Cylchlythyr CorCenCC

Rhifyn 7: Hydref 2016

Corpws Cenedlaethol Cymraeg Cyfoes
National Corpus of Contemporary Welsh

Cyfarchion gan y Prif Ymchwilydd

Croeso i seithfed rhifyn cylchlythyr CorCenCC. Nawr bod tymor y cynadleddau wedi dirwyn i ben, rydym yn dechrau troi ein sylw at dasg a allai gael ei hystyried fel un o rai pwysicaf y prosiect: casglu data. Yn ystod yr wythnosau a'r misoedd nesaf, bydd aelodau o'r tîm yn gweithio'n galed i recriwtio cyfranogwyr a dechrau casglu'r 10 miliwn o eiriau llafar, ysgrifenedig a digidol (e-iaith) Cymraeg a fydd yn llenwi corpws CorCenCC [os ydych am ein helpu ni gyda hyn rhowch wybod i ni!]. Yn y cylchlythyr hwn, rydym yn cyflwyno Ap Torfoli CorCenCC. Bydd yr ap yn ein helpu wrth i ni fynd i'r afael â chasglu data iaith trwy ganiatáu i ddefnyddwyr recordio eu hiaith lafar eu hunain. Bydd yr ap yn cael ei lansio'n swyddogol yn ystod yr wythnosau nesaf.

Yn y cylchlythyr hwn, rydym hefyd yn rhoi'r newyddion diweddaraf o 5 prif becyn gwaith y prosiect, yn ogystal â rhannu newyddion cyffrous am gytundeb diweddar a wnaed rhwng CorCenCC ac S4C. Yn olaf, yn ein hadran 'cwrdd â'r tîm' rydym yn cyflwyno Tess Fitzpatrick, sy'n Athro Ieithyddiaeth Gymhwysol yn y Ganolfan Ymchwil Iaith a Chyfathrebu ac sy'n un o Gyd-Ymchwilwyr CorCenCC.

Mwynhewch! Dr Dawn Knight (Prifysgol Caerdydd)

Newyddion

Mae tîm CorCenCC yn falch o gyhoeddi bod cytundeb wedi ei lofnodi rhwng y prosiect ac S4C, gan eu gwneud yn bartneriaid ar y prosiect. Bydd y bartneriaeth hon yn ein galluogi i ddefnyddio data o'r sianel yn CorCenCC (o raglenni teledu i recordiadau o'u cyfarfodydd), ac yn helpu gyda gweithgareddau a digwyddiadau ymgysylltu â'r cyhoedd. Rydym yn awyddus i ddiolch yn fawr iawn i S4C am eu cydweithrediad – rydym yn edrych ymlaen at weithio gyda nhw ar y prosiect!

S4C

www.s4c.cymru/cy/

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)

Dros y misoedd diwethaf rydym wedi bod yn gweithio'n galed er mwyn cwblhau ein fframwaith samplo. Gwaith pwysig tu hwnt yw hwn, achos mi fydd y penderfyniadau a wnaed wrth lunio'r fframwaith yn dylanwadu ar y ffordd rydym yn mynd ati i gasglu data, ac felly hefyd ar gynnwys y corpws ei hun. Rydym ar fin mynd ati i gasglu data iaith Gymraeg, ac mae'r tîm i gyd yn frwdfrydig dros ben am hyn. Os ydych chi am gyfrannu eich data ysgrifenedig, llafar neu e-iaith, cysylltwch â ni ar bob cyfrif!

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

Mae anodwyr yn gweithio'n galed i neilltuo tagiau (o'n set tagiau rhannau ymadrodd) â llaw i ddetholiad bach o destunau Cymraeg ar hyn o bryd. Mae'r set dagiau wedi ei hadeiladu'n bwrpasol at y prosiect, a diben y gwaith anodi â llaw yw creu corpws 'safon aur'. Byddwn yn defnyddio'r corpws hwn yn y misoedd nesaf i werthuso'r offer prosesu sydd hefyd yn cael eu datblygu ar gyfer y pecyn gwaith hwn.

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

Yn WP3, mae gwaith llunio lecsicon semantig Cymraeg ar raddfa fawr ar gyfer y tagiwr semantig Cymraeg ar y gweill. Hyd yn hyn, rydym wedi casglu dros 329,000 o eiriau Cymraeg o amrywiaeth o adnoddau corpws. Darganfuwyd i'r casgliad hwn gwmpasu 97.63% o'r geiriau mewn corpws prawf.

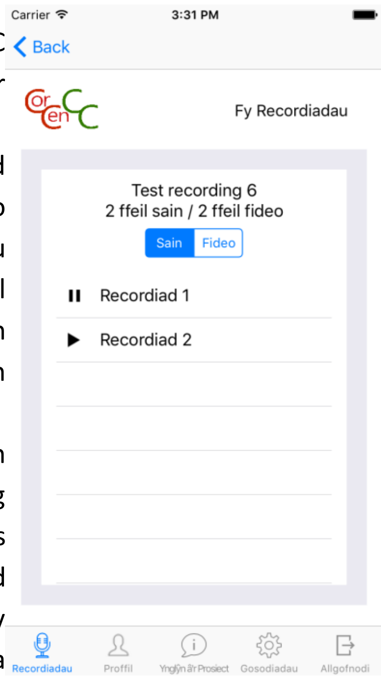
Nod: Cwmpasu/llunio'r pecyn cymorth pedagogaidd ar-lein (arweinydd: Enlli Thomas)

Lansiwyd fersiwn ar-lein ein holiadur ym mis Medi, ac rydym wedi llwyddo i gyrraedd athrawon, hyfforddwyr, darlithwyr a thiwtoriaid o ystod eang o gyd-destunau – o'r rhai sy'n addysgu mewn ysgolion cynradd (cyfrwng Cymraeg a chyfrwng Saesneg) i'r rhai sy'n addysgu oedolion. Llawer o ddiolch i bawb sydd eisoes wedi ymateb – bydd eich mewnbwn yn ein helpu i deilwra ein pecyn cymorth pedagogaidd yn ôl eich anghenion ac anghenion eich myfyrwyr. Os nad ydych chi wedi cael cyfle i gwblhau'r holiadur eto, dyw hi ddim yn rhy hwyr! Dilynwch y ddolen hon http://arolwg.corcencc.cymru/holiadur/holiadur-i-addysgwyr/ i rannu'ch gwybodaeth a'ch profiad â ni.

Nod: Adeiladu seilwaith i gynnal CorCenCC a chreu'r corpws (arweinydd: Irena Spasić)

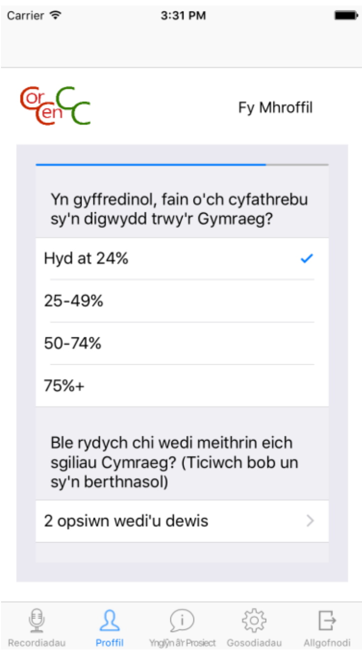
Bellach, mae Ap Torfoli CorCenCC wedi ei greu ac mae e ar fin cael ei ryddhau i'w brofi'n dechnegol. Rydym yn bwriadu lansio'r ap yn yr wythnosau nesaf, a'i ryddhau i'r gymuned er mwyn iddo gasglu eich data Cymraeg!

Ap Torfoli CorCenCC



Y mis hwn rydym wedi gorffen gwaith ar fersiwn beta gyntaf Ap Torfoli CorCenCC – neu'r 'ap' o hyn allan – sydd bellach yn cael ei brofi cyn inni ei ryddhau i'r cyhoedd am y tro cyntaf.

Mae'r ap, sydd yn gweithio ar iOS (mae cynlluniau i greu fersiwn Android ar y gweill), yn caniatáu i siaradwyr Cymraeg recordio eu sgyrsiau ar draws ystod o gyd-destunau er mwyn iddynt gael eu cynnwys yn y corpws terfynol. Un o'r apiau cyntaf o'i fath ar gyfer y diben yma yw hwn, gan fod casglu data corpws trwy ddull torfoli yn gyfeiriad cymharol newydd i'r maes. Wrth gwrs rydym yn mynd i ddefnyddio'r dull hwn law yn llaw â dulliau mwy traddodiadol o gasglu data iaith.



Mae'r ap yn gwneud y broses o gasglu iaith lafar yn hawdd ac yn golygu y gall siaradwyr Cymraeg ymgysylltu â phrosiect CorCenCC pan fo'n gyfleus iddyn nhw ei wneud. Mae hyn yn golygu bod cyfrannu at y corpws yn brofiad llawer mwy personol, gan roi i ddefnyddwyr berchnogaeth a rheolaeth ar eu data iaith eu hunain, yn hytrach na chael unrhyw un arall yn ymyrryd yn y broses mewn unrhyw ffordd. Drwy wneud y broses yn ddi-dor a thrwy sicrhau bod y siaradwyr eu hunain yn gyfrifol am eu cyfraniadau at y corpws, gall siaradwyr fod yn siŵr bod y recordiadau maent yn eu rhannu wir yn adlewyrchu eu Cymraeg pob dydd yn y ffordd fwyaf naturiol posibl.

Bydd yr ap ar gael i'w ddefnyddio yn rhad ac am ddim a bydd y fersiwn gyntaf yn cael ei lansio'n swyddogol yn yr wythnosau nesaf, ar ôl i'r cyfnod prawf ddirwyn i ben. Gallwch gadw golwg ar lansiad yr ap trwy ein sianeli arferol (hynny yw Facebook, Twitter a'r cylchlythyr) ac mae croeso i chi gysylltu â ni am fwy o wybodaeth os oes gennych ddiddordeb mewn cyfrannu eich data Cymraeg trwy ddefnyddio'r ap.

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 Yr Athro Tess Fitzpatrick, sy'n Gyd-Ymchwilydd ar brosiect CorCenCC ym Mhrifysgol Caerdydd

Proffil: Tess Fitzpatrick

Yn 2011, gwnes i gyflwyniad ym Mhrifysgol Newcastle am sut mae dysgwyr ieithoedd yn gwneud cysylltiadau rhwng y geiriau maen nhw'n eu gwybod a'r geiriau maen nhw'n eu dysgu. Dawn Knight gynhaliodd y digwyddiad, ac mewn caffe wedyn, trafodon ni gynllun i greu dyfais dysgu geirfa ginesthetig i blant ac awgrymais i ein bod ni'n ei pheilota mewn ysgolion Cymraeg. (Mae'r prosiect yna gyda'r teitl dros dro "matiau dawnsio/twister" ar y gweill o hyd.) Ar y trên adre, ces i decst gan Dawn: 'bydd angen i ni ddefnyddio'r corpws Cymraeg – oes modd i ti anfon lync ata i?'. Ymgynghorais i â Steve Morris, roeddwn i'n gweithio gyda fe yn Abertawe, ac fe anfonon ni nifer o ddolenni at Dawn, ond fe ddaeth yn amlwg yn syth nad oedd corpws cynhwysfawr, hygyrch o'r Gymraeg y byddai modd ei chwilio, yn bodoli. Dechreuodd egin gweledigaeth CorCenCC flaguro ar orwelion Tyneside a Bae Abertawe.

"mannau cychwyn stori CorCenCC ..."

Ar ôl gweithio'n achlysurol gyda Dawn a Steve, roeddwn i'n ymwybodol nid yn unig o'u harbenigedd (mewn ieithyddiaeth corpws ac iaith/pedagogeg y Gymraeg), ond hefyd o'r egni, yr ymrwymiad a'r hiwmor parod sydd ganddyn nhw wrth iddyn nhw fynd at eu gwaith – ac rwy'n credu y byddai'r tri ohonon ni'n cytuno, bum mlynedd yn ddiweddarach, fod yr un olaf o'r tri hyn wedi bod yn rhan hanfodol o sut rydyn ni'n ymwneud â'n gilydd. I mi, pe bai corpws Cymraeg yn cael ei greu, roedd yn ymddangos mai nhw oedd y bobl i wneud i hynny ddigwydd. Mewn cyfarfod hanner diwrnod o'r tri ohonon ni mewn caffe yng ngorsaf Temple Meads, Bryste, cafwyd cadarnhad fod sgìl, gwybodaeth (a hiwmor) y tîm craidd a gweledigaeth CorCenCC yn dod i'r fei; ar ôl dwy flynedd arall o drafodaethau rhithiol, a chyfarfodydd mewn lleoliadau gwahanol (cantîn yn Southampton, cyntedd gwesty ym Mryste, fy nghegin i), roedd tîm prosiect llawn gyda ni ac fe wnaethon ni gais am grant – roedd y weledigaeth yn glir.

"dyma ble cafodd rysáit CorCenCC ei hysgrifennu......"

Mae CorCenCC yn gwneud synnwyr i mi mewn ffyrdd sy'n ymwneud â'm profiad fy hun fel athrawes iaith, dysgwraig iaith, ieithydd cymhwysol – a rhywun sy'n byw yng Nghymru. Rhwng 1989 a 2003 pan oeddwn i'n gweithio fel athrawes iaith a hyfforddwraig athrawon Saesneg, addysgu aml-gyfrwng a deunyddiau hunan-gyrchu oedd offer safonol fy ngwaith (a llawer ohonynt wedi'u cynllunio ar sail corpws). Sylweddolais i nad oedd gan bob math o addysgu iaith yr un faint o adnoddau. Des i i ddeall hynny o'r galwadau am lyfrau cwrs sbâr gan wledydd sy'n datblygu, ac, yn gynnar yn y 1990au, o'r cyrsiau hyfforddi a arweiniais ar gyfer athrawon yn nwyrain Ewrop oedd yn diwallu'r galw newydd, brys am ddosbarthiadau Saesneg. Ar ddiwedd y 1990au dechreuais astudio'n rhan amser ar gyfer PhD mewn Ieithyddiaeth Gymhwysol, ac yn sydyn, roedd gen i bersbectif hollol newydd ar addysgu a dysgu iaith. Un canlyniad oedd ychwanegu'n sylweddol at fy nehongliad o'r ymadrodd "ag adnoddau llawn". Gellir crynhoi hynny fel "wedi'i gefnogi gan offer a deunyddiau sydd yn eu tro wedi'u seilio ar dystiolaeth ymchwil", ac mae hyn, i mi, yn diffinio'r cyfraniad y gall CorCenCC ei wneud i addysgu'r Gymraeg. Rwyf yn aelod o dîm y pecyn gwaith pedagogaidd, ac o wybod sut mae fy mhrofiad fy hun yn greiddiol i'm hymchwil mewn caffael ail iaith, rwy'n llawn cyffro am greu "pecyn cymorth" sydd wir yn cael ei hysbysu gan y rhai sy'n ei ddefnyddio.

Symudais i Gymru o Loegr yn fy 20au cynnar; ces i fy magu ym mherfeddion Caint, a symudais i Lundain yn 18 oed. Tyfodd fy ymwybyddiaeth o werth byw mewn cymuned ddwyieithog yn raddol iawn, a heb fawr o sylw yn y lle cyntaf – roeddwn yn byw yma am lawer o flynyddoedd cyn mynd i ddosbarth Cymraeg. Yn fwyfwy, rwyf wedi cael fy hudo gan gymhlethdod y cysylltiadau rhwng iaith a chymuned, dylanwad cynnil y dirwedd ddwyieithog, a'r synnwyr o allu cyrchu peuoedd newydd sydd wedi bod yn rhan o'm profiad fel dysgwraig, o'm hymdrechion sigledig cyntaf i siarad Cymraeg (rhyw "paned o de" araf, gofalus wrth i Mark Stonelake ofyn "beth ti'n moyn" mewn caffe – roeddwn i eisiau coffi ond ces i bwl o banig). Rwy'n teimlo ei bod yn fraint byw mewn rhan ddwyieithog o'r byd a chael gweithio gyda thîm (amlieithog) CorCenCC.

"Yn archwilio tirluniau eraill Cymru"

Tess Fitzpatrick, FitzpatrickT@caerdydd.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

Cardiff University Prifysgol Caerdydd · Swansea University Prifysgol Abertawe · Lancaster University · Prifysgol Bangor University · E·S·R·C Economic & Social Research Council · Arts & Humanities Research Council