

Gottskálf Jensson

Ummyndanir klassískra orðasafna í rafrænum spunavélum samtímans

1 Langlíf orð dauðra tungumála

Klassískar bókmenntir, og þá á ég við latneskar og forngrískar, eru samkvæmt gamalli skilgreiningu þær bókmenntir sem staðist hafa tímans tönn. Þótt orðasöfn unnin upp úr klassískum textum séu vissulega forgengileg — já, í raun þegar forgengin er þau koma út — því langt er síðan latína og forngríska voru lifandi tungumál, mætti einnig og af sömu ástæðu kalla þau endanlegri og í þeim skilningi sígildari en mörg önnur orðasöfn enda bætist ekki við orðaforðann ef tungumálið er dautt. Auðvitað eru ekki öll klassísk orðasöfn fullkomlega ítarleg eða skýringar þeirra svo óskekikular að ekki megi gera betur auk þess sem það torveldar iðulega vinnslu þeirra að fjölmargir textar á fornmaalunum eru óaðgengilegir í handritum og hafa aldrei komið út á prenti. Mikið ritað efni á latínu og grísku leynist vissulega í handrita- og skjalasöfnum víða um heim og sumt felst áreiðanlega enn í jörðu og bíður fornleifauppgötvana framtíðarinnar.

Vinnu við nýjar latneskar og forngrískar orðabækur vindur hægt fram á okkar tímum en framfarir eru þó greinanlegar. Um það mun ég nefna dæmi síðar. En þrátt fyrir það má segja að ekki verði vart mikilla breytinga á hinu nánast stöðuga klassíska orðamengi. Gæði þess fílólógiska og lexíkógrafíska starfs, sem unnið hefur verið, segjum frá því á 18. öld, eru almennt mikil og verkin hafi staðist svo vel tímans

tönn að gamlar orðabækur á sviðinu koma enn að ágætum notum. Til dæmis má nefna *Nucleus Latinitatis* eða *Kleyfsa*, frá 1732, með íslenskum orðskýringum (endurútgefinn á vegum Orðabókar Háskólans 1994) sem ég nota mikið sjálfur og hef fregnað að aðrir noti nú í skorti á annarri og nýrri latnesk-íslenskri orðabók. *Kleyfsi* gagnast auðvitað sérstaklega þeim sem gera íslenskar þýðingar á latínutextum frá síðari öldum. Aðrar klassískar orðabækur hafa komið að góðum notum fyrir hið bandaríska verkefni á veraldarvefnum, *The Perseus Project*, sem er hýst hjá Tufts University í Boston, vegna þess að þær eru komnar úr höfundarrétti vegna aldurs þótt þær séu fullgildar enn. Hér má nefna hina ágætu latnesk-ensku orðabók Lewis and Short, *A Latin Dictionary* (Oxford 1879), og grísk-ensku orðabækurnar Liddell and Scott, *A Greek-English Lexicon* (Oxford 1940), og Georg Autenrieth, *A Homeric Dictionary* (New York 1891). Allar þrjár hafa verið skráðar í gagnagrunn og samtengdar forngrískum og latneskum textasöfnum með hjálp véla sem greina orðmyndir og finna jafnframt „flettiorð“ þeirra. Þannig má fá beygingarfræðilega greiningu hvers orðs í forngrískum og latneskum textum safnsins, „fletta upp“ enskri orðabókarskýringu þess, auk þess sem kalla má fram enskar þýðingar textans (flestar gamlar úr höfundarrétti) og birta athugasemdir úr skýringarritum. Perseus-verkefnið hófst árið 1987 og einskorðaðist útgáfan í upphafi við CD-ROM-diska. Frá árinu 1995 hefur Perseus-verkefnið verið á veraldarvefnum, *The Perseus Digital Library*, <www.perseus.tufts.edu>. Þar er safnað bæði latneskum og forngrískum textum en auk þess enskum endurreisnarritum og einhverju öðru efni. Eins og áður sagði nýtir Perseus sér þá staðreynd að fjölmargar textaútgáfur og handbækur í klassískum fræðum og fornámálum eru komnar úr höfundarrétti og orðnar almannaeign. Verkefnið sýnir hve vel gamlar prentaðar orðabækur liggja við þess háttar endurvinnslu í gagnaveitum á veraldarvefnum.

Orðaforðinn í klassískum orðabókum (lat. *copia uerborum, uocum*) — vart er hægt að ræða um efnið án þess að gera það í grísk-latneskum hugtökum — er sem sagt allstöðugt mengi en hið sama er ekki endilega hægt að segja um það form eða mynd (*forma, idea, morfe*) sem þetta efni tekur á sig, sérstaklega á tímum þegar stafrænar spunavélar vefa okkur hverfula vefi á rafrænum skjáflötum. Hér er allt í stanslausri ummyndun og engin formgerð orða- og textasafna stendur lengur við en í nokkur ár, á meðan einstakar síður eru spunnar á augnabliki sem

svör við fyrirspurnum notandans og verða því líklega aldrei birtar aftur með nákvæmlega sama hætti. Yfirborðið er hér allt. Hinar sjáanlegu myndbirtingar á skjánum eða í útprentinu. Hvað gerist í raun inni í vélinni þarf notandinn ekki að skilja nema að litlu leyti, ekki frekar en lesandinn þarf að vita nokkuð um samsetningu þess pappírs eða bleks sem myndar efnislegan veruleika prentaðrar blaðsíðu. Stafræna spunavélin vefur umfram allt texta (lat. *textus*) í sinni eiginlegu merkingu, það er að segja ‘vefnað’ eða ‘vef’ úr orðum.¹

Á þeim síðum, sem fylgja hér á eftir, ætla ég að freista þess að sýna með latnesk-íslensku dæmi hvers konar rannsóknir verða mögulegar með samtengingu texta og orðasafna í gegnum greiningarvélar orðmynda í textunum. Í framhaldi af því ræði ég fornt hugmyndaknippi sem með sérkennilegum hætti upplýsir og auðveldar skilning á viðtækum kennilegum afleiðingum hinnar nýju tækni og loks mun ég stuttlega velta vöngum yfir mögulegum verkefnum, einkum í tengslum við latínuskrif eftir Íslendinga og um íslenskt efni.

2 Líkamlegar ummyndanir ólíkamlegra orða (dæmi: *tímans tönn*)

Flakki hins forngríska og latneska orðamassa, sem svo lítið breytist sjálfur í þeim sífelldu ummyndunum klassískra orðasafna sem nú ber fyrir augu í netvæddum upplýsingaveitum og þá á ég einkum við hinar tvær stóru á sviði klassískra fræða, *The Perseus Project*, sem áður var nefnd, og *Thesaurus Linguae Graecae* <www.stephanus.tlg.uci.edu>, sem nánar mun kynnt síðar, ætla ég að freistast til þess að líkja við flæking eilífra sálna á milli síbreytilegra líkama samkvæmt ævaforðri kenningu sem rómverska skáldið Publius Ovidius Naso, eða Óvíð eins og hann heitir á dönsku umritaðri á íslensku, útlistaði í merku kvæði.

Í fimmtánda og síðustu bók sögulgjóðsins *Metamorphoses* (titillinn mætti þýða á íslensku sem *Ummyndanir* eða *Umformanir*) leggur skáldið langa ræðu í munn grísk-ítalska spekingnum Pýþagórasi (6. öld f.o.t.) sem ætlað er að útskýra hverfuleika mynda og ásýnda heimsins, að allt í veröldinni sé breytingum undirorpið, allir líkamar ummynd-

¹Sjá t.d. DVD-disk Alison Marek, *Text & Textile: An introduction to Wool Working for Readers of Greek and Latin*, 2004.

ist — nema sjálfar sálirnar í líkömunum sem aldrei eyðist og alltaf séu eins þótt þær færi sig úr einum líkama í annan. Þessi ævafora kenning er kölluð á grísku *metempsychosis* eða *palingenesia* en á latínu *transmigratio animarum*. Ellin, eða öldrun manna, er eitt þeirra dæma sem Óvíð tekur í grátbroslegri útlistun sinni á breytileika þeirra líkama sem fyrir ber í heiminum. Þýðing mín á dæminu er svohljóðandi (frumtextann set ég neðanmáls):

Fer þá og Helena fagra að gráta' er hún lítur í spegli
ferlega hrukkað sitt andlit á gamalli konu og undrast
hvers vegna karlmenn sig hafi í tvígang þá numið á brottu.
Hlutanna átvagl, Tíminn, og Ellin, sú andstyggð! þau drepa
eyðingu alls mjög á langinn, sem nagað er **tímans af tönnum**.²

Það er frá síðasta versinu hér að ofan sem *tímans tennur* hafa komist inn í nútímamál, að vísu alltaf í eintölu af einhverri ástæðu sem ég kann ekki að skýra. Elstu íslensku dæmin, frá miðri 19. öld, virðast sýna að þegar *tímans tönn* var fyrst beitt í íslensku máli þekktu notendur hennar hugmyndaknippið forna sem hún tilheyrði. Að óreyndu hefði ég giskað á að íslenska hefði fengið sína *tímans tönn* ekki beint úr latínu, heldur úr millimálinu dönsku, *tidens tand*, sem þangað gæti hafa borist úr þýsku, *der Zahn der Zeit*.³ En af samhengi orðræðunnar þar sem orðtakið kemur fyrst fyrir í 3. árgangi hálfsmánaðarritsins *Þjóðólfs*, 21. desember 1850, er kannski ekki langsótt að ætla að hún sé tekin beint úr *Metamorfósum* Óvíðs, eða að höfundur textans, Sveinbjörn Hallgrímsson, stofnandi og þá enn ritstjóri *Þjóðólfs*, hafi í það minnsta vitað hvar hún átti uppruna sinn og þekkt til eiginlegs samhengis tannarinnar við hugmyndina um stöðugt innrabyrði og síbreytilegt ytrabyrði heimsins fyrirbæra.

Orðtakið kemur fyrir í byrjun hugmyndaríkrar samræðu, að þeirra tíma hætti, sem ber yfirskriftina „Andi Ingólfs og Tíðarandinn tala saman“:

²*Met.* 15, 232-6: flet quoque, ut in speculo rugas adspexit aniles, / Tyndaris et secum, cur sit bis rapta, requirit / tempus edax rerum, tuque, invidiosa vetustas, / omnia destruitis vitiatque **dentibus aevi** / paulatim lenta consumitis omnia morte.

³Enskt dæmi er að finna frá endurreisnartímanum, í leikriti Shakespeares *Measure for Measure* eða Líku líkt, IV. þætti 3. sviði, þótt orðalagið *tooth of time* sé ekki notað í nútímaensku. Til gamans má geta að fjall eitt í Nýju Mexíkó í Bandaríkjunum, ævafornt storkubergsinnskot eða storknuð bergkvika í eldgíg, sem risið hefur upp úr jarðlögum þegar annað mýkra berg veðraðist í burtu og skagar nú eins og gríðarmikil tönn upp úr landslaginu, heitir Tooth of Time. Sjá t.d. Wikipedia <http://en.wikipedia.org/wiki/Tooth_of_Time>.

Tíð: Heill sjertu mikli landnámsmaður! Það er mjer fögnuður að finna þig að máli, og sjá hina **formmannlegu ásýnd** þína, yfirbragðið og augnaráðið, eins og það var á fyrri dögum; og hverju er það að þakka öðru, en **hinu staðfasta og eilífa sem inni fyrir er**, og sem **tímans tönn** getur ekki unnið á?

Andi Ing: Kom þú sæll á minn fund, lokkumkrýndi tíðarandi! Það er mjer einnig fögnuður, að sjá þína **breytilegu ásýnd**, og svipinn núna svo frjálslegan og bjartan yfirlits, sem vitnar um hina góðu kosti, sem þú býrð yfir. [Leturbreyting mín.]

Hér er því að vísu haldið fram að innrætið ráði nokkru um útlitið, hugmynd sem Óvíð er ekki ókunnugt um í *Metamorfósunum*, en vandlega er greint þar á milli. Bæði hin „formmannlega ásýnd“ Ingólfs og hin „breytilega ásýnd“ Tíðarandans mótast af þeim eiginleikum og kostum sem inni fyrir eru.

Hitt dæmið er að finna í kvæði Benedikts Gröndals, „Hugfró“, ort að sögn í Þýskalandi 1858, þar sem höfundur ávarpar skáldastéttina og útlistar varanleika orðsins listar í samanburði við aðrar listir:⁴

Og loksins þér, sem ljóða vekið hreim,
hvað léði yður drottins guðdómshönd?
Ei Mozarts list, ei málverks brögðin vönd,
né marmarann að setja lífs í heim —
hún veitti yður það, sem öllu ræður,
allsherjar mynda og lista sameining,
sem enginn myndað fær, en geystar glæður
gínandi snarka fram um stjörnuhring;
því *orðið* **má ei deyja, æ það lifir**
og allt því lýtur, bæði mynd og hljóð;
lýðirnir hverfa, en það er öllu yfir,
og **ekki dofnar neitt um tímans flóð**.
Mörg liggja spor frá myrkri Hómers tíða:
Myndirnar sukku niðr í aldar hrönn,
marmarinn eyddist fyrir tímans tönn:
Tigið og **eilíft stendur kvæðið fríða**,
því *orðið*, það er drottins eiginn hljómur,
en enginn spegill, það er lífsins fjör,
og á því vinnur enginn dauðadómur,
þó dróttir veifi bitrum feigðar hjör. —
[Leturbreyting mín.]

⁴Benedikt Gröndal, *Ritsafn*, 1. bindi, bls. 236, og skýringar bls. 544.

Hér virðist ljóðlist fornaldar, samsömuð sálunum í síbreytilegum lík-
ömum fyrirbæranna, ferðast óbreytt sjálf í gegnum tíma og rúm. Í
skýringum með kvæðinu, þegar það var fyrst prentað í ársritinu *Gefn*
1870, segir skáldið um þessa ljóðrænu orðræðu sína um varanleika
hinna skáldlegu orða:

Allir listamenn eru skáld. En í eiginlegum skilningi kalla menn
skáld þann, sem vinnur listina fyrir *orðsins* kraft. **Orðið er ekki
bundið við hin líkamlegu meðul sem hinir aðrir þurfa; mynd-
ir, litir og hljóð [...]** fjötra það ekki grand, en þó hefur það allt
þetta í valdi sínu og **takmarkast hvorki af rúmi, tíma né lík-
amlegum hlutum [...]** Þó vér *ritum*, og þó vér hljótum að sjá
eða *heyra* orðið, þá er það samt **fullkomlega andlegt og öld-
ungis ólíkamlegt**, en *samt* nær það út yfir allt og faðmar jafnt
allan hinn líkamlega heim sem hinar huldu tilfinningar andans.
[Leturbreyting mín.]⁵

Þessi elstu dæmi um *tímans tönn*, sem skráð eru í ritmálsskrá Orða-
bókar Háskólans, gætu bent til þess að það hafi verið skólastrákar á
Bessastöðum, heimilisfastir og uppaldir hjá Sveinbirni Egilssyni á Ey-
vindarstöðum, sem gerðu orðtakið gjaldgengt í íslensku máli og virð-
ist jafnframt ekki ósennilegt að þeir hafi þekkt hið upprunalega sam-
hengi þess í *Metamorfósum Óvíðs*.⁶

Til þess hef ég verið að grafa hér upp þennan fróðleik um *tímans
tönn* að mér þótti áhugavert að vita meira um sögu þessa orðtaks, sem
fyrir kemur í heiti málþingsins þar sem stofninn að þessari grein var
upphaflega fluttur, en einkum og sér í lagi til þess að sýna hvers konar
rannsóknir eru mögulegar með hjálp stafrænna orðabóka með dæma-
söfnum annars vegar og rafrænna textasafna hins vegar. Án sögulegra
dæma, tilvitnana í samhengið þar sem orðin koma fyrir, veitir einföld
orðabók afar litlar upplýsingar um orðin. Dæmin í *Þjóðólfi* og „Hug-
fró“ fann ég með því að leita fyrst í ritmálsskrá Orðabókar Háskól-
ans <<http://lexis.hi.is/cgi-bin/ritmal/leitord.cgi?adg=innsl>> að *tönn*
og finna síðan staðinn í *Þjóðólfi* á Tímarit.is <<http://www.timarit.is/?issueID=313809&pageSelected=0&lang=0>>. Að vísu varð ég að leita í

⁵Tilvitnun eftir Benedikt Gröndal, *Ritsafn*, 1. bindi, bls. 549-550.

⁶Sveinbjörn Hallgrímsson (f. 1814) var systursonur nafna síns Egilssonar og í föstri
hjá honum á Eyvindarstöðum á Álftanesi frá 10. ári. Hann lauk stúdentsprófi frá
Bessastaðaskóla 1834 og bjó eftir það í fimm ár hjá fjölskyldu Sveinbjarnar Egilssonar.
Benedikt Gröndal, sonur Sveinbjarnar, var tólf árum yngri og alinn upp á Eyvindar-
stöðum.

prentaðri bók að kvæði Benedikts Gröndals. Þessi tvö gagnasöfn, notuð saman, orðasafnið og textasafnið, gera nýjar rannsóknir mögulegar á orða- og hugmyndasögu nútímamálsins, sem áður voru ill- eða ógerlegar. Þetta er aðeins einn af mörgum möguleikum sem samþætt- ing orðasafna og textasafna á veraldarvefnum bjóða upp á.

Það má kalla sérstakt lán fyrir þessa litlu tilraun mína að Benedikt Gröndal skyldi í „Hugfró“ og skýringum með kvæðinu eigna *orðinu* einmitt sömu eiginleika og sálin hefur hjá Pýþagórasi, samkvæmt Óvíð. Benedikt gerir það „eilíft“ og „fullkomlega andlegt og öldungis ólíkamlegt“ þótt birtingarmyndir þess geti verið margvíslegar. Þessi skilgreining Benedikts á orðinu fellur vel að þeirri andstæðu sem ég vildi hér draga upp á milli klassískra orðasafna og þeirra fjölbreytilegu mynda sem þau geta nú tekið á sig í upplýsingaveitum á vefnum og greiðir um leið fyrir leik mínum að fornum orðum sem endurheimta fyrri tengslavef (lat. *contextus*) og öðlast nýjan á stafrænum vefstólum nútímans.

3 Hin miklu afrek fyrr og síðar á sviði klassískra orðasafna

Alltaf er verið að gefa út gamla latneska og forngríska texta í nýjum útgáfum og vissulega uppgötvast enn nýir textar á þessum málum, textar sem annaðhvort hafa fundist í rykföllum *codicibus* á hillum gamalla bókasafna, eins og páfabréfið um Skálholt sem Jónas Gíslason gaf út í XXIII. hefti *Sögu* árið 1985, eða komið hafa í ljós í fornleifauppgreftri, eins og 58. kvæði Saffóar frá Lesbey um ellina og ljótleika hennar sem grafið var úr söndum Egyptalands og útgefið í hitteðfyrri (West 2005). Eins og áður sagði liggur mikið af óútgefnum textum á latínu og grísku á handrita- og skjalasöfnum, svo ég minnst ekki á papýrusbrotin fjölmörgu í geymslum ýmissa stofnana og einstaklinga um allan heim. Þetta óútgefna efni telst ekki til nútíma textasafna og er því ekki orðtekið þegar unnið er að slíkum söfnum.

Annað vel þekkt vandamál er að vinnan við orðasöfnun og ritun orðabóka er mjög tímafrek. Aðstandendur stærsta latneska orðasafnsins, *Thesaurus Linguae Latinae*, sem ætlað er að ná yfir allt það sem skrifað var á latínu frá 3. öld f.Kr. til 600 e.Kr., hafa sýnt mikla þrautseigju í gegnum tíðina. Vinna við þessa miklu orðabók hófst ár-

ið 1894 og helstu hvatamenn verkefnisins voru Berlínarprófessorarnir Theodor Mommsen og Ulrich von Wilamowitz-Moellendorff. Fyrsta heftið kom út árið 1900 og nú hafa tveir þriðju hlutar safnsins verið prentaðir. Hverri færslu fylgir dæmasafn með öllum fornum dæmum sem skipta máli. Á síðasta áratug, þegar farið var að huga að því að koma orðasafninu á rafrænt form, voru seðlarnir orðnir tíu milljónir talsins. Enn vantar þó í N-P og algjörlega Q-Z. Um tveir tugir klassískra fræðimanna frá ýmsum löndum starfa nú að verkefninu sem rekið er við Bayerische Akademie der Wissenschaften í München.

Latína er tungumál sem notað var sleitulaust frá fornöld og fram á 19. öld. Margt var auðvitað skrifað á latínu eftir 600, þegar *Thesaurus Linguae Latinae* setur mörkin í orðasöfnun sinni. Til eru allnokkrar sérstakar orðabækur yfir miðaldalatínu og nú á síðari tímum hafa menn jafnframt verið að taka saman orðabækur yfir nýlatínu, þ.e.a.s. latínu endurreisnarinnar og síðari alda. Einnig er til í það minnsta eitt orðasafn yfir samtímalatínu.

Af orðabókum yfir miðaldalatínu eru margar til og sumar afmarkaðar við bókmenntir ákveðinna landa. Helstar eru *Dictionnaire latin-français des auteurs chrétiens*, fyrst tekin saman af Albert Blaise (Turnhout 1954–1967) en nýlega endurskoðuð og leiðrétt á vegum Paul Tombeur (2005); *Firmini Verris Dictionarius, Dictionnaire latin-français de Firmin le Ver 1440*, útgefin af Brian Merrilees og William Edwards (Turnhout 1994) með leiðréttingum ritstjóra; og þá hið mikla rit, *Glossarium mediae et infimae latinitatis*, sem á að upphafsmanni Charles du Fresne (1610–1688) en var aukið af Benediktsmunkum og ýmsum fleiri og loks prentað með viðaukum í átta binda nýrri útgáfu í ritstjórn Léopold Favre á árunum 1883–1887. Þessar þrjár eru allar aðgengilegar á vefnum hjá hollenska forlaginu Brepolis <brepolis.net> en því miður er Landsbókasafn – Háskólabókasafn ekki áskrifandi að því efni. Handhæg er latnesk orðabók J. F. Niermeyer, *Mediae latinitatis lexicon minus*, með orðskýringum á frönsku og ensku sem nú er í annarri útgáfu. Þessi upptalning er engan veginn tæmandi og á aðeins að þjóna þeim tilgangi að gefa sýnishorn af því helsta sem til er.

Af orðabókum yfir nýlatínu, þ.e.a.s. latínu eftir 1300, mætti nefna René Hoven, *Lexique de la prose latine de la Renaissance*, sem hann gerði í samvinnu við Laurent Grailet. 10. útgáfan er endurskoðuð og töluvert aukin, hún hefur einnig verið þýdd á ensku af Coen Maes og sú þýðing endurskoðuð af Karin Renard-Jadoul (Leiden: Brill, 2006). Einnig

má nefna nýlatneskan orðalista, *Neulateinische Wortliste* (NLW), sem upphaflega var gefinn út árið 1998 á veraldarvefnum <<http://www.lrz-muenchen.de/~ramminger/>>. Þessi orðalisti er uppfærður reglulega og gerður út af Johann Ramminger fræðimanni á *Thesaurus Linguae Latinae*. Í febrúar 2007 voru í NLW nálægt 8000 sérstaklega nýlatnesk orð, sem safnað hafði verið úr 1234 höfundum og 2243 verkum. Eins og orðabók Hovens er NLW afar gagnlegt tæki fyrir sagnfræðilegar, málvísindalegar og bókmenntafræðilegar rannsóknir á latneskum endurreisnartextum. Einnig er að finna nýlatneska orðalista í ýmsum útgefnum bókum og tímaritum (t.d. *Instrumentum lexicographicum* sem gefið er út á hverju ári í heftum tímaritsins *Humanistica Lovaniensia*).⁷

Lexicon recentis Latinitatis (Orðabók yfir latneskt nútímamál) geymir yfir 1500 orð og er gefin út af félaginu Latinitas. Félagið var stofnað með páfabréfi árið 1976 af Páli páfa VI. Tilgangurinn er meðal annars sá að taka saman og smíða latnesk nýyrði til þess að þýða orð í ítölsku sem ásamt latínu er opinbert tungumál í Páfagarði. Í þessari orðabók má meðal annars finna latneska þýðingu á hinu algenga fasta orðalagi *a priori* sem greinilega er ekki álitnið skiljanleg latína. *A priori* er þýtt svo á latínu: *ex antecapto iudicio!*

Mesta stórvirki íslensks lexíkógrafs á þessu sviði er latnesk og grísk orðabók Páls Árnasonar, eða Paul Arnesen eins og hann nefndist í Danmörku, þar sem hann varði allri sinni starfsævi. Páll fæddist á Breiðabólssstað í Snæfellsnessýslu árið 1776. Ættingi hans tók hann með sér til Danmerkur þegar hann var fjórtán ára og kom honum fyrir í latínuskólanum á Helsingjaeyri. Hann kom afar sjaldan til Íslands og síðast árið 1800 en hélt tryggð við Íslendinga og hjálpaði að sögn mörgum íslenskum námsmanni í Danmörku. Grísk-danska orðabókin hans, *Græsk-Dansk Ordbog til Brug for den studerende Ungdom*, sem prentuð var 1830, var frumherjaverk, um 1900 síður að lengd, og auðvitað sárgrætilegt frá sjónarhóli Íslendinga að hún skyldi ekki hafa verið með íslenskum skýringum. Þó skrifar Páll á eftir nafni sínu á titilsíðu skammstöfunina *Isl.*, sem stendur fyrir 'islending', eins og væri það lærdómstitill. Páll gaf einnig út *Ny Latinsk Ordbog til brug*

⁷Yfirlitið um nýlatneskar orðabækur í þessari efnisgrein byggir á og endursegir efni úr Marc van der Poel, ritdómur um „René Hoven, *Lexique de la prose latine de la Renaissance*. English translation by Coen Maes. Leiden: Brill 2006“. Bryn Mawr Classical Review 2007.03.12. Slóðin er: <<http://ccat.sas.upenn.edu/bmcr/2007/2007-03-12.html>>.

for den studerende Ungdom. Þetta verk er enn meira að vöxtum, hálft fjórða þúsund síður. Bæði þessi verk hef ég notað allmikið í mínum rannsóknum enda nákvæm og óvenju örlát á orðmyndir og skýringar.

4 Sundurleysi stafrófsraðarinnar og merkingar- bært samhengi

Stafrófsröðin, sem notuð er í hefðbundnum orðasöfnum, er nánast sjálfsagt afsprengi stafrófsins þar sem stafirnir eiga sér fasta röð (*alef-bet-gimel; alfa-beta-gamma*) og getur því ekki verið miklu eldri en þrjú þúsund ára eða frá þeim tíma er fyrsta semíska stafrófið er talið fundið upp. Föst röð stafanna varð óhjákvæmileg í fornum stafrófum, t.d. hinu hebreska og gríska, þegar bókstafirnir voru einnig látnir tákna töluorðin í stað sérstakra tölustafa sem ekki tíðkuðust. Stafrófsröðin byggist í raun á því að lesa talnagildi bókstafanna í upphafi orða og raða þeim svo í samræmi við það. Menn áttuðu sig á hagnýti stafrófsraðarinnar þegar í fornöld og var hún talsvert notuð einnig á miðöldum en sannaði þó fyrst í nútímanum hvers megnug hún var. Röðun orða í orðasöfnum eftir stafrófsröð er gagnleg af því að hún auðveldar mönnum að finna orð. Þeir þurfa ekki að skoða öll orðin heldur geta þeir flýtt sér til þess staðar í röðinni sem við á. Þetta geta menn af því að þeir hafa frá barnsaldri lagt á minnið röð stafanna. Notkun stafrófsraðarinnar sparar mikla vinnu og eykur því afköst manna. Stafróf vélheilanna — ASCII (American Standard Code for Information Interchange) eða nýrri kóðar eins og Unicode sem byggja á ASCII en fara langt umfram 128 tákni — fylgir á köflum röð bókstafanna í stafrófinu en þyrfti ekki að gera það. Skjótvirka örgjörva rafeindaheilanna munar ekkert um að „skoða“ hvert einasta orð og það gera þeir í hvert sinn sem þeir eiga að finna einhverja tiltekna samstöfu. Þessa vinnu gætu menn einnig framkvæmt en það myndi taka þá ærinn tíma. Stafrófsröð bókstafa og annarra tákna skiptir ekki máli fyrir vélarnar, nóg er að sérhverju rittákni samsvari ákveðið stafrænt bitamynstur til þess að rafheilinn geti leitað uppi gefnar táknarunur í orða- og textasöfnum. Örgjörvinn getur ekki aðeins leitað að stöfum stafrófsins heldur að sérhverju rittákni og þar með töldum greinarmerkjum.

Stafrófsröðin er orðræðunni sem slíkri óeiginleg. Orðum verður ekki raðað í stafrófsröð án þess að slíta þau úr skiljanlegu samhengi. Auk þess verður að smætta fjölda orðmynda niður í eitt flettiorð til þess að hægt sé að raða því gagnlega í stafrófsröð. Fyrir mörgum árum kom ég að skiptinema á Nýja-Garði sem gat með engu móti fundið í símaskránni föður vinkonu sinnar, sem var *Arnardóttir*, og vissi þó heimilisfangið. Hann leitaði auðvitað í öfugum enda stafrófsins sem sýnir vel hversu tillært athæfi um er að ræða. Því mætti kannski halda fram að raðað í stafrófsröð séu orðin aðskilin frá líkama (lat. *corpus*) orðræðunnar að nokkru leyti eins og sálin er aðgreind frá líkamanum í kenningu Pýþagórass samkvæmt Óvíð. Þegar Benedikt Gröndal semur skýringar sínar yfir kaflann í „Hugfró“ sem vitnað var til að ofan og heldur því fram að hið skáldlega orð sé „fullkomlega andlegt og öldungis ólíkamlegt“ er hann að ástunda það hugarfar sem nefna mætti lexíkógrafíska frumspeki. Uppröðun og orðskýringar orðabóka hafa mótað hugsun manna um tungumálið. Þess vegna er mikilvægt að gera sér grein fyrir að tilhögun orðabókarinnar er annarleg og fjarlæg allri venjulegri málnotkun. Það er í sjálfu sér þversögn að menn þurfi að raða orðum í sundurlausar stafrófsraðir til þess að geta fundið þau aftur en vélar geti fundið orðin í náttúrlegu samhengi orðræðunnar. En sú þversögn hverfur að mestu þegar haft er í huga að menn hafa sjálfir kennt vélunum þessa kúnt til þess að þurfa ekki lengur að raða orðum í stafrófsraðir. Menn hafa lengi vitað hvað stafrófsröðin hefur kostað siðmenninguna. Löngu fyrir tölvuöld þekktu menn annarlegar aukaverkanir stafrófsraðarinnar og þráðu að raða orðasöfnum á nýjan leik.

5 Gagnsemi uppstokkunar orðasafna fyrir tíma stafrænna vefstóla

Þessar dálítið skrýtnu vangaveltur mínar kalla á dæmi um möguleika og mikilvægi breytilegrar uppröðunar jafnvel á tímum áður en stafrænar spunavélar fóru að tíðkast. Sveinbjörn Egilsson rannsakaði, sem kunnugt er, skáldamálið norræna og samdi orðabók yfir það, *Lexicon poeticum*, með latneskum skýringum. Þetta orðasafn varð yfir 1000 smáletraðar síður þegar það var prentað, að Sveinbirni látnum, af hinu Konunglega norræna fornfræðafélagi, *Societas regia antiquari-*

orum septentrionalium, í fimm heftum á árunum 1855-60. Flettiorðin í þessu riti voru norræn eða íslensk, tekin bæði úr skáldamálinu og prósa, en orðskýringar allar voru á latínu. Um 60 árum síðar endurskoðaði Finnur Jónsson prófessor orðabók Sveinbjarnar og þýddi meðal annars orðskýringarnar á dönsku. Þegar danski málfræðingurinn Rasmus Kristian Rask hafði unnið sambærilega endurskoðun á Orðabók Björns Halldórssonar, á öndverðri nítjándu öld, bætti hann við orðskýringum á dönsku en fjarlægði ekki þær latnesku. Þannig hefur sú orðabók skýringar á tveimur málum, latínu og dönsku. Því miður, liggur mér við að segja, vann Finnur Jónsson ekki svona, heldur fjarlægði hann allar latneskar skýringar úr hinni nýju útgáfu sinni á *Lexicon Poeticum*. Finnur, sem jók orðasafn Sveinbjarnar, hefði auðvitað getað samið nýjar skýringar á latínu. Hann var menntaður í klassískum fræðum við Hafnarháskóla og hafði lokið latneskum skýringum Jóns Sigurðssonar á skáldatali í Snorra-Eddu-útgáfu Árnanevndar. Þetta síðasta hefti útgáfunnar var prentað 1887. Með aðstoð stafrænna spunavéla nútímans væru ýmsir möguleikar til þess að ummynda slíkt verk ef það væri til. En við höfum að minnsta kosti fyrstu útgáfu *Lexicon Poeticum* með latnesku orðskýringunum frá hendi Sveinbjarnar. Sonur hans, Benedikt Gröndal, spann reyndar og ummyndaði orðabók föður síns, þótt ekki væri það á stafrænum rafeindarokki, þegar hann tók saman lykil sinn yfir skáldamálið, *Clavis Poetica*. Tilgangur þeirrar orðabókar, sem höfundur kallar póetískt apparat (*apparatus poeticus*), var að raða dæmunum í *Lexicon Poeticum* á nýjan leik eftir latneskum flettiorðum „til þess að auðveldlegar mætti sjá, hvílk og hve mikil og ótæmandi hafi verið auðlegð hins forna norræna skáldamáls“ (*quo facilius adpareat, qualis et quam immensa et inexhausta fuerit copia antiqui sermonis poetici septemtrionalis*) eins og segir í formálsorðum bókarinnar. Með því að stokka efnið svona upp á nýtt röðuðust t.d. öll mannsheitin og -kenningarnar saman undir latneska flettiorðinu *vir*, ‘karl’, og fylltu þrjátíu og þrjá og hálfan dálk sem vissulega sýndi ákveðna auðlegð skáldamálsins. En úr varð einnig latnesk-íslensk orðabók með afar skáldlegum orðskýringum.

Mikilvægi uppröðunar verður einnig ljóst af nýlegri endurútgáfu *Klyfsa*, eða *Nucleus Latinitatis* frá 1738, á vegum Orðabókar Háskólans. Þessi ágæta latnesk-íslenska orðabók til notkunar í Skálholtskóla, eins og segir á titilsíðu, raðar vissulega flettiorðum í stafrófsröð en aðeins sönnunum orðum eða etýmólógískum orðum — gríska orð-

ið *etymos* merkir ‘raunverulegur’ eða ‘sannur’. Afleidd og skyld orð eru færð undir „sönn orð“. Þótt gagnlegur sé var *Kleyfsi* aldrei frumleg smíð heldur þýddi Jón Árnason biskup í Skálholti latnesk-danska orðabók sem kom út undir sama titli, *Nucleus Latinitatis*, í Kaupmannahöfn (fyrsta útgáfa sennilega 1712). Sú bók var til notkunar í latínuskólum Danmerkur og Noregs og hafði aftast gagnlegt registur yfir dönsk orð. Verkið hefur verið eignað prófessor Hans Gram við Hafnarháskóla en það er útgefið nafnlaust enda líklega þýtt úr öðru verki þýsku eða frönsku. Hin etýmólógíska uppröðun flettiorða í *Nucleus Latinitatis*, sem átti að kenna nemendum að hugsa í latneskum stofnum og grunnorðum, var helsti ókostur bókarinnar í notkun. Þess vegna var það afar vel til fundið af endurútgefendum bókarinnar árið 1994 að taka saman 86 síðna einfalda skrá yfir latnesk orð (fjórir dálkar á síðu) með blaðsíðutali. Allir venjulegir notendur *Kleyfsa* byrja á því að leita í þessari skrá og fletta síðan upp á síðunni sem þar er gefin upp. Hinn endurútgefna *Nucleus Latinitatis* hafa, eins og áður sagði, íslenskir fræðimenn töluvert notað við þýðingar og rannsóknir á íslenskum latínuritum.

6 Ummyndanir klassískra orðasafna á veraldarvefnum

En snúum nú aftur, í lokin, að klassík í rafrænum spuna. Skipulögð beiting tölvutækninnar á klassísku málin hófst árið 1968 þegar David Packard, annar af stofnendum stórfyrirtækisins Hewlett-Packard, hóf með hjálp frumstæðrar reiknivélar í kjallara Harvard Science Center að framleiða orðstöðulykil fyrir öll orð sem koma fyrir hjá rómverska sagnfræðingnum Livius.⁸ Þessi orðstöðulykill og fleiri álíka voru prentaðir og þótti hið gagnlegasta tæki við textarannsóknir. Í kjölfar þessarar tilraunar fylgdu þrjú frumherjaverkefni. Fyrsta og kannski merkasta verkefnið hófst árið 1972. Þetta er *Fjársjóður grískrar tungu*, betur þekktur undir sínu latneska heiti, *Thesaurus Linguae Graecae* (TLG), sem frá upphafi var ætlað að verða stafrænt textasafn yfir allar forngrískar bókmenntir. Aldarfjórðungi síðar hefur verkefnið náð markmiði sínu, að koma öllum útgefnum grískum bókmennt-

⁸Samantekt þessi byggir á og endursegir á köflum Gregory Crane et al. 2007, bls. 4-9.

um fram til 600 e.Kr. á stafrænt form, og hefur nú víkkað út svið sitt til þess að ná yfir býsantískar og síðari bókmenntir. Stuðst er við nýjustu og bestu útgáfur hvers texta en lesháttarafrigðum sleppt.

Annar viðburður í stafrænum spuna klassískra mála varð á áttunda áratugnum þegar þessum sama David Packard tókst að þróa vélbúnað sem gat ekki aðeins ráðið við textasafn eins og TLG heldur einnig við grískt letur og ritvinnslu á grísku. Kerfið var fyrst sýnt á aðalfundi Samtaka amerískra textafræðinga (American Philological Association) í Boston árið 1979. Þessi sérhæfða tölva, sem byggð var á almennu módeli, The Hewlett Packard Minicomputer, var heitin í höfuðið á forngríska skáldinu Íbýkosi og ein slík stóð á Háskólabókasafninu þegar ég var við nám í Háskóla Íslands á níunda áratugnum.

Í þriðja lagi hafa, síðan Ibicus-tölvann var hönnuð, verið þróuð fjölmörg kerfi til þess að vinna úr TLG-textasafninu. Á vefsíðum TLG eru talin upp fleiri en 12 slík kerfi. Síðan um árþúsundamótin hefur TLG-textasafnið aðeins verið aðgengilegt á veraldarvefnum þótt fram að því hefði það verið fánlegt á CD-ROM-diskum. Textasafnið á vefnum taldi í fyrri 91 milljón orða. Nýjung á TLG er að hægt er að tengja textana þar við orðmyndagreiningarforrit og grísk-enskar orðabækur á *The Perseus Digital Library* til þess að auðvelda nemendum og fræðimönnum lestur hinna forngrísku texta. Þannig er hægt að fletta upp í orðabókum 91 milljón orða í textum TLG með því einu að „klikka“ á þau. Þótt TLG-textasafnið hafi frá upphafi verið styrkt bæði af opinberum sjóðum og einkasjóðum og vinnan við það hafi kostað tugi milljóna Bandaríkjadala er það nú aðeins aðgengilegt þeim sem borga áskriftargjald. Starfsmenn Háskóla Íslands eru svo heppnir að hafa aðgang að TLG vegna þess að Stofnun Árna Magnússonar í íslenskum fræðum greiðir fyrir netáskriftina.

Margir klassískir fræðimenn hafa einnig aðgang að miklu stafrænu latnesku textasafni á CD-ROM-diskum sem tekið var saman af The Packard Humanities Institute en hefur ekki haldið áfram. Þó geymir PHI 5.3-diskurinn nánast alla latneska texta frá því fyrir árið 200 e.Kr. Auk þess selur *Biblioteca Teubneriana* aðgang að sínu víðfeðma textasafni á veraldarvefnum. Þetta textasafn var upphaflega ritröð fyrir grískar og latneskar textaútgáfur hins virta austur-þýska Teubnerforlags í Leipzig sem eftir fall Berlínarmúrsins hefur lagt sig mjög eftir nútíma útgáfuaðferðum. Einnig eru einstaklingar og ýmsir hópar óþreytandi við að bæta í veraldarvefinn latneskum og grískum text-

um. Sem dæmi um slíkt textasafn er *The Latin Library*, <www.thelatinlibrary.com>, þar sem hægt er að lesa flesta klassíska texta á latínu í heilu lagi. Það verkefni var allt unnið í sjálfböðavinnu.

Þessi textasöfn, sem hægt er að fínkempa í leit að sérstökum orðum og orðmyndum, hafa nú um allnökkurt skeið verið vinnutæki klassískra fræðimanna og haft bein áhrif á textafræðilegar rannsóknir þeirra. Sem dæmi um nýtt svið innan klassískra fræða sem beint má rekja til þess háttar orðaleita eru nákvæmar rannsóknir á merkingarsviðum einstakra lykilhugtaka, s.s. tilfinninga og ástriðna, í textasamhengi. Slíkar rannsóknir hafa með fræðilegum og öguðum aðferðum og tíðum vísunum í textadæmi leyst af hólmi eldri innsæisskrif um hugarfar og ástriður fornmanna.

7 Lokaorð

Helstu niðurstöður þessarar athugunar minnar eru þær að ummyndanir forngrískra og latneskra orðasafna á veraldarvefnum séu ekki einfalt framhald á orðabókargerð fyrri alda þótt þær byggist á þeim vísindalega grunni sem fyrri kynslóðir textafræðinga hafa lagt, heldur byltingarkennd nýjung í orðabókargerð og textanotkun sem hafi þegar, og muni enn frekar í framtíðinni, markað húanísk fræði á varanlegan hátt. Á þessu sviði hafa klassískir fræðimenn, einkum bandarískir, unnið brautryðjendastarf sem nýtast mun fræðimönnum á sviðum annarra tungumála og bókmennta með margvíslegum hætti. En til þess að hin nýju vísindi fari að gagnast fræðimönnum í sem flestum greinum (og undirgreinum) er auðvitað mikið starf óunnið.

Ljóst er að hin nýju orða- og textasöfn spunnin margvíslega á stafrænum vefstólum bjóða upp á mikla endurvinnslu gamals efnis. Sjálfur vildi ég sjá í framtíðinni fullkomið safn allra íslenskra latínutexta, og þar með taldar allar latneskar þýðingar á íslenskum miðaldertextum, tengt við orðmyndagreiningarvél og þær orðabækur sem Orðabók Háskólans hefur gefið út á prenti á undanförunum árum og á því á rafrænu formi, s.s. *Kleypsa* og Björn Halldórsson. Með slíkt tæki í höndunum gæti íslenskur fræðimaður með latínukunnáttu ofið merkingarvefi sem aldrei fyrr.

Heimildir

- Autenrieth, Georg. 1891. *A Homeric Dictionary for Schools and Colleges*. New York: Harper and Brothers.
- Benedikt Gröndal. 1948–1954. *Ritsafn*. Ritstjóri: Gils Guðmundsson. Reykjavík: Ísafoldarprentsmiðja.
- Crane, Gregory, David Bamman, Alison Babeu. 2007. ePhilology: when the books talk to their readers. Sjá: <http://repository01.lib.tufts.edu:8080/fedora/get/tufts:facpubs:gcrane-2006.00003/bdef:TuftsPDF/getPDF> (Sótt 29. júní 2007.)
- Hallgrímur Sveinbjörnsson. 1850. Andi Ingólfs og Tíðarandinn tala saman. *Þjóðólfur* 3. ár (21. desember), bls. 221–223.
- Jónas Gíslason. 1985. Lengi er von á einum. Áður óprentað páfabréf um Skálholt komið í leitirnar. *Saga* XXIII:187–194.
- A Latin Dictionary*. 1879. (2. útg. aukin og endurbætt, endurpr. 1984.) Ritstjórar: Charlton T. Lewis og Charles Short. Oxford: Clarendon Press.
- A Greek-English Lexicon*. 1940. (8. útg. aukin og endurbætt, endurpr. 1989.) Ritstjórar: Henry George Liddell og Robert Scott. Oxford: Clarendon Press.
- Marek, Alison, Susan Edmunds, Prudence Jones og Gregory Nagy. 2004. *Text & Textile: An introduction to Wool Working for Readers of Greek and Latin*. New Jersey: Department of Classics, Rutgers University. [DVD-diskur. Upplýsingar og viðbótarefni, sjá: http://classics.rutgers.edu/text_textile/textile1.htm (Sótt 29. júní 2007).]
- Orðabók Háskólans, Ritmálsskrá*. Sjá: <http://lexis.hi.is/cgi-bin/ritmal/leitord.cgi?adg=innsl>. (Sótt 29. júní 2007.)
- The Perseus Digital Library*. Ritstjórar: Gregory Crane. Boston: Tufts University. Sjá: <http://www.perseus.tufts.edu>. (Sótt 29. júní 2007.)
- van der Poel, Marc. 2006. René Hoven, *Lexique de la prose latine de la Renaissance*. English translation by Coen Maes. Leiden: Brill 2006. *Bryn Mawr Classical Review* 2007.03.12. Sjá: <http://ccat.sas.upenn.edu/bmcr/2007/2007-03-12.html> (Sótt 29. júní 2007.)
- Thesaurus Linguae Graecae, A Digital Library of Greek Literature*. University of California, Irvine. Sjá: <http://www.stephanus.tlg.uci.edu>. (Sótt 29. júní 2007.)

Timarit.is. Sjá: <http://www.timarit.is/>. (Sótt 29. júní 2007.) Reykjavík: Landsbókasafn Íslands, Háskólabókasafn. <http://www.timarit.is/>
West, Martin L. 2005. A New Sappho poem. *Times Literary Supplement* 5334 (27. júní).

Lykilorð

klassískar orðabækur, klassískir textar, textafræði, íslenska, tölvuvinnsla í hugvísindum

Keywords

classical dictionaries, classical texts, philology, Icelandic, computing in the humanities

Abstract

Most Ancient Greek and Latin *lexica*, as well as classical text editions, are still valid scholarly works, over 70 years after they were published. Which has given classical scholars, unentangled by copyright issues and the related extra costs, a considerable advantage in applying the latest hardware technologies, and developing new software for the transformation of this old—although not out-of-date—printed material, into electronic textual bodies, for the purpose of scholarly research. Thus, paradoxically, computing in the Ancient Greek and Latin classics, especially in North American universities, has been cutting-edge for decades. This paper describes the most important existing classics sites on the Internet (The Perseus Project and TLG), discusses in some detail the theoretical implications of the new technology, provides a demonstration of the kind of research made possible through connecting texts and *lexica* by means of parsing scripts, and speculates about possible future projects in the context of Icelandic literature in Latin.

Gottskálf Jenson
Háskóla Íslands
Aðalbyggingu
IS-101 Reykjavík
gthj@hi.is