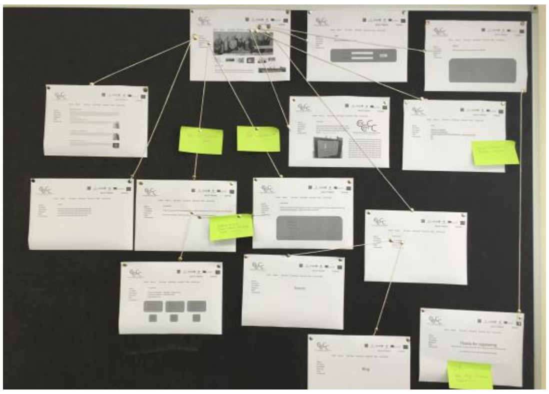
Cylchlythyr CorCenCC

Rhifyn 5: Awst 2016

Corpws Cenedlaethol Cymraeg Cyfoes
National Corpus of Contemporary Welsh

Cyfarchion gan y Prif Ymchwilydd

Croeso i 5ed rhifyn cylchlythyr CorCenCC. On'd yw amser yn hedfan pan fyddwch chi'n cael hwyl! Mae'r rhan fwyaf o'r haf hwn (hyd yma) wedi'i neilltuo at wneud cynnydd ar dasgau sy'n perthyn i becynnau gwaith unigol y prosiect (WPs – ceir diweddariadau isod) yn ogystal â lledu'r gair am waith CorCenCC mewn gwahanol gynadleddau academaidd a digwyddiadau cyhoeddus. Rwy'n falch iawn o ddweud bod gyda ni dros 100 o bobl sydd wedi tanysgrifio i'r cylchlythyr a/neu sydd wedi cytuno i gyfrannu data i'r prosiect yn barod – ac mae ein hymdrechion recriwtio newydd ddechrau! Rydym hefyd wrthi'n rhoi gwefannau swyddogol prosiect CorCenCC (www.corcencc.org a www.corcencc.cymru) at ei gilydd, a gobeithiwn y bydd y rhain yn helpu adeiladu a chryfhau cymuned CorCenCC ymhellach byth, a gobeithiwn hefyd y byddant yn sbardun ar gyfer cyfnod casglu data'r prosiect. Rydym yn gobeithio lansio'r safleoedd erbyn diwedd mis Medi fan bellaf – felly cadwch eich llygaid ar agor am ddiweddariadau/wybodaeth bellach ynglŷn â hyn.



Bwrdd darlunio gwefannau CorCenCC

Yn ôl i'r rhifyn presennol: y mis hwn rydym yn dod â newyddion diweddaraf tîm CorCenCC; adolygiadau o gynadleddau a digwyddiadau diweddar; diweddariadau ar y WPs, ac rydym hefyd yn cyflwyno aelod arall o dîm prosiect CorCenCC, sef Mark Stonelake, Cyd-Ymchwilydd sy'n gweithio ym Mhrifysgol Abertawe.

Mwynhewch! Dr Dawn Knight (Prifysgol Caerdydd)

Newyddion

Hoffem gymryd y cyfle hwn i gyflwyno aelod mwyaf newydd tîm CorCenCC, sef Lorena Varona. Bydd Lorena yn gweithio fel Cynorthwyydd Gweinyddol rhan-amser hyd ddiwedd prosiect CorCenCC a bydd hi'n gweithio o Brifysgol Caerdydd. Meddai Lorena: *'mae gwell dealltwriaeth o'n hiaith yn hanfodol er mwyn gwybod am ein diwylliant a diogelu ein gwreiddiau. Rwy'n falch iawn o fod yn rhan o dîm prosiect CorCenCC a helpu darganfod mwy am y defnydd o'r iaith Gymraeg. Diolch!'* Croeso i'r tîm, Lorena!

Cynadleddau a digwyddiadau

Cynhadledd y Ganolfan Dysgu Cymraeg Genedlaethol – 8fed Gorffennaf 2016

Gwahoddwyd arweinydd y prosiect, Dawn Knight, i roi anerchiad ynghylch CorCenCC yng nghynhadledd agoriadol y Ganolfan Dysgu Cymraeg Genedlaethol, a gynhaliwyd yng ngwesty'r Marriott yng Nghaerdydd ar 8fed Gorffennaf 2016. Enw'r cyflwyniad oedd 'Corpora and Pedagogy: developing the community-driven National Corpus of Contemporary Welsh', a'r nod oedd rhoi gwybod i'r cynrychiolwyr yno am gynlluniau cyffredinol y prosiect ac amlinellu sut y bydd y corpws o ddefnydd iddynt. Edrychwn ymlaen at gyd-drafod â'r Ganolfan drwy gydol y prosiect ac at roi diweddariadau iddynt ynglŷn â'n cynnydd. Llawer o ddiolch i Efa Gruffudd Jones (Prif Weithredwr y Ganolfan) a Helen Prosser (Cyfarwyddwr Strategol) am wahodd Dawn i siarad yn y digwyddiad hwn!

WISERD – 13^eg^-14^eg^ Gorffennaf, Prifysgol Abertawe

Torheulio yn y Bae

Roedd yr haul yn bendant yn tywynnu wrth i Dîm Rheoli CorCenCC (Dawn Knight, Tess Fitzpatrick a Steve Morris) deithio i gampws newydd Prifysgol Abertawe yn y Bae i roi papur o'r enw 'CorCenCC – Corpws Cenedlaethol Cymraeg Cyfoes (The National Corpus of Contemporary Welsh)' yng nghynhadledd flynyddol WISERD. Mae WISERD, Sefydliad Ymchwil Gymdeithasol ac Economaidd, Data a Dulliau Cymru, wedi'i leoli o fewn pum Prifysgol yng Nghymru: Abertawe, Aberystwyth, Bangor, Caerdydd a De Cymru, ac mae'n gweithio mewn partneriaeth â holl Brifysgolion y DU. Cynhadledd Gwyddorau Cymdeithasol fwyaf Cymru yw cynhadledd flynyddol WISERD, ac felly mae'n llwyfan berffaith ar gyfer trafod y draweffaith gymdeithasol, economaidd ac addysgol y bydd prosiect CorCenCC yn debygol o'i chael.

Eisteddfod Genedlaethol, Y Fenni

Yn ystod wythnos gyntaf mis Awst, ymwelodd aelodau tîm CorCenCC â'r Fenni, cartref Eisteddfod Genedlaethol Cymru 2016. Rhoddodd Cyd-ymchwilydd CorCenCC, Steve Morris, a dau Gynorthwyydd Ymchwil, Gareth Watkins a Mair Rees, gyflwyniadau ym mhebyll Prifysgolion Caerdydd ac Abertawe, i gyflwyno cysyniad CorCenCC ac amlinellu'r cynnydd hyd yma. Ychwanegodd un o lysgenhadon CorCenCC, Nia Parry, at y cyflwyniadau, gan siarad yn angerddol am ba mor arloesol mae'r prosiect a pha mor fuddiol y bydd e i siaradwyr a dysgwyr Cymraeg fel ei gilydd.

Yn y misoedd nesaf, byddwn ni'n barod i ddechrau casglu data iaith ar gyfer y corpws, ac elfen bwysig iawn fydd sgyrsiau a thestunau gan aelodau o'r cyhoedd. Darparodd yr Eisteddfod gyfle perffaith i brosiect CorCenCC gyrraedd cynulleidfa eang o siaradwyr Cymraeg, ac felly ymunodd Cydymaith Ymchwil, Steve Neale, a Chynorthwyydd Ymchwil, Jennifer Needs, â Gareth a Mair, er mwyn dosbarthu gwybodaeth ac ennyn diddordeb ymysg Eisteddfodwyr.

Dyma oedd profiad cyntaf Steve Neale o'r Eisteddfod, ac fe sy'n crynhoi ein hymweliad â'r Eisteddfod orau: *O safbwynt rhywun sy'n dysgu Cymraeg – ac un sy newydd ddechrau ei dysgu – roedd yr Eisteddfod yn lle hynod groesawgar. Mae'r balchder y mae pobl yn ei ddangos wrth ddefnyddio a hybu eu hiaith yn gwbl amlwg, fel y mae eu cefnogaeth i'r rhai sy'n ymddiddori mewn archwilio diwylliant Cymru a'i hiaith. Roedd ymatebion pobl i brosiect CorCenCC yn adlewyrchu'r teimlad hwn – mae yna ddiddordeb gwirioneddol ac amlwg mewn offer a syniadau a all roi'r Gymraeg o fewn cyrraedd i gymaint o bobl â phosib.*

Y flwyddyn nesaf, byddwn ni yn yr Eisteddfod unwaith eto – y tro hwn er mwyn dangos sut gallwch chi gyfrannu at y corpws drwy ap ar eich ffôn symudol. Felly os colloch chi ni yn y Fenni, dewch i'n gweld ni yn Ynys Môn yn 2017!

Diweddariadau ar Becynnau Gwaith (WPs 1-5) CorCenCC

Mae'r gwaith ar brosiect CorCenCC wedi'i ddosbarthu ar draws 6 phecyn gwaith cydgysylltiedig, y mae gan bob un ohonynt dasgau, nodau ac amcanion penodol. Mae WP0 yn cynnwys gweithgareddau dylunio, cwmpasu a gweithgareddau hyfforddi, ac yn cynnwys holl aelodau tîm y prosiect. Ceir diweddariadau byr ar Becynnau Gwaith 1-5 isod.

Nod: Casglu, trawsgrifio ac anonymeiddio enwau o'r data (arweinydd: Steve Morris)
Bu llawer o gynnydd yn y pecyn gwaith hwn dros yr ychydig fisoedd diwethaf. Mae rhestr o gonfensiynau trawsgrifio, ynghyd â rhestr o gonfensiynau anonymeiddio i'w defnyddio wrth adeiladu corpws CorCenCC wedi eu drafftio, eu dosbarthu i arbenigwyr yn y maes ar gyfer adborth, a'u cwblhau. Ar ben hynny, mae ein gwaith ar y fframwaith samplo, o ran ymateb i adborth gan arbenigwyr yr iaith Gymraeg ac arbenigwyr corpws, bron wedi'i gwblhau.

Nod: Datblygu'r set tagiau/tagiwr rhannau ymadrodd (arweinydd: Dawn Knight)

Yn WP2, rydym wedi bod yn sefydlu tasg anodi i ddarparu data wedi'i wirio â llaw a fydd yn sail i waith gwerthuso ein hoffer prosesu Cymraeg yn nes ymlaen.

Nod: Datblygu tagiwr semantig i'r Gymraeg a thagio ystyr yr holl ddata (arweinydd: Paul Rayson)

Mae WP3 wedi canolbwyntio ar ddatblygu cynffurf y tagiwr semantig Cymraeg, gan gynnwys adeiladu lecsicon semantig Cymraeg a fframwaith meddalwedd. Mae'r lecsicon semantig yn cynnwys 16,416 o lemata Cymraeg ar hyn o bryd, sydd wedi'u dosbarthu yn ôl categorïau semantig USAS sy'n cynnwys tua 174,780 o ffurfiau ffurfdroadol geiriau Cymraeg. Mae gwerthusiad cychwynnol yn dangos bod y lecsicon semantig cyfredol yn cwmpasu tua 72.42% o'r geiriau mewn testunau Cymraeg.

Nod: Cwmpasu/llunio'r pecyn cymorth pedagogaidd ar-lein (arweinydd: Enlli Thomas)

Peilotwyd holiadur gyda nifer fach o ymarferwyr ym maes addysgu Cymraeg ac, ar sail eu hadborth, datblygwyd holiadur newydd ar gyfer cynulleidfa fwy. Dosbarthwyd hwn i athrawon a thiwtoriaid Cymraeg mewn dwy gynhadledd ym mis Gorffennaf, a bydd fersiwn ar-lein hefyd yn cael ei gynhyrchu. Bydd yr arolwg hwn – ynghyd â grwpiau ffocws yn nes ymlaen eleni – yn ein helpu i adnabod y blaenoriaethau ar gyfer pecyn cymorth pedagogaidd CorCenCC.

Nod: Adeiladu seilwaith i gynnal CorCenCC a chreu'r corpws (arweinydd: Irena Spasić) Yn WP5, rydym wedi bod yn gweithio ar ddylunio'r isadeiledd cefndirol ar gyfer ein corpws, sefydlu prif wefannau ein prosiect, a pharatoi at ryddhau ein ap torfoli.

Cwrdd â'r tîm

Bob mis byddwn yn rhoi sylw i aelod gwahanol o'r tîm CorCenCC estynedig yn ein cylchlythyr. Bydd hyn yn rhoi cyfle i bawb ddweud ychydig wrthych chi am eu cefndir; beth maen nhw am ei weld o CorCenCC a sut maen nhw'n credu y gallai gyfrannu at eu gwaith eu hunain, neu yn fwy cyffredinol, at waith eraill yng Nghymru. Y mis yma mae'r sbotolau ar Mark Stonelake, sy'n Gyd-Ymchwilydd ym Mhrifysgol Abertawe.

Proffil: Mark Stonelake

Mae gwir angen am hyn. Pam nad yw e wedi cael ei wneud cyn hyn? Dyna'r hyn feddyliais pan glywais fod Corpws Cymraeg cynhwysfawr ar y gweill. Bues i wrth fy modd pan ofynnwyd i fi chwarae rhan fach yn y prosiect.

Cefais fy ngeni a magu yn Aberdâr yng nghymoedd De Cymru. Des i i Abertawe ym 1974 yn 17 oed, i fynychu'r Coleg Celf. Yn dilyn ymdrech aflwyddiannus i ddod yn artist, gwnes i nifer o swyddi 'go iawn' fel gwas fferm, tirmon, garddwr ac arborydd, cyn dychwelyd i addysg ffurfiol fel myfyriwr hŷn i wneud gradd yn y Gymraeg ym Mhrifysgol Abertawe

ym 1986. Dechreuais fel tiwtor Cymraeg i Oedolion rhan-amser tua'r un adeg. Wedi cyfnod byr fel ymchwilydd, des i'n diwtor/drefnydd llawn amser ym 1993, gan arbenigo mewn datblygu, trefnu a dysgu cyrsiau Cymraeg dwys iawn. Rhwng 1996 a 2014, bûm hefyd yn diwtor neu'n diwtor arweiniol ar gyrsiau blynyddol a drefnir yn yr Unol Daleithiau gan Gymdeithas Madog. Yn 2006 cefais fy mhenodi fel Swyddog Cwricwlwm ac Adnoddau yng Nghanolfan Cymraeg i Oedolion De-orllewin Cymru. Mewn ymarfer ailstrwythuro'r llynedd, fe'm penodwyd fel Rheolwr Cwricwlwm a Rhaglen. Yn dilyn ailstrwythuro pellach, oherwydd newidiadau mawr yn y maes eleni, dechreuaf swydd newydd fel Rheolwr Datblygiad Proffesiynol ac Ansawdd y mis hwn yng nghanolfan Dysgu Cymraeg Ardal Bae Abertawe yn Academi Hywel Teifi. Bydd y cwricwlwm yn rhan o'm gwaith o hyd, yn ogystal â threfnu'r rhaglen o gyrsiau Cymraeg, rheoli ansawdd a hyfforddiant.

Y to nesaf

Dros y blynyddoedd, rwyf wedi datblygu ystod eang o adnoddau ar gyfer tiwtoriaid Cymraeg a dysgwyr, gan gynnwys cyrsiau ar-lein a chyrsiau cyfunol ar bob lefel. Bûm yn rhan o wahanol brosiectau gyda nifer o wahanol sefydliadau megis: Y Lolfa, Y BBC a S4C, gan gynhyrchu adnoddau amrywiol o werslyfrau a llyfrau darllen graddedig i DVDau a chryno-ddisgiau. Yn ddiweddar, bûm yn datblygu deunyddiau ar gyfer ap S4C a ddefnyddir gyda rhaglen *Dal Ati* ar gyfer dysgwyr y Gymraeg. Gobeithio y bydd fy mhrofiad mewn datblygu adnoddau yn ddefnyddiol i'r tîm wrth greu'r pecyn cymorth pedagogaidd.

Mae gan y Corpws newydd y potensial i newid, nid dim ond cynnwys cyrsiau Cymraeg, ond hefyd y ffordd mae'r iaith yn cael ei dysgu. Bydd rhestrau amledd yn galluogi blaenoriaethu'r geiriau ac ymadroddion mwyaf eu defnydd a'r rhai mwyaf defnyddiol. Gellir datblygu cyrsiau fydd yn targedu grwpiau oedran arbennig, grwpiau cymdeithasol, gweithfeydd ac ardaloedd daearyddol arbennig , ar sail ymchwil fanwl gywir, a gall dysgwyr weld sut y defnyddir geiriau yn y cyd-destun cywir. Bydd y gallu i greu cyrsiau ar sail yr iaith lafar a ddefnyddir, yn hytrach na cheisio dyfalu'r hyn a ddefnyddir, yn gam enfawr ymlaen ym myd Dysgu Cymraeg i Oedolion. Bydd y Corpws yn adnodd gwych ar gyfer dylunwyr cyrsiau, ymchwilwyr, tiwtoriaid a dysgwyr. Ond ar ben hynny i gyd, pan glywaf ryw ben bach hunanbwysig yn dweud,"So ni'n gweud e fel 'ny", gallaf ddweud, "Ydych" a phrofi'r peth. Dyna beth yw cynnydd!

Mark Stonelake, e.m.stonelake@swansea.ac.uk

CorCenCC ar-lein

Mae'r wybodaeth ddiweddaraf am ddatblygiadau'r prosiect hefyd ar gael drwy Facebook www.facebook.com/CorCenCC/; Twitter https://twitter.com/corcencc (gallwch ein trydar @CorCenCC). Gallwch hefyd gysylltu â ni drwy anfon neges i gyfeiriad ebost y prosiect: corcencc@caerdydd.ac.uk neu ewch i'n gwefan, sef: http://sites.cardiff.ac.uk/corcencc/

Mae CorCenCC yn brosiect ymchwil a ariennir gan ESRC/AHRC (Grant Rhif ES/M011348/1). Mae tîm CorCenCC yn cynnwys y **Prif Ymchwilydd** - Dawn Knight; **y Cyd-Ymchwilwyr** - Tess Fitzpatrick, Steve Morris, Irena Spasić, Paul Rayson, Enlli Thomas, Mark Stonelake a Jeremy Evas; **y Cynorthwywyr Ymchwil** - Steven Neale, Jennifer Needs, Mair Rees, Scott Piao a Gareth Watkins; **Ymgynghorwyr** - Kevin Donnelly, Kevin Scannell, Laurence Anthony, Tom Cobb, Michael McCarthy a Margaret Deuchar; **Grŵp Ymgynghorol y Prosiect** – Colin Williams, Karen Corrigan, Llion Jones, Maggie Tallerman, Mair Parry-Jones, Gwen Awbery, Emyr Davies (CBAC-WJEC), Gareth Morlais (Llywodraeth Cymru), Owain Roberts (Llyfrgell Genedlaethol Cymru), Aran Jones (Saysomethingin.com) ac Andrew Hawke (Geiriadur Prifysgol Cymru).

Os oes gennych unrhyw sylwadau neu gwestiynau am gynnwys y cylchlythyr hwn, cysylltwch â Dr Dawn Knight: KnightD5@caerdydd.ac.uk