

Frá máltilfinningu um málheildir til mállíkana

Eiríkur Rögnvaldsson uppgjafaprófessor í íslenskri málfræði

RASK-RÁÐSTEFNA

ÍSLENSKA MÁLFRÆÐIFÉLAGSINS OG MÁLVÍSINDASTOFNUNAR 25. JANÚAR 2025



∌H

- Íslenska málfræðifélagið
 - Ritstjóri Íslensks máls
 - Formaður

- Málvísindastofnun
 - Fulltrúi stúdenta í stjórn
 - Forstöðumaður tvisvar



UM ORĐARÖÐ Í ÍSLENSKU

Fyrirlestur á vegum Íslenska málfræðifélagsins 8. nóvember 1982

1. Grundvallarorðaröð

- 1.1 Tvær merkingar: a) Algengasta eða "hlutlausasta" orðaröð í yfirborðsgerð (Greenberg); b) Orðaröð <u>allra</u> setninga í djúpgerð (Chomsky).
- 1.2 Grundvallarorðaröð Íslensku. Ummæli nokkurra málfræðinga: "Einföldust er sú röð orða, að fremst sé frumlagið, en þá umsögnin (sögnin), og síðan aðrir hlutar setningar (þiggjandi, andlag, o.s.frv.) ... (Jakob Jóh. Smári 1920).

Frumlag - sögn er "The simple and most common word order" og "normal order" (Stefán Einarsson 1949).

"The normal or most usual word order in declarative sentences in Icelandic is <u>subject + verb</u> ..." (Jón Friðjónsson 1978).

"Das Isländische hat im allgemeinen die Wortreihenfolge: Subjekt, Verb, indirektes Objekt, direktes Objekt" (Magnús Pétursson 1978).

Röðin frumlag - sögn - andlag er "neutral" (Höskuldur Þráinsson 1979).

"the most usual, or unmarked, word order in declarative sentences is subject + finite verb + ..." (Maling 1980).

- $S \longrightarrow Ny+Sy(Ay)$ (Jon Gunnarsson 1973).
- $S \longrightarrow Nl+Sl+(Al)$ (Kristján Árnason 1980).
- S -> Nl Hl ((ao) ao) Sl (Fl) (Fl) (Höskuldur Þráinsson 1982).



EIRBUR RÖGNVALDSSON

Um orðaröð og færslur í íslensku

Málvisindastofnun Háskóla Íslands Reykjavík 1990





Á ráðstefnu 1958:

"Any natural corpus will be skewed. Some sentences won't occur
because they are obvious, others because they are false, still others
because they are impolite. The corpus, if natural, will be so wildly
skewed that the description would be no more than a mere list."

Aspects of the Theory of Syntax, 1965:

• "[...] though "probability of a sentence (type)" is clear and well defined, it is an utterly useless notion, since almost all highly acceptable sentences (in the intuitive sense) will have probabilities empirically indistinguishable from zero and will belong to sentence types with probabilities empirically indistinguishable from zero. Thus the acceptable or grammatical sentences (or sentence types) are no more likely, in any objective sense of this word, than the others."

BYLTING Á SKRIFSTOFUNNI!

Ef þú vinnur mikið með ritvél, blað og blýant eða skæri og ljósritunarvél, þá er fullkomið ritvinnslukerfi það sem þig vantar!

COMPUCORP OMEGA

er eina fullkomna ritvinnslukerfið á Íslandi.

Compucorp getur unnið á næstum hvaða tungumáli sem er, án breytinga.

Innbyggð leiðréttingarorðabók. Fáanlegar orðabækur fyrir flest tungumál Evrópu. Nú má prentvillupúkinn fara að vara sig!

Samtenging margra véla möguleg.

Auðvelt að halla skjá til að auðvelda aflestur

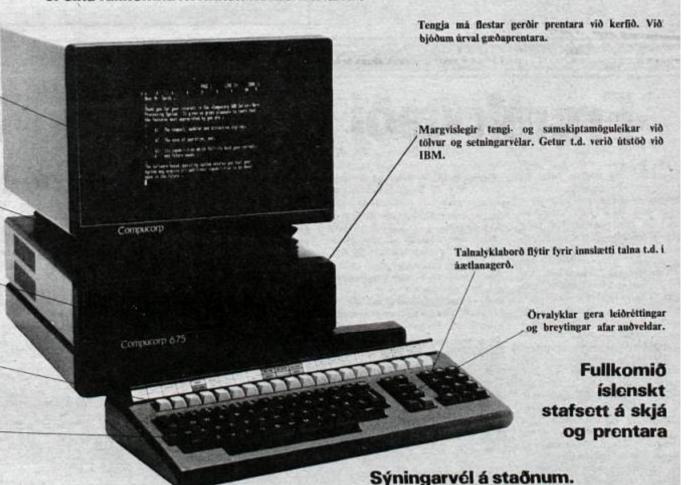
Hraðvirk diskdrif. Velja má um mismunandi geymslurýmd.

Allar helstu aðgerðir eru framkvæmdar með sérmerktum lyklum.

> Staðlað laust ritvélaborð. Allir stafir á réttum stað.

MICRO

Mest selda ritvinnslukerfi á Norðurlöndum!





Síðumúla 22 105 Reykjavík Sími 83040

Rafreiknir með íslenska útgáfu af Word Perfect

Hugbúnaðarfyrirtækið Rafreiknir hefur hafið sölu á íslenskri útgáfu af ritvinnsluforritinu Word Perfect, sem þykir
eitt öflugasta ritvinnsluforritið
sem til er á markaðinum og það
sem náð hefur hvað mestri útbreiðslu í Bandaríkjunum, að
því er fram kemur í frétt frá
Rafreikni. Forritið er fullþýtt á
íslensku, bæði skjámyndir, allar
skipanir til tölvunnar svo og
handbækur.

Ritvinnslu- og máltölvunarfyrirtækið Ritmál hf. var fengið til að þýða forritið úr ensku og höfðu

þýðendurnir nána samvinnu um þýðinguna bæði við starfsmenn Rafreiknis og Satellite Software, framleiðanda Word Perfect, Í frétt Rafreiknis kemur fram að mikil áhersla hefur verið lögð á að þýða þetta forrit á sem flest tungumál og fæst það t.d. nú á öllum Norðurlandamálunum. Í frétt Rafreiknis segir, að það hljóti að vera sérstakt fagnaðarefni að erlent fyrirtæki hvetji til slíks, enda í samræmi við ríkjandi stefnu hér, þar sem Íslendingar hafi jafnan lagt metnað sinn í að nota móðurmálið á öllum sviðum tækni og vísinda. Flestir eigi

auðveldara með að tileinka sér kerfi, sem hafi verið þýdd á gott daglegt mál, heldur en þau sem eru á erlendum málum.

Í frétt Rafreiknis eru raktir ýmsir kostir þess að nota ritvinnslukerfið. Bent er á að þeir sem vélriti mikið þekki vandamálið sem upp komi við leiðréttingar og breytingar. Iðulega burfi að vélrita sömu síðuna hvað eftir annað, þótt aðeins sé breytt örfáum orðum. Með því að nota ritvinnslu á tölvu hverfi slík vandamál því að þar sé hægt að kalla textann fram á skjá, leiðrétta og prenta eins oft og þarf. Þeir sem eigi svokallaðar einkatölvur eða PC, geti notað Word Perfect í stað ritvélar og eigi textinn að fara í prentsmiðju, þurfi ekki að setja hann aftur, heldur sé nóg að flytja hann af tölvudiskinum í setningavélina.

Ennfremur er í frétt Rafreiknis bent á, að í jafn fullkomnu ritvinnslukerfi og Word Perfect er, megi gera ýmislegt sem annars bekkist aðeins í dýrum og fullkomnum tölvum. Með þessu forriti sé unnt eins og í flestum öðrum slíkum kerfum að skipta um stafastærðir, setja spássíur, línubil, dálka, pappírslengd og annað er hæfi góðri ritvinnslu. Til viðbótar megi síðan nefna að með forritinu sé unnt að setja í dálka líkt og í dagblöðum, reikna, tölusetja kafla sjálfkrafa og raða í íslenska stafrófsröð. Ef til vill sé bó möguleikinn á vélrænni villuleit í textanum mikilvægasti kosturinn við ritvinnsluna, en hjá Ritmáli hf. sé nú unnið að gerð tölvutæks íslensks orðasafns er muni fylgja forritinu. Það leiti að stafavillum í textanum og leggi til ritmyndir í stað þess sem rangritað er auk þess að koma í stað stafsetn-

ingarorðabókar.



Aðstandendurnir – Þessir fjórmenningar stóðu að íslenskun ritvinnsluforritsins Word Perfect – t.v. Pétur Friðriksson og Lúðvík Friðriksson frá Rafreikni og Vilhjálmur Sigurjónsson og Eiríkur Rögnvaldsson frá Ritmáli hf.



Eiríkur Rögnvaldsson og Vilhjálmur Sigurjónsson Ritmál hf.

UM ORÐASAFN MEÐ WORDPERFECT OG ÍSLENSKA ORÐTÍÐNI

Ritmál hf. hefur gert íslenskt orðasafn til nota með ritvinnsluforritinu WordPerfect, sem Rafreiknir hf. hefur umboð fyrir. Slík orðasöfn eru þegar til á fjölmörgum málum, og mun íslenska safnið vera hið þriðja stærsta; hefur að geyma um 106 þúsund orðmyndir. Hér fer á eftir greinargerð um tilurð safnsins.

1. Textar

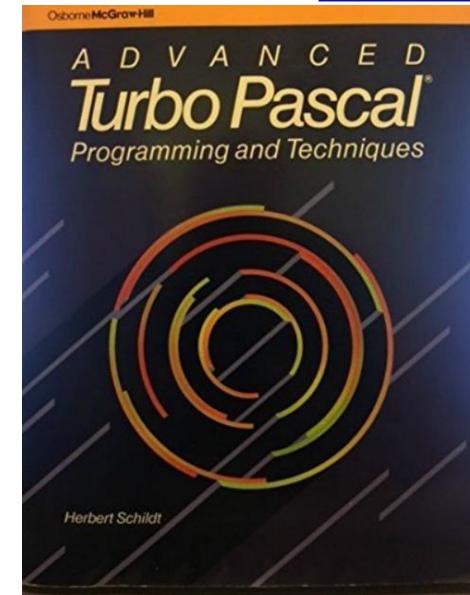
Safnið byggist á umfangsmikilli orðtíðnikönnun, sem tekur til um 2,5 milljónar lesmálsorða. Þar af voru tíðnitölur úr textum með um 700 þúsund orðum fengnar hjá Íslenskri málstöð, en að öðru leyti var könnunin gerð af starfsmönnum Ritmáls. Leitast var við að hafa textana sem orðteknir voru sem fjölbreyttasta, og fer hér á eftir yfirlit um helstu efnisflokka sem textarnir teljast til:

Blaðafréttir og -greinar um ýmis efni Bókmenntafræði, bókmenntasaga, gagnrýni Dýrafræði Guðfræði, trúarrit Heimspeki, hugmyndasaga Iðnaður, ýmsar greinar Íbróttir, ýmsar greinar Leikritun, leikdómar Lögfræði Matur og matargerð Málfræði Sagnfræði, ævisögur Skáldsögur, innlendar Skáldsögur, þýddar Tíska og fatnaður Tölfræði Tölvur og tölvun Uppeldis- og kennslufræði Útgerð og fiskvinnsla Verslun og viðskipti

```
begin
 clrscr;
 writeln (' -----');
 writeln ('
               SKEYTI
 writeln ('
 writeln ('
 writeln (' ------');
 writeln (' -----');
 lowvideo;
 gotoxy(4,6);
 writeln ('Með þessu forriti má velja úr tíðniskrá öll dæmi um tiltekið viðskeyti sem');
 gotoxy (4,7);
 writeln ('má vera 1-4 stafir að lengd, sama hvaða ending fer á eftir. Einnig er hægt');
 gotoxy (4,8);
 writeln ('að klippa beygingarendinguna af og steypa saman beygingarmyndum sama orðs.');
 gotoxy (4,12);
 write ('Nafn inntaksskrár...');
 gotoxy (4,13);
 write ('Nafn úttaksskrár....');
 gotoxy (4,15);
 write ('Fjöldi stafa í viðskeyti..');
 gotoxy (4,16);
 write ('Viðskeyti.....');
 gotoxy (4,17);
 write ('Orðflokkur....');
 gotoxy (4,18);
 write ('Kyn...');
 gotoxy (20,18);
 write ('Beyging...');
 gotoxy (4,20);
 write ('Klippa beygingarendinguna af?');
 gotoxy (4,21);
 write ('Steypa beygingarmyndum saman?');
end;
```

procedure intro;





NOKKUR VIÐSKEYTI OG TÍÐNI ÞEIRRA

the affect of go ; anything and

ings of element at order of days;, ands taker Absorder Shannowers

jud rest i aversi Die Suffice im Reliedischen (1997), Harr neur

himse him tegar sin each og sinaloge me date om text fell som totalt vikkigti. Hitter or

applying till som abridget och flav for

estands residings as had before I estands residings as had before I estands have actum as were by tald had sideacyti Historiagos and

nd mar of all authors marking bear of taking of children and takin and children any respect-

receivings of a through og ekapun, am er obsesse vidensti. Dad er

varts oolder regree & hjánaháttar.

or kilmukurar, om at dani di

tecil. Ro skenlaga er betta mylel-

Jesam desces surum of soon

orthics son emisked strankelin

excerned var on had next bur

septe tajut ratur ob viblogii. Es successi sit africian vi-

skeyle fideal citie take I fed of shearts must take shall similaber

offerential and the state of th

Off or bud refriers pro at ag-

ings the threeful og street we

menne vid eitine at har er old.

all with her benefits. Life to the district

egral in forms statife relatived.

does her, gland switched merk-

vil atom take from all property of colonger contractions, number or least (riched over 1000 for each of the riches). loked. My feet unnick and homes it an article of went platform and were hisverkam andenfurnar vikur og mencagar". Nivarlagif acpir sir ingt kopp å ad gena sagt frå helste mikt som jodern selekseta om ble albandblum hår å beskart råd- om til sernede sen storsjon til som stefan, Wee befor het ekki gefist timi til mi meda nidnestodor era sel og skylth, og þei má ekki taka þar ibler sem ög kem hár með af hátillegu; har geta átr eftir að breytast eltikvað við námani atlagun hüttig geni ehti nid. nornetri ger, og nins Henier Tyrir að þar memi milde. Einnig, menst þaf norn og orðið dámur mi viana mildo meira ür þeim ekt og sir. Vladienar markli þá tillum sem fyrir liggja, en þar er "viarlegar dómar", blóta breylir, earna sagun, að til þens hefur ekkil þetta og «démar gjatar siret upp-

1. Textarnir

Astroday fort jud at og hill jesses thing films up to be destroyed ländam erkið af tenturs (a icknistr allmanti ale Lit rilitie ka militarios. Tri no pura ingressed are magraf må safna að maðalakikkingu or á tilinn 61—80 þómað karnila-20 alies. Mig larged til ad ekto Jesus laris til sinheres eniferen legra stingara og þá á bani við af karna viðnosjanstöra. Við-nosjál em ekki þall reling, að tilska bego natively or at their how by much pass at do tak had bet med. Bellen, you far kenn cryster out or and an had better off along offers their variantil on takin metti ut-

Tectivar was affron tagt Next. carper of easily ogene, and intendeere og spiller, og eierg ser tæste i labout af elte år augbbotan av arbet arcum which I (I). li Tegradir teatac

Shouthettir og greiner tire jords Standards bikmantanga

gagarjan. Dýrostováli Godfradi, teluviti kimpiki, hapnyokongo Lectur, Femar grosser, Lectures, industrial

gefred, reage hikbogir, mlenke kakhogur, giddar. This or farmer.

Wear og thear. lapside og becentaftente

mitta of sionetic Teaturer vore and vager chis streaming valder the at note I posun tigung og vissungs tid tirta um kverst deergoelir beb sis. I ritoke tëshur, begar an er al rech fugut sibshaply gotte comlast: souldt beidamsystes labwert. (In the the excitors is an order of any olds are both ceffer of all fally ris and like to all like to all like. earth ages for spart man knice for terrig with skull bests, the alone recession. A c. i. have freelighed, ad fit alsociate knowled to all strongs boats premat mil er yelle o. i medica a tibeles. Lif og gefangelehere of olls can per or finish and Eta a so adje syndrom other titreibale teria, perseguel tel capthology berget Herbetta est kagi dala.

Fig still having all attractions at on low all a set goes stabilities. alls arhabri to Thantestown, par sen alar orienyake kana iya.e i sahirinda I passan 1930 pinana karribodar sepakat wen 111 passed missessell orderects When togeth by a sarries Suck til all with he been not real divoluce. mushalarage and or side

2. Vidskeytin can be brief han egger shoe can be enall benneger tip ift median etc. 2. Vidskeytin

I jenne crimii utta igud segis — x of gots spattu stelgrasings o — (scokeled, as jone ad tata act frá klema sem ég kel gert und- - steleytan. Eg yaga þó sí fleiri - stelekti millege, eks estrepreter-mitaril A cikin nekkurra við- - mil væld i estadir meskurga og — levik þenni velam sí bil að negl skogus i blesska millimannill. Eg - linnar klema (1946), av segi að - er só degra sædi við agrir ag freel Parent viters of deal nept or at deept and sept and real septe and with anythr ay bus til nafnerd, str. gefandt av sugar og sall om algreg vilskosti. l'étique et a l'étique d'ob-ur cetté dans, de gétique. Seult cape. Es adi mayet af pei som burns ar after haven by all forest williant of ball citizen sit on diense og better. Um fel fyrre talar Hallitz Hallitzson galse i Connected an all olaha shall area. -the, one wind to extend at the two green [1276], per som namn syne fram å aft frik omforset skinner som sicytina -m-, sée po aff de rétra: select he half imprintings world ni-ng vilkknytine 4, i idlere bestern tieriere og vilegere fleri æ færtt at daily to all the og vibriosyti sie. merry we came at version marchenter qui per engin de haflogu ravidrgo i vitare wrewta-

Vill val vidagette skepeti for filter. ments with read an groves fel evabeautier is relieve obligar. ulleine karn, fyer i befiten oeben-ug üg inki, sikti serk, ba. sec. rolnell Mr. 4 by leady oblice. rallectedor grain irá co, og hvar so Loss of note i promeader bigger aftenhousel syst. Re length there's early Kharell beer vig. at licks med of a city on all man lique:

Ath Jose takentekatu cakentagur s Audir noldad zu videlayna Augthat or ad you salet early over the are roots or eith large of bears of color old manual bases. By ratall of erecht form; tid all know had then staturations at 1 electrica deals of heffiger sodiegie kenn d efte. Variety of the all videoptics are store all formitt will order yaden. hoffing oreing solou budgement. og derfall orde pel it i crosses der Richte obs. og henne og stellomie kom kan også det sed joken farmet och at beste it at our or layther entiry I been tinks be senter alded borest his beliad elitivate keep tood of orders were olde turk eliberance vehicept I may og wyn Durray ar eritti af elikaka url ura uy eterkarit fan

Guiffrnesor	ar (1927)	this at arra real feet we be			
Del	Nest WA	1966	Mahi	Depart	
Sees.ic.	fac	44	Seed		
TOE	Th-	180-	- Illian		
endowal	EX.L	-0.0	grifted-	554	
brakta.	100	# Z2:	0.000000	100	
1000	rest.	-3-	mold		
1541	30	d.	2.00	-87	
6828	20 Aug	-3	2.0		
Ref.	1764	1.3.	20		
ATTO BY	1000 m	-Ante-	antiblists		
64,504	84	1877	#16000 2000 100 100		
Kd=	69-	+	100		
60	3	1	Es.		
hari	100		hori		
kuring	7000	Atte	berthy:		
plant .	430	450	Altri-	4	
glowers	241	100	20.65	-6.7	
glegymaka	200	-m-do-	Fig. With-	-	
40kg	1/06/	4.	16.00		
Inches.	344	145	formal.	4.5	
Hittidae	25-	4-9-	Hal-	4.8	
tiel in	201	-	tokin	48.	
660	14	· din	84	4	
derests.	Bir-	-m-di-	firm	4.	
mylen.	78.86+	400	totale-	46	
100	760	at .	I/M:		
Marta	2.00	4	tjus	460	
getager.	1000	-CP	pitte.	CE.	
RIGHT.	100	4.5	Child		
talangt	16	422	1850	1	
Sees	350	42	Johns	-3	
1405.9	3.5	- 4	48.5		
acter	Ext. 167		awker.		
Mar	6- due		éthe		
or elevelor.	108- RT-	4	Inc. 8:20-	-8"	

History ore grain) nanemenantmanagered of fare paymen also men mitar I printerpolarization con abelieux officia az enfu de belev aleffit appears, shere, Speri of heldoducedam Dall igger i Segun executions, believely by an about a some med og villakert som ens viss dichtett, þeger orðil beggilter. Dara eins stafter er rejög erftit að en í valstivísan flokkaði úg orða. firm we be and affect but or our negaritate on the parameter of the parameter and of the parameter of the pa ager beggspravendinger & etter, im-peux ad twee wildowski. Eg nic (et jam kut od cegga flestim eften a Pulcestone Bayman bald by hour. of or skept juite distant of Mean admiture; first justice, words, refritop take rist og ordgefraring till rik i rikknomik. Bekke mindelen ingen ur edelogiid de, ann og l helmir, hijódritt, speld soufre, feurteen á ad adtolja jadtalt sana. eibelocstiftedu ekki. Swidylgja regud. ar bus vandracii til vidbotar, ad badbannio all bar cerbur bub ebbi mednokkm rokt fredål å ef nouse lidt. by fell part of the gogen of a same in the same in absolut tolve og by lagh. skill I had I herta skipet artik. Me i taki ig order ellefesti, jeftere helt tilk for med tell vilkdosett som en i skipetning hen skanskan i bign elle alein olt kjöl, 4 sin og í sín fósia, og -a, sar. Sigló, púlkó coufre thomas para telar arm i begings, og file after motter sends. Vidasoyti i ordennen. Hir gilde þó á l og á, hannig að þest vildaseyti sanbetað tilð senna og ero viða, að van Ukolakya uudbusta. Paž ar iku kiturail znad sigi ad tefa elestão viduleyti og hvað ekid. Vil out a lea file-fg eingings efter

santinulegan etkun, en riki sogolagues. Daring poter word I namen thicken of hed som by take came children's of digulage and of terms on large, og olog, of hel one og tall het valskept at augulage att etti og hið surra, ögni viðdosti eru al i first se erom ortflokki, c.e. sk, ner er'l teòl (yangaredar, islendur og unfaretur, islendu, klep en it i fær er ma kyr, ca, eng og ing heli (a. og led fismegar, blendinger, sejling, drettmagt, sar i til av kyr, er jamerikani, skipan ilhani, og etuig ero daere pro ad curso videbegetë of notad hardi med stocket og sedari. beygegu i sana kyai. Par mi nafaa -kakur, -kaki bambekur, sana-I other become children toldi de sidnicythe toll of a bept, offic jed nonno Mr., term -lefter, -lefti var-iald sitting some vivelegald, vegrahave all her bale often order oft. basi wies og sterle, beggings til skeptis. Europ vore sjender og -jajutur, sire og i verknafter og stifnulter tolir jul soms, enla oft id her i mili Serakaha ara comleyrorchidostin -an oc -un tain asm

elt, enda seu par eiterig virl felip-an, skipun). Eerfronze wedar al

gots been, all vamaleges one aski dil menia are vibrosytif-log- in Swing-atorium; eine og monn vita er það

mjog skyrng i siskaælan ika Dar a mili er på örskyrker að gretas

chrim de santingé, seo cê ég good has organ tileaune til flokkun-I shareefedri lake Aleganders 16harmsoner, Die Suffice im Idändischen en talls alle top 188 edukesti Enincky tula bja mar ward 48. Denoue tiller ore bit olde. away sambori agar, but an standore teler Alexander had eitt vidskeytt sem og tid ten og öfuge. Ot þan viðskogti som og kummaði sem talm bil Alexander neres brit: -hierar, wer the var rolls, og avo dans ng dati. Detto era nying teknolo-noryti, en af Shangurov/labikinsi ad duess are bas order ad eschwerts logil frist ingresselve order indensite.

Frod offer neckands 1984. Dur corner reside from beyginger-eninger to from ad and med videkoption degree green, on I beingererbern. From dir, and can op tables.

Degre all pic com ad firm gegreen

been dir in al lifts internal picon.

Desper all pic com ad firm gegreen

been dir in al lifts internal picon.

Desper dir in com ad firm gegreen

been directable and have app dent

en 1888 journal reform. Dat blazer and picon.

Desper directable and the size of the siz

tuntil Jeg- karner gakkan fyrir doestiru -andi. Verga er all telja - skeyti. Next loora een två vilsdoesti the obdestion degar, or see both begingswellings in the en-tended of the gatest normal whitesybyerkyrmathords, sing og um-String action is an arrangement as a submitted r tan wegar akki alilaf -ak, or l fineria anti-vih-keyto-ingtinda tenggia skelle villeknyel vilitar - bangt ad gall skrimm, har som orden olkof tulavant aff abskotangban ad - ern in nambengia. Oft var betta adur I kuriken marhordure. See er eftir. Noblear vehicesti korea see olde fore north renes vid vice meri

> hoose officed gull leads trying of dohnan biyour be surfishers of Your or het will an instal an oldoen, kerner brengt arnare engler byer. Vil getern tekiñ ordil metmega sid nagairmar silta og versa en hrad er tege i begrind? Og sen mætti begri telja og alik ikklamili borno con viò nesstani invert elegata ager, t.d., er han segár léléger orðusíðusejti, er mignikgar erdag-ar ljungararfa eru mikki flem en approra orbanea, nema hy bakh on-hare alcollà rik meta gaga pel ad-

2 Niderstäder

dag 9697 -m 5257

databas)

chapter.

2850

distant Hill distant

1207 404

1499 -cp-2

3085 - Admin

4400

-trg-i

411 -bitt-gr

385 -dosper 580 -dosper 530 -dol(na)

275 -mm.

100 - word 215 - km i

523 -4 Years

103 -may-er 103 -mai

156 -40am

100 -cag

75 -cir. 65 -cir. 47 -cir.

163 -cd1(a)

-and

En na er mal til herid ad fra diffit i niturationmar in literary irri. Per ná gà từ i [i], þar ein ains affectulars so fit managementality manual striken.

> 2018 2018 1366 AND ST

> > titles.

-tagel -tard-t

dies.

-02K

sup-ar

420

-ul-(hac.)

-old back

-titles

Shell:

46.16.1

-mig-s

sale citizent sidekeyti ad geyma, in fyrn histina, rolia som setti ad-

and one or well-point to fillets retu, on local or moist-? Energy as a relies

ord are og vitend og verund; vid

tycomet grains har vibilizatio and

og tengine fort Notara aufreid-

elelogti i illam diore tibilere

manotentar grinde carder ret eg

andwine til mens og ero vide, so

fearer reder perefer degt reat relds:

Arcedatian (10 binard ordainma.

less esgra era dillorer denr og i below asserts or enterport radial efforburn der 3 tenturum 3429 teismannel ordnynde son incindes somm ordny des og fillagedoft-ribetette deg. Boynese er må in. milvisiodestelleren, bik ribdoutió deg. Begniar er mê briggis efun nebanna hér só sama. ficzenia orbib siepta tra serii Cwide, ar rebiil regrar û ribbert. Vibblogiin -of, -skarur, -semi, -dómur og sine en joud ingrible et diffe-un à under sander et ai of jess orb eru mitter beer ekke notud i fleir-

tolar. Day or built of popper serial skint begangganguskin hvere ords og hvert uppflettiont þet abeins takit often order occia are onto tad one. (Campleda ere of afakaplega constalever Seel interconnell ord med

-don-er

-E82-18

-morl

dender

-trg-l -tap-tr

-sloa

MALE

-000013

-keti -keti

-Ittioh-m

alldal

-tind

-12.5

-101

or Tea actuation at the hot some on how means Acceptant or of long sector his mater I profes sent; Fig. 6 minutes and order mod free ja val-storyti ar 1593 -0.24E 444(00.)

videleyturum id	hir hokkumii t
(weenstanninger)	ner, Dudmänn
on lengt dollar a	re invert belts.
section mand on	cititized area!
dg tal samt salt	bette cigi od s
maleigud van etitel	s. Stream within
i william	
3 App 41	dissan-lag-cri-
2-142	Octident
2 -41	gods me
4 chie	(DAYS-40-CF)
L ages	Machite-lay-u
(6 - m)d	losser-o.
7 (44)	(Secretaria)
8 31(16)	(Best-Bil)
9 41	(gret-raft)
13.4efour	(manteh-m)
16-16-(G) de ce	(norbinadous)
12/2/1	60
H-l-toni-	(regards).
100000	
18 -tap-us	(undergrap)
14 sugar	(Start-langua)
15 - Ing-er	ting lagran
te la	(called-full
17 -kg-i	found-ingvi
26 -4-1(n.) 19 -akso-er	indi-6-6
20 -40-4	(Kinker-abseptor (Kinker-abseptor (Kinker-abseptor
	(Katha-en-er
21 -eerd 22 -kell-kr	(gunns-eess)
22 400	(intertar)
24 -406	(Athend)
IB-orne	Diomento-fem
06 of (100)	(Manual)
27 (464	(Dat-Int-II)
25 -total	(leif-indl)
29 -mm-ar	thorings our
30 100-1	(non-emission)
Bi -do-se	(hei-littren)
82 -aH/ce I	
23 VSv.10	(customen
9.0 All (1987)	5-50 BM 57

-0.00 -0.00 -0.00 -0.00 I fromate different ora vibility in a dia variough (dags orderers losses 25-4-4-1(b) (special-)

ali beggiagarromär al leverja ordi
agneriague. Alfordavert er aff
afterderifelosytif -sice hefer 105
executandi ordesendo, en nóviro 5
interestiff Soggregoriyasian
kirs. Pad täknar ad all fluotum ort-
minds into theodotte but form m
of en boygingament.

but out we anticked data are you also mad un write whiteytanna. Par or notirioga occi-sinada. slide emisken vilskeyte i mikra. bilikabeteypa earron@kenbesyteparrowten sina orde, our og nicebook 5386 ordereds beerig damen per sen bort offshipt:

ng i delbeter à trèse, on fierte og rassarinfredictoform, rikisatofoun coufes Hagreyndis d hale við betta er of, að virkni viðunesta klóti all delecte leys at expo marks of hell has redrigers magnificated rot-Secular or have city here unroctor ad reeds; sum vibilityti bestbeing from es will used all bount. our natificipa vid terms valuosopti og led your staypt aureun orders sine ogvicindalegue, velicindalegue ost. 2v. Erander vilkkepti sees orr chelege i hanza dallorenn detta The fig. 6 filtrary britses one logged under under both 1 persons. Suprage day, Day or both at copye service or mod (fidagenories) shoots 40. seen or had spitto office i formate. different his dollar had mitter t des test (44) pe her wegge namenden sen, bei al hal kernr beggegeregeit hen er ob fitte | mangateter inserarchen, beinafrana. Dit brogder avo vid all two III, mildill og thill, en af twent þeir

ir sareistaispar. Ro ale syntr-legar gituriega yttrhandynmideleytics leg or real. Roule of degar gittings.ytt-lyingarushmideleytics legun but band; or left lighter describered boð tengin ekki ner mörgum rótum, on har exter ere him worse follow

Til ab reyas no at th servers heldersynd of been olds surran by ig til framer alvins opp är þessem litin ar þannar gorð að og gaf orð eig fyrir sam þess í hverri skrá, þennig að fyrsta sæmð gaf ekk stig. arrab too make, Silve mindi of

	44	(liceratio)	333
8	dia dijesi	Describ!	AL THE STATE OF THE PARTY OF TH
œ	el lefo-ar	(get-of)	
ö	4efour	(man-leib-ari	-34
ø	(6) de la	(work-mad-m)	-
я			
ú	- Hans	(organida).	32
03			-
в	ung-ar	(safegras)	_
ø	ang-ar.	(\$000-mag car)	130
ø	- lings or	ting lagran.	-
90	ing-er is kg-i	(called full	B. C.
×	-kg-i	found-ingvi	-
51	44401	Indi-8-4	800
п	-Maner	(Kikac-aksep-art)	200
п	dos.	((0)(73-40-4)	
93	-week -lakis-ke	(grant-eres)	
83	Assists.	(begreen-latting)	-314
3	ene ene	(rek-etter)	333
в	126	(vib and)	
8	dict.	Discussion (
а	(fee) (b)	Direct !	
v	(d. (es)	(fixe-lat-i):	
ĸ,	-8061	(leif-indl)	
93	-MERIEE	thorongic-ornered	
96	100.9	(0.07-0.000-0)	th.
	-to-ex	(hal-lim-er)	
23	-aHine I	(stat-skil)	
a	district.	(strationer)	
8	-EII-A	(121-621-4)	M
뤈	- SECTION	(pent-hit-er) (pert-arch to) (per-al-fr) (paks-herif)	
8	-00000-47	Smorney (n)	
刨	Metal	Southern)	- 15
剩	-30081	Quality-feet()	

(News) ---Same relies 42-4-MA-4 | Dell' erroll |Estitled| |Perlevill 47-4-44 Industr

Dad sirike jo ad kange daerste i uð þetta enduð unungu tið ris-mil og niðundiðumur jeðu bregglegs, eletimate tolkengthene of talvall and others. Area or like at vidskrytið -0 mynni þá hadda allmjög á skránni. Skildra hagnigad am ports and fit of green Signatur John sorus (1984), but homes from all i Slangererdabékimi, sen nela mi all endamporali taimati are adv 114 ord sted 4, were or rirelega belowinglimetra en med ling, sin 30 finel. es med aus, og 6 form en med aris Gerar edukayti sam era man al-gunguri I Mangumerbabbicinal en beauth being hiller ellericrates. nddayid ogin, san 17 dani en un i Shagurordabik; sit stasta carri factat um bad i beier trotter ove og dorkdi. Að eke kristl og ekki eð kami urkredu ritmynni þem, -times on umeal or levert ofth of all tota and som purely one christial

4. Loksont

Eg ball namera skit rekit neim un jetta al egg. I jeste eigi valkire. He lippe for beimbil chistère og jad er jost af hage er að sinns ir benara i ýran vegt. Pegar ritralisabri. Defablica: His ekilore, scotter abgreging 1 bilantockri mynd må athaga hvenar extra and or har on there may breath with skeeti, og ejd namely beset earsband. Ster å redd Glishe i redding og tille. i tenta. Dur mi coung ge uldar book olda og stgetta dærne tan tæst tid og rekje þannig að en bærja egti ossikeli visslehne sideleying allage because his against opposite the best of the property of the control of brigarn NO begar blendeguese smar oru til i bilsutaissi mest weri the many forwarding of pure sum-terdage about it pairs, og sjå twentig histfallið mill strættera viðcorts befor bresst. Delta og mange fleire, vicei hoogt no going an mix fafur skill media tiral til þam sen, bread now piles works

conde J'Ammeson, 1971, bie Suffan-te Mandischen, Polyke Arbeite Hadro-tes 1975, Sephjerick, ider Kanner, 1984 Samert unfrech und Annerta Sir Vienner dag mit Bescheit. El Eliar Landelsy 2, oktober 1884, St., 1884-1984, St.,

50-15t, Oaks re Golffanske, 1800, lebent militeret reds scales og drungs. Riterarysgie.

inchentic. 18th Orbitation of the State of t

1970 Falling there is a Ratio States A Stop temporarie Mady, Nacolska studies

Set Salesta orderjedna. Margin hosta I. Provincia Benerit van Sagonii ng March

121: 121

The department of the control of the cont

ORÐABÆKUR · IÐUNNAR

ORĐA

EIRÍKUR RÖGNVALDSSON



fer Oddur með hinn tuttuganda mann uns hann kemur skammt frá bæ þeim upp liggja í dalinn frá Völlum uns þeir koma í Hólmsland og töluðust sig skjótt og síðan ríða þau uns þau koma til þings. Unnur gekk og fyrir norðan jökul og svo uns þeir komu til Svínafells. Flosi sendi eg eiga rétting allra orða minna uns eg kem máli mínu til réttra ráða skyldu. Einar fór suður aftur uns hann kemur heim. Nokkuru síðar hafði . Býr Þorbjörn skip sitt og fer uns hann kemur í Brattahlíð. Tekur Eiríkur Fara hvorutveggju með hinu vestra landinu uns þeir koma fyrir nes það er og upp í heiði til Gönguskarðs uns þeir koma svo upp að ein í kvöld og bíðið mín þar uns eg kem á morgun, nema eg . Nú dregur þó saman með þeim uns hann kom upp í ásinn. Þá og fara svo vestur til sveita uns hann kom til Helgafells, því að Þeir ríða nú út eftir héraði uns þeir koma á Arneiðarstaði. Er gengið að ánni og upp eftir nesinu uns þeir komu á Víðivöllu. Þá var til komu. En þau Gísli fara uns þau koma í Friðarey til Styrkárs fara dag þann og um nóttina uns þeir koma í dal þann er ljær þeim hestana og ríða þeir uns þeir komu á Mosvöllu og þaðan hann í Geirþjófsfjörð og er þar uns haustar. Þá fer hann enn til af þessu máli. Sveinarnir fara nú uns þeir koma í Geirþjófsfjörð og liggja hluti vel. Leið nú svo fram uns Þorfinns var heim von. mann. Þeir voru fimm saman, riðu uns þar til er þeir komu á varðhöld á sér. Hann fór nú uns er hann kom í Vatnsfjarðardal og góðu föruneyti og fóru alla leið uns þau komu í Noreg. Tóku skildu. Gekk Gunnar þá leið sína uns hann kom heim. Urðu hans menn skip reiður mjög og fara nú uns þeir koma þar að er Geir Indriðastíg hjá Þyrli og beið þar uns Þorsteinn fór til blóthúss síns sem hana á baki og ríðið svo uns þér komið fram að Faxalæk. Hann þeir Sámur þá allt að einu uns þeir koma á heiðarbrúnina. Sá hann Þetta ráð taka þeir, fara uns þeir koma í Saurbæ. Tók Bersi Karlsson vex upp með móður sinni uns hann var tólf vetra gamall og honum og fór hann um sveitir uns hann kom til Óttars og bað mikla á hávu skafti. Hann fer uns hann kemur í sauðahús og rekur þeir nú sem þeir mega mest uns þeir koma mjög svo í öndverðan ríður undan sem mest getur hann uns hann kemur heim til þverár og eftir þeim. Ríða þeir leið sína uns þeir koma á Miklabæ í Óslandshlíð. yfir Lagarfljót og upp með fljótinu uns hann kom í Atlavík snemma morguns.



4. Lokaorð

Hér hefur verið drepið á nokkur atriði sem gæta þarf að þegar fornir textar eru notaðir sem setningafræðilegar heimildir. Meginatriðin eru þessi:

Í fyrsta lagi er ekki hægt að nota kveðskap sem heimild um setningagerð, enda þótt það hafi stundum verið gert. Samanburður við nútímamál sýnir okkur að þar eru jafnvel grundvallarreglur um setningagerð óbundins máls þverbrotnar í einföldum og alþýðlegum kveðskap, án þess að það torveldi skilning.

Í öðru lagi verður að varast að bera saman fornt ritmál og talmál nútímans, þegar verið er að skoða breytingar á setningagerð. Við vitum að í nútímamáli er munur talmáls og ritmáls töluverður, og ekki annað hægt en gera ráð fyrir því að svo hafi einnig getað verið að fornu.

Í þriðja lagi getur verið nauðsynlegt að líta framhjá einstöku setningum sem fyrir koma í fornum textum, og telja þær málfræðilega rangar; þ.e., í ósamræmi við málkerfi textans. Hér verður þó að fara með mikilli gát, og taka mið af eldri stigum málsins og síðari þróun, eftir því sem hægt er.

Í fjórða lagi er stundum hægt að beita setningafræðilegum prófum til að átta sig á einstökum setningagerðum. Þessi próf gefa þó yfirleitt mun óvissari niðurstöðu en í nútímamáli, vegna þess að þau byggjast flest á því að hægt sé að bera saman tækar setningar og ótækar.

Í fimmta lagi verður að fara mjög varlega í að draga ályktanir af þögn varðveittra texta um tiltekið atriði. Það er mjög hæpið að gefa sér að tiltekin setningagerð hafi ekki verið til í fornu máli, þótt engin dæmi finnist um hana; og enn hæpnara er að byggja kenningar á því að þessi setningagerð hafi ekki verið til.



1.2 Efniviður

Eftirtalin rafræn söfn sem öll eru opin á netinu hafa einkum verið nýtt til dæmaleitar:

- (4) a. Íslensk sendibréf frá 19. öld (http://brefasafn.arnastofnun.is/)
 - b. Tímarit.is (http://timarit.is/)
 - c. Íslenskt textasafn (http://corpus.arnastofnun.is/)
 - d. Bækur.is (http://baekur.is)
 - e. Ritmálssafn Orðabókar Háskólans (http://www.arnastofnun.is/ page/ritmalssafn)
 - f. Sögulegi íslenski trjábankinn, IcePaHC (http://linguist.is/icelandic_treebank)
 - g. Mörkuð íslensk málheild (http://mim.arnastofnun.is)
 - h. Ordbog over det norrøne prosasprog (http://onp.ku.dk/)
 - i. Netútgáfan (http://www.snerpa.is/net/)

Einnig var leitað rafrænt í vesturfarabréfum (Böðvar Guðmundsson 2001, 2002), bókum úr ritröðinni *Sýnisbók íslenskrar alþýðumenningar* (Bragi Þorgrímur Ólafsson 2004, Davíð Ólafsson og Sigurður Gylfi Magnússon 2001, Sigrún Sigurðardóttir 1999),² og uppskriftum fjölmargra viðtala við Vestur-Íslendinga.³ Auk þess hefur verið leitað í ýmsum öðrum sérstaklega völdum heimildum, bæði rafrænum og prentuðum.



GSLT

GSLT ▶ Kurser

Sveriges nationella forskarskola i språkteknologi

info@gslt.hum.gu.se

Svensk version saknas



GSLT's courses 2001 announcement

During the autumn term 2001 the new Swedish national Graduate School of Language Technology (GSLT) will offer two introductory graduate level courses in language technology: Speech Technology and Natural Language Processing. You can find detailed information about the courses on the school's webpages at http://www.gslt.hum.gu.se. The courses will be given in three parts. An intensive introduction given in Goeteborg (Speech Technology 17/9-19/9, NLP 19/9-21/9), distance components conducted via the web, and a final meeting of one or two days towards the end of the semester (the final session for Speech Technology will take place at KTH in Stockholm, and for NLP in at Vaexjoe University).

The courses are open not only to the school's own graduate students but also to others who meet the following requirements:

- 1. you have completed an undergraduate degree or have completed equivalent studies.
- 2. you have completed a half year of full time undergraduate study (20 Swedish credit points) in computational linguistics or speech related studies or have equivalent experience
- 3. you must name a academic advisor who is willing to support your application and whom we may contact if necessary

It may be necessary for us to limit numbers for practical reasons.

Money provided by NorFA's Nordic language technology programme is available to pay reasonable travel and living expenses for students from Scandinavia (including Sweden and Finland) who wish to take these courses.

If you are interested in taking the courses, please fill out the form provided on http://www.gslt.hum.gu.se by 5th September. We will let you know by 10th September whether there is a place for you on the course.

The Icelandic μ-TBL Experiment:

μ-TBL Rules for Icelandic Compared to English Rules

Eiríkur Rögnvaldsson

University of Iceland and GSLT

January 2002

This paper reports on a work in progress. Auður Þórunn Rögnvaldsdóttir, Eiríkur Rögnvaldsson, Kristín Bjarnadóttir, and Sigrún Helgadóttir are experimenting on the possibilities of using Torbjörn Lager's μ -TBL program to build a PoS tagger for Icelandic. Kristín writes about the preparation of the corpus and the tagsets; Auður writes about the templates for Icelandic; Sigrún writes about the testing of the tagsets and the templates; and Eiríkur writes about the rules learned by the program. Although each paper presents work that has primarily been carried out by the author, the four authors have worked closely together on this experiment. Each paper should be more or less self-contained, but they all refer heavily to work described in one or more of the other papers.

```
fpken
            hann
                       hann
s f q 3 e b o átti
                       eiga
nheo
            afmæli
                       afmæli
a o
nkeo
                       dagur
            dag
            oq
                       og
nkeng
            hvolpurinn
                       hvolpur
                       Vaskur
n k e n - m
            Vaskur
sfq3eb
            var
                       vera
            afmælisgjöf afmælisgjöf
nven
```



```
tag:sfg3eb>sfg1eb <- tag:fp1en@[-1,-2] o
tag:cn>c <- tag:svg3en@[1,2] o
tag:cn>c <- tag:svg3eb@[1,2] o
tag:af>fplfn <- wd:við@[0] & tag:sfglfn@[1] o
tag:sfg3en>sfg1en <- tag:fp1en@[-1,-2] o
tag:cn>c \leftarrow tag:sfg3eb@[1,2] o
tag:af>fp1fn <- wd:við@[0] & tag:sfg1fb@[1] o
tag:sfg3eb>sfg1eb <- tag:fp1en@[1] o
tag:svg3eb>svg1eb <- tag:fp1en@[-1] o
tag:fpken>fpkeo <- tag:af@[-1] o
tag:cn>c <- tag:sfg3en@[1,2] o
tag:sfg3en>sfg2en <- tag:fp2en@[-1,-2] o
tag:ssg>sbghen <- wd:var@[-1,-2] o
tag:foheb>lhebsf <- wd:einu@[0] & wd:í@[-1] o
tag:fahen>faheo <- tag:af@[-1] o
tag:af>fplfn <- wd:við@[0] & tag:sfglfn@[-1] o
```

Notkun textasafna við setningafræðirannsóknir



- "Í stað þess að lesa allt ÍSTAL-safnið, eða fara gegnum á sjöunda þúsund dæma um það, dugir því að skoða þessi 30-40 dæmi. Í stað margra tíma þreytandi yfirlegu þar sem hætta á mistökum er veruleg kemur 10-20 mínútna vinna þar sem tækifæri gefst til að skoða hvert dæmi vandlega og meta hvort það falli undir þá setningagerð sem leitað er að."
- "[W]e have shown that a corpus that has been tagged using a rich tagset based on morphosyntactic features can fruitfully be used in the search for a number of syntactic constructions, and hence is a valuable tool in studying syntactic variation and change."

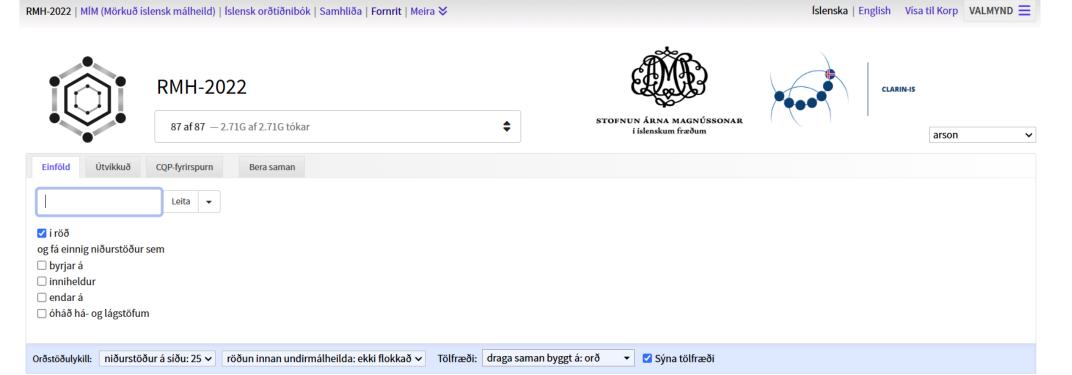
Nú leita þeir um skóginn og finna *Gísla* **eigi** — now search they about the-forest and find Gisli not — 'Now they search through the forest and don't find Gisli' er hann dræpi *Þórð* **eigi** — og förunauta hans when he killed Thord not — and companions his 'if he didn't kill Thord and his companions' og fundu *Þórð* **eigi** — sem von var að and found Thord not — as expectance was at 'and not surprisingly, they didn't find Thord'

Mörkuð íslensk málheild | Orðtíðnibók | Fornrit] Leitarlína 1 Leitarval Bæta við orði Eyða út orði





Bæta við leitarlínu 2



Charles J. Fillmore, "Corpus linguistics" or "Computer-aided armchair linguistics", 1992



- "Armchair linguistics does not have a good name in some linguistics circles. A caricature of the armchair linguist is something like this. He sits in a deep soft comfortable armchair, with his eyes closed and his hands clasped behind his head. Once in a while he opens his eyes, sits up abruptly shouting, "Wow, what a neat fact!", grabs his pencil, and writes something down. Then he paces around for a few hours in the excitement of having come still closer to knowing what language is really like."
- "Corpus linguistics does not have a good name in some linguistics circles. A caricature of the corpus linguist is something like this. He has all of the primary facts that he needs, in the form of a corpus of approximately one zillion running words, and he sees his job as that of deriving secondary facts from his primary facts. At the moment he is busy determining the relative frequencies of the eleven parts of speech as the first word of a sentence versus as the second word of a sentence."
- My conclusion is that the two kinds of linguists need each other. Or better, that the two kinds of linguists, wherever possible, should exist in the same body.

Máltilfinning, málheildir og mállíkön



- Þróun síðustu ára í gervigreind er mjög áhugaverð
 - munu risastór mállíkön leysa máltilfinningu málhafa alveg af hólmi?
- Chomsky hafnaði því í viðtali 2004
 - Corpus linguistics doesn't mean anything. It's like saying [...] suppose physics and chemistry decide that instead of relying on experiments, what they're going to do is take videotapes of things happening in the world and they'll collect huge videotapes of everything that's happening and from that maybe they'll come up with some generalizations or insights. Well, you know, sciences don't do this.
- en bætti svo við:
 - But maybe they're wrong. Maybe the sciences should just collect lots and lots of data and try to develop the results from them.
- Erum við komin þangað núna, tuttugu árum síðar?



Takk fyrir áheyrnina!

eirikur@hi.is