Holmen um efnið frá sjónarhóli minnihlutahópa í Danmörku í greininni „Does the Nordic concept of parallel language use apply to minority children in Denmark?“. Hins vegar fjallar Birna Arnbjörnsdóttir um samband þjóðtungunnar, dönsku og ensku á Íslandi og í Færeyjum í greininni „Four Languages in a Changing Nordic Linguistic Environment“. Hún beinir einkum sjónum að því hversu gott vald íslenskir og færeyskir háskólanemar hafa á dönsku og ensku, hvaða áhrif það hefur á nám þeirra að þurfa að nota önnur mál en móðurmálið í námi sínu og hvernig megi auka færni þeirra í því samhengi. Þessar greinar, einkum grein Birnu, eru áhugaverðar í samanburði við grein Ara Páls Kristinssonar og Haraldar Bernharðssonar í þessu hefti þar sem þeir fjalla um notkun íslensku og ensku í háskólum á Íslandi í ljósi málstefnu Norðurlanda en þeir nálgast efnið úr svolítið annarri átt.

Sigurður Jónsson, Christer Laurén, Johan Myking & Heribert Picht.
Parallellspråk og domene / Parallelsprog og domæne / Parallellspråk och domän / Samhliða mál og umdæmi / Rinnakkaiskieli ja domeeni. Oslo: Novus forlag, 2013. (196 bls.) ISBN 978-82-7099-726-8.

Í bókinni er umræða um málstefnu Norðurlanda, bæði í heild og innan einstakra ríkja eða svæða. Sérstök áhersla er lögð á hugtökin „samhliða mál“ (d. *parallelsprog*) og „umdæmi“ (d. *domæne*), merkingu þeirra og notkun, einkum í tengslum við sérfræðimál, t.d. í háskólum. Í því sambandi er m.a. gerð grein fyrir stöðunni í hverju málsamfélagi fyrir sig. Höfundarnir eru norrænir iðorðafraeðingar.

Ný rit í nafnfræði

Talsvert kom út um norræna nafnfræði á árinu 2013 eins og fyrri ár, ekki síst í tímaritum en einnig í útgáfum *Norrænu nafnfræðisamtakanna* (NORNA, Nordiska samarbetskommittén för namnforskning). Á árinu 2013 komu út NORNA-rapporter 89 og 90 og hafa að meðaltali komið út um tvö hefti á ári síðan samtökin voru stofnuð 1971. Í NORNA-rapporter eru einkum gefin út greinasöfn sem tengjast ráðstefnum og þingum samtakanna.

Nøvn i strandamentanini/Nøvne i kystkulturen. Forelæsninger fra det 41. NORNA-symposium i Tórshavn 2.-4. juni 2011. Redaktion Tina K. Jakobsen, Kristin Magnussen, Anfinnur Johansen og Eivind Weyhe.

(NORNA-rapporter 89.) Uppsala: NORNA-förlaget. 2013. (380 bls.) ISBN 978-917-276-088-2.

Í ritinu eru 19 greinar sem tengjast nöfnum og strandmenningu á einn eða annan hátt. Greinarnar fjalla bæði um örnefni og mannanöfn í strandmenningunni á Norðurlöndum, frá Finnlandi í austri til Íslands í vestri en nær líka allt suður til Hollands. Fjallað er bæði um nútímanöfn og nöfn frá miðöldum og síðmiðöldum. Við sögu koma m.a. nöfn á seglskútum við Finnlandsstrendur (Anita Schybergson), nöfn á veiðistöðum við Ísland (Svavar Sigmundsson) og nöfn sem koma fyrir í Knýtlinga sögu (Alexandra Petrulich). Guðrún Kvaran á greinina „Íslandske person- og gadenavne tilknyttet havet“. Þar er sagt frá eiginnöfnum fólks eins og *Ægir* og *Unnur* og samsetningum með *Haf*, *Mar*- og *Sæ*- og einnig frá götum sem draga heiti sitt af sjónum. Þá fjallar Peder Gammeltoft um stöðu örnefnarannsókna í nútímanum og skyggnist inn í framtíðina. Ritinu lýkur með samantekt yfir umfjöllunarefnið eftir Mats Wahlberg. Greinar eru á dönsku, norsku og sænsku með útdráttum á ensku eða þýsku. Ritið kallast á við NORNA-rapporter 41 frá 1989 sem ber titilinn *Stadnamn i kystkulturen*.

Namn i stadsmiljö. Handlingar från NORNA:s 42 symposium i Helsingfors den 10–12 november 2011. Redigerade av Leila Mattfolk, Maria Vidberg och Pamela Gustavsson. (NORNA-rapporter 90.) Helsingfors: Institutet för de inhemska språken og NORNA-förlaget. 2013. (270 bls.) ISBN 978-952-544-682-1.

Í ritinu eru sextán greinar eftir jafnmarga höfunda. Greinarnar fjalla á einhvern hátt um nöfn í borgum og öðru þéttbýli, hvort sem um er að ræða nöfn á stöðum eða fólki. Þannig er þarna að finna greinar sem fjalla sérstaklega um rannsóknasviðið þéttbýlisnöfn (*urban names, urbane namn*) (Gunnstein Akselberg), um slangur í Helsinki (Mona Forsskåhl), um mannanöfn í sænskum bæjum á 17. öld (Katharina Leibring og Kristina Neumüller), um hvernig þéttbýlisnöfn eru notuð í auglýsingaskyni (Staffan Nyström) og um óformleg nöfn sem verða til meðal barna í þéttbýli í Gautaborg, m.a. í fleirtyngdu umhverfi (Maria Löfdahl og Sofia Tingsell). Þá fjallar Marianne Blomqvist um þær leifar sem kýr hafa skilið eftir í örnefnum í strandhéruðum Finnlands frá miðöldum til nútíma. Greinarnar eru á norsku og sænsku en með enskum útdráttum.

Upplýsingar um hvernig hægt er að panta ritin er að finna á heimasíðunni www.norna.org.

Ainiala, Terhi, Minna Saarelma og Paula Sjöblom. *Names in Focus. An Introduction to Finnish Onomastics.* Þýð.: Leonard Pearl. (Studia Fennica Linguistica 17.) Helsinki: Suomalaisen Kirjallisuuden Seura. 2012. (287 bls.) ISBN 978-952-222-387-6.

Þriðja ritið sem hér verður nefnt ber titilinn *Names in Focus. An Introduction to Finnish Onomastics*. Það kom upphaflega út á finnsku árið 2008. Höfundarnir

eru Terhi Ainiala, Minna Saarelma og Paula Sjöblom, allar háskólakennarar í nafnfræði í Finnlandi. Skrifað er um efnið frá finnskum sjónarhóli en skírskotunin er miklu víðari en svo og má þetta teljast ágætt inngangsrít að fræðunum. Sambærileg rit eru til á hinum skandinavísku málunum en þau eru flest komin talsvert til ára sinna. Sjö aðalkaflar eru í bókinni. Fyrsti kafli er um fræðilegan bakgrunn nafnfræðinnar og kenningar. Fjallað er um skilgreininguna á nöfnum annars vegar og orðum hins vegar en það er síður en svo einfalt mál. Annar kafli er um finnskar heimildir og finnska rannsóknasögu. Aðalkaflarnir tveir eru um örnefni (kafli 3) og manna-nöfn (kafli 4). Í örnefnakaflanum er fjallað um málfræði örnefna og merk-ingarfræði, um orðmyndun meðal örnefna, orðsifjafraði, formleg nöfn og óformleg og nöfn í þéttbýli. Fróðlegur er kafli um mun á nafnasafni (*onomastikon*) barna og fullorðinna. Kaflinn um mannanöfn hefst á inngangi um mannanöfn, um tengsl nafna og menningar og um mismunandi nafnakerfi. Fjallað er sérstaklega um nafnakerfi í Evrópu, um finnska mannanafna-fræði, um viðurnefni og gælunöfn og nöfn minnihlutahópa. Kaflar 5–7 eru um dýranöfn, nöfn úr viðskiptalífínu og nöfn í bókmenntum. Af þessum þremur síðastnefndu er 6. kafli, um viðskiptanöfn, ekki síst áhugaverður. Þar er fjallað er um nöfn á fyrirtækjum, vörum, vörumerkjum og öðru því sem tengist atvinnulífi og auglýsingastarfsemi í nútímanum (e. *commercial names*). Þessum flokki nafna hefur ekki mikið verið sinnt á Íslandi og þar liggur mikið efni órannsakað. Ritaskrá og atriðisorðaskrá fylgja með. Utgefandi er Finnish Literature Society í Helsinki (www.finlit.fi).

Hallgrímur J. Ámundason
hja@hi.is

Ný tíðniorðabók

Quasthoff, Uwe, Sabine Fiedler & Erla Hallsteinsdóttir (ritstj.). *Frequency Dictionary Icelandic. Íslensk tíðniorðabók. ISL.* (Frequency Dictionaries 3.) Leipzig: Leipziger Universitätsverlag. 2012. (x + 109 bls. ásamt geisladiski.) ISBN 978-3-86583-656-4.

Þetta verk lætur ekki mikið yfir sér við fyrstu sýn, ríflega hundrað síður í fremur litlu broti. Prentaði hlutinn geymir þó ekki nema brot af sjálfri tíðniorðabókinni því sjálf er hún að stærstum hluta á meðfylgjandi geisladiski. Í bókinni er inngangur að verkinu þar sem gerð er grein fyrir efniviðnum og úrvinnslu hans og birtar ýmsar áhugaverðar tölfræðilegar upplýsingar um orðaforðann, t.d. um tíðni bókstafa og atkvæðagerðir, lengd orðmynda, samspil tíðni og lengdar o.fl. Þar er auk þess skrá yfir þúsund algengustu orðmyndirnar í tíðniröð og önnur yfir 10 þúsund algengustu orðmyndir í

stafrófsröð. Meginhluti sjálfrar tíðniorðabókarinnar er svo á geisladiskinum, skrá yfir eina milljón orðmynda sem eru þær algengustu í málheildinni sem liggur að baki ásamt tíðniupplýsingum um þær. Skráin er annars vegar birt í stafrófsröð á pdf-formi (í þremur hlutum) og hins vegar í þremur textaskjölum þar sem efninu er í fyrsta lagi raðað í stafrófsröð, í öðru lagi í tíðniröð og í þriðja lagi í öfugri stafrófsröð. Rafræna efnið má nota á margvíslegan hátt með svokölluðu „creative commons“ leyfi (CC-BY) sem heimilar notendum að hlaða efninu niður og vinna frekar úr því.

Orðmyndirnar og tíðniupplýsingarnar byggjast á risastóru textasafni með tæplega 34 milljónum setninga og 545 milljónum lesmálsaða. Textarnir eru frá tímabilinu 2001–2011 og þeir eru að stærstum hluta fengnir úr íslensku vefsíðusafni Landsbókasafns – Háskólabókasafns og úr íslenskri útgáfu *Wikipedia* en einnig er í safninu allmikið efni úr dagblöðum. Efnið endurspeglar því íslenska nytjatexta í upphafi 21. aldar, einkum ritmálið eins og það birtist í vefmiðlum. Öll úrvinnsla efnisins var rafræn og er aðferðunum lýst í inngangi verksins. Einungis sá hluti efnisins sem birtist í prentaða hlutanum, þ.e.a.s. skrár um 10 þúsund algengustu orðmyndirnar, var yfirfarinn handvirkt. Tíðniupplýsingarnar snúa því að orðmyndum fremur en orðum. Mismunandi myndir sama orðs, hvort sem þær tengjast ólíkum rithætti eða beygingu, eru ekki tengdar saman og orðmyndum með stórum og litlum upphafsstaf er haldið aðgreindum. Þannig birtist t.d. orðið *adrenalín* í þrettán myndum: *Adrenalin*, *adrenalin*, *Adrenaline*, *adrenaline*, *adrenalín*, *Adrenalín*, *adrenalíni*, *Adrenalíni*, *adrenalínið*, *Adrenalínið*, *adrenalíninu*, *adrenalíns* og *adrenalínsins*; í stafrófsröðuðu listunum birtast myndirnar þó yfirleitt í röð eða a.m.k. í námunda hver við aðra svo notendur hafa yfirsýn yfir þær. Einstakar orðmyndir geta sömuleiðis verið tví- eða margræðar, t.d. *Gestur* (sérafafn eða samnafn (í upphafi setningar)), *hafi* (so. *hefur* eða no. *haf*) o.s.frv. Orðabókin er systurverkefni *Íslensks orðasjóðs* (http://wortschatz.uni-leipzig.de/ws_isl/) sem byggir á sama textagrunni og með því að nota hann samhliða má fá hugmynd um dreifingu og að nokkru leyti merkingu orðmyndanna í *Íslenskri tíðniorðabók*.

Íðorð og sérfræðimál

Tölvuorðasafn. Íslenskt-enskt, enskt-íslenskt. 5. útgáfa, aukin og endurbætt. Orðanefnd Skýrslutæknifélags Íslands tók saman. Reykjavík: Orðanefnd Skýrslutæknifélags Íslands. 2013. (570 bls.) Rafræn útgáfa (pdf: http://tos.sky.is/site_media/TO_utgafa5.pdf) og vefútgáfa (<http://tos.sky.is/>).

Nýlega kom út fimmta útgáfa af *Tölvuorðasafni*. Fyrsta útgáfa þess er frá 1983 og síðan kom það út 1986, 1998 og 2005, aukið og endurbætt með hverri nýrri útgáfu. Í formála Sigrúnar Helgadóttur, formanns orðanefndarinnar, er gerð ítarleg grein fyrir efni útgáfunnar. Þar kemur m.a. fram að hugtökum hafi

fjölgað um tæplega 4% frá síðustu útgáfu og samkvæmt því eru í henni u.þ.b. átta þúsund hugtök. Auk þess hafi eldra efni verið yfirfarið og endurskoðað. Nýja útgáfan er einungis birt rafrænt á vef Skýrslutæknifélagsins. Þar má bæði leita að íslenskum og enskum heitum í vefútgáfu safnsins og sækja orðasafnið í heild sem pdf-skjal til útprentunar eða notkunar í eigin tölvu. Einnig má sækja orðasafnið á sérstöku sniði (TBX 'Term Base eXchange') sem hentar fyrir þýðingarminni og til nota í máltækniiverkefnum.

Orðasafn í líffærafræði. I. Stoðkerfi. Enska – latína – íslenska. Ritstj.: Jóhann Heiðar Jóhannsson. (Íðorðarit Stofnunar Árna Magnússonar í íslenskum fræðum 3.) Reykjavík: Stofnun Árna Magnússonar í íslenskum fræðum. 2013. (56 bls.; rafræn útgáfa (pdf: <http://www.ordabanki.hi.is/files/stodkerfi.pdf>). ISBN 978-9979-654-278.

Þetta er fyrsti hluti nýs orðasafns í læknisfræði. Það byggir að hluta á *Íðorðasafni lækna* sem kom út í heftum á árunum 1986–1989 og á alþjóðlegu líffæraheitunum, *Terminologia Anatomica*. Orðasafnið er skipulagt þannig að fyrst kemur flokkaður orðalisti með enskum uppflettiorðum og jafnheitum á latínu og íslensku ásamt stuttri skilgreiningu og er það meginhluti verksins. Á eftir fylgja svo tveir stafrófsraðaðir orðalistar, annar með enskum uppflettiorðum og íslenskum jafnheitum og hinn með íslenskum uppflettiorðum og jafnheitum þeirra á ensku.

Orðasafnið er gefið út í prentuðu kveri en það má einnig sækja sem pdf-skjal á vef Orðabanka Íslenskrar málstöðvar (<http://www.ordabanki.hi.is>). Orðasafnið er þriðja ritið í nýrri ritröð, Íðorðarit Stofnunar Árna Magnússonar í íslenskum fræðum. Hin fyrri, sem bæði komu út 2012, eru orðasöfn í stjórnmálafræði og faraldsfræði.

Ari Páll Kristinsson & Jóhannes B. Sigtryggsson (ritstj.). *Juridisk sprog i Norden*. Rapport fra en nordisk konference om klarsprog, Reykjavík 11.–12. oktober 2011. Nordisk sprogkoordination. 2012. (111 bls.) Rafræn útgáfa: http://nordisksprogkoordination.org/dokumenter-til-download/Juridisk_sprog_i_norden.pdf. ISBN: 978-87-89410-50-0.

Þetta er ráðstefnurit frá 6. norrænu ráðstefnunni um málskýrð sem haldin var í Reykjavík haustið 2011. Þema hennar var lögfræðimál á Norðurlöndum með áherslu á þýðingar lögfræðilegra texta og margmála lagatexta. Í ritinu eru birtar átta greinar sem byggjast á fyrirlestrum af ráðstefnunni auk útdráttu úr öllum 13 fyrirlestrunum sem haldnir voru og fjórum veggspjöldum sem þar voru kynnt. Útdráttunum fylgja stutt ágríp af ferli fyrirlesaranna. Ritið er einungis gefið út rafrænt á vef Skrifstofu um norrænt tungumálasamstarf (d. *Nordisk sprogkoordination*).

Nýtt afmælisrit

Guðrún Kvaran. *Glíman við orðin*. Ritgerðasafn gefið út í tilefni sjö-tugsafmælis höfundar 21. júlí 2013. Reykjavík: Stofnun Árna Magnússonar í íslenskum fræðum. 2013. (xiv + 374 bls.) ISBN 978-9979-654-25-4.

Í ritinu er safnað saman úrvali greina eftir Guðrúnu Kvaran sem gefa yfirlit yfir viðfangsefni hennar og skrif á löngum starfsferli. Greinarnar eru alls níttján og þeim er skipað í fimm flokka eftir efni: Orðabókafræði, nafnfræði, orðfræði og orðsifjar, Biblíumál og loks málfræðiskrif fyrri alda. Sem dæmi um greinar í ritinu má nefna „Andans kona og orðabókarpúl“ sem fjallar um orðabókastörf Bjargar C. Þorláksdóttur og framlag hennar til *Íslensk-danskrar orðabókar* Sigfúsar Blöndals, „Söfnun Þórbergs Þórðarsonar úr mæltu máli“ sem fjallar, eins og nafnið gefur til kynna, um orðasöfnun hans og yfirlitsgrein um íslenskar Biblíuþýðingar, „Almúganum til sæmdar og sáluhjálpar. Um íslenskar Biblíuþýðingar“. Allar hafa greinarnar birst áður, í tímaritum, ráðstefnuritum eða afmælisritum, og þær komu upphaflega út á árunum 1990–2011. Tvær greinar um nafnfræði eru skrifaðar á dönsku en aðrar eru á íslensku. Framan við greinarnar er, auk efnisyfirlits og heilla-óskaskrár (Tabula gratulatoria), formáli ritnefndar, þar sem ferill Guðrúnar er rakinn í stuttu máli og gerð grein fyrir efni og efnisskipun í ritinu. Í bók-arlok er birt ítarleg ritaskrá höfundar þar sem talið er allt sem birst hefur eftir Guðrúnu á prenti, frá árinu 1967 til 2013. Ritnefnd bókarinnar skipuðu þau Gunnlaugur Ingólfsson, Jóhannes B. Sigtryggsson og Þóra Björk Hjartardóttir.

Málvísindi og íslensk málfræði

Höskuldur Þráinsson & Matthew Whelpton (ritstj.). *Chomsky. Mál, sál og samfélag*. Reykjavík: Hugvísindastofnun og Háskólaútgáfan. (385 bls.) ISBN 978-9935-23-016-4.

Í tilefni af heimsókn Noams Chomskys á 100 ára afmæli Háskóla Íslands 2011 var stofnað til þverfræðilegs meistaranámskeiðs þar sem fyrirlesarar voru úr ýmsum deildum Háskóla Íslands auk gestafyrirlesara frá Háskólanum á Akureyri og Háskólanum á Bifröst. Fyrirlestrararnir fjölluðu um framlag Chomskys til vísinda, samfélagsmála og stjórnmálaumræðu og sjálfur hélt hann sérstakan, opinn fyrirlestur um fræðilega afstöðu sína í málvísindum í tengslum við námskeiðið auk opinbers hátíðarfyrirlesturs um stjórnmál og samfélag. Báðir fyrirlestrar Chomskys eru birtir í ritinu, hvor um sig sem inngangur að hinum tveimur hlutum verksins: „Maður, mál og sál“ og „Maður og samfélag“. Í fyrri hlutanum eru auk greinar Chomskys, „The ‘Generative Enterprise’: Its Origins, Goals, and Prospects“, ellefu greinar sem

byggjast á fyrirlestrum úr námskeiðinu. Þær eru flestar um málvísindaleg efni en einnig á sviði sálfræði og félagsvísinda. Í heild veita þessar greinar ágætt yfirlit yfir hugmyndir og kenningar Chomskys á sviði hugvísinda og áhrif þeirra á rannsóknir á Íslandi, einkum á íslensku máli og málfræði, og í sumum þeirra er einnig vikið að öðrum kenningum í hugvísindum og þær bornar saman við hugmyndir Chomskys. Áhrif hans birtast á margvíslegan hátt, þ.á m. í vali á viðfangsefnum, sjónarhorni og fræðilegum grundvelli rannsókna. Síðari hlutinn hefst á grein Chomskys, „The Two 9/11s: Their Historical Significance“. Titillinn vísar annars vegar til valdaránsins í Chile 1973 og hins vegar til árásarinnar á tvíburaturnana í New York 2001 en báðir atburðirnir áttu sér stað 9. september. Auk þess eru í þessum hluta fjórar greinar sem hver á sinn hátt fjalla um samfélags- og stjórnmálaumræðu og áhrif Chomskys á því sviði. Flestar greinarnar í ritinu eru á íslensku með útdrætti á ensku en greinar Chomskys og ein grein til viðbótar í hvorum hluta eru birtar á ensku með útdrætti á íslensku. Auk þess er inngangur ritstjóra birtur á báðum málunum.

Höskuldur Þráinsson, Ásgrímur Angantýsson & Einar Freyr Sigurðsson (ritstj.). *Tilbrigði í íslenskri setningagerð. I. Markmið, aðferðir og efniviður*. Reykjavík: Málvísindastofnun Háskóla Íslands. 2013. (127 bls.) ISBN 978-9979-853.

Í ritinu er gerð grein fyrir rannsóknarverkefninu „Tilbrigði í íslenskri setningagerð“ og niðurstöðum þess og er þetta fyrsta bindi af þremur. Verkefnið naut öndvegisstyrks frá Rannsóknasjóði 2005–2007 og fól í sér viðamikla rannsókn á setningafræðilegum tilbrigðum í nútímamáli, einkum með þremur skriflegum yfirlitskönnunum en jafnframt með athugunum sem byggðust á textagögnum úr talmáli og óformlegu ritmáli. Ritið skiptist í fimm kafla sem fjalla fyrst og fremst um markmið, aðferðir og efnivið í rannsókninni í heild en í síðustu tveimur köflunum er auk þess gerður samanburður á niðurstöðum úr ólíkum hlutum rannsóknarinnar. Í seinni bindunum tveimur, sem væntanleg eru á næstunni (efnisyfirlit þeirra er birt með efni fyrsta bindis), verður svo gerð nánari grein fyrir niðurstöðunum í heild (2. bindi) og sérathugunum sem byggja á niðurstöðum um einstök setningafræðileg fyrirbæri (3. bindi).

Ráðstefnur 2013

Málþing um tungumál og alþjóðavæðingu í háskólastarfi

Málþing um tungumál og alþjóðavæðingu í háskólastarfi fór fram í sal Þjóðminjasafnsins í Reykjavík dagana 18.–20. júlí 2013. Þingið fór að mestu fram á ensku undir yfirskriftinni *Languages and Internationalization in Higher Education: Ideologies, Practices, Alternatives*. Um 40 sérfræðingar víða að tóku þátt í þinginu.

Málþingið var einkum helgað kostum og göllum þeirrar þróunar að enska verður stöðugt útbreiddari sem kennslutungumál á háskólastigi og sem tungumál fræðilegrar útgáfu hvarvetna í heiminum. Norðurlöndin voru í brennidepli en auk þess komu fyrirlesarar úr mjög mörgum áttum, s.s. frá Aserbaídsjan, Indlandi, Írlandi, Kanada, Kasakstan, Kýpur, Lúxemborg, Mexíkó, Spáni og Þýskalandi. Sem fyrr segir var sérstök áhersla á umfjöllun um Norðurlönd og norræna háskóla enda var málþingið sérstaklega tengt undirbúningi bókarinnar *English at Nordic universities: ideologies and practices* (ritstj. Anna Kristina Hultgren, Frans Gregersen og Jacob Thøgersen) sem er væntanleg frá John Benjamins 2014 í ritröðinni *Studies in World Language Problems*. Meðal frummælenda var Ari Páll Kristinsson sem kynnti væntanlegan bókarkafila um hugmyndafræðilega þætti íslenskrar málverndar og samspil þeirra við dulda og ljósa málstefnu íslenskra háskóla.

Málþingið var haldið að frumkvæði The Center for Research and Documentation on World Language Problems, í Rotterdam, sem einnig annaðist fjármögnunina og naut framlaga frá Esperantic Studies Foundation og frá John Benjamins Publishing Company. Samstarfsaðilar á Íslandi voru Stofnun Árna Magnússonar í íslenskum fræðum, Íslensk málnefnd og Stofnun Vigdísar Finnbogadóttur í erlendum tungumálum. Ari Páll Kristinsson kynnti fyrir málþingsgestum starfsemi Árnastofnunar og Guðrún Kvaran starf Íslenskrar málnefndar.

The Center for Research and Documentation on World Language Problems stendur af og til fyrir alþjóðlegum málþingum af þessu tagi um málstefnu og málnotkun í alþjóðlegu samhengi. Málþingaröðin er kennd við japanska menntafrömuðinn og diplómatann Inazo Nitobe (1862–1933).

Ari Páll Kristinsson
aripk@hi.is

Tólfta ráðstefna Norræna orðabókarfræðifélagsins í Osló 13.–16. ágúst 2013

Ráðstefna Norrænna orðabókarfræðifélagsins (Nordisk forening for leksikografi), hin 12. í röðinni, var haldin í Osló dagana 13.–16. ágúst 2013. Framkvæmd ráðstefnunnar var í höndum norska málráðsins, Språkrådet, og Institutt for lingvистiske og nordiske studier við Oslóarháskóla. Ráðstefnan var haldin í nýjum og glæsilegum húsakynnum vísindagarða háskólans, Forskningsparken. Þáttakendur voru um 90 talsins frá öllum Norðurlöndunum.

Ráðstefnan fór fram í tveimur málstofum og alls voru flutt um 40 erindi. Efni erindanna var margvíslegt og varðaði bæði fræðilega og hagnýta hlið orðabóka og orðabókargerðar, m.a. efnisafmörkun og val flettiorða, gerð og notagildi málheilda í orðabókarlegu samhengi, gagnagrunna sem greiningarumhverfi og nýjungar í gerð stafrænna orðabóka. Þá var gerð grein fyrir ýmsum verkefnum á veggspjöldum, þ. á m. var lýsing á ISLEX sem Halldóra Jónsdóttir og Þórdís Úlfarsdóttir stóðu að. Norsk orðabókargerð var að vonum fyrirferðarmikil, en nú fer senn að ljúka hinu stóra og mikla orðabókarverki um nýnorsku og norskar mállýskur, *Norsk ordbok*, og verið er að móta yfirgripsmikið verk um bókmálið. Aðalfyrirlesarar voru Arleta Adamska-Sałaciak frá Háskólanum í Poznań í Póllandi en erindi hennar bar yfirskriftina „Sameness of meaning in a bilingual dictionary“, Lars Trap-Jensen frá Det Danske Sprog og Litteraturselskab í Kaupmannahöfn en hans erindi nefndist „Leksikografisk tradition og fornyelse: digital revolution eller organisk reformisme?“ og Victoria Rosén frá Háskólanum í Bergen sem nefndi erindi sitt „Samspillet mellom leksikon og syntaks“.

Þátttakendur frá Íslandi voru venju fremur fáir að þessu sinni. Framlag til dagskrárinnar af Íslands hálfu var auk kynningar Halldóru og Þórdísar erindi Jóns Hilmars Jónssonar sem hann nefndi „Semantisk nærhet og synonymi i parataktisk og paradigmatiske perspektiv“. Þá ber að nefna að Anna Helga Hannesdóttir frá Háskólanum í Gautaborg flutti erindi um val og hlutverk notkunardæma í ISLEX.

Aðalfundur orðabókarfræðifélagsins var haldinn í lok ráðstefnunnar þar sem kosin var ný stjórn félagsins. Nýkjörinn formaður, Henrik Hovmark frá Nordisk Forskningsinstitut við Háskólann í Kaupmannahöfn, boðaði síðan til næstu ráðstefnu félagsins í Kaupmannahöfn vorið 2015.

Osló skartaði sínu fegursta þessa sumardaga og skipuleggjendur lögðu sig fram um að gera ráðstefnugestum dvölinna sem eftirminnilegasta.

Jón Hilmar Jónsson
jhj@hi.is

Á döfinni

2014

Ráðstefnur um nafnfræði

Tvær ráðstefnur eru fyrirhugaðar á árinu á vegum NORNA. Fyrri ráðstefnan ber titilinn *Scandinavian Names and Naming in the medieval North Atlantic Area* og er 44. ráðstefna NORNA. Hún verður haldin í Caen í Frakklandi dagana 23.–25. apríl í samvinnu við Centre Franco-Norvégien en Sciences Sociales et Humaines, Office Franco-Norvégien d’Echanges et de Coopération og Centre de Recherche Inter-langues sur la Signification en Contexte (CRISCO).

Seinni ráðstefnan, sem verður 45. ráðstefna NORNA, hefur yfirskriftina *Navn og navnebærer*. Hún verður haldin á Skagen á Jótlandi dagana 1.–4. október 2014. Ráðstefnan er haldin í samvinnu við Københavns Universitet og Nordisk Forskningsinstitut.

Nánari upplýsingar eru á www.norna.org.

Alþjóðleg máltækniráðstefna á Íslandi

Níunda alþjóðlega LREC-ráðstefnan (*Language Resources and Evaluation Conference*) verður haldin í Reykjavík dagana 26.–31. maí 2014. Ráðstefnan sjálf stendur í þrjú daga (28.–30.5.) og aðrir þrjú dagar fyrir og eftir ráðstefnuna eru lagðir undir sérstakar málstofur. Ráðstefnan er á vegum ELRA-samtakanna (<http://www.elra.info/>) í samvinnu við Háskólann í Reykjavík, Háskóla Íslands, Stofnun Árna Magnússonar í íslenskum fræðum og Stofnun Vigdísar Finnbogadóttur í erlendum tungumálum.

Nánari upplýsingar eru á vefsíðunni <http://lrec2014.lrec-conf.org/en/about>.

Evrópuráðstefna um orðabókafræði

Sextánda alþjóðlega EURALEX-ráðstefnan verður haldin í Bolzano á Ítalíu dagana 15.–19. júlí 2014 undir yfirskriftinni *The User in Focus*. Á undan ráðstefnunni verður boðið upp á þrjú mismunandi vinnustofur. Ráðstefnur EURALEX eru haldnar annað hvert ár í nafni evrópsku orðabókafræðisamtakanna, European Association for Lexicography (<http://www.euralex.org/>), og draga jafnan að sér mikinn fjölda þátttakenda.

Nánari upplýsingar eru á vefsíðunni <http://euralex2014.eurac.edu/en/Pages/default.aspx>.

2015

Norræn ráðstefna um orðabókafræði

Prettánda ráðstefnan um norræn orðabókafræði verður haldin á vegum NFL í Kaupmannahöfn dagana 20.–22. maí 2015. Nýlega var auglýst eftir tillögum að erindum og skulu útdrættir berast fyrir 15. janúar á næsta ári. Tekið er við fyrirlestrum um hvaðeina sem tengist orðabókum og orðabókagerð en að þessu sinni verður lögð áhersla á notendur rafrænna orðabóka („Den digitale ordbogsbruger“).

Fylgjast má með fréttum af ráðstefnunni á vefsíðu norræna orðabókafræðifélagsins, NFL: <http://nordisksprogkoordination.org/nfl>.

Norræn íðorðaráðstefna

Nordterm 2015 verður dagana 9.–12. júní 2015 í Reykjavík. Norrænu íðorðasamtökin *Nordterm* gengast fyrir ráðstefnum, námskeiðum, vinnustofum og fundum um íðorð og íðorðastarf og eru *Nordterm*-dagarnir haldnir annað hvert ár, til skiptis í aðildarlöndunum fimm.

Nordterm 2015 hefst 9. júní með námskeiði um íðorðafræði og vinnustofu um hagnýt úrlausnarefni í íðorðastarfi. Dagana 10. og 11. júní verður ráðstefna um norræna og almenna íðorðafræði. Loks verður aðalfundur *Nordterm* 12. júní. Dagana á undan verða einnig fundir í stjórnarnefnd og vinnunefndum *Nordterm*-samtakanna.

Öllum sem áhuga hafa á íðorðamálum er heimil þátttaka. Ráðstefnuboð með nánari upplýsingum verður sent út haustið 2014, ásamt slóð á vefsíðu með upplýsingum um skráningu. Um starfsemi *Nordterm* og tengla á fyrirhugaða viðburði samtakanna, sjá veffangið <http://www.nordterm.net/>.