GÃ¶gn: lÃfÃ¦Ã° mÃ¡ltÃ¦kni og gervigreindar

Allar mÃ¡ltÃ¦knilausnir, hvort sem Ã¾aÃ° eru spjallmenni, Ã¾Ã½Ã°ingarvÃ©lar, talgervlar, talgreinar eÃ°a mÃ¡lrÃ½nihugbÃºnaÃ°ur, byggjast Ã¡ miklu magni fjÃ¶lbreyttra mÃ¡lgagna. Og mÃ¡ltÃ¦knin er aÃ° breyta lÃfi okkar og samfÃ©lÃ¶gum um Ã¾essar mundir. Spurningin er lÃ¶ngu hÃ¦tt aÃ° snÃºast um Ã¾aÃ° hvort viÃ° eigum aÃ° Ã¾rÃ³a og nota mÃ¡ltÃ¦kni. Ã• staÃ° Ã¾ess snÃ½r hÃºn aÃ° Ã¾vÃ- hvernig landslag er aÃ° teiknast upp Ã mÃ¡ltÃ¦kni og gervigreind. Er Ã¾aÃ° tÃ¦kni sem talar Ã-slensku? Er Ã¾aÃ° tÃ¦kni sem hentar alls konar fÃ³lki viÃ° mismunandi aÃ°stÃ¦Ã°ur og kemur til mÃ³ts viÃ° Ã¾arfir manneskjunnar? EÃ°a er Ã¾aÃ° einsleit tÃ¦kni sem Ã¾vingar notendur til Ã¾ess aÃ° aÃ°laga mÃ¡l sitt og jafnvel suma Ã¾Ã¦tti daglegs lÃfs aÃ° tÃ¦kninni? Er Ã¾aÃ° tÃ¦kni sem er Ãºtilokandi eÃ°a inngildandi? SvÃ¶rin viÃ° Ã¾essum spurningum velta aÃ° miklu leyti Ã¡ Ã¾vÃ hvaÃ°a gÃ¶gn eru til reiÃ°u til Ã¾ess aÃ° Ã¾rÃ³a tÃ¦knilausnir.

MÃ¡ltÃ¦kni fyrir Ãslensku Ã¾arf aÃ° Ã¾jÃ¡lfa Ã¡ Ãslenskum mÃ¡lgÃ¶gnum. Ã?Ã³ aÃ° viÃ° njÃ³tum gÃ³Ã°s af Ã¾rÃ³un algrÃma og Ã¾jÃ¡lfunaraÃ°ferÃ°a, og einnig af grunnlÃkÃ¶num sem Ã¾jÃ¡lfuÃ° eru Ã¡ Ã¶Ã°rum tungumÃ¡lum, Ã¾Ã¡ talar og skilur tÃ¦knin ekki Ãslensku nema hafa sÃ©Ã° Ã-slensk mÃ¡lgÃ¶gn.

Eftir Ã¾vÃ sem tÃ¦knin aÃ° baki mÃ¡ltÃ¦knilausnum Ã¾rÃ³ast og notkun Ã¾eirra verÃ°ur Ãºtbreiddari, Ã¾vÃ meiri krÃ¶fur verÃ°um viÃ° aÃ° gera til Ã¾ess aÃ° lausnir henti notendum og aÃ°stÃ¦Ã°um hverju sinni. ViÃ° viljum aÃ° tÃ¦knin skilji unga sem aldna, konur, kvÃ¡r og karla, VestfirÃ°inga, NorÃ°lendinga, Ã¾au sem tala Ãslensku sem annaÃ° mÃ¡l o.s.frv. HÃºn Ã¾arf aÃ° skilja formlegt mÃ¡l og Ã³formlegt, skapandi skrif og skÃ½rslur, skilja stÃ¶k orÃ° og stuttar yrÃ°ingar Ã samtali og samhengi Ã lÃ¶ngum textum. Ã?aÃ° Ã¾arf aÃ° skilja meira en orÃ°anna hljÃ³Ã°an â?? Ã¾aÃ° getur til dÃ¦mis veriÃ° verkefni aÃ° meta hvort texti eÃ°a tal lÃ½sa Ã¡nÃ¦gju eÃ°a Ã³Ã¡nÃ¦gju, Ã hvaÃ°a samhengi eitthvaÃ° stendur, til dÃ¦mis varÃ°andi sÃ¶guna, frÃ¦Ã°asviÃ°, stjÃ³rnmÃ¡laskoÃ°anir eÃ°a annaÃ°. SÃ¶muleiÃ°is viljum viÃ° aÃ° tÃ¦knin getiÃ° â??tjÃ¡Ã° sigâ?? Ã¡ fjÃ¶lbreyttan hÃ¡tt en skili ekki sama mÃ¡lsniÃ°i Ã Ã¶llum aÃ°stÃ¦Ã°um.

Ã?ekkjum gÃ¶gnin

ViÃ° Ã¾urfum aÃ° vanda okkur sÃ©rstaklega vel Ã¾egar viÃ° lÃ¡tum tÃ¦knina taka Ã¡kvarÃ°anir fyrir okkur Ã¡ grundvelli greiningar Ã¡ gÃ¶gnum sem hÃºn hefur yfir aÃ° rÃ¡Ã°a. GÃ¶gnin geta innihaldiÃ° bjaga sem endurspegla til dÃ¦mis rÃ¡Ã°andi umrÃ¦Ã°u eÃ°a hlutdrÃ¦gar Ã¡kvarÃ°anir sem teknar hafa veriÃ° Ã gegnum tÃÃ°ina t.d. varÃ°andi rÃ¡Ã°ningar Ã stÃ¶rf. ViÃ° Ã¾urfum aÃ° Ã¾ekkja gÃ¶gnin sem viÃ° notum og vera meÃ°vituÃ° um hvaÃ°a gildrur gÃ¦tu leynst Ã Ã¾eim.

Til Ã¾rÃ³unar Ã¡ mÃ¡ltÃ¦kni sem getur fengist viÃ° talaÃ° og ritaÃ° mÃ¡l Ã¾arf bÃ¦Ã°i tal- og textagÃ¶gn, sem og myndbandsupptÃ¶kur Ã¾ar sem hÃ¦gt er aÃ° greina samhengi milli tals og varahreyfinga auk andlits- og lÃkamstjÃ¡ningar.

GagnasÃ¶fn eru Ãºtbúin Ã¡ mismunandi hÃ¡tt eftir tilgangi Ã¾eirra. StÃ³ru mÃ¡lÃkÃ¶nin og nÃ½justu taltÃ¦knilÃkÃ¶n vinna meÃ° grÃÃ°arlegt magn gagna. HÃ©r erum viÃ° aÃ° tala um Ã¾Ãºsundir eÃ°a tugÃ¾Ãºsundir milljarÃ°a tÃ³ka af textum og hundruÃ° Ã¾Ãºsunda klukkustunda af tali. Til samanburÃ°ar inniheldur flaggskip Ãslenskra gagnasafna, RisamÃ¡lheildin, um 2,5 milljarÃ°a orÃ°a, lÃklega um einn tvÃ¶ Ã¾Ãºsundasta af Ã¾jÃ¡lfunargÃ¶gnum stÃ³ru mÃ¡lÃkananna.

MeÃ° samstilltu Ã¡taki allra, Ã¾aÃ° er opinbera geirans og einkafyrirtÃ¦kja, almennings og sjÃ¡lfstÃ¦Ã°ra tal- og textaframleiÃ°enda, mÃ¦tti Ã¾Ã³ lyfta grettistaki Ã Ã¾vÃ aÃ° fÃ¦ra Ãslenska mÃ¡ltÃ¦kni upp Ã¡ nÃ¦sta stig, stig Ã¾ar sem skortur Ã¡ gÃ¶gnum er ekki helsti akkillesarhÃ¦llinn. TÃ¦knin er til en hÃºn er gagnslÃtil Ã¡n gagna.

Ã• undanfÃ¶rnum Ã¡rum hefur Ã¡ Ã•slandi byggst upp gÃ³Ã° Ã¾ekking Ã¡ Ã¾vÃ aÃ° ÃºtbÃºa vÃ¶nduÃ° gagnasÃ¶fn Ã¾ar sem gÃ¦Ã°i eru metin fram yfir magn. Vinnan hefur ekki aÃ°eins snÃºiÃ° aÃ° Ã¾vÃ-aÃ° fÃnÃ¾jÃ¡lfa lÃkÃ¶n til Ã¡kveÃ°inna nota, til dÃ¦mis til Ã¾ess aÃ° tala Ãslensku eÃ°a til Ã¾ess aÃ° standa sig sÃ©rlega vel Ã mÃ¡lrÃ½ni, heldur einnig aÃ° Ã¾vÃ aÃ° prÃ³fa og meta gÃ¦Ã°i mismunandi lausna. LÃkan sem stendur sig best Ã Ã¡kveÃ°inni tegund verkefna er ekki endilega besta lÃkaniÃ° fyrir Ã¶ll verkefni og allar aÃ°stÃ¦Ã°ur. HÃ©r Ã¾arf ekki eingÃ¶ngu aÃ° taka tillit til gÃ¦Ã°a Ãºtkomu heldur Ã¾arf oft aÃ° lÃta til Ã¾Ã¡tta eins og hraÃ°a og hagkvÃ¦mni. Ã?aÃ° er Ã¾vÃ nauÃ°synlegt aÃ° hafa aÃ°ferÃ°ir til Ã¾ess aÃ° meta gagn og gÃ¦Ã°i mismunandi lausna Ã¡ mismunandi verkefnum.

TÃ¦knilegt sjÃ¡lfstÃ¦Ã°i

Ã• samhengi tÃ¦kninnar erum viÃ° ekki eyja. Mikil Ã¡hersla hefur veriÃ° lÃ¶gÃ° Ã¡ Ã¾aÃ° undanfarin Ã¡r aÃ° stÃ³ru tÃ¦knifyrirtÃ¦kin Ã BandarÃkjunum hugi aÃ° Ãslensku og Ã¶Ã°rum smÃ¦rri tungumÃ¡lum Ã sinni tÃ¦kni, en viÃ° Ã¾urfum einnig aÃ° lÃta mun meira til EvrÃ³pu. Ã?aÃ° er Ã erfÃ°amengi EvrÃ³pu aÃ° huga aÃ° fjÃ¶lbreytni, Ã¾ar sem til dÃ¦mis Ã¶ll tungumÃ¡l Ã¡lfunnar eiga aÃ° fÃ¡ sitt plÃ¡ss. Ã• meÃ°an hefur Ã¾aÃ° veriÃ° aÃ°alsmerki bandarÃsku fyrirtÃ¦kjanna aÃ° huga fyrst og fremst aÃ° Ã¾eim tungumÃ¡lum sem flestir tala eins og ensku, kÃnversku og spÃ¦nsku. Ã•slenska hefur Ã¾Ã³ fengiÃ° mjÃ¶g mikiÃ° plÃ¡ss miÃ°aÃ° viÃ° hÃ¶fÃ°atÃ¶lu, Ã¾Ã¶kk sÃ© Ã¾eirri vinnu sem fariÃ° hefur fram hÃ©r heima bÃ¦Ã°i Ã¡ tÃ¦kni- og samskiptasviÃ°inu.

ViÃ° Ã¾urfum aÃ° huga aÃ° Ã¾vÃ aÃ° vera ekki of hÃ¡Ã° stÃ³rfyrirtÃ¦kjunum og vinna Ã sameiningu aÃ° Ã¾vÃ aÃ° efla Ãslenska og evrÃ³pska mÃ¡ltÃ¦kni sem getur Ã¾jÃ³naÃ° okkar mÃ¡lsvÃ¦Ã°um Ã¡ eigin forsendum. Ã?aÃ° er strategÃsk Ã¡kvÃ¶rÃ°un aÃ° halda sjÃ¡lf um taumana, ekki bara elta heldur aÃ° vera Ã fararbroddi Ã Ã¾vÃ aÃ° Ã¾rÃ³a tÃ¦kni Ã hÃ¦sta gÃ¦Ã°aflokki okkur til hagsbÃ³ta. Og Ã©g veit aÃ° Ã©g er farin aÃ° endurtaka mig: grunnurinn liggur Ã gÃ¶gnum.

ViÃ° Ã¾urfum aÃ° byggja upp ferla og innviÃ°i til Ã¾ess aÃ° sem flestir sjÃ¡i hag sinn Ã Ã¾vÃ, eÃ°a a.m.k. aÃ° Ã¾aÃ° sÃ© ekki ÃÃ¾yngjandi, aÃ° leggja til gÃ¶gn til Ã¾rÃ³unar Ãslenskrar mÃ¡ltÃ¦kni. Grunnurinn hefur veriÃ° lagÃ°ur gegnum Ãslensku mÃ¡ltÃ¦kniÃ¡tlunina. SÃº vinna skilaÃ°i mikilli Ã¾ekkingu Ã¡ Ã¾essu sviÃ°i og viÃ° hÃ¶fum nÃº reynslu af tilreiÃ°ingu fjÃ¶lbreyttra gagnasafna.

SÃÃ°an Ã¾arf aÃ° setja vÃ©lina Ã gang, ef svo mÃ¡ segja, og halda Ã¡fram aÃ° safna. LÃfÃ¦Ã°in Ã¾arf stÃ¶Ã°ugt aÃ° renna. Ã?relt gÃ¶gn er sama og Ãºrelt tÃ¦kni. Ã?g hvet Ã¾ig, lesandi gÃ³Ã°ur, til Ã¾ess aÃ° skoÃ°a hvaÃ°a gÃ¶gn Ã¡ Ã¾Ãnum vegum gÃ¦tu komiÃ° aÃ° gagni Ã Ã¾vÃ verkefni aÃ° stuÃ°la aÃ° lifandi, manneskjuvÃ¦nni tÃ¦kni til framtÃÃ°ar.

Hvernig getur Ã¾Ãº eÃ°a Ã¾itt fyrirtÃ¦ki lagt til gÃ¶gn?

AlmannarÃ³mur hefur sett upp *HeimildagÃ¡tt* til Ã¾ess aÃ° taka viÃ° textagÃ¶gnum frÃ¡ fyrirtÃ¦kjum: https://almannaromur.is/thin-islenska-er-malid

Ã•rnastofnun og Grammatek vinna um Ã¾essar mundir aÃ° umfangsmikilli sÃ¶fnun talgagna, til dÃ¦mis Ãºr hlaÃ°vÃ¶rpum og Ã¶Ã°ru fjÃ¶lmiÃ°laefni. TengiliÃ°ur hjÃ¡ Grammateki er Anna BjÃ¶rk NikulÃ¡sdÃ³ttir (anna@grammatek.com) og hjÃ¡ Ã•rnastofnun Einar Freyr SigurÃ°sson (einar.freyr.sigurdsson@arnastofnun.is), sem auk Ã¾ess getur tekiÃ° viÃ° hvers kyns fyrirspurnum varÃ°andi mÃ¡lgÃ¶gn.

Til upplÃ½singar

UnniÃ° upp Ãºr fyrirlestri Ã¡ mÃ¡lÃ¾ingi European Language Data Space (LDS), VerÃ¶ld, hÃºs VigdÃsar, 9. oktÃ³ber 2025.

Mynd

Anna BjÃ¶rk NikulÃ¡sdÃ³ttir.