

Gwell offer technoleg cyfieithu ar gyfer y Diwydiant Cyfieithu yng Nghymru: Arolwg dadansoddol

Adroddiad Terfynol

2009

Project Rhif.: HE 06FSP 1014

Paratowyd gan
Uned Technolegau Iaith, Canolfan Bedwyr, Prifysgol Bangor

Delyth Prys, Gruffudd Prys, Dewi Bryn Jones



Cynnwys

- 0 Crynodeb gweithredol
- 1 Rhagarweiniad
- 2 Nodau a chyd-destun
- 3 Methodoleg
- 4 Canfyddiadau allweddol
 - 4.1 Trosolwg o'r Diwydiant Cyfieithu yng Nghymru
 - 4.1.1 Maint y Diwydiant Cyfieithu
 - 4.1.2 Rhaniad Merched/Dynion o fewn y Diwydiant
 - 4.1.3 Ieithoedd Cyfieithu
 - 4.1.4 Cymhwysedd Iaith Frodorol
 - 4.1.5 Cymwysterau Cyfieithwyr
 - 4.1.6 Cyflogaeth Cyfieithwyr
 - 4.1.7 Swyddogaeth yr Uned Technolegau Iaith
 - 4.2 Y Defnydd o Dechnoleg Cyfieithu yng Nghymru
 - 4.2.1 Diffiniad o Dechnoleg Cyfieithu
 - 4.2.2 Y Defnydd Rhyngwladol o Dechnoleg Cyfieithu
 - 4.2.3 Cymariaethau gyda'r Defnydd yng Nghymru
 - 4.2.4 Gwahaniaethau Iaith-benodol yn y Defnydd o Dechnoleg Cyfieithu
 - 4.2.5 Offer Technoleg Cyfieithu Penodol
 - 4.2.6 Diffyg Arbenigedd Technegol ymhlith Cyfieithwyr
 - 4.2.7 Offer Technoleg Cyfieithu a Ddatblygwyd yng Nghymru
 - 4.3 Y Diwydiant Cyfieithu ac Economi Cymru
 - 4.3.1 Arwyddocâd Economaidd y Diwydiant Cyfieithu
 - 4.3.2 Arbedion Ariannol drwy Ddefnyddio Technoleg Cyfieithu
 - 4.3.3 Effaith Anuniongyrchol ar Economi Cymru
- 5 Casgliadau
- 6 Argymhellion

Atodiadau (yn Saesneg)

- Atodiad 1: Astudiaethau Dichonolrwydd
 - Astudiaeth Dichonolrwydd 1: Cymhorthion hygyrchedd
 - Astudiaeth Dichonolrwydd 2: Adnoddau cynnwys

Astudiaeth Dichonolrwydd 3: Adnoddau meddalwedd

- Atodiad 2: Trafodaethau Sgopio
- Trafodaeth Sgopio 1: Cyfieithwyr llawrydd
 - Trafodaeth Sgopio 2: Gwasanaethau cyfieithu'r sector cyhoeddus
 - Trafodaeth Sgopio 3: Cwmnïau cyfieithu preifat
 - Trafodaeth Sgopio 4: Cyfieithwyr anabl
 - Trafodaeth Sgopio 5: Cwsmeriaid y diwydiant cyfieithu

Atodiad 3: Cynllunio ar gyfer y Dyfodol – Papur Trafod

0. Crynodeb gweithredol

- 0.1 Mae'r diwydiant cyfieithu yn sector sylweddol yn economi Cymru ac mae'n gyflogwr pwysig ar gyfer merched, gan ddarparu cyflogaeth sy'n gofyn am sgiliau uwch mewn ardaloedd trefol, lled drefol a gwledig.
- 0.2 Mae'r diwydiant yn cynnal ac yn cefnogi gwasanaethau sy'n cael eu darparu'n ddwyieithog yng Nghymru, ac yn cynorthwyo sectorau eraill o economi Cymru i allforio eu nwyddau a'u gwasanaethau yn y farchnad fyd-eang.
- 0.3 Mae'r defnydd o offer technoleg cyfieithu, beth bynnag fo'r ieithoedd a gyfieithir, ar hyn o bryd yn annigonol yn y diwydiant cyfieithu yng Nghymru. Byddai dod â'r diwydiant i fyny i lefel yr ymarfer rhyngwladol gorau cyfoes yn fanteisiol i'r diwydiant cyfieithu ei hun a hefyd i economi Cymru yn gyffredinol, yn enwedig ei busnesau allforio.
- 0.4 Manteision creiddiol defnyddio technoleg cyfieithu a gweithio'n agos gyda chwsmeriaid y diwydiant cyfieithu yw:
- cynyddu'r gallu i gyfieithu o 40%, ac arbed 20% mewn amser gweinyddol, heb unrhyw gynydd mewn lefelau staffio, drwy ddefnyddio technoleg cyfieithu'n briodol
 - twf pellach o 50% yn y sector drwy gynyddu'r gallu i gyfieithu i gyfarfod â'r galw domestig, a thwf o 300% mewn denu busnes cyfieithu o'r tu allan i Gymru
 - cynyddu cyfleoedd allforio i gwsmeriaid presennol a phosibl o 19% drwy wneud defnydd priodol o wasanaethau cyfieithu ac amlieithog.
- 0.5 Nid yw'r diwydiant cyfieithu yng Nghymru, hyd yn oed ar gyfer y pâr ieithoedd Cymraeg-Saesneg, yn gweithredu yn ynysig y tu allan i'r farchnad gyfieithu fyd-eang. Mae darparwyr gwasanaethau cyfieithu sydd wedi'u lleoli yng Nghymru yn eu cael eu hun fwyfwy yn cystadlu gyda chwmnïau o rannau eraill o Brydain, yr Undeb Ewropeaidd a'r byd sydd wedi buddsoddi mewn technolegau cyfieithu er mwyn gwasanaethu marchnadoedd ac ieithoedd mwy lleol. Felly ceir bygythiad i'r diwydiant cyfieithu yng Nghymru ar hyn o bryd oherwydd bod cwmnïau cyfieithu mawr o'r tu allan sy'n fwy technolegol-ymwybodol yn dechrau treiddio i'r farchnad. Yr unig ffordd i wrthweithio hyn yw drwy ddarparu'r modd i'r diwydiant o fewn Cymru ddod yn fwy technolegol gymwys ei hunan.
- 0.6 Ymhlith argymhellion ar gyfer gweithredu yn y dyfodol mae:
- sefydlu canolfan arddangos ar gyfer y diwydiant cyfieithu a'i gwsmeriaid yng Nghymru er mwyn arddangos y manteision o ddefnyddio technoleg cyfieithu
 - ymgymryd â rhaglen o goladu offer technoleg cyfieithu perthnasol ar gyfer 'bocs tŵls cyfieithwyr' i'w ddefnyddio gan y diwydiant cyfieithu yn Nghymru
 - gweithio gyda chwsmeriaid presennol a phosibl y diwydiant cyfieithu i'w galluogi i gymryd rhan yn llawnach yn yr economi amlieithog, fyd-eang
 - adnabod nodau mwy tymor hir ar gyfer ymchwilio a datblygu technoleg cyfieithu newydd yng Nghymru.

1. Rhagarweiniad

1.1. Mae'r adroddiad hwn yn rhoi manylion yr astudiaethau dadansoddol a gynhaliwyd rhwng Hydref 2008 a Chwefror 2009 yn y project *Gwell Offer Technoleg Cyfieithu ar gyfer y Diwydiant Cyfieithu yng Nghymru*. Ariannwyd y project hwn gan Lywodraeth Cynulliad Cymru dan y Rhaglen Datblygu Capasiti ac Ecsbloetio Gwybodaeth, Arbenigedd Academiaidd i Fusnesau (A4B).

1.2. Briff y project oedd ymchwilio i ddichonolrwydd ecsbloetio gwybodaeth o'r gronfa wybodaeth ym Mhrifysgol Bangor i'r diwydiant cyfieithu yng Nghymru, gan gymryd ystyriaeth lawn o natur y diwydiant yng Nghymru, y cyfartaledd uchel o unig fasnachwyr a micro SMEs sydd yma a'i swyddogaeth allweddol fel darparwr gwasanaethau cyfieithu i ddiwydiant Cymru yn gyffredinol. Mae anghenion diwydiant Cymru yn bodoli yng nghyd-destun y farchnad amlieithog, fyd-eang, ac yn cynnwys anghenion allforwyr a buddsoddiad o'r tu allan.

1.3. Mae Gwell Offer Technoleg Cyfieithu ar gyfer y Diwydiant Cyfieithu yn golygu datblygu a darparu adnoddau megis terminoleg gymwys, geiriaduron, gwirwyr sillafu a gramadeg, systemau ac ychwanegion cof cyfieithu, meddalwedd cyfieithu gyda chymorth cyfrifiadur, systemau rheoli llif gwaith a'r holl gymhorthion cyfieithu technolegol newydd.

1.4 Mae defnydd priodol o dechnoleg cyfieithu yn dod ag arbedion sylweddol mewn effeithiolrwydd ac ansawdd i'r diwydiant cyfieithu a'i gwsmeriaid. Byddai bocs tŵls cyfieithydd yn dwyn ynghyd adnodau allweddol i wella'r broses gyfieithu, o ran gwella cynhyrchiant, rheoli ansawdd, a bodlonrwydd cwsmeriaid.

1.5 Mae'r adroddiad hwn yn berthnasol i gyfieithwyr proffesiynol a'r diwydiant cyfieithu yng Nghymru, ond hefyd i'w cwsmeriaid a'u cwsmeriaid posibl yn niwydiant Cymru yn gyffredinol. Mae potensial gwasanaethau iaith a chyfieithu i gefnogi gweithgareddau allforio a marchnata diwydiant Cymru yn y farchnad amlieithog, fyd-eang, wedi cael eu tanbriso yn sylweddol yn y gorffennol. Bydd galluogi'r diwydiant cyfieithu yng Nghymru i ddarparu'r gwasanaethau hyn i weddill diwydiant Cymru yn gyfraniad sylweddol i lwyddiant economaidd Cymru.

2. Nodau a chyd-destun

2.1 Amcanion penodol y project oedd:

- Archwilio i ddichonolrwydd rhaglen ddatblygu gydlynol o gymhorthion cyfieithu a fyddai'n arwain at focs tŵls cyfieithu ar gyfer y diwydiant cyfieithu yng Nghymru.
- Adnabod a blaenoriaethu cydrannau allweddol megis pyrth ar gyfer geiriaduron electronig, meddalwedd cof cyfieithu, meddalwedd cyn- ac ôl-gyfieithu, a systemau rheoli dogfennau dwyieithog, a'u darparu ar gyfer y sector darged.
- Adnabod unrhyw fylchau yn y ddarpariaeth, eu gwerthuso o ran eu heffaith yn gwella proffidoldeb y diwydiant cyfieithu yng Nghymru a darparu cyfleoedd i'r diwydiant meddalwedd ddatblygu cynnyrch newydd.
- Gweithio gyda Chymdeithas Cyfieithwyr Cymru (CCC), y corff proffesiynol ar gyfer cyfieithwyr yng Nghymru a gweithio gyda'i haelodau mewn ymarferiad Sgopio i fapio allan anghenion y diwydiant a chynhyrchu cynllun datblygu gwasanaeth.
- Ystyried bocs tŵls cyfieithu fel dull o ddwyn ynghyd adnoddau allweddol i wella'r broses gyfieithu, o ran gwella cynhyrchiant, rheolaeth ansawdd a bodlonrwydd cwsmeriaid.

2.2 Cyd-destun:

Mae cymhorthion cyfieithu ar gyfer y diwydiant cyfieithu yng Nghymru ar hyn o bryd yn wasgaredig a thameidiog, ac nid ydynt yn asio gyda'i gilydd yn amgylchedd cydlynol neu lif gwaith esmwyth. Mae cyfran sylweddol o'r diwydiant yn gwmnïau bach iawn neu yn unigolion llawrydd yn gweithio ar eu liwt eu hun. Yn aml iawn nid oes ganddynt lawer o gefnogaeth ac arbenigedd dechnolegol, ac mae'r offer sydd ar gael naill ai'n rhy gymhleth neu'n rhy ddrud, neu nid ydynt wedi cael eu datblygu ar gyfer amgylchedd penodol y diwydiant cyfieithu yng Nghymru.

Awgrymodd ymchwil blaenorol nad oedd y diwydiant cyfieithu yng Nghymru yn defnyddio llawer ar offer technoleg cyfieithu. Bydd datblygu ac integreiddio'r offer angenrheidiol yn gwella cyflymder, cywirdeb ac effeithiolrwydd y diwydiant cyfieithu. Gallai hefyd gadw'r costau i lawr ar gyfer cwsmeriaid, lle mae costau cyfieithu uchel yn rhwystr i ddatblygu eu busnes dwyieithog ac amlieithog.

3. Methodoleg

- 3.1 Gweithiodd y project yn agos gyda Chymdeithas Cyfieithwyr Cymru wrth ymchwilio'r project hwn. Rhoddodd y Gymdeithas fynediad at ei haelodaeth broffesiynol ei hun, a gweithredu hefyd fel cynghorwyr arbenigol ar y diwydiant cyfieithu yng Nghymru. Ymdrechwyd hefyd i gysylltu â darparwyr gwasanaeth cyfieithu yng Nghymru nad oeddent yn aelodau o CCC, yn arbennig y rhai oedd yn cynnig gwasanaethau cyfieithu i nifer o ieithoedd gwahanol a chyfuniadau anarferol o ieithoedd. Ceisiwyd cyngor ychwanegol gan gwsmeriaid y gwasanaeth cyfieithu, yn ogystal â phobl mewn diwydiant a busnes nad ydynt ar hyn o bryd yn defnyddio gwasanaethau cyfieithu, ond y gallai eu busnesau elwa ar eu defnyddio.
- 3.2 Cynlluniwyd y project i gynnwys tair astudiaeth dichonolrwydd, un yn canolbwyntio ar adnoddau cynnwys ar gyfer cyfieithwyr, un ar adnoddau meddalwedd, a'r drydedd ar gymhorthion hygyrchedd ar gyfer cyfieithwyr anabl. Yr oedd y project hefyd yn cynnwys pum trafodaeth sgopio, un yr un gyda chyfieithwyr anabl, cyfieithwyr llawrydd, gwasanaethau cyfieithu sector gyhoeddus, swyddfeydd cyfieithu preifat, ac un gyda chwsmeriaid presennol a phosibl gwasanaethau cyfieithu yng Nghymru.
- 3.3 Defnyddiodd y project holiaduron wedi'u targedu, cyfweliadau wyneb yn wyneb a dros y ffôn, ymchwil desg, ymweliadau personol a chyflwyniadau grŵp fel dulliau ymchwil. Gyda chymorth CCC, sefydlodd y project dri arddangosiad technoleg cyfieithu, a'u treialu gyda chyfanswm o 78 mynychwr mewn dwy gynhadledd, un yng ngogledd a'r llall yn ne Cymru.
- 3.4 Casglwyd gwybodaeth am gyfieithwyr yng Nghymru, beth bynnag oedd eu hieithoedd cyfieithu, o gyfeirlyfrau cyhoeddus o gyfieithwyr, hysbysebion, proffiliau ar-lein, gwefannau personol a chyswllt uniongyrchol.
- 3.5 Ni ofynnwyd am restrï o fanylion cyfieithwyr yn uniongyrchol gan asiantaethau cyfieithu gan fod yr asiantaethau hynny yn ystyried y rhestrï cyswllt hyn fel asedau masnachol hynod o werthfawr.
- 3.6 Cafwyd manylion am gyfieithwyr rhwng y Gymraeg a'r Saesneg o wefan CCC, ac o ffigurau cyffredinol a darparwyd yn uniongyrchol gan y Gymdeithas ei hun.
- 3.7 Gan nad oes cymdeithas benodol yng Nghymru ar gyfer cyfieithwyr sy'n cynnig parau iaith eraill ond sy'n gweithio o fewn Cymru, lle bynnag roedd modd casglwyd eu manylion o gyfeirlyfrau cymdeithasau Prydain-gyfan megis yr *Institute of Chartered Linguistics* a'r *Institute of Translators and Interpreters* (ITI). Lle roedd modd canfod bod cyfieithwyr yn gweithio yng Nghymru, casglwyd manylion hefyd o farchnadoedd cyfieithu ar-lein lle mae cyfieithwyr yn hysbysebu eu gwasanaethau.
- 3.8 Gan fod cadw cyfrinachedd a phreifatrwydd cyfieithwyr a chwmnïau cyfieithu yn hanfodol i sicrhau eu bod yn cyfrannu'n agored ac yn onest i'r project, nid yw'r adroddiad ar y trafodaethau Sgopio yn cynnwys unrhyw fanylion a allai olygu bod modd adnabod cyfieithwyr unigol neu ddarparwyr gwasanaeth cyfieithu.

4. Canfyddiadau allweddol a thrafodaeth

4.1 Trosolwg o'r diwydiant cyfieithu yng Nghymru

4.1.1 Maint y Diwydiant Cyfieithu

Y mae'r diwydiant cyfieithu yn ddiwydiant cymharodd newydd yn Nghymru, a phrofodd dwf aruthrol yn y 1990au a dechrau'r 2000au o ganlyniad i basio Deddf yr Iaith Gymraeg a sefydlu Cynulliad Cenedlaethol Cymru (gan greu galw yn y farchnad ddomestig) a thwf y rhyngwyd a'r economi fyd-eang (gan greu galw rhyngwladol amlieithog).

Gwnaed amryw o ymdrechion i fesur maint y diwydiant yng Nghymru. Gan nad oes unrhyw ofyniad cofrestru statudol fel y cyfryw ar aelodau'r diwydiant, mae cael hyd i ateb awdurdodol wedi bod yn broblemus yn y gorffennol. Fodd bynnag, mae arolwg o faint y diwydiant iaith yn Ewrop yn cael ei gynnal ar hyn o bryd ar ran y Comisiwn Ewropeaidd, gyda'r nod o ddeall maint y diwydiant yn yr Undeb Ewropeaidd, o ran gwerth economaidd, strwythur a phrif yrwyr¹.

Yn 2005, gyrodd *Arolwg o'r Diwydiant Cyfieithu yng Nghymru* (Menter a Busnes ar gyfer CCC a Bwrdd yr Iaith Gymraeg 2005) 900 holiadur i randdeiliaid allweddol yng Nghymru. Derbyniwyd 171 ateb, gan gyfieithwyr yn y sectorau preifat, cyhoeddus a gwirfoddol yn ogystal â chomisiynwyr cyfieithu. Adroddodd yr arolwg hwn yn unig ar gyfieithwyr oedd yn gweithio gyda'r pâr ieithoedd Cymraeg a Saesneg, ac ni chyfeiriwyd at gwmnïau cyfieithu yn gweithio yng Nghymru gyda ieithoedd ar wahân i Gymraeg a Saesneg². Fodd bynnag, nid oes arolwg cyffredinol arall o'r diwydiant cyfieithu wedi'i gynnal yng Nghymru hyd y gwyddom, felly mae'r nifer o 900 rhanddeiliad allweddol yn y diwydiant cyfieithu yn weithgar yng Nghymru wedi'i gymryd fel gwaelodlin ar gyfer yr astudiaeth gyffredol.

At ddibenion yr astudiaeth hon *Gwell Offer Technoleg Cyfieithu ar gyfer y Diwydiant Cyfieithu yng Nghymru*, casglwyd gwybodaeth oddi wrth gyfanswm o 505 cwmni cyfieithu yn cynnwys unig fasnachwyr, yn gweithio yng Nghymru. Mae hyn yn cynnwys cwmnïau yn gweithio gyda phob iaith heb gyfyngiad. Nid yw'r rhestr hon yn gwbl gynhwysfawr; roedd hi'n anodd cael manylion rhai cyfieithwyr sy'n gweithio i gwmnïau preifat, yn arbennig os nad ydynt yn aelodau o gymdeithas gyfieithu broffesiynol, gan eu bod yn masnachu o dan enw'r cwmni yn hytrach nag fel unigolion. Fodd bynnag, y mae'r rhestr yn cynnwys y rhan fwyaf o unig fasnachwyr yng Nghymru sy'n hysbysebu eu gwasanaethau yn gyhoeddus.

Yn ychwanegol at unig fasnachwyr, canfuwyd cyfanswm o 17 darparwr gwasanaethau cyfieithu preifat a oedd yn cyflogi timau cyfieithu mewnol. Unwaith eto y mae hwn yn ffigur ceidwadol, gan ei bod hi'n bosib na lwyddwyd i adnabod rhai ohonynt, am y gall hi fod yn anodd yn aml gwahaniaethu rhwng cwmnïau bychain sy'n cynnwys dau neu dri cyfieithydd a chyfieithwyr sy'n gweithio ar eu pen eu hun.

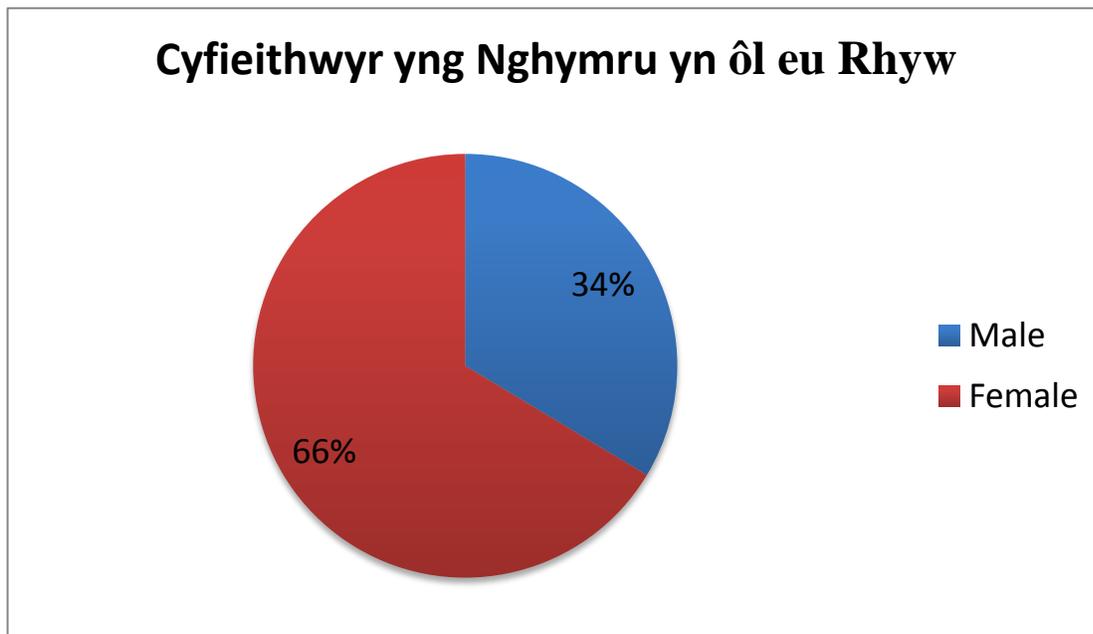
¹ Cafwyd manylion ar wefan LTC: <http://www.langtech.co.uk/content/view/82/1/>

² *Arolwg o'r Diwydiant Cyfieithu yng Nghymru* (Menter a Busnes ar gyfer CCC a Bwrdd yr Iaith Gymraeg 2005) <http://www.byig-wlb.org.uk/cymraeg/cyhoeddiadau/>

Casglwyd hefyd wybodaeth ynghylch cyfieithu yn y sector gyhoeddus, gan fod nifer helaeth o gyfieithwyr yn ymwneud â chyfieithu Cymraeg/Saesneg ar gyfer cyrff sector gyhoeddus, gan gynnwys awdurdodau lleol. Cafwyd bod unedau cyfieithu mewn deuddeg allan o'r 22 awdurdod lleol yng Nghymru. Cafwyd bod llawer o gyrff eraill megis awdurdodau heddlu ac ymddiriedolaethau lleol hefyd yn defnyddio gwasanaethau cyfieithwyr, naill ai yn fewnol neu'n allanol.

4.1.2 Rhaniad Merched/Dynion o fewn y Diwydiant

Lle roedd modd, casglwyd data ar nifer y merched a'r dynion oedd yn cyfieithu. O'r 455 cyfieithydd yr oedd eu rhyw yn hysbys, roedd 66% yn ferched a 34% yn ddynion.



Ffigur 1: Cyfieithwyr yng Nghymru yn ôl eu Rhyw

Ceir sawl rheswm am y cyfartaledd uwch o ferched yn gweithio yn y proffesiwn. Adroddodd nifer o gyfieithwyr eu bod wedi dod yn aelodau o'r proffesiwn cyfieithu ar ôl dod yn famau, gan ei fod yn eu galluogi weithio o'u cartref a gofalu am eu plant, ac ennill cyflog a oedd yn adlewyrchu eu cymhwyster academiaidd (gweler isod Cymwysterau am ffigurau ar hyn) a'u harbenigedd mewn ieithoedd.

Ffactor bosibl arall am y cyfartaledd uchel o gyfieithwyr benywaidd yw'r cyfartaledd uwch o ferched sy'n astudio ieithoedd yn gyffredinol. Er enghraifft, mae merched yn cyfrif am 66% o ddsbarthiadau TGAU mewn Ieithoedd Modern Tramor yng Nghymru, ac mae'r anghydbwysedd hwn yn bwydo trwodd i Addysg Uwch³.

Nodwedd arall am gyfieithwyr benywaidd yng Nghymru oedd y nifer o gyfieithwyr a ganfuwyd oedd ag enwau cyntaf a oedd yn ymddangos fel rhai o darddiad tramor, ond gyda chyfenwau

³ *MFL in Wales: Building on Bilingualism*, CILT Cymru (gw. <http://www.ciltcymru.org.uk>)

nodweddiadol Gymreig. Yn y mwyafrif o achosion yr edrychwyd arnynt, canfuwyd fod y cyfieithwyr dan sylw wedi symud i Gymru o wledydd eraill ar ôl priodi gwŷr o Gymru ac ymgartrefu yng Nghymru. Gellir gwneud gwaith cyfieithu o bell ar gyfer y cwsmer, felly mae gan gyfieithwyr (cyfieithwyr llawrydd yn arbennig) yn aml fwy o ryddid i adleoli na gwŷr neu wragedd sydd mewn proffesiwn lle maent yn gaeth i leoliad arbennig, megis swyddfa neu ffatri. Rheswm arall mai'r cyfieithydd yw'r cymar sy'n adleoli yw eu bod yn medru cadw eu cwsmeriaid yn eu gwlad wreiddiol er eu bod yn symud cartref, ac y byddant hefyd yn medru ennill cwsmeriaid newydd yn eu cartref newydd. Gall yr hyblygrwydd yma hefyd symbylu rhai gwŷr neu wragedd priod sy'n symud i Gymru ac sy'n meddu ar gymwysterau iaith i sefydlu'u hunain fel cyfieithwyr, er nad oes ganddynt efallai brofiad proffesiynol blaenorol o gyfieithu.

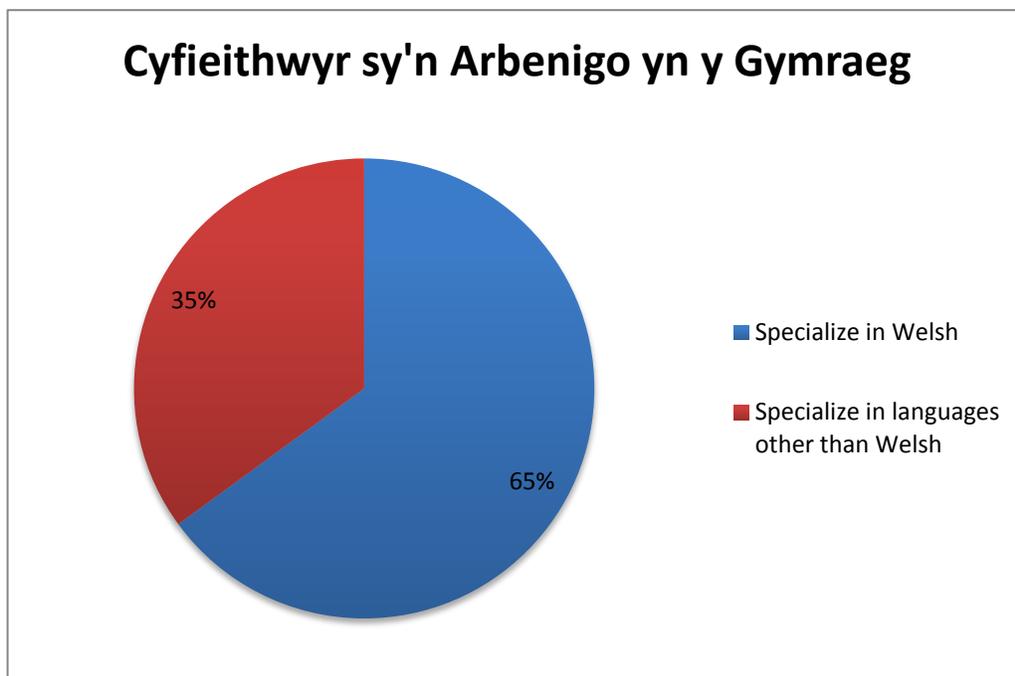
Beth bynnag yw'r rhesymau am y cyfartaledd uchel o ferched yn y diwydiant cyfieithu yng Nghymru, y mae ganddo oblygiadau ar gyfer y defnydd o gymhorthion technolegol i gyfieithu, am fod nifer o astudiaethau wedi dangos fod merched yn llai tebygol na dynion i ddangos diddordeb mewn TG ac astudiaethau cyfieithu⁴.

4.1.3 Ieithoedd Cyfieithu

Yn ystod y project hwn, canfuwyd bod mwyafrif cyfieithwyr Cymru yn gwasanaethu'r pâr ieithoedd Cymraeg a Saesneg. Mewn gwirionedd, cafwyd fod cyfartaledd y cyfieithwyr a oedd yn arbennig yn y pâr hyn yn 65% o holl gyfieithwyr Cymru. Gellir priodoli hyn i'r galw am gyfieithu Cymraeg ymhlith cyrff yn y sector gyhoeddus. Fodd bynnag yn ystod y blynyddoedd diweddar gwelwyd mwy o alw am wasanaethau cyfieithu oddi wrth gwmnïau preifat, a allai gael ei gyfyngu eto os caiff Deddf yr Iaith Gymraeg ei chryfhau ymhellach.

⁴ Er enghraifft, mae Estynyn adrodd fod bron 3 gwaith y nifer o fechgyn ag o ferched yn cofrestru am Astudiaethau Cyfrifiadur a TG i safon AS ac A (gw. http://www.estyn.gov.uk/publications/cy_Remit_45.pdf)

Cyfieithwyr sy'n Arbenigo yn y Gymraeg



Ffigur 2: Cyfieithwyr sy'n arbenigo yn y Gymraeg

Yr ieithoedd tramor yr adroddwyd sy'n cael eu cyfieithu yng Nghymru, a'u poblogrwydd cymharol, yw Ffrangeg, Almaeneg, Sbaeneg, Pwyleg, Lithiwaneg, Arabeg, Rwsieg, Tseineeg, Eidaleg, Groeg, Iseldireg, Twrceg, Romaneg, Ffineg, Tsiec, Portiwgeeg, Swedeg, Malaeg, Hindi, Ffarsi, Bengaleg. Adroddodd cyfieithwyr sy'n byw ac yn gweithio yng Nghymru eu bod yn gweithio gyda phob un o'r ieithoedd uchod.

Sylwer, fodd bynnag, ei bod hi'n anodd gwybod yn sicr a yw asiantaeth gyfieithu sydd â'i swyddfa yng Nghymru, ac sy'n hysbysebu ei gwasanaethau yn cyfieithu i Ffarsi, er enghraifft, yn defnyddio cyfieithydd sy'n byw yng Nghymru, neu yn gosod y gwaith i gyfieithydd yn byw yn Iran, neu yn rhywle arall. Yn yr un modd, mae'n debyg bod y nifer o gyfieithwyr sy'n byw yng Nghymru ac yn gweithio gyda ieithoedd tramor yn cael ei dan-adrodd, gan y gallant fod yn gweithio i asiantaethau cyfieithu sydd â'u swyddfeydd mewn gwledydd eraill.

Er bod nifer helaeth o ieithoedd yn cael eu cynnig gan y diwydiant cyfieithu yng Nghymru, gall i nifer o gyfieithwyr sydd ar gael i wasanaethu ieithoedd penodol fod yn isel. Achos felly yw'r iaith Hindi, gydag ond un cyfieithydd yn cael ei adnabod fel un sy'n arbenigo mewn Hindi, er gwaethaf amlygrwydd economaidd cynyddol India ar lwyfan y byd.

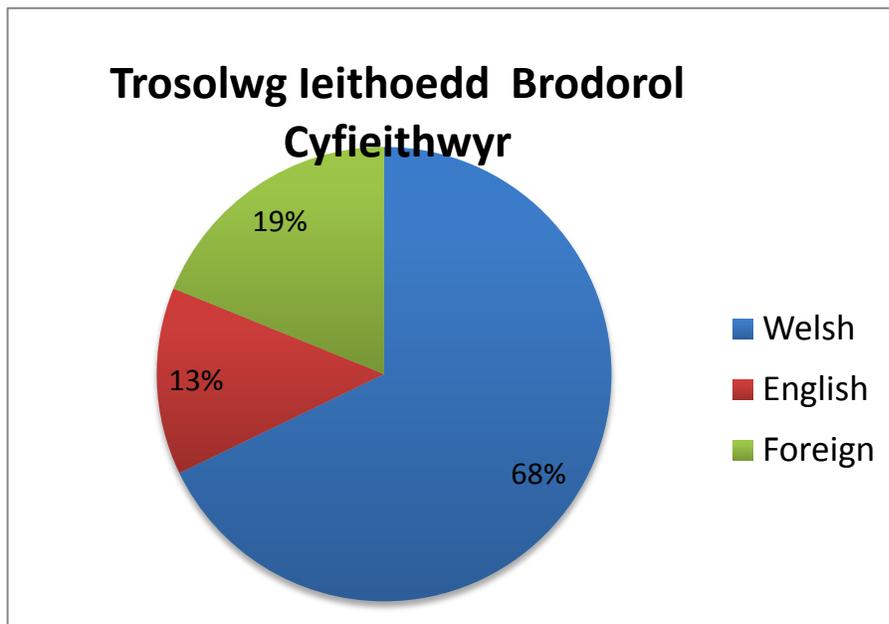
Mae'r grŵp ieithoedd Ffrangeg, Eidaleg, Almaeneg a Sbaeneg (FIGS), sef prif ieithoedd economaidd Ewrop, yn cael eu gwasanaethu'n dda. Mae hyn yn adlewyrchu'r ffaith mai'r rhain yw'r ieithoedd tramor traddodiadol sy'n cael eu dysgu yn ein hysgolion, ac y gall dinasyddion y gwledydd hyn setlo a byw yn rhydd yn y Deyrnas Unedig dan gyfraith yr Undeb Ewropeaidd.

Fodd bynnag mae'r faith fod llai o wasanaethau cyfieithu ar gael i farchnadoedd sy'n datblygu, megis clwstwr BRIC Brasil, Rwsia, India a Tsieina, yn wendid yn y ddarpariaeth gyfieithu yng

Nghymru os yw Cymru i fanteisio ar fynediad at y marchnadoedd hyn yn y dyfodol, yn enwedig gan mai'r rhain o bosibl fydd y marchnadoedd allforio pwysicaf maes o law.

4.1.4 Cymhwysedd Iaith Frodorol

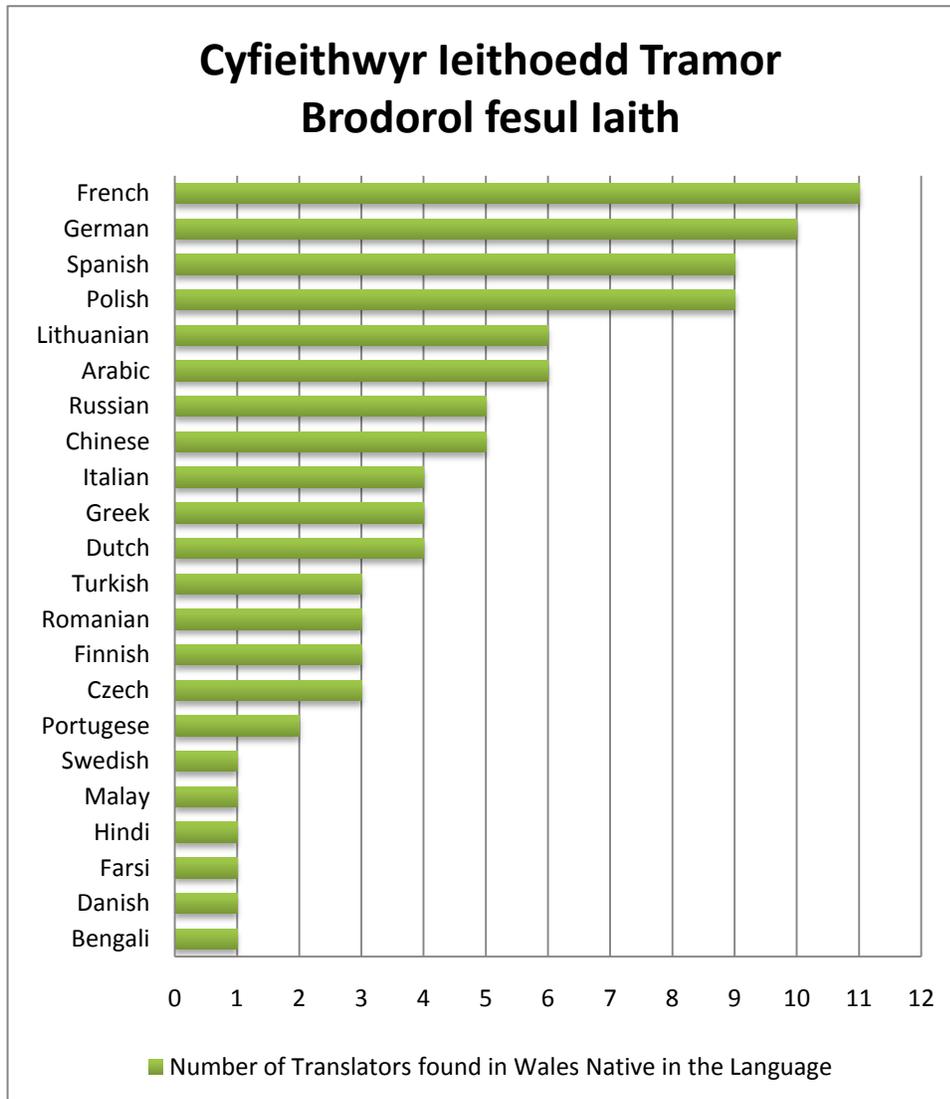
Derbynnir yn gyffredinol mai'r ymarfer gorau mewn cyfieithu proffesiynol yw i gyfieithwyr gyfieithu i mewn i'w hiaith frodorol o iaith y maent wedi'i dysgu yn hwyrach yn eu bywyd. Yr eithriad i'r rheol hon yw'r cyfieithydd dwyieithog, gan gynnwys y mwyafrif o gyfieithwyr Cymraeg/Saesneg (bydd y rhai sydd wedi'u hadnabod yn y siart isod fel siaradwyr Cymraeg brodorol hefyd yn gyfatebol frodorol yn Saesneg).



Ffigur 3: Trosolwg ieithoedd brodorol cyfieithwyr

Gan fod y mwyafrif llethol o'r cyfieithwyr hynny a gafodd eu hadnabod fel siaradwyr Cymraeg brodorol yn cyfieithu rhwng y Gymraeg a'r Saesneg, mae'r graff uchod hefyd yn dangos gymaint o gyfieithu Cymraeg/Saesneg a geir yng Nghymru o ran niferoedd cyfieithwyr.

Casglwyd y data hwn ynghylch iaith frodorol cyfieithwyr a gaed yn gweithio yng Nghymru fel arwydd o'r amrediad eang o ieithoedd sy'n cael eu cynnig i safon cynhwysedd brodorol yng Nghymru. Mae cyfieithwyr sydd â chymhwysedd iaith frodorol yn Saesneg ond sy'n meddu ar sgiliau ieithoedd tramor yn aml yn denu buddsoddi i mewn i Gymru pan fydd cwmnïau tramor yn eu cyflogi i gyfieithu eu cynnyrch i'r Saesneg. Mae hyn yn cynnwys cyfieithu llenyddiaeth megis llyfrau a chylchgronau o'r Saesneg i ieithoedd megis Ffineg ar gyfer tai cyhoeddi mewn marchnadoedd tramor megis y Ffindir.



Ffigur4: Cyfieithwyr ieithoedd brodorol tramor fesul iaith

Yn ogystal â gallu brodorol yn yr iaith dramor, mae cyfieithwyr o dramor yn fwy tebygol o fod yn ymwybodol i ddiwylliant, arferion a materion sensitif marchnadoedd tramor na rhywun nad yw'n frodor o'r wlad honno. Mae hyn yn help i amddiffyn y cwsmer rhag camgymeriadau poenus a drudfawr wrth fynd i mewn i farchnad dramor. Trafodir hyn yn fanylach yn Atodiad 2: Trafodaeth Sgopio 5.

4.1.5 Cymwysterau Cyfieithwyr

Er nad oes unrhyw ofyniad ffurfiol ar gyfieithwyr i gael cymwysterau proffesiynol er mwyn cael bod yn ymarferwyr, y mae gan y diwydiant yn gyffredinol weithlu sy'n meddu ar sgiliau uchel, wedi'u haddysgu i lefel gradd brifysgol neu uwch. O sampl o 310 aelod o CCC, adroddodd 84% fod ganddynt Radd Baglor, 23% fod ganddynt Radd Meistr, a 4% fod ganddynt Ddoethuriaeth.

Casglwyd manylion hefyd am gymwysterau academaidd cyfieithwyr nad oedd yn aelodau o Cymdeithas Cyfieithwyr Cymru. Roedd y sampl hwn yn cynnwys yn unig y rhai a oedd wedi nodi rhyw fath o gymwysterau (60 allan o 195) ac felly nid oes modd gwneud cymhariaeth uniongyrchol gyda'r canlyniadau uchod ar gyfer aelodau CCC. O'r cyfieithwyr eraill yng Nghymru, nododd 92% fod ganddynt Radd Baglor, 53% â Gradd Meistr, a 5% â Doethuriaeth.

Mae'r canlyniadau hyn yn dangos bod cyfieithu yn broffesiwn sy'n cyflogi gweithlu â sgiliau uchel iawn, gyda'r mwyafrif yn meddu ar raddau lefel bagloriaeth, a chyfran sylweddol hefyd yn meddu ar gymwysterau ôl-raddedig.

Yn ychwanegol at gymwysterau addysgol cyffredin, gall cyfieithwyr fod â chymwysterau neu achrediad penodol fel cyfieithwyr. Trafododd y *Strategaeth Genedlaethol ar gyfer datblygu'r Proffesiwn Cyfieithu Cymraeg/Saesneg*, a gyhoeddwyd gan Fwrdd yr Iaith Gymraeg a CCC yn 2006⁵, ffyrdd o reoleiddio'r diwydiant cyfieithu yng Nghymru. Mae'r strategaeth hon fodd bynnag yn canolbwyntio ar hyfforddi ac achredu ac nid yw'n edrych ar faterion technoleg cyfieithu mewn unrhyw fanylder. Mae CCC wedi cyflwyno system o arholiadau yn arwain at aelodaeth o'r gymdeithas, er mwyn darparu lefel dderbyniol o sicrwydd ansawdd i gwsmeriaid⁶.

Mae Bwrdd yr Iaith Gymraeg ar eu gwefan, mewn adran o'r enw "Sut ydw i yn comisiynu cyfieithu" hefyd yn nodi:

“Gan gofio fod cyfieithu yn grefft sy'n gofyn am gryn arbenigedd mae'n hollbwysig eich bod yn ystyried defnyddio cyfieithydd sy'n aelod o gymdeithas cyfieithu proffesiynol”⁷.

Ar wahân i CCC, y cymdeithasau proffesiynol eraill sy'n recriwtio cyfieithwyr yng Nghymru yw'r *Institute of Translation and Interpreting*⁸ a *The Chartered Institute of Linguists*⁹, ac mae'r ddwy gymdeithas hon yn agored i gyfieithwyr sy'n gweithio gyda ieithoedd ar wahân i Gymraeg a Saesneg, ac mae rhai cyfieithwyr yn aelodau o fwy nag un gymdeithas. Mae gan y tair cymdeithas ofynion arholi llym tebyg ar gyfer darpar aelodau.

4.1.6 Cyflogaeth Cyfieithwyr

Mae data a dderbyniwyd oddi wrth Cymdeithas Cyfieithwyr Cymru yn dangos bod y mwyafrif o gyfieithwyr Cymraeg eu hiaith yn gweithredu fel cwmnïau preifat neu gyfieithwyr llawrydd (59% o sampl o 310) o'u cyferbynnu â'r rhai sy'n gweithio'n uniongyrchol i'r sector gyhoeddus (41% allan o sampl o 310).

⁵ <http://www.byig-wlb.org.uk/cymraeg/cyhoeddiadau/pages/publicationitem.aspx?puburl=/cymraeg/cyhoeddiadau/cyhoeddiadau/4286.pdf>

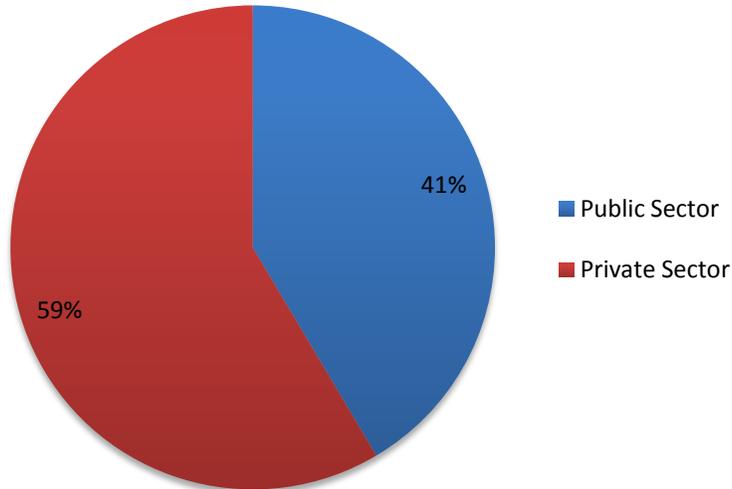
⁶ http://www.welshtranslators.org.uk/full_membership-35.aspx

⁷ <http://www.byig-wlb.org.uk/Cymraeg/defnyddio/Pages/Sutmaecomisiynucyfieithydd.aspx>

⁸ <http://www.iti.org.uk/indexMain.html>

⁹ <http://www.iol.org.uk/>

Dadansoddiad o Gyflogaeth Aelodau Cymdeithas Cyfieithwyr Cymru

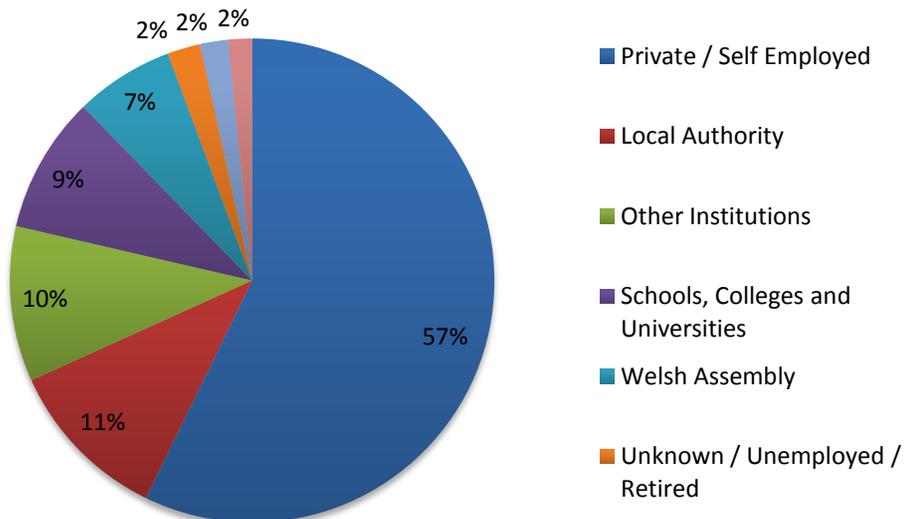


Ffigur 5: Dadansoddiad Cyflogaeth Aelodau CCC

Sylwer bod llawer o gyfieithwyr sy'n gweithio yn y Sector Breifat yn gweithio i gwmnïau sy'n gwasanaethu anghenion y Sector Gyhoeddus (gweler Atodiad 2).

Mae'r canlynol yn ddadansoddiad manylach o'r ffigurau a ddarparwyd gan CCC.

Dadansoddiad Cyflogaeth Manylach Aelodau CCC



Ffigur 6: Dadansoddiad Manylach Cyflogaeth Aelodau CCC

Yn ystod oes y project, ni chafwyd unrhyw gyfieithwyr ar wahân i rai yn cyfieithu rhwng y Gymraeg a'r Saesneg oedd yn cael eu cyflogi yn fewnol gan y sector gyhoeddus yng Nghymru. Fodd bynnag, mae llawer o awdurdodau lleol yn cyflogi cyfieithwyr allanol yn ôl yr angen i ddelio gyda thasgau cyfieithu mewn ieithoedd eraill. Cafwyd galw cynyddol yn ddiweddar am arbenigwyr yn ieithoedd Dwyrain Ewrop megis Pwyleg, yn hytrach na ieithoedd Asiaidd rhai cymunedau ethnig traddodiadol.

Dim ond 6 cwmni preifat a ganfuwyd a oedd yn gweithio gyda ieithoedd ar wahân i Gymraeg ac sy'n cyflogi nifer o gyfieithwyr mewnol, ac roedd y rhain yn tueddu i gyflogi llai o staff. Ymddengys mai'r arfer ymhlith cwmnïau preifat yw is-contractio'r gwaith cyfieithu mewn ieithoedd llai aml eu defnydd yn ôl y galw. Gall y cyfieithwyr llawrydd y maent yn eu defnyddio fod yn gweithio iddynt o unrhyw le yn y byd.

4.1.7 Swyddogaeth yr Uned Technolegau Iaith

Ceir nifer o weithgareddau sy'n berthnasol i gyfieithu ym Mhrifysgol Bangor. Yn eu plith ceir:

- Addysgu ac ymchwilio i astudiaethau cyfieithu yn yr Ysgol Ieithoedd Modern.
- Gwasanaeth cyfieithu mewnol Prifysgol Bangor. Mae'n cyflogi 5 cyfieithydd llawn amser yn gofalu am gyfieithu ysgrifenedig a chyfieithu ar y pryd, dyma'r gwasanaeth cyfieithu prifysgol mwyaf mewn unrhyw brifysgol yn y Deyrnas Unedig.
- Gwasanaeth cyfieithu a golygu masnachol Prifysgol Bangor, Cymraeg Clir. Mae hwn i'r Gymraeg yr debyg i'r hyn yw *Plain English* i'r iaith Saesneg, ac mae'n galluogi busnesau a chyrrff i gyfathrebu yn glir a diamwys gyda'r cyhoedd yn Gymraeg.
- Mae'r Uned Technolegau Iaith (UTI) hefyd yn cynnwys y Ganolfan Safoni Termau. Prif swyddogaeth yr Uned yw ymchwilio a datblygu adnoddau a thechnolegau iaith ar gyfer y Gymraeg a sefyllfaoedd amlieithog. Mae'r Uned yn darparu technegau iaith ac ymchwil pwrpasol ar sail comisiwn i unrhyw gorff.

Mae'r tair uned olaf wedi'u lleoli o fewn Canolfan Bedwyr ym Mhrifysgol Bangor. Mae gan yr UTI hefyd gyswllt gyda gweithgareddau Astudiaethau Cyfieithu yn yr Ysgol Ieithoedd Modern, ac i ymchwil yn yr Ysgol Ieithyddiaeth a'r Ysgol Cyfrifiadureg.

Ers iddi gael ei sefydlu yn 2001, mae gweithgareddau'r UTI wedi dod â hi i gysylltiad â'r defnydd o dechnolegau iaith yn y diwydiant cyfieithu yng Nghymru a ledled y byd. Mae rhai o'i aelodau hefyd yn aelodau o'r cyrrff rhyngwladol y mae'r diwydiant yn eu cydnabod fel rhai sy'n gyfrifol am safoni yn y farchnad gyfieithu fyd-eang.

4.2 Y Defnydd o Dechnoleg Cyfieithu yng Nghymru

4.2.1 Diffiniad o Dechnoleg Cyfieithu

Mae technoleg cyfieithu yn derm ambarél sy'n cynnwys yr holl gymhorthion cyfieithu o fewn amgylchedd electronig, cyfrifiadurol neu ar y we. Enghreifftiau o gymhorthion felly yw geiriaduron electronig, offer gwirio ar gyfer sillafu a gramadeg, meddalwedd cof cyfieithu, ac

offer amgylchedd cyfieithu (TEnt yn Saesneg), sydd weithiau yn cael eu galw gyda'i gilydd yn offer cyfieithu trwy gymorth cyfrifiadur (CAT yn Saesneg). Mae technoleg cyfieithu hefyd yn cynnwys cyfieithu awtomatig (peirianyddol), sydd yn aml yn cael ei dalgyrru yn MT yn Saesneg. Yn gynyddol, mae cymhorthion gweinyddol i'r broses gyfieithu, megis systemau rheoli dogfennau dwyieithog neu amlieithog, rheolwyr llif gwaith, a chynhyrchwyr anfonebau, yn cael eu cynnwys dan y pennawd technoleg cyfieithu. Mae cynhyrchion soffistigedig sy'n dwyn ynghyd nifer o nodweddion creiddiol ar gyfer amgylchedd cyfieithu cynhwysfawr yn cael eu creu, ac mae'r maes yn datblygu'n barhaus.

4.2.2 Y Defnydd o Dechnoleg Cyfieithu yn Rhyngwladol

Gwnaed yr arolwg rhyngwladol diweddaraf ar Dechnoleg Cyfieithu gan Elina Lagoudaki yn 2006¹⁰. Canolbwyntiodd ei *Translation Memories Survey* 2006 yn benodol ar y defnydd o systemau cof cyfieithu, ac agweddau cyfieithwyr tuag atynt. Fodd bynnag, yr oedd yr adroddiad hefyd yn cynnwys gwybodaeth werthfawr ar amgylchedd ac arferion gwaith y cyfieithwyr, a'r canfyddiad o dechnoleg cyfieithu yn gyffredinol, ac felly yn rhoi cymhariaeth ddefnyddio i arferion a chanfyddiadau cyfieithwyr yng Nghymru. Cafwyd 874 ateb gan weithwyr proffesiynol ym maes cyfieithu i'r arolwg rhyngwladol mewn 54 gwlad ar draws y byd. Ymhlith ei hymatebwyr, roedd y cyfartaledd o unigolion a oedd yn adrodd eu bod yn defnyddio technoleg amgylchedd cyfieithu (TEnt) yn 82.5%.

Gwnaeth yr awdur y pwynt fod y mwyafrif o ymatebwyr (61%) wedi nodi cynnwys technegol fel prif faes eu harbenigedd, a chadarnhaodd ganlyniadau o nifer o arolygon eraill fod y rhai sy'n arbenigo mewn testunau technegol yn fwy tebyg o ddefnyddio offer cof cyfieithu. Canfyddiad creiddiol arall o'r arolwg hon sy'n berthnasol i'n hastudiaeth ni o'r sefyllfa yng Nghymru yw mai po fwyaf hyddysg oedd y cyfieithwyr proffesiynol yn y defnydd o gyfrifiaduron, y mwyafrif tebygol oeddynt o ddefnyddio system cof cyfieithu.

O'r 17.5% a adroddodd nad oeddynt yn defnyddio offer cof cyfieithu, dim ond 11% a adroddodd nad oeddynt yn gwybod am fodolaeth offer o'r fath, a'r rheswm mwyaf cyffredin (28%) a roddwyd am beidio defnyddio offer o'r fath oedd eu bod yn anaddas ar gyfer natur eu gwaith hwy. O'r rhai nad oedd yn defnyddio cof cyfieithu, adroddodd 71% y byddent yn fodlon prynu neu roi cynnig ar ddefnyddio system cof cyfieithu yn y dyfodol agos.

Edrychodd yr arolwg rhyngwladol hefyd ar gyfeiriadau posibl yn y dyfodol, gan ofyn i ymatebwyr am eu hagweddau ar offer technoleg cyfieithu eraill, a'r blaenoriaethau ar gyfer ymchwil a datblygu. Daeth nodweddion megis gwirwyr sillafu integredig yn uchel ar eu rhestr o'r hyn a ystyrir yn bwysig, a nodwyd hefyd yr angen i ddatgloi potensial adnoddau iaith sy'n bodoli ar CD-ROM a'r rhyngwryd, yn arbennig fel y gallent ddefnyddio "storfa adnoddau unigol ac unffurf".

Mae'n werth dyfynnu dau o'r sylwadau a gafwyd yn y cyd-destun hwn yn llawn eto oherwydd eu perthnasedd i'r arolwg presennol:

¹⁰ Elina Lagoudaki *Translation Memory systems: Enlightening users' perspective*. Imperial College London (2006), ar-lein yn <http://www3.imperial.ac.uk/portal/pls/portallive/docs/1/7294521.PDF>

“[An ideal TM system would be] a feature-rich, single small-scale tool that is highly-adaptable to individual translators' needs, [...] that places preference on new features that enable a translator to get started quickly on any format [and] to access additional reference materials more directly (CD dictionaries, major online dictionaries, bilingual parallel texts).”

“One full-fledged app[lication] please, but make it modular, so you can keep tasks apart when necessary, and with a well documented and powerful API, so that add-ons like e. g. language-specific assembly algorithms can be made and plugged in. I don't mind paying, but I do mind closed source lock-in effects and strongly resent platform lock-in.”

4.2.3 Cymariaethau gyda Defnydd yng Nghymru

Yn 2005 adroddodd *Arolwg o'r Diwydiant Cyfieithu yng Nghymru*¹¹ ar y defnydd o dechnoleg cyfieithu yng Nghymru. Danfonwyd holiaduron i tua 900 cwmni cyfieithu yng Nghymru a derbyniwyd 371 ateb. Pan holwyd hwy am eu defnydd o dechnoleg cyfieithu, dim ond 20% a adroddodd eu bod yn defnyddio offer amgylchedd cyfieithu, sy'n sylweddol lai na'r 82.5% defnydd a adroddwyd yn rhyngwladol gan Lagoudaki (2006). Ni ofynnodd arolwg 2005 gwestiynau ynghylch natur y cynnwys oedd yn cael ei gyfieithu, ond hyd yn oed a chaniatáu am y faith y gallai fod yn llai technegol ei natur ac felly o bosib yn llai ailadroddus na gwaith y mwyafrif o ymatebwyr i'r arolwg rhyngwladol, ymddengys fod y nifer sy'n defnyddio offer amgylchedd cyfieithu yng Nghymru yn rhyfeddol o isel.

Fodd bynnag, fel ofynnodd yr arolwg yng Nghymru gwestiynau am adnoddau iaith electronig eraill a ddefnyddid gan gyfieithwyr proffesiynol, a chafwyd fod 74% o ymatebwyr yn defnyddio Cysill, gwirydd sillafu a gramadeg Cymraeg, gyda 59% yn defnyddio CysGair, geiriadur electronig dwyieithog Cymraeg/Saesneg. Yr oedd 48% o ymatebwyr yn defnyddio'r compendiwm meddalwedd mwy newydd Cysgliad, sy'n cynnwys Cysill a chyfres o eiriaduron termau dwyieithog Cymraeg/Saesneg yn ogystal â CysGair. Yr oedd 62% o ymatebwyr yn defnyddio geiriadur termau dwyieithog electronig arall, *Y Termiadur Ysgol*. Yr oedd yr adnoddau hyn i gyd wedi cael eu cynhyrchu gan yr Uned Technolegau Iaith yng Nghanolfan Bedwyr, Prifysgol Bangor.

Mae canlyniadau tebyg wedi cael eu canfod gan Tegau Andrews, myfyriwr PhD ym Mhrifysgol Bangor, lle dangosodd ei hymchwil mai dim ond 47% o gyfieithwyr proffesiynol yng Nghymru oedd yn defnyddio unrhyw fath o offer cyfieithu trwy cymorth cyfrifiadur¹².

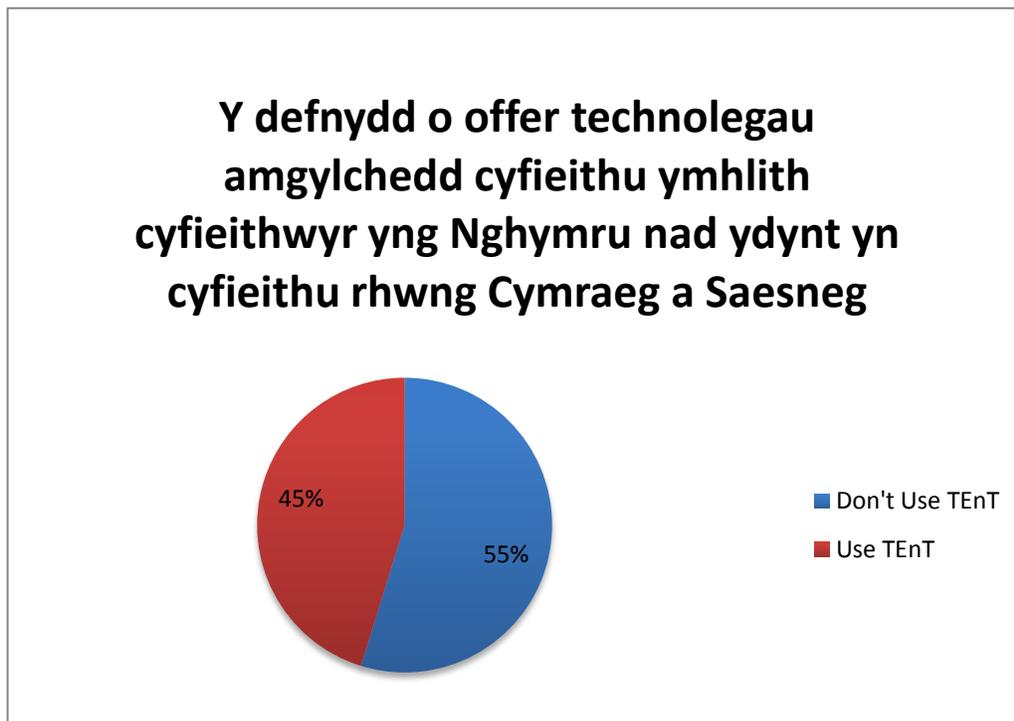
Mae'r canlyniadau hyn yn cael eu cadarnhau ymhellach yn yr Astudiaethau Dichonolrwydd a'r Trafodaethau Sgopio a wnaed ar gyfer yr adroddiad presennol, lle dywedodd llai na 50% o'r rhai a oedd yn bresennol y grwpiau ffocws eu bod yn defnyddio meddalwedd cof cyfieithu.

¹¹ *Arolwg o'r Diwydiant Cyfieithu yng Nghymru*, Menter a Busnes, cyhoeddwyd gan CCC a Bwrdd yr Iaith Gymraeg (2005) ar gael o wefan ByIG <http://www.byig-wlb.org.uk/>

¹² Mae'r PhD hwn a noddwyd dan gynllun Amcan 1, o'r enw *A Study of Current Practice in Web Localization and its Application to Welsh Business* i fod i gael ei gwblhau yn ystod haf 2009. Rhoddwyd canlyniadau cyntaf yr astudiaeth yn bersonol i awduron yr adroddiad presennol.

4.2.4 Gwahaniaethau Iaith-benodol yn y Defnydd o Dechnoleg Cyfieithu

Roedd yr ymatebwyr i'r *Arolwg o'r Diwydiant Cyfieithu yng Nghymru* (2005) yn gweithio'n bennaf gyda'r pâr ieithoedd Cymraeg/Saesneg, ac mae canlyniadau Tegau Andrews hefyd yn ymwneud yn bennaf gyda chyfieithwyr sy'n gweithio gyda'r ddwy iaith hon. Ceisiodd yr astudiaeth bresennol felly ganfod a oedd yna unrhyw wahaniaeth o bwys yn y defnydd o offer amgylchedd cyfieithu gan gyfieithwyr proffesiynol yn gweithio yng Nghymru gyda chyfuniadau eraill o ieithoedd. Syndod oedd canfod mai ychydig iawn o wahaniaeth oedd yna. O sampl o 115 cyfieithydd proffesiynol a oedd yn gweithio gyda ieithoedd ar wahân i Gymraeg a Saesneg, dim ond 45% ohonynt adroddodd eu bod yn defnyddio system cof cyfieithu.



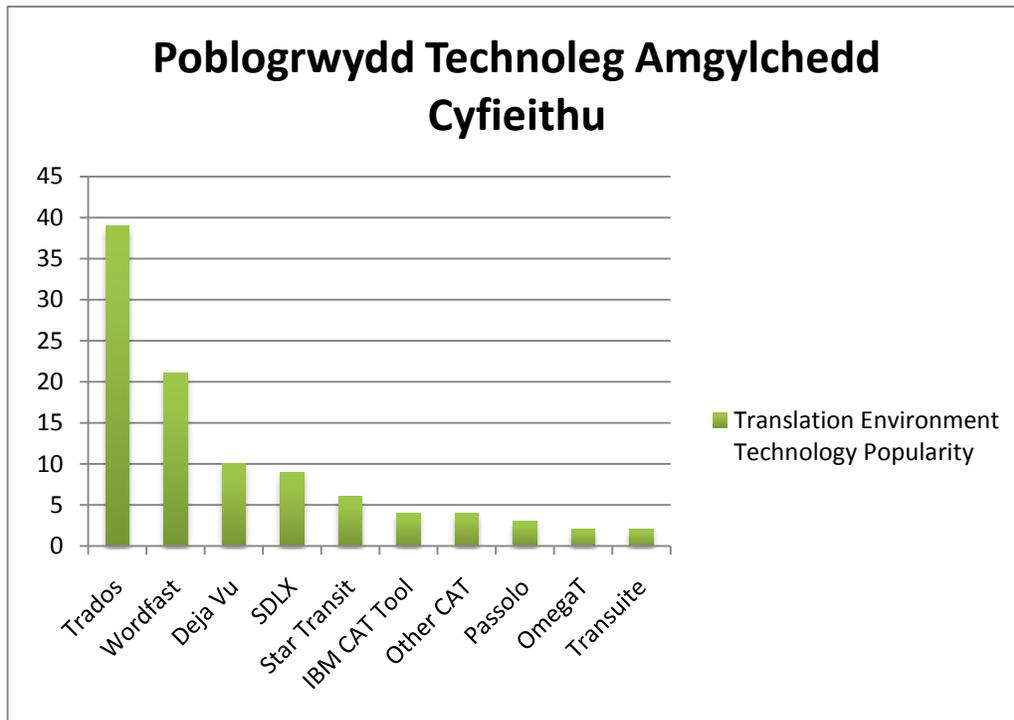
Ffigur 7: O sampl o 115 cyfieithydd proffesiynol

Mae'r tebygrwydd rhwng ffigurau Tegau Andrews ar gyfer cyfieithwyr rhwng y Gymraeg a'r Saesneg (47%) a'r ffigurau a gasglwyd yn ystod y project hwn ar gyfer cyfieithwyr parau iaith eraill (45%) sy'n defnyddio offer amgylchedd cyfieithu yn dangos bod y diffyg defnyddio offer amgylchedd cyfieithu ymhlith cyfieithwyr yn nodwedd o ddiwydiant cyfieithu Cymru yn gyfan, ac nad yw'n gyfyngedig i gyfieithwyr sy'n gweithio gydag un pâr arbennig o ieithoedd.

Mae hyn yn dangos fod y defnydd o offer amgylchedd cyfieithu ymhlith cyfieithwyr yng Nghymru, beth bynnag fo'r ieithoedd sy'n cael eu cyfieithu, yn sylweddol is na'r hyn a geir ymhlith y proffesiwn cyfieithu yn rhyngwladol, a dylai fod yn achos pryder gwirioneddol ar gyfer y diwydiant cyfieithu yng Nghymru.

4.2.5 Offer Technoleg Cyfieithu Penodol

Casglwyd gwybodaeth ar gyfer yr arolwg hwn ynghylch yr amgylcheddau cyfieithu a ddefnyddir gan gyfieithwyr yng Nghymru, ar wahân i'r rhai sy'n cyfieithu rhwng Cymraeg a Saesneg, er mwyn cynorthwyo mwy ar ddatblygu adnoddau cynnwys ar gyfer meddalwedd offer amgylchedd cyfieithu yn y dyfodol. O'r 52 cyfieithydd nad oedd yn aelodau o CCC a oedd yn defnyddio offer amgylchedd cyfieithu, roedd y rhan fwyaf (39) yn defnyddio'r offer amgylchedd cyfieithu Trados, yn cael ei ddilyn gan Wordfast (21), Déjà Vu (10) a SDLX (9), sydd bellach yn rhan o'r pecyn SDL Trados.



Mae'r canlyniadau hyn yn dangos mai'r tri h pecyn meddalwedd technoleg amgylchedd cyfieithu sydd bellach yn cael eu cefnogi yw SDL Trados (yn cynnwys Trados a SDLX), Wordfast, a Déjà Vu, ac maent yn cyd-fynd â chanlyniadau'r arolwg rhyngwladol *Translation Memories Survey* (2006).

Dangosodd canlyniadau arolwg Tegau Andrews sydd ar fin eu cyhoeddi (2009), gyda sampl o 73, hefyd mai Trados, Wordfast a Déjà Vu oedd y tair system fwyaf poblogaidd. Fodd bynnag, ymhlith y cyfieithwyr Cymraeg/Saesneg yn ei harolwg, cafodd mai Déjà Vu oedd y mwyaf poblogaidd, gyda Wordfast yn ail, yn cael eu dilyn gan Trados. Mae'n bosibl fod y dewis o Déjà Vu yn cael ei ddylanwadu gan y ffaith mai dyma'r feddalwedd y mae Gwasanaeth Cyfieithu Llywodraeth Cynulliad Cymru yn ei ddefnyddio. Ymhlith cyfieithwyr llawrydd fodd bynnag, ymddengys mai Wordfast oedd y mwyaf poblogaidd, mae'n debyg oherwydd ei bris cymharol isel wedi'i gyplysu gyda set nodweddion cymharol lawn.

Mae'r adnoddau meddalwedd technoleg amgylchedd cyfieithu hyn yn annibynnol ar iaith, ac felly yr un mor ddefnyddiol i gyfieithwyr pa ieithoedd bynnag maent yn gweithio â hwy.

Mae adnoddau cynnwys ar y llaw arall, yn eu plith geiriaduron iaith gyffredin a geiriaduron termau arbenigol, offer prawfddarllen (yn enwedig gwirwyr sillafu a gramadeg) a chofion cyfieithu wedi'u poblogi'n barod neu gorpora cyfochrog o gyfieithiadau, yn eu hanfod yn ddibynnol ar ieithoedd penodol. Mae poblogrwydd rhai o'r adnoddau cynnwys Cymraeg/Saesneg a enwyd uchod yn 4.2.3, a chanlyniadau tebyg ar gyfer y defnydd o'r adnoddau hyn yn cael eu cadarnhau gan grwpiau ffocws o'r ymchwiliadau cyfredol.

Er ei bod hi'n anoddach mesur y defnydd o adnoddau cynnwys ar gyfer y cyfuniadau eraill o ieithoedd y mae cyfieithwyr proffesiynol yn eu cynnig yng Nghymru, tybir mai'r gymuned/cymunedau iaith cartref sy'n arwain ar datblygu'r adnoddau cynnwys yn yr ieithoedd hynny. Felly byddai rhywun yn disgwyl i adnoddau cynnwys Ffrangeg gael eu datblygu yn bennaf mewn gwledydd lle siaredir Ffrangeg, yn union fel y byddid yn disgwyl i adnoddau cynnwys Cymraeg gael eu datblygu yn bennaf yng Nghymru.

4.2.6 Diffyg Arbenigedd Technegol ymhlith Cyfieithwyr

O'r rhwystrau i ddefnyddio mwy ar dechnoleg cyfieithu yng Nghymru, yr un a ddyfynnwyd amlaf gan gyfieithwyr mewn grwpiau ffocws oedd y diffyg mynediad hawdd at wybodaeth ynghylch yr hyn oedd ar gael. Er enghraifft, ymhlith yr adborth oddi wrth y grŵp ffocws a gaed yng Nghaernarfon oedd y sylw:

“Hoffwn ddod wyneb yn wyneb â rhywun (ym Mangor) a allai gyflwyno pecyn o raglenni addas ar fy nghyfer (fel rhywun sydd heb eto ddefnyddio unrhyw raglen fel Cysill ar fy nghyfrifiadur)”.

Roedd hyn yn arbennig o wir am y nifer mawr o unig fasnachwyr sy'n cyfrif am lawer o'r diwydiant yng Nghymru, gan fod y rhain yn llai tebygol o gael mynediad at gefnogaeth dechnegol na'r rhai sy'n gweithio o fewn amgylchedd cwmni. Efallai hefyd fod arbenigedd technegol yn cael ei weld fel byd dynion, a bod y gwahaniaeth rhwng dewis merched i astudio ieithoedd a dewis bechgyn i astudio pynciau technolegol megis cyfrifiadureg, hyd yn oed ar lefel TGAU (a drafodir yn 4.1.2 uchod) yn cario drosodd i ddiffyg brwdfrydedd dros fabwysiadu technoleg cyfieithu ar y lefel broffesiynol.

4.2.7 Offer Technoleg Cyfieithu a Ddatblygwyd yng Nghymru

Mae bron y cyfan o'r gwaith datblygu offer technoleg cyfieithu sydd wedi digwydd mor belled yng Nghymru wedi digwydd ym Mhrifysgol Bangor. Hyd yn oed felly, gellir dadlau bod offer megis Cysgliad (yn cynnwys y gwirydd sillafu a gramadeg Cysill a'r compendiwm geiriaduron Cysgeir) wedi cael eu datblygu fel cymhorthion iaith cyffredinol a heb eu hanelu yn benodol at gyfieithwyr proffesiynol.

Yn 2006, cyhoeddodd Bwrdd yr Iaith Gymraeg ddogfen strategaeth *Technoleg Gwybodaeth a'r Gymraeg*¹³. Roedd y ddogfen hon yn cynnwys adran ar gyfieithu a therminoleg, yn edrych ar faterion cyfieithu drwy gymorth cyfrifiadur, rheoli termau, technoleg alinio, llif gwaith cyfieithu a chyfieithu peirianyddol. Yn ogystal â rhoi disgrifiadau byr o'r technolegau, mae'r adroddiad yn rhoi targedau ar gyfer ystyried potensial dichonolrwydd technolegau o'r fath yn benodol ar gyfer yr iaith Gymraeg. Cyhoeddwyd adroddiad blwyddyn gyntaf ar ei weithredu yn 2007¹⁴, ond nid ymddangosodd adroddiad ar gyfer 2008 hyd yn hyn. O'r pynciau hyn, dim ond mewn cyfieithu peirianyddol y cafwyd gwahoddiad i dendro, ac nid aethpwyd ymlaen gyda'r project hwn.

Mae'r Uned Technolegau Iaith yn PB wedi derbyn nifer o gomisiynau oddi wrth Fwrdd yr Iaith Gymraeg i ddatblygu offer technoleg iaith ar gyfer y Gymraeg. Mae'n cydnabod fod datblygu offer o'r fath yn gofyn am fuddsoddi helaeth ac yn ceisio gwneud y mwyaf o adnoddau drwy ymestyn ei gylch gorchwyl i ddatblygu offer sy'n medru cynnal nifer o ieithoedd, gan ostwng y gost ar gyfer pob iaith newydd sy'n cael ei chynnal. Mae ei lwyfannau meddalwedd newydd, megis Maes T, a ddatblygwyd ar gyfer golygu a chyhoeddi geiriaduron electronig a therminolegau arbenigol wedi'u cynllunio i fod yn generig ac yn addasadwy ar gyfer pob amgylchedd iaith. Mae gan y Gymraeg lawer i'w ennill drwy gael ei normaleiddio o fewn prif ffrwd datblygiad technoleg iaith rhyngwladol.

Mae'r UTI hefyd yn gweithio gyda'r sector breifat i ddatblygu offer iaith a chyfieithu er mwyn gwella eu gallu i gystadlu'n economaidd. Mae wedi defnyddio'r Rhaglen Trosglwyddo Gwybodaeth (KTP) fel mecanwaith i helpu diwydiant. Mae'r rhaglen KTP wedi galluogi cwmnïau cyfieithu preifat yng Nghymru i ddatblygu offer i ateb eu hanghenion technolegol eu hun. Galluogodd y KTP cyntaf o'r fath, gyda Cymen, cwmni cyfieithu sydd â'i bencadlys yng Nghaernarfon, y cwmni i gyflwyno rheolaeth termau ac amgylchedd technoleg cyfieithu wedi'u llunio'n unswydd, i'w amgylchedd ei hun. Mae KTP ar hyn o bryd gyda Testun Cyf., cwmni isdeitlo a chyfieithu sydd â'i bencadlys yng Nghaerdydd, yn ymchwilio i integreiddio technoleg lleferydd ac isdeitlo i'r feddalwedd isdeitlo y mae'n ei ddefnyddio i gyfieithu ac isdeitlo rhaglenni teledu i'r Gymraeg a'r Saesneg fel ei gilydd.

4.3 Y Diwydiant Cyfieithu ac Economi Cymru

4.3.1 Arwyddocâd Economaidd y Diwydiant Cyfieithu

Amcangyfrifwyd fod y diwydiant cyfieithu a darparwyr gwasanaethau iaith perthnasol yn werth \$12.1UD biliwn yn 2007. Yn seiliedig ar y llinell gogwydd dros y pum mlynedd diwethaf, amcangyfrifir y byddai twf y farchnad yn rhoi diwydiant gwerth UD\$24 biliwn erbyn 2012¹⁵. Dyfynnwyd fod diwydiant cyfieithu Ewrop yn werth £3.5 biliwn yn 2006, gyda'r Deyrnas

¹³ Technoleg Gwybodaeth a'r Iaith Gymraeg: Dogfen Strategaeth (2006) gw: <http://www.byig-wlb.org.uk/Cymraeg/cyhoeddiadau/Cyhoeddiadau/3964.pdf>

¹⁴ Technoleg Gwybodaeth a'r Iaith Gymraeg: Adroddiad Gweithredu'r Flwyddyn Gyntaf <http://www.byig-wlb.org.uk/cymraeg/cyhoeddiadau/pages/publicationitem.aspx?puburl=/cymraeg/cyhoeddiadau/cyhoeddiadau/5158.pdf>

¹⁵ O ffigurau a roddwyd gan Renato S. Beninatto a Donald A. DePalma yn *Ranking of Top 25 Translation Companies*, cyhoeddwyd gan y *Common Sense Advisory* (2008) ar gael ar y we http://www.commonseadvisory.com/members/res_cgi.php/080528_QT_2008_top_25_lsps.pdf

Unedig fel y chwaraewr mwyaf ynddi, a'r prif gwmnïau cyfieithu yn y DU yn unig gyda'i gilydd yn darparu gwerth £146.87 miliwn o waith cyfieithu y flwyddyn¹⁶.

Fe'i cafwyd yn anodd cael hyd i adroddiadau cynhwysfawr ar werth y diwydiant cyfieithu yng Nghymru, gan nad oes ffigurau annibynnol yn cael eu cyhoeddi. Fodd bynnag, y mae ein hamcangyfrifon ein hun o'r diwydiant, sy'n tynnu ar wybodaeth gyfrinachol a roddwyd gan nifer o gwmnïau bach a chanolig eu maint ac unig fasnachwyr am eu trosiant blynyddol, yn awgrymu ei fod yn werth o gwmpas £45 miliwn yn 2007. Yn unol â thueddiadau byd-eang, gallai hyn ddyblu i £90 miliwn erbyn 2012. Gallai cynyddu'r gallu i drin mwy o waith o fewn y diwydiant, er mwyn cyfarfod â'r galw yn y farchnad gartref nad yw'n cael ei ateb ar hyn o bryd, ac er mwyn denu busnes newydd o'r tu allan i Gymru, fwy na threblu'r ffigur hwn.

Fel diwydiant sy'n cyflogi graddedigion â sgiliau uwch, a gan bod unig fasnachwyr a chwmnïau bach yn medru gweithio o'u cartrefi ac o swyddfeydd gwledig, os yw cysylltiadau rhyngwryd boddhaol ar gael, y mae'n ychwanegiad gwerthfawr i economi gwledig Cymru. Mae'n cyflogi nifer arwyddocaol o ferched (gweler 4.1.2 uchod), ac mae gweithwyr cartref yn adrodd bod cyfieithu yn gyfuno'n dda gyda dyletswyddau teuluol eraill.

4.3.2 Arbedion Economaidd drwy Ddefnyddio Technoleg Cyfieithu

Mae'r dadleuon dros fabwysiadu technoleg cyfieithu i helpu'r broses gyfieithu yn cynnwys yr arbedion amser sylweddol y gellir eu hennill felly, gwelliannau yng nghysondeb a chywirdeb y defnydd o dermau ac yn safon y cyfieithu ei hun¹⁷. Fodd bynnag, o'r safbwynt economaidd, yr arbedion sylweddol mewn cost yw'r hyn sy'n dadlau gryfaf dros fabwysiadu technoleg cyfieithu.

Mae costau cyfieithu yn faich ychwanegol ar wasanaethau cyhoeddus a busnesau preifat fel ei gilydd. Oni bai bod ffyrdd o alluogi mwy o waith i gael ei gyfieithu, gan gadw costau cyfieithu o fewn terfynau ar yr un pryd, bydd cyfieithwyr yn eu prisio eu hunain allan o'r farchnad. Wrth i waith cyfieithu, yn enwedig ar gyfer rhai cyfuniadau o ieithoedd, cael eu ffermio allan i gyfieithwyr rhatach dramor, mae'r bygythiad i ddarparwyr gwasanaethau cyfieithu gartref yn dod yn fwyfwy real. Mae defnyddio technoleg cyfieithu i gynyddu trosiant yn un ffordd o barhau i ddarparu gwerth am arian i gwsmeriaid.

Mae gwahanol amcangyfrifon wedi'u rhoi ar yr arbedion amser a geir drwy'r defnydd o dechnoleg cyfieithu. Mewn gwirionedd mae'r arbedion yn dibynnu ar y dechnoleg a ddefnyddir a'r math o ddogfennau a gyfieithir (gweler yr Astudiaethau Dichonolrwydd yn Atodiad 1). Ystyrir bod cyfartaledd arbediad o 40% mewn amser a chost yn rhesymol yn y diwydiant ar brojectau cyfieithu ailadroddus, os defnyddir yr offer a'r adnoddau cyfieithu technegol

¹⁶ Dyfynnwyd o Richard Michie, *Hidden cost of translation and interpretation to the UK Government* (2006) http://www.freepressreleases.co.uk/Press_Releases/Government/Hidden_cost_of_translation_and_interpretation_to_the_UK_Government_200612192423/

¹⁷ Er enghraifft: "When users were asked why they voluntarily used a TM tool, the greatest proportion of respondents reported using a TM tool because it saved them time (86%), because the consistency in terminology improved (83%) and because the tool helped in improving the quality of the translation output (70%), quoted from Elina Lagoudaki *Translation Memory systems: Enlightening users' perspective*. Imperial College London (2006), ar-lein ar <http://www3.imperial.ac.uk/portal/pls/portallive/docs/1/7294521.PDF>

perthnasol. Manylwyd ar hyn mewn un astudiaeth achos, gan ddefnyddio un technoleg amgylchedd cyfieithu penodol:

“The company has seen a marked increase in efficiency after going over to Translator's Workbench. When localising updated versions of products, as much as 80% of the material may be re-used and with Translator's Workbench's facilities for searching and "automatic translation" the amount of work is considerably decreased (no more cutting and pasting). Miller estimates that on average the speed of translation has increased by about 50%. However this is not a pure saving since the administration and technical overhead has increased by around 10-15%, since introducing Translator's Workbench. (Indeed this is one of the few disadvantages he identified with introducing the system). So a more realistic figure for efficiency gains would be in the region of 40%. As TRADOS products continue to improve (and customers also update their own versions of Translator's Workbench) he can foresee even greater efficiency gains.¹⁸”

Mae creu busnes newydd drwy gryfhau'r gadwyn gyflenwi leol, a chynyddu'r gallu i ymgymryd â chyfieithiadau amlieithog o fewn Cymru, yn hytrach na gorfod ei yrru allan i gyfieithwyr mewn gwledydd eraill, yn debyg o arwain at gynnydd o 50% mewn busnes cyfieithu o fewn Cymru, a gallai denu busnes cyfieithu o wledydd eraill i mewn i Gymru arwain at gynnydd o 300% mewn busnes cyfieithu o fewn Cymru. Gallai effaith cynyddu effeithiolrwydd cyfieithu a'r gallu i gymryd mwy o waith cyfieithu drwy ddefnyddio rhaglenni ac adnoddau technolegol gyda'i gilydd felly fod yn werth £307.8 miliwn yn flynyddol o fewn y diwydiant yng Nghymru. Mae hyn yn ychwanegol at yr effaith anuniongyrchol ar economi Cymru yn gyffredinol.

4.3.3 Effaith anuniongyrchol ar economi Cymru

Faint bynnag y bydd gwell darpariaeth a defnydd o dechnoleg cyfieithu yn gwella iechyd economaidd y diwydiant cyfieithu ei hun yng Nghymru, galluogi diwydiant Cymru i gyfathrebu a chystadlu'n well yn y farchnad fyd-eang, amlieithog fydd yn dod â'r gwerth mwyaf i economi Cymru.

Mae'r *Welsh Assembly Government -Language Skills Capacity Audit* (2002) yn datgan bod cwmnïau o Gymru yn colli busnes mewn marchnadoedd rhyngwladol oherwydd diffyg sgiliau ieithyddol a chyd-ddealltwriaeth ddiwylliannol. Adroddwyd fod hyn yn achosi colled o 19% mewn busnes oherwydd diffyg sgiliau ieithyddol. Roedd defnyddio gwasanaethau cyfieithu yn cynnig “immediacy of results and a means of dealing with highly responsible tasks, such as translation of contracts” (adran 2.5 yn y *Welsh Assembly Government -Language Skills Capacity Audit* 2002).

Byddai ennill y 19% o fusnes coll yn rhoi hwb i economi Cymru o hyd at £8,417 miliwn yn flynyddol, yn seiliedig ar faint economi Cymru yn 2007.

Mae'r drafodaeth sgopio gyda chwsmeriaid y diwydiant cyfieithu yng Nghymru (Atodiad 2) yn rhoi manylion ychwanegol am ffyrdd y gallai busnes yng Nghymru elwa o wneud gwell defnydd o wasanaethau cyfieithu. Mae hefyd yn enwi diffyg canllawiau a dealltwriaeth o'r broses

¹⁸ O'r project Euromap, gweler http://est.dk/euromap/CASE_STUDIES/EN_bowne.htm

gomisiynu ar gyfer cyfieithu fel rhwystr difrifol i'r defnydd effeithiol o wasanaethau cyfieithu gan sectorau eraill o ddiwydiant Cymru.

Disgwylir hefyd y bydd arbedion effeithiolrwydd o fewn y sector cyfieithu hefyd yn cael traweffeithiau eang ar economi Cymru. Ymhlith y rhain mae mwy o allu i ateb anghenion domestig gweinyddiaeth gyhoeddus ddwyieithog, lle mae'r galw ar hyn o bryd yn fwy na'r cyflenwad, a lle mae angen cadw costau dan reolaeth heb aberthu safon. Byddai arbedion bychan o 10% mewn costau cyfieithu i'r sector gyhoeddus yn rhoi arbedion blynyddol o £8 am bob mil o eiriau a gyfieithir, yn seiliedig ar gostau cyfieithu 2007 Swyddfa Cymru (a adroddwyd mewn ateb Hansard 14 Mawrth 2007).

Dylai ystyriaethau o werth am arian yn gysylltiedig ag allbynnau hefyd gynnwys y disgwyliad fod y project yn disgwyl osgoi o leiaf un ffiasgo tebyg i un diweddar system TG newydd Libra yn Whitehall lle bydd yn rhaid gwario £4miliwn yn ychwanegol oherwydd fod y gallu i gyfieithu gwysion llys i'r Gymraeg wedi cael ei adael allan o'r fanyleb wreiddiol drwy amryfusedd¹⁹.

¹⁹ Gweler <http://www.walesonline.co.uk/news/wales-news/2009/02/13/whitehall-s-welsh-translation-upgrade-costs-4m-91466-22919064/>

5 Casgliadau

Casgliadau'r project yw fod y diwydiant cyfieithu yng Nghymru ymhell ar ei hôl hi o ran mabwysiadu'r ymarfer gorau rhyngwladol a defnyddio'r atebion technolegol. Mae angen y rhain ar fyrder er mwyn cynyddu maint y gwaith y gellir ei gyfieithu ac er mwyn cadw costau i lawr heb gyfaddawdu ar ansawdd.

Mae'r sector yn dal i gadw rhywfaint o'i flas 'diwydiant y bwthyn' gyda chyfieithwyr yn meddu ar gymwysterau uchel, yn dod o gefndir ieithyddol yn hytrach na thechnegol, ac heb gadw i fyny gyda datblygiad cyflym technoleg cyfieithu. Mae maint bychan cwmnïau hefyd yn golygu yn aml nad oes ganddynt yr adnoddau i brynu cymorth a chyngor arbenigol i mewn, neu, gan fod mwy o alw nag y gellir ei gyflenwi, nad oes ganddynt yr amser i ymchwilio'n ddigonol i'r dechnoleg. Mae'r diffyg mynediad at offer cyfieithu a diffyg cyfleoedd ar gyfer cyfnewid gwybodaeth gydag arloeswyr technolegau yn ddwy broblem sydd yn rhaid delio â hwy.

Yn ogystal â rhaglen ar gyfer datblygu 'bocs tŵls cyfieithu' ar gyfer y diwydiant yng Nghymru mae angen brys am ganolfan arddangos i wasanaethu anghenion cwmnïau bach a chanolig a'u cwsmeriaid, yn arbennig unig fasnachwyr a chwmnïau bach iawn sy'n teimlo ar hyn o bryd nad oes ganddynt unman i droi am gymorth.

Mae angen rhoi sylw ar frys i'r gefnogaeth y gall y diwydiant cyfieithu ei roi i fusnesau yng Nghymru yn eu hymdrechion i gyfathrebu ac allforio mewn amgylchedd amlieithog, byd-eang. Mae angen cymorth ar ddiwydiant Cymru i wneud y mwyaf o'r gefnogaeth hon, ac mae angen cryfhau'r gadwyn gyflenwi leol o fewn Cymru.

Fodd bynnag, byddai rhaglen o gymorth i'r diwydiant cyfieithu a'u cwsmeriaid, yn datblygu, arddangos a defnyddio technoleg cyfieithu, yn debyg o gael effaith sylweddol ar economi Cymru. Mae'r effaith yn un driphlyg:

1. Cynyddu capasiti o fewn y diwydiant cyfieithu yng Nghymru
2. Ateb y bygythiad o gael eu cymryd drosodd gan gystadleuaeth ryngwladol sy'n fwy ymwybodol o dechnoleg
3. Cynorthwyo sectorau eraill o ddiwydiant Cymru i wneud y mwyaf o'r potensial ar gyfer gwerthiant, marchnata a gwasanaethau amlieithog, mewn marchnad fyd-eang.

O'r rhain, mae'n debyg mai'r effaith ar alluogi sectorau eraill o ddiwydiant Cymru i weithredu'n llawn yn y farchnad fyd-eang, amlieithog yw'r un fwyaf. Fodd bynnag, ni ellir gwneud hyn ond ar sail diwydiant cyfieithu iach, technolegol flaengar. Byddai darparwyr gwasanaeth cyfieithu a'u cwsmeriaid fel ei gilydd yn elwa o gryfhau swyddogaeth offer technoleg cyfieithu yn y diwydiant cyfieithu yng Nghymru.

6 Argymhellion

- 6.1 Sefydlu canolfan arddangos ar gyfer y diwydiant cyfieithu a'u cwsmeriaid yng Nghymru lle gallant weld manteision technoleg cyfieithu a'r defnydd ohoni.
- 6.2 Ymgymryd â rhaglen o goladu a datblygu rhaglen o offer cyfieithu perthnasol ar gyfer 'bocs tŵls cyfieithwyr' i'w defnyddio gan y diwydiant cyfieithu yng Nghymru.
- 6.3 Datblygu rhaglen o weithio gyda'r cwsmeriaid a darpar gwsmeriaid y gwasanaeth cyfieithu i'w galluogi i rhan yn well yn yr economi amlieithog, fyd-eang.
- 6.4 Dod â'r nodau uchod at ei gilydd mewn Canolfan Technoleg Cyfieithu Genedlaethol, gyda chylch gorchwyl i lywio datblygiad technoleg cyfieithu yng Nghymru mewn ffordd gynaliadwy gyda golwg ar fod yn hyfyw a thyfu dros dymor hir.

Appendix 1: Feasibility Studies Report: Components for a coherent ‘translator’s toolkit’

Feasibility Study 1: Accessibility aids

Feasibility Study 2: Content resources

Feasibility Study 3: Software resources

Feasibility Study 1: Accessibility aids

1 Background

Disability is a serious issue for workers and employers in all industries, and the translation industry is no exception.

Translators typically sit in front of a computer for most of their working day. This may make it a suitable career for linguistically qualified professionals who are disabled, especially given the multiplicity of aids now available to enable disabled users to fully access computing resources. However, the heavy use of computers and sedentary workstations poses its own risks to health, and the preventative aspects of any tools to protect translators from such risks as repetitive strain injury (RSI) need also be considered. The health and safety of all translators using a computer workstation needs to be assessed and issues of accessibility are relevant to all, not only to those translators who report themselves as disabled in any way.

Accessibility needs generally fall into two categories: mobility and vision. Mobility includes problems with fine motor skills needed for typing and controlling a mouse, as well as using walking aids or wheelchairs. Vision issues include deteriorating eyesight and eye strain, as well as visual impairment and blindness. Given that different solutions are required for the two categories of disabilities, and so they are discussed separately below. A common feature however is that these solutions are not in the main specific to translators, and are in general suitable to all computer users.

Further details on the needs of disabled translators may be found in Appendix 2, Scoping Discussion 4: Disabled translators.

2 Hardware Devices

These include ergonomic devices designed to cause less strain to the body than standard equipment when used for long periods. They are suitable for some mobility-impaired translators and those wishing to avoid RSI. They include:

Alternative Keyboards

- Split 'wave' keyboards, considered more ergonomic by some
- Dvorak – a keyboard layout calculated to be the most efficient for typists of English
- Colemak – an alternative keyboard layout to Dvorak, more similar to QWERTY and therefore, it is claimed, easier to learn
- Handykey – a keystroke combination-based single-handed keyboard (marketed by www.handykey.com)

Alternative Mice

- Trackballs
- Touch pads

- Graphics tablets
- Joysticks

The use of multiple input devices for different computer-related tasks is sometimes advised as it is less likely to induce strain than the use of a single device.

Advice on suitable hardware is often available from disability organisations, from occupational therapists and health and safety personnel in larger organisations. However, the high proportion of freelancers employed within the translation industry means that many translators are responsible for their own health and safety and do not receive any guidance from an employer. Currently, information targeted directly at translators is hard to find or nonexistent.

Technology development of these hardware devices is well established, and translators' needs are in general catered for in the wider context of computer devices for disabled users. However the benefit of using these devices needs to be highlighted to translators and their employers, and discussion and demonstration of available aids should be mainstreamed within translation technology. Translators in general would benefit from hands-on demonstrations of these devices, in order to raise awareness and to give individuals the opportunity to try out a variety of devices in an unpressured environment.

Mobility issues to be considered also include adjustment to desk and chair height, provision of footrests, correct posture and use of good quality office chairs, and are relevant to all translators, irrespective of whether they report themselves as disabled or not.

3 Software Typing Aids

Software solutions can also ease the strain of typing. They include:

Predictive text applications

These lessen the number of keystrokes required by translators whilst translating. Commercial applications exist, e.g. Type Booster (see www.typebooster.com). Such applications usually include massive built-in vocabularies of most frequently used words, e.g. Type Booster has an English vocabulary derived from the 100 million British National Corpus of English. They have the facility to learn new words, and language specific data could be made available relatively easily for unsupported languages where corpora or lexicons are available.

Autocomplete algorithms for writing shorthand

These may be set by individual translators or their support personnel. For example, translators could set "Welsh Assembly Government" to be entered into a document every time a shorthand such as "wag*" is typed.

Lists of such lengthy terms common in documents needing to be translated could be generated from bilingual corpora such as translation memories, and standardized

shorthand established, especially for full forms of acronyms (by inserting an asterisk following the acronym, for example).

Translation memory systems

Although reducing typing strain is not the primary intention of a TMS, it is a useful added benefit of such systems. For example, when previously translated sentences are identified for reuse, these can be utilized using a single keystroke rather than a keystroke for every character found within the sentence.

New developments in commercial TM systems are now addressing the needs of disabled users. This may happen as part of a general drive to increase productivity. One such feature is the ability to leverage phrases from the TM as the translator types (see SDL Trados Studio 2009 Autosuggest – “Smart Suggestions as you Type”).

4 Speech Recognition

Speech recognition software converts the user’s speech into editable text on the user’s computer, and can work in conjunction with software such as Microsoft Word and some TEnT such as Trados (although some issues have been reported). This technology is the one that is most in demand from translators unable to type, and from those wishing to avoid RSI.

For translators with severely impaired mobility, speech recognition software can also be used to navigate the system and software menus, allowing access to the software’s functionality without the use of traditional input methods.

The market leader for speech recognition is Nuance’s Dragon Naturally Speaking, although Microsoft Vista also includes extensive speech recognition capabilities. For optimum performance, such systems need to be trained to recognise the particularities of the user’s speech, individual voice, manner of speaking and accent. The quality of their output also generally improves when the system has been trained to recognise the individual user’s voice.

However, these systems are language-dependent, and only English and a handful of other commercially attractive languages are currently catered for. This works well for translators working *into* English, and other major European languages (Dutch, French, German, Italian and Spanish in the case of Dragon Naturally Speaking). It is not yet developed sufficiently to aid translating into Welsh, which accounts for a large proportion of translation work in Wales.

Integration of Speech Recognition into TEnT

A recent research project at the LTU at Bangor University has developed basic speech recognition resources for Welsh as part of a project funded by the Welsh Language Board. This will provide a basis for the development of commercial software to supply the needs of translators in need of speech recognition for Welsh and other less-resourced languages. Although the demand for speech recognition from translators of other less-resourced languages within

Wales is minimal, such projects which seek to provide low-cost solutions will aid other language communities in developing appropriate resources for their own languages.

Good quality speech recognition for Welsh is a key component currently missing from the disabled translator's toolkit in Wales.

5 Text to speech technology

Text to speech software allows blind or visually impaired users to have access to the text found on computers and on the internet. Electronic text in documents, emails, calendars etc is read aloud using a synthetic voice. This allows enables visually-impaired translators to access source texts and type the translation using either touch typing or a Braille keyboard. In a translation office environment, other members of staff may be visually-impaired, and text-to-speech technology also allows administrators and other members of the office to access written text.

While text-to-speech software is language-dependent, it is available for a wider range of languages than speech recognition. The LTU at Bangor University has developed a diphone voice for Welsh, which is available under free and commercial licences (see http://www.e-gymraeg.org/wispr/index_en.htm). This voice is currently available for screen readers and has been integrated into a basic word-processing programme, EdGair (a Welsh version of Manchester University's EdWord programme). An improved commercial version of the voice has appeared on numerous Welsh public sector websites through licensing agreements with companies specializing in speech enabling online content such as Readspeak and Texthelp.

6 Other aids for visually impaired translators

Low cost adjustments

Whilst commercial companies and academic researchers like to concentrate on high end technological solutions, some small-scale adaptations within existing computer workstation environments can make a significant difference to the individual translator. These include applications with features for:

- Increasing the contrast of the computer screen and interface design
- Increasing the size of icons
- Magnifying applications (Zoomtext for example)
- Catering for colour blindness

Adherence to guidelines on standards for screen display, such as the Web Content Accessibility Guidelines (W3C) also help visually impaired translators. These need to be taken into consideration when designing any new technology aids, e.g. terminology portals and electronic dictionaries.

Monitor Screens

Many translators based in Wales reported using computer screens that are small by today's standards. Larger screens enable the translator to increase the onscreen size of documents whilst still allowing enough of the document to be viewed without scrolling. This lessens the strain on the translator's eyes, which is especially important as vision naturally deteriorates with age.

The use of a pair of screens was very uncommon, despite the benefits to the translation's productivity of having the translation appear on one screen whilst being able to refer to additional resources such as online dictionaries, reference works or the original text on the other screen.

Freelance translators reported reluctance to spend money on what they considered to be non-essential purchases, whilst both private and public translation offices seemed unaware of the health and safety benefits of using larger or multiple screens. Monitor screens are not expensive to buy considering the time translators spend using them, and offer a low cost solution to improving productivity and health in the workplace.

Demonstrating the benefits of larger, multiple monitors should be a priority in any project showcasing translation technology.

7 Remote Simultaneous Translation Facilities

Simultaneous translation is the use of an interpreter to translate words spoken in one language into another language whilst the original speaker is still speaking. Translation equipment consisting of microphones, translator's panel and headsets provide an easy and unobtrusive solution to the problem of holding meetings and conferences in a bilingual and multilingual environment. This technology is very mature in Wales and widely used at local and national level. Equipment is also manufactured within Wales (for example see the Cyflais system at http://www.idb-tech.com/interp_e.htm).

While the interpreter usually travels to the venue where the meeting to be translated takes place, discussions within CCC have come up with the intriguing possibility of remote provision of simultaneous translation, where the interpreter translates from his or her home, or other convenient location, thus removing the need to travel to meetings, with the original sound and the oral translation being relayed electronically. This technology has yet to be fully developed, although simultaneous translation through video-conferencing facilities in Wales is already being used. It has been suggested that further development of this technology would be welcomed by mobility-impaired interpreters. Companies already manufacturing translation equipment should be encouraged to research and develop remote translation devices. This would help mobility impaired interpreters, as well as those based in rural and remote locations or those not wishing or unable to travel far. It would also help cut down on the cost of interpretation services, where travel costs often outweigh the actual per hour charge for the simultaneous translation.

8 Further information

A number of organisations provide information and guidance relevant to disabled translation aids, including:

- The Health & Safety Executive <http://www.hse.gov.uk/>
- The Shaw Trust <http://www.shaw-trust.org.uk/home>
- Disability Wales: <http://www.disabilitywales.org/>
- RNIB Cymru
http://www.rnib.org.uk/xpedio/groups/public/documents/code/public_rnib003463.hcsp
- Wales Council for the Blind <http://www.wcb-ccd.org.uk/>
- Wales Council for Deaf People <http://www.wcdeaf.org.uk/>
- Access Group Resources <http://www.accessgroupresources.co.uk/>
- Web Accessibility Initiative <http://www.w3.org/WAI/>

Feasibility Study 2: Content resources

1 Definition

Content resources for translators are linguistic assets referred to during the translation process. They include monolingual, bilingual and multilingual glossaries, dictionaries, and specialist terminology dictionaries (together called lexical resources); proofing tools for spelling and grammar checking; and previously translated text in the form of pre-populated translation memories, concordances, or aids to automatic translation.

These resources not only ensure the quality of the translation; they also increase productivity by helping the translator to solve linguistic issues that would otherwise prevent or delay translation, allowing the translator to complete a translation with its attendant proofreading, editing and revision tasks in the shortest possible time without loss of quality.

Some of these content resources have been available traditionally in paper form, as a collection of dictionaries, grammar books and published style guides or manuals. However, their availability in various electronic formats has opened out the field to new and improved content resources, expanding beyond what was traditionally available and automating a number of translation processes. The provision of *content* resources is therefore closely entwined with the *software* resources available to translators. The software acts as containers or delivery media for the content, but an important difference is that content resources are language dependent, whereas software is, or can be, language independent.

Content resources therefore need to be discussed according to the needs and requirements of individual languages that translators work with. More commercially attractive languages, such as the major European languages of the computer-literate Western world, are well-endowed with content resources, whereas the languages of the third world, where computer technology has traditionally been disadvantaged, are less abundant, although there are currently numerous initiatives to remedy this deficiency.

Welsh, as a minority language in a developed Western nation, sits in an intermediate position in terms of content resources available. On the one hand it does not have the range or choice of resources available for the major European languages, but on the other hand many of the resources it has have already made the transition into electronic media and the knowledge base within Wales itself possesses the capacity to respond to translators' needs.

2 Lexical resources

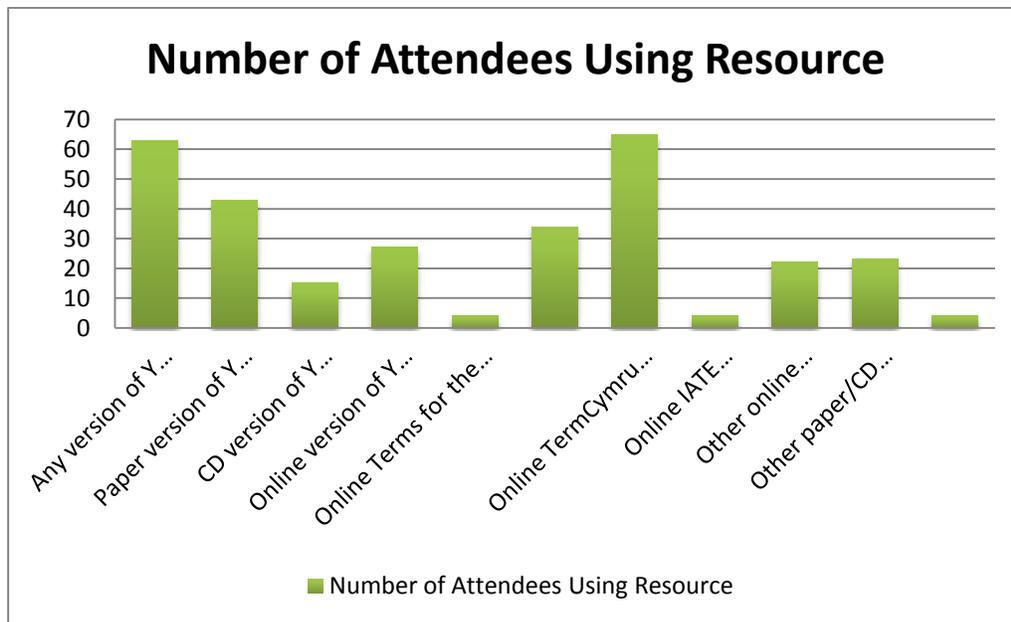
While general language dictionaries provide the baseline of translation tools, professional translators also need access to the specialist terminologies used in their specific domains. These may concern education, tourism, health, law, the car industry or any number of other specialisms. Specialist terminology dictionaries may be produced by academia, private publishing houses, industry associations or other bodies. On a European level, the European Union provides on-line electronic resources through its IATE (Interactive Terminology for Europe) termbase (<http://iate.europa.eu/>), and the Eurovoc Thesaurus (see

<http://europa.eu/eurovoc/>) which covers the fields in which the European Communities are active. These resources are available mainly in the 22 official languages of the European Union, and are freely accessible without charge through the internet. They are the most comprehensive multilingual resources publicly available world-wide. Given, however, that they are of limited usefulness to the translation industry in Wales where the dominant language pair is English/Welsh, adding Welsh to the Eurovoc database would be one way of remedying this situation. The parliaments of Russia, Ukraine and Albania have paid for Eurovoc to be translated into their languages, as has the provincial Government of Bizkaia into Basque within the European Union.

A selection of the main interactive on-line specialist English/Welsh terminology dictionaries which are available on the web are listed below:

- Y Termiadur: <http://geiriadur.bangor.ac.uk/termiadur/>
(Standardized terminology for the National Curriculum in Wales, published by ACCAC, the Qualifications, Curriculum and Assessment Authority for Wales)
- Dictionary of Terms for the Creative Industries: <http://termau.org/cyfryngau/>
(published by the Centre for Welsh Medium Higher Education)
- TermCymru: <http://www.termcymru.cymru.gov.uk/>
(published by the Welsh Assembly Government)
- The Welsh National Database of Terms <http://www.e-gymraeg.co.uk/bwrdd-yr-iaith/termau/Default.aspx>
(published by the Welsh Language Board)

Y Termiadur is available in 3 other formats in addition to the on-line version: a traditional paper dictionary, a CD and a downloadable version for mobile phones. A sample of 78 translators in two workshop sessions arranged by CCC in November 2008 were asked for their use of these resources, as well as the IATE multilingual database. The results are given in the table below.



The most popular resources by far are the TermCymru and Y Termiadur. This is to be expected as they are both comprehensive in scope and authoritative as sources. IATE's low score reflects the fact that it Welsh does not feature as one of its languages.

Asked what other content resources translators would like to have on-line, 83% of the workshop attendees replied that they would like to have an on-line Welsh language thesaurus available. A thesaurus, although monolingual, helps translators find the exact word, or 'mot juste' that they are looking for when they have an approximate translation, but believe that a near equivalent or synonym would better fit the specific context. A Welsh thesaurus is available on CD within the Cysgliad compendium published by the LTU, Bangor University and could be adapted to fit an on-line environment and be packaged with terminology resources.

A similar percentage of the workshop attendees, 83%, noted that they would like an on-line Welsh/English general language dictionary to aid them in their translation work. Again, such a dictionary (called *Cysgair*) exists within the CD Cysgliad compendium. If adapted to fit an on-line environment, it could be packaged with other terminology resources. The Welsh Language Board is currently engaged in a project to provide an on-line version of the *Welsh Academy Dictionary*, the largest and most comprehensive general English-Welsh dictionary in existence. It is an invaluable aid to translators as it includes idioms, phrases and additional information relevant to the translation process. The digitization work is being undertaken by the LTU at Bangor University, using its own *Maes T* on-line terminology and lexicography editing tool.

Place name dictionaries are another class of valuable resources for translators. Place-names in Wales can cause difficulties for translators as they may be unsure if a place-name found in a document is a bilingual one with an alternative in the language of the translation, as is the case with Swansea (Abertawe) and Cardiff (Caerdydd). Added confusion is caused by the fact that many places in Wales possess the same name in one language, but the names in the other language do not correspond such as in the cases of Newport (Casnewydd) in Gwent and Newport (Trefdraeth) in Anglesey and Pembrokeshire respectively, and of the many versions of

Tal-y-bont found in Wales, only one of them having the equivalent name of Buttington in English. The Welsh Language Board-funded *Enwau Cymru* on-line database (<http://www.e-gymraeg.co.uk/enwaucymru/>) provides translators with translations and information on place names in Wales. A particularly valuable feature of this database is its links to various on-line maps showing the location of place names. Therefore, it is possible to see the location of the various ‘Newports’ in Wales and match it with the correct place name in Welsh. Translators would benefit from the integration of this database into other translation environment tools.

Apart from TermCymru, all the above resources have been developed by the LTU at Bangor University and conform to the international SOAP: Simple Object Access Protocol standard (<http://www.w3.org/TR/soap12-part1/>) which enables them to be federated with each other and with other SOAP-conformant terminology dictionaries.

Federating such resources, and bringing them together in a one-stop shop environment has consistently been cited by translators in Wales as one of the main improvements they would like to see to their translation environment, saving time and improving quality together. This has been refined as the idea of a web-based Terminology Portal. Again 83% of respondents at the CCC workshops noted that they would like to see a terminology portal to bring together the resources they most used into one location. One of the respondents noted:

“Mae hyn yn gwbl angenrheidiol er mwyn gofalu am ddyfodol cyfieithu fel diwydiant sydd yn un allweddol yng Nghymru ac yn Ewrop.”
 (“This is absolutely necessary in order to safeguard the future of translation as a key industry in Wales and in Europe”).

Such a portal would not need to be language dependent as it could federate terminology and allied resources from any languages. Resources that are already available and used by the industry in Wales should be federated first, with any gaps in provision being identified for future projects. The need for databases to be SOAP-compliant before they can be federated needs to be emphasized however, and should be a core requirement for any new terminology or related project.

2 Proofing Tools

Spelling and grammar checkers

Electronic spellcheckers are now routinely included in all leading wordprocessing packages such as Microsoft Word and OpenOffice.org. A different spellchecker is needed for each language, and extension packs for languages not supplied with the original software may be obtained either free of charge or for an additional fee. A Welsh spellchecker is available for both Microsoft Word and OpenOffice.org. A similar spellchecker is available in the Cysill programme, sold by Bangor University as part of the Cysgliad compendium. All translators questioned in the focus groups reported using spellcheckers, with a nearly universal use of Cysill for Welsh spelling and grammar checking. In some cases however, it was noted that Cysill was being used only at a very basic level, with higher features such as the ability to create and edit a personal dictionary not being used, or not being well understood. A demonstration of the higher functions of Cysill

was very much appreciated by the focus groups, with requests made for further sessions of this nature.

The point was made that Cysill and similar tools are general purpose language aids and are not targeted specifically at translators. Many suggestions for additional functionality were made, especially in the integration of Cysill etc. into coherent workflows within translation environment tools. At present Cysill has to be run over a document as a further proofing step after correcting and revising the initial translation. This is time consuming, and is problematic for those using TEnTs (translation environment tools).

Integration of proofing tools such as Cysill into TEnTs was seen as the main translators' need. Proofing tools for major languages were already included in the leading commercial TM systems, and the absence of Welsh in such commercial products was criticized as a major weakness. If existing software producers were unable or unwilling to include Welsh in their TM proofing tools, it was suggested that an alternative TM system be developed that could more easily include Welsh content. This was seen as a solution for other less-resourced languages as well.

For languages such as English, which possess complex spelling conventions and are non-phonetic in nature, a spellchecker may be a higher priority than a grammar checker. English has less complex morphology, and therefore grammar checkers are not seen as so important by many who translate into English. For Welsh and other more morphologically complex languages, however, grammar checkers are seen as essential translation aids. Further work is being undertaken in the development of the grammar rules within Cysill, but again, the priority for translators is the integration of such tools within their TM or translation environment.

Further demands

Amongst other items on translators 'shopping lists' were:

- Spellchecking dictionaries for technical terms (it was noted that some some TEnT allow the use of multiple spellchecking dictionaries, and that the possibility of producing additional subject specific dictionaries could be explored for commonly translated subject fields).
- Modules to identify deprecated, non-standard terms, and point the translator to the appropriate standard term.
- 'Do not suggest' lists (spellcheckers often employ lists of words that should not be suggested as corrections. However, translators need occasionally to disable such lists, e.g. when translating vulgarities in literary texts).
- Modules highlighting stylistic inconsistencies, for example deviations from a company's own house style.
- Features to allow the setting of formal, informal, colloquial and other language registers in individual languages.

It was pointed out that a feature to alert the user to an unsuitable style register could also be used by translators' customers to prepare text in clear, unambiguous language before translation. This would be especially valuable for industry customers such as exporters who need text translated from English into multiple languages.

3 Pre-populated Translation Memories and other pre-translated text

Translation Memories as *systems* are discussed in the study on software resources. However, the content that make such systems useful are parallel, equivalent text in two or more different languages. A TM becomes more useful to a translator the more reusable content it contains. It is sometimes possible to obtain pre-prepared translation memories that give translators a head start when using such systems. The value of such memories varies according to the number of matches the translator is likely to obtain. They are more likely to offer translated segments if they are in a related limited domain (e.g. car manuals or food menus) rather than general all-purpose ones. Public bodies who regularly use translation services may wish to share such resources freely with others, but commercial translation agencies regard their memories as valuable commercial assets and are unlikely to distribute them freely. Customers in the private sector may build up their own TM content and use it to keep down translation costs, only needing to pay for new segments that need translation.

Pre-prepared content will be needed for a demonstrator of TM technology, and this could provide the opportunity for collecting material which is often retranslated. This would include:

- Standard business letters
- Common idioms
- Software interfaces
- Software licences
- Terms and conditions of use
- Food packet texts
- Laws and regulations
- Names of organisations

Provided copyright clearance is obtained, such a resource could be a valuable asset to translators in Wales.

ELDA (Evaluations and Language resources Distribution Agency) exists in Europe to facilitate the distribution of language resources in general (see <http://www.elda.org/>). However, their services are aimed at researchers and developers rather than at the ordinary translator. The needs of translators may best be served by a dedicated translation technology centre.

Feasibility Study 3: Software resources

1 Background

The purpose of this feasibility study is to investigate the software based translation tools available to translators. These include software aids that help the translator during the translation process itself, as well as software that facilitates the management of the translator's resources from project to project. They also include software that administer and manage the translation process, and software that may be used by customers of the translation industry as they engage with translators and translation tasks.

1.1 Definition of Translation Software Resources Software programs or applications that are designed to help translators with the task of translation are generally known as a Computer-Assisted Translation (CAT) tools. These aid the translator in delivering efficient, consistent and high quality translation. CAT tools range from commercial products costing thousands of pounds to free software available for download from the web. They include small, single purpose utilities as well as whole translation environments that attempt to integrate many diverse translation tools into one integrated Translation Environment. This feasibility study will follow the practice established by Jost Zetzsche of referring to these environments as Translation Environment Tools, or TEnTs, whilst the term 'CAT tools' will be used to refer to translation tools in general, whether or not they are packaged as part of TEnTs such as Trados, Wordfast or Déjà Vu.

CAT tools and their components generally include the following:

- Translation environments (TEnTs)
 - File Segmentation
 - Integration of various CAT tools
 - File Filters
- Proofing Tools
 - Spellcheckers
 - Grammar Checkers
 - Quality Assurance (QA)
 - Terminology QA
 - Well-formed code QA
- Terminology Aids
 - Terminology Resource Interfaces
 - Paper
 - CDs
 - Websites
 - Portals

- Servers
 - Terminology Extraction Utilities
 - Monolingual Term Extraction Utilities
 - Bilingual Term Extraction Utilities
 - Terminology Standardization Systems
- Translation Memory Systems
 - Segment Alignment
 - TM Management
 - TM Anonymizing
- Concordancers
- Online Resource Portal

These tools are further explained below:

- **Translation Environment Tool (TEtT)**, which form a working environment for the translator and usually include various CAT functionalities, especially Translation Memory systems.
- **Spellcheckers**, either as part of word processing software or translation environments, or as stand-alone programs.
- **Grammar Checkers**, as separate programs, or as built into word processors such as Microsoft Word or translation environments.
- **Dictionaries** on CD-ROM, or online, including both multilingual and monolingual dictionaries.
- **Terminology databases** on CD-ROM or online.
- **Translation Memory** managers (TMM), which store translated text segments in a database alongside the source language text segments so that they may be reused during translation if similar segments of text are encountered.
- **Text Alignment** utilities for aligning previously translated document segments with their corresponding segments in the original documents for the purpose of creating Translation Memories.
- **Terminology Extraction** utilities for the identification of a document's terminology prior to its translation so that key terms may be identified and standardized in advance.
- **Bilingual Terminology Extraction** utilities to create bilingual terminology lists from previously translated materials and their original counterparts.
- **Terminology Managers** which allow translators to manage and exploit their own terminology lists electronically. This ranges from simply using tables created in the translator's word processing software or spreadsheet, using a simple database, or using tailored solutions specifically designed for managing terminology, including the sharing of terminology between a team.

- **Concordancers** for searching existing texts such as monolingual or multilingual corpora and translation memory to discover the context and usage of a word or phrase, or find translation strategies that have been previously employed.
- **Filters** that allow translators to translate the language content of files within their translating environment (with all its associated aids) whatever the file formats, and without having access to the specialist software used to create the documents. Specific filters must be created for each individual file formats, but many are available for download.
- **Translation Management Systems** that allow project managers to manage multiple translation jobs, the availability, specialization, and language skills of their translators, and the details and requirements of their clients.
- **Quality Assurance** utilities for automated quality checks – terminology, well formed code.

In addition, the following more general tools are of use to the translator:

- Simple **Project Management** software that allow freelancers or very small teams to structure complex translation projects, assign the various tasks to different people, and track the progress of each of these tasks.
- **Search tools**, which allow the user to query already translated texts or reference documents, including internet search engines, but that also include specialized index tools developed for the translation industry.

1.2 Translation Environment Tools (TEnT)

Translation Environment Tools (TEnT) such as Déjà Vu, SDL Trados and Wordfast are integrated software environments designed to assist the translator when translating by bringing together numerous complementary CAT functionality in one interface.

These generally include Translation Memory (TM) systems, spellcheckers and terminology databases (also known as termbases or TBs). In addition some also include more specialized translation tools such as concordancers and automated QA checks.

Increasingly some example based machine translation (EBMT) is also to be found in the latest releases of some commercial TEnT systems. These make use of the translator's TMs and TBs to suggest automated translations, but are largely rule based and therefore language dependent.

1.3 List of TEnT tools

The following is a list of widely available Translation Environment Tools.

Commercial Translation Environment Tools

The three most widely used commercial Translation Environment Tools both internationally and in Wales (according to Lagoudaki 2006 and Andrews 2009) are:

- SDL Trados
- Déjà Vu
- Wordfast

Other commercial Translation Environment Tools include:

- Across
- AppleTrans
- Cafetran
- CatsCradle
- ENLASO Localization Tools
- Fusion Translate CMT
- Heartsome Translation Suite
- IBM Translation Manager
- iLocalize
- Lingobit Localizer
- Logoport
- Logovista
- MemoQ
- Metatexis
- Multilizer
- Multitrans
- ProMemoria
- SIMILIS
- TrAID
- Transolution
- Transware Ambassador
- Visual Localize
- WebBudget
- Wordfisher

Even though this list is not exhaustive, it is apparent from the above list that a wide variety of Translation Environment Tools are available.

New Translation Environment Tools are regularly launched, reflecting the fact that TeNT software is seen as a growth area. Many such programs fail to sufficiently penetrate the market, and a small but significant proportion of TeNT software packages were identified as no longer being supported. This confusion and breadth of choice compounds the difficulty faced by translators considering adopting Translation environment Tools.

SDL Trados, Déjà Vu, and Wordfast account for 90% of the international market according to Lagoudaki (2006), and these three systems are also the market leaders in Wales, where it is estimated that the figure would be similar or higher.

In addition to commercial TeNTs, there is an increasing number of free TeNTs available for use. Some of these are closed code (i.e. other software developers may not adapt and further develop them without permission) and some are open source (permitting other developers to adapt and develop them with varying degrees of restriction or with no restrictions imposed).

Free commercial Translation Environment Tools include:

- AidTransStudio Basic
- Across Freelance

Free Translation Environment tools featuring an open source licence include:

- OmegaT
- Anaphraseus
- Virtaal

Open source software may, with due regard to any licensing restrictions, be used to adapt and develop various utilities for the translation industry in Wales, both for demonstration purposes and for dissemination amongst practitioners.

2 Translation Memory (TM) systems

Translation Memory systems are one of the main components of translation environment tools (TEnTs). TM systems store translated text segments in a TM database alongside the corresponding source language text segments, so that they may be reused during translation if similar segments of text are encountered.

Basic TM systems are language independent. They require no knowledge of a language's grammar rules or morphology to function, relying solely on comparing the statistical similarity of translation segments to function.

TM systems do not produce new original translations, and do not replace the translator. By allowing the reuse of similar or identical existing translations, they greatly increase productivity and consistency of the translator.

In his academic study of translation practices, Bowne (1999) reported that using a TM system whilst translating repetitive texts resulted in up to a 40% saving. Figures of between 20% and 80% savings have also been reported. During a recent 85,000 word software translation project at Bangor University, the use of a TM system to translate the highly repetitive text files caused an immediate 20% (or 17,000 word) reduction in the translation work required, starting with an empty memory with no previous input. This reduction was

achieved solely by identifying instances of identical text strings within the software's various language files, and requiring only the first instance to be translated. Subsequent identical examples were then translated automatically, building the TM content as the project progressed.

Estimates based on a current average cost of translation from Welsh to English is £80 per 1000 words gives a general cost saving of:

$$17,000 \text{ words} \times \text{£}80 \text{ per } 1000 \text{ words} = \text{£}1360$$

Note that the above figure includes only savings from the identification and reuse of 100% percent matches found within the text awaiting translation.

It does not include the further significant savings from the identification and use of partial (less than 100%) matches, nor does it include the potentially considerable savings from the use of pre-existing TMs and terminology lists or glossaries. Nor does it include the considerable time saving gained in using such a TM system.

Taken together, these could increase the potential savings in efficiency from 20% to 50-60%.

This figure corresponds to the savings experienced by one translator (specializing in translation from English into Russian) by reusing translations from pre-existing TMs:

“For example, when working with one of my long-term clients, Volvo Cars International, about 60% of the sentences in the brochure on the Volvo S60 were sentences I had already translated working on the Volvo S/V70, S80, S/V40 brochures. In such cases, which are not at all rare, a Translation Memory saves me a lot of time and ensures perfect consistency of terminology, which is appreciated by the client.”

Andrei Gerasimov

Source: <http://www.proz.com/translation-articles/articles/220/1/An-Effective-and-Inexpensive-Translation-Memory-Tool>

2.1 Sharing of TMs

Translation Memories may be shared between translators, allowing multiple translators to contribute to and benefit from the TM. Sharing translation memories in this way can greatly increase the size of the TM and consequently the chance of matches being found in the TM for sentences requiring translation.

Owing to the amount of translation work stemming from the public sector in Wales, methods of facilitating the sharing of TMs between various public sector bodies and their translation service providers should be developed where appropriate as these could lead to significant financial savings and increased consistency through the pooling of resources.

2.2 Standardizing TMs

The sharing of Translation Memories raises issues of standardization. Where a standardized language is desirable, such as in many official documents, repetitions from the TM can be maximized by establishing guides to be followed by all contributing translators for the following linguistic features:

- Style (for instance, the use of short or long forms of the verb in Welsh)
- Register (formal vs informal, for instance)
- Dialect (different words may be used for the same thing in different regions)
- Spelling convention (for instance *organize* and *organise*)

Although such issues are often addressed within organization style guides, their further development within a proofing tool similar to a grammar checker are outlined in Sections 3.4 and 3.5 below.

The standardization of text is not always appropriate. The automatic conversion of TM segments from one style/register/dialect/spelling to another would be of benefit in cases where documents are required to be translated into different styles, registers, dialects or spelling conventions.

For example, Welsh words for *milk*, *rake*, *up*, *mole* and so on vary between certain Welsh regions (primarily between the north and south), and methods of automatically localizing TM segments would be useful when translating texts targeted at specific areas. Such variations are especially useful in local marketing campaigns or promotion of regional foods.

A grammar checking element would need to be included to deal with maintaining grammatical agreement in the case of gender change and the exchange of an initial vowel or consonant for a consonant (in the first instance) or vowel (in the latter).

2.3 Integration of Machine Translation in Translation Memory Systems

Machine Translation is increasingly being used to combine information held in terminology databases with TM systems to increase their efficiency. However, whilst TM systems have historically been language independent, these new advances require the use of some information regarding the grammar rules and morphology of the supported languages and are therefore language dependent. Further discussion on Machine Translation technology may be found at the end of this Feasibility Study.

2.4 Text Alignment Utilities

Text Alignment Utilities are designed to convert documents that were translated before the adoption of Translation Memory systems into Translation Memories. This is

particularly important for translators who are in the process of adopting to using a TEnT for the first time. Translators who are recent adopters of TEnT technology may not possess a Translation Memory of any significant size. However, Translation Memory systems require large pre-existing Translation Memories before their benefits are fully experienced. Although TMs are gradually created during translation using a TEnT, some time may pass before the TM of a new adopter becomes large enough for its benefit to the translator to outweigh the inconvenience of having to learn a new way of working. By allowing a new adopter to convert documents which were translated prior to the adoption of a TEnT into TM format, Text Alignment Utilities can ensure that new adopters immediately gain from the full benefit of TM systems. This decreases the chance that new adopters revert to using their previous methods of translation due to their feeling that the cost of using Translation Environment Tools in terms of time investment and complexity outweigh any productivity and consistency benefits.

Text Alignment Utilities work by converting pairs of source and target text document into Translation Memories by aligning segments of texts in the translated document with the corresponding segments in the original source language document. These created Translation Memories can be a valuable time saving resources for any translators (not just recent adopters of TEnTs) when translating similar documents in the future. However, most existing Text Alignment Utilities need significant manual input from the translator and can therefore be a time-consuming and costly process.

Despite this expense, many translators in Wales, including private companies and public sector translator units stated that they had undertaken Manual Text Alignment prior to adopting the use of TEnT, and that it had been a worthwhile endeavour.

Examples of existing Manual Text Alignment Software are Trados's Winalign utility and the ID-based alignment utility and Alignment Verification editor which are a part of the Okapi Framework.

2.5 Automated Text Alignment Utilities

Automated Text Alignment Utilities such as MultiTrans's Alignment Agents do exist, but are not fully automatic. These attempt to decrease the need for human correction by using language independent pattern recognition and statistical techniques. However human input is still required, as the processes is not always infallible. The addition of language dependent information for both source and target language is believed to have great potential for evaluating the success of statistical or pattern based alignment methods.

For Welsh, further work needed includes lemmatizing the words found in both texts and using bilingual dictionary information to check for the existence of equivalent terms in the source and target language segment. Alignment results gained from statistical and pattern based techniques can then be further validated.

Positive matches can be kept and added to the Translation Memory, whilst negative matches can simply be discarded - the vast amount of text which can be correctly aligned

using automatic techniques removes the need for every single segment to be retained. Such an utility has many valuable applications for the translation industry in Wales including:

1. Increasing the number of TEnT adopters who permanently convert to using the software by ensuring they have sufficiently large and relevant TMs to ensure that the change in working methods renders them more productive despite the learning curve faced when using unfamiliar software.
2. The quick creation of valuable TMs for organizations retaining original documents and their corresponding translations from the period prior to having adopted TEnTs.
3. The creation of TMs from the documents found on the computers of retired or retiring translators who did not adopt the use of TEnT software.

2.6 Anonymizing TMs

Although the ability to share TMs between multiple translators working internally and externally can lead to significant savings derived from increased efficiency (as well as improved consistency), care needs to be taken that confidential information is not inappropriately distributed. The use of software to automatically anonymize sensitive information such as names and addresses should be investigated to facilitate the management of sharing TMs between organizations, and between organizations and external translators.

2.7 Translation Memory Editor

The regular editing and managing of TMs is important to ensure that inadvertent errors and inconsistencies are not propagated throughout future translations, and that the names of entities such as organizations and companies which may change due to mergers or rebranding can be updated.

This task, however, can be difficult, especially when managing multiple TMs containing thousands of segments.

TMs may need to be edited to ensure that:

- Use of terminology is consistent
- Organization names are up to date
- Spelling and grammar are correct

Most commercial TMs contain some form of mechanism for editing TMs. However these do not contain language dependent functionality or resource lists that would enable problems arising from the above issues to be identified.

As these are closed source software programs, only a limited amount of adaptation for issues arising from the Welsh context is feasible without gaining the input and support of the commercial developers.

However, open source TM editors such as Okapi Olifant are potential candidates for which software processes and linguistic resources could be developed to enable the Translation Industry in Wales to quickly and efficiently improve the quality of their TMs.

Linguistic resources such as lists of problematic or generally confused terms could be produced for use with the editors' filtering mechanisms, allowing terminology inconsistencies in the TM to be easily identified and corrected.

2.8 Translation Memory Manager

Whilst many translators may use only one large TM, most TEnT vendors recommend that multiple TMs for separate subjects and clients be maintained. The reason for this is twofold:

- Translation Memory systems may struggle to cope with TMs beyond a certain size
- By selecting a TM from the same domain as the translation, the relevance of the matches is likely to be higher.

Effectively managing and tracking multiple TMs can be difficult.

Useful aids for managing TMs include the following:

- The automated tagging of TMs by field/subject/domain by recognizing key terms found within a document.
- Tools for joining and splitting TMs
 - a customer such as the Welsh Assembly Government could build their own translation memory, tag them, and export relevant segments to translators who are given translation tasks (e.g. in the field of Social Services)
- Version control

A locally installed, cut-down version of the open source Globalsight Translation Management System could help in managing different TMs.

3. Proofing Tools

Proofing tools include:

- Spell checkers, either as part of word processing software or translation environments, or as standalone programs;

- Grammar checkers, as separate programs, or as built into word processors such as Microsoft Word or translation environments;
- Quality Assurance utilities for automated quality checks – terminology, well formed code.

3.1 Spelling and grammar checkers

Spellcheckers are extremely important aids for translators as poorly spelled or mistyped words are widely considered to be indicators of poor translation and immediately obvious to clients in a way that slightly mistranslated sentences are not.

It is important that these technologies are also available to specialist text editing software such as TEnTs, as these are key to industries such as translation. Translators are increasingly given work to translate in specialist file formats such as Photoshop, Dreamweaver, Illustrator, Acrobat. These pose their own challenges to translators, and any problems are compounded by lack of integrated spelling and grammar checkers for them in languages such as Welsh.

3.2 Existing Welsh Language Proofing Tools

Microsoft Office Welsh Proofing Tools

The LTU developed the Welsh language spellchecker for Microsoft. It is available as a free download for owners of Microsoft Word or Microsoft Office. Some commercial software TEnT such as Déjà Vu, Wordfast and Trados can utilize the Microsoft Office spellchecker if installed, so its continued development is of vital importance to the translation industry in Wales.

Welsh Proofing Tools for Open Source Software

The LTU developed a Welsh language version of the MySpell open source spellchecker used by many open source programs including Firefox, Thunderbird, OpenOffice.org.

Myspell has since been superseded by Hunspell, an improved spellchecker engine which is backward compatible with MySpell dictionaries. Hunspell is superior to MySpell in terms of its support for languages in need of UTF-8 encoding and its support for languages with complex morphology, both of which are required in the case of Welsh.

The LTU is currently in the process of creating a Welsh version of Hunspell. It has addressed many of the issues experienced in some software programs with the Myspell implementation, including the change to UTF-8 character encoding which enables the spellchecker's use with Omega T, the open source TEnT.

Encouraging the integration of the Hunspell engine within commercial print and web design packages such as those produced by Adobe (Dreamweaver, Illustrator, Photoshop) is desirable so that spell checkers for less resourced languages are available to designers and sign makers who may not speak or understand the languages they are called to work with. This would help to reduce the instances of misspellings introduced post-translation during the production of websites, posters and signs.

Hunspell has already been adopted by Google as the spelling checker engine for its browser, Chrome, with the search engine giants utilizing its expertise to add new words that have appeared in many languages to its wordlists.

Grammar Checker for Open Source Software

Work is currently underway at the LTU to enable the integration of the unit's grammar checking engine into open source software, beginning with OpenOffice.org.

Cysill

Cysill is the LTU's flagship proofing tool. It includes both spellchecking and grammar checking functionality and forms part of the unit's commercial Cysgliad package.

3.3 Further work needed on spellcheckers

Integration with text editing programs, including TEnT

Cysill uses Windows's COM interface to exchange text data with other programs such as Microsoft Word. Whilst Cysill installs a button on the Microsoft Word toolbar that initiates the spelling and grammar checking processes, Cysill is in fact a standalone program. The button serves merely to launch the program and transfer the text which requires checking to the Cysill interface, at which point the checking process is started. Upon completion, the validated text is transferred back to the program of origin, again using the COM interface.

Most other text editing programs that are compatible with Windows's COM interface can also check the spelling and grammar of their text in this way, substituting the use of a toolbar button with the use of a user defined shortcut key combination (Ctrl + Alt + W is the default).

Improving Spellchecker Suggestions

Frequency counts derived from the number of times a word appears in a corpus should be used to improve the accuracy of suggestions by suggesting commonest words first. A larger text corpus is urgently needed for Welsh, to improve this functionality.

Using n-grams would improve the relevance of suggestions by referring to the words surrounding the misspelling in order to probabilistically determine the most appropriate correction.

Performance

Currently, Cysill must be run as a process for spelling and grammar checking to occur. Having the program run continuously as you type (as is the case with the red and green squiggles in Word) would improve functionality.

Such a feature would allow translators using TEnT software to check the spelling and grammar of a segment prior to moving on to the next, at which point the previous segment is submitted to the translation memory.

Many Translation Environments allow the spelling and grammar to be checked following the completion of translation.

However, unless the TM segments have been updated following any corrections, the segments sent to the TM will still contain any original errors, despite the errors having been corrected in the translated document.

Many translators are unaware of this, or, unless they follow a designated workflow pattern, do not always remember to update their TM segments following the final proofing. This inevitably leads to misspellings and typing errors being propagated by the TM system.

Interface Improvements

Improvements to the interface could include:

- Enabling hot keys to be activated by keyboard or by voice
- Better support for visually impaired and disabled users
 - Large icon mode
 - High contrast mode

Terminology Integration

- Enhanced terminological information, including displaying the status of a term:
 - Standardized (recommended)
 - Candidate (no standard term, accepted)
 - Deprecated (not accepted) etc.

Other improvements to consider:

- preferred terms suggested first

- non-preferred terms identified as incorrect with preferred terms being suggested in their place
- deprecated terms within the text flagged as errors of higher severity

Multilingual Document Support

Improvements could include:

- The ability to add English words to an English dictionary, but exempt English words that are similar to Welsh misspellings
- Build a dictionary of words from other languages –click on 'Learn as English' to add word to an 'English' personal dictionary
- Work with other languages' Hunspell (eg. English Hunspell)

3.4 Grammar Checkers

As grammar checking is dependent on words having been spelt correctly, for best results grammar checking should occur prior to spellchecking or in conjunction with spellcheckers.

Grammar rules

Improvements could include:

- Advanced/personal features for proofing
 - Specify translator's own personal register
 - Specify translator's own choice of spelling for words with multiple possible spellings
- Ability to disable grammar rules
- Ability to add translator's own simple rules
- A feature for sharing modifications or additions to lexicon/grammar rules with others
 - LTU
 - Other users
- Digital Rights Management of resources to allow proprietary data to be licensed for others to use
- Tokenizing features – especially with numbers which may trigger different grammar rules in different languages
- Contextual Spell Checking (as the blue squiggles in MS Office 2007 – see <http://blogs.msdn.com/naturallanguage/archive/2006/06/19/637359.aspx>)

- For example: + 'He bought a pear of shoes' (pair) + He was loosing too much time (losing) + You cannot associate this account on more then one mailbox (than)

3.5 Style Checker

The functionality of the Grammar Checkers may be expanded to allow the checking of adherence to Style Guidelines, e.g. for the house style of specific organisations, for specific language registers such as Cymraeg Clir, and for preparing for Machine Translatability.

User Created Guidelines

- To aid in maintaining an individual translator's own consistency

Thesaurus

- The ability to suggest synonyms from the thesaurus if the same word is used twice in close proximity within the same sentence.

4 Terminology Software

4.1 Terminology Managers

Terminology Managers allow translators to manage and exploit their terminology lists electronically. The most advanced allow terminology to be managed from within the translation environment, enable terminology lists to be created automatically prior to translation, and permit terminology databases to be shared between translators.

Glossaries of general words and phrases can also be kept using translation environment's terminology tools.

However as many words possess multiple meanings in general use, their unconsidered inclusion can cause problems if automated facilities such as automatic term propagation and example based machine translation are to be employed.

Automatic Term Propagation

Automatic term propagation allows for the automatic translation of words or phrases found in a document if a corresponding entry and translation exists in the translator's glossary or terminology database.

Example based machine translation such as that found in Déjà Vu combines information from the translator's terminology database with segments from the translation memory to generate better matches. Here, the inclusion of non-subject specific terminology could lead to a generated sentence referring to a mammalian mole, rather than the chemical unit. Therefore any specifications should include the ability to create and manage separate subject specific terminology databases rather than a single all-encompassing one. The possibility of generating general glossaries featuring only monosemic words (i.e. words with only one meaning) should be explored. Translation systems should be

enabled to give priority to using technical terms specific to the document's subject if there are alternate meanings for the word in general usage or in other subject domains.

Links to customer specific glossaries could also be provided, or even hosted, at the customers' request. Customers could specify that their terminology takes precedent over from other sources.

Exchanging Termbases – TBX

Most Translation Environment Tools use different file formats for managing terminology databases. To facilitate the exchange of terminology databases from one translation program to another the Localization Industry Standards Association (LISA), established the TBX standard.

TBX standing for TermBase eXchange, allowing their use by standard compliant translation environments. For software that manages termbases and glossaries in a nonstandard way, a CSV (comma separated value) file could be prepared.

Bilingual translators translate language pairs in both directions. They require the capability to flip direction of glossaries, e.g. from EN-GB/CY-GB to CY/GB/EN-GB whilst managing these files. Issues arising from this include mutated forms/inflections, where the form of a word in a document may be very different to the form of a word used as a headword in a traditional dictionary. Glossaries in translation tools therefore need to be able to cope with mutations and inflections. Some market leading TEnT software currently allow 'fuzzy matching' to find inflected forms, but this does not work well for heavily inflected languages, and language specific solutions need to be developed.

Glossaries need to be in formats that translators can use, whatever the software they employ.

4.2 Terminology Sharing, Exchange and Dissemination Tools

Whilst many Terminology Management Software packages allow the sharing of terminology, alternative methods are needed for sharing and exchanging terminology resources on an inter-agency or national scale.

To ensure that translators can utilize such content resources quickly and efficiently, the resources need to be:

- Collated, with one search being able to display the results from multiple resources
- Accessible and searchable using the translator's computer, preferably:
 - In the appropriate format for use with the translators software of choice
 - As an integrated part of the translator's workflow
 - Available at appropriate parts during the translator's workflow

Terminology Portal

A single central online portal for multilingual terminology resources for the translation industry in Wales could include:

- searchable online terminology databases
- downloadable terminology databases, exportable from Maes-T in various formats such as:
 - TBX
 - CSV
 - .ger (for integration into Cysgeir)
 - .multiterm and TermbaseStructureDefinition.xdt (for integration into Trados)
 - Déjà Vu Terminology Database format
 - .wfst (for integration into Wordfast)

Copyright issues need to be resolved before embarking on any such project. Copyright holders are sometimes happy to allow their material to be federated, provided they can retain control by allowing data to be disseminated and shared but not edited or adapted. The Eurotermbank model provides an example of how this may be done.

An all-Wales solution could be a terminology web server version of the LTU's Maes-T software, to which the TEnT software of individual translators would connect and receive any required terminology as needed from within their translation environments.

Other benefits of the web server method (as opposed to downloading terminology databases) is that the terminology used by the translator using the web server are always current and up to date.

At the GUM3C held in Bangor University 2008, Martin Wunderlich of Salesforce.com described the use of TinyTM as a web server for OmegaT. (see on-line proceedings at <http://www.gum3c.org/english/index.html>). A similar webserver should therefore be feasible for terminology, rather than translation memory.

Existing Terminology Portals in Various Languages

The following are examples of various monolingual and multilingual terminology portals that could be linked to or federated from a terminology resource portal for the translation industry in Wales:

The Welsh National Database of Terms <http://www.e-gymraeg.co.uk/bwrdd-yr-iaith/termau/Default.aspx>

Slovene Terminology Web Portal and the TBX-Compatible Simplified DTD/schema
http://www.lrec-conf.org/proceedings/lrec2008/pdf/553_paper.pdf

Eurotermbank
www.eurotermbank.com

EURALEX - European Association for Lexicography
<http://www.euralex.org/>

German Terminology Portal
<http://www.iim.fh-koeln.de/dtpEN/terminologieEN.html>

IATE – European Terminology Portal
www.iate.europa.eu

Sámi Terminology Portal
www.risten.no

Lexilogos – French TP
www.lexilogos.com

Termcat
www.termcat.net

TIS – Terminology Database of the Council of the European Parliament
<http://tis.consilium.eu.int/utfwebtis/frames/introEN.htm>

DocuTradso – Spanish TP
<http://www3.uva.es.DocuTradSo/#1>

TechWriter.de
www.tw-h.de

Exchange formats International standards exist for the exchange of resources such as Translation Memories between different translation software. These ensure that translators are not locked in to using a specific translation environment tool once purchased.

These include:

TMX, the LISA standard for Translation Memory Exchange, enabling the exchange of translation memories between different translation environment tools.

TBX, a corresponding LISA standard for the exchange of terminology databases between different translation environment tools.

OLIF, an open lexicon interoperability framework to allow data exchange between different language processing systems.

Whilst TMX is well supported by the developers of TEnT, TBX support is less common, and not as fully implemented.

A former Tech director at the European Language Resources Distribution Agency had this to say:

“TBX, as well as the OLIF project (open lexicon for interoperability between different language processing systems) has met a different type of resistance that is difficult to overcome. If the language tool vendors (Terminology tools, MT tools, etc) show how their tools categorize the entries, then they risk showing the secret of how they do things. This is much more risky than the TMX initiative from a business point of view for software development and distribution.”

Jeff Allen on ProZ.com

Plans are afoot to combine these specifications:

“If it turns out to be feasible to avoid inventing new wheels, then TBX will consist of OLIF and a LISA subset of MARTIF.”

OSCAR Advances, Alan Melby, Brigham Young University (2009)

http://www.lisa.org/globalizationinsider/1998/01/oscar_advance.html?printerFriendly=yes

4.3 Terminology Extraction Utilities

Terminology Extraction Utilities are designed to automatically extract terminology from a text or corpus. Several methods are in use, ranging from simple frequency based models coupled with stop lists of general words, to more advanced language dependent methods which make use of part of speech tagging to identify likely terms.

Monolingual Term Extraction Utilities

Monolingual Terminology Extraction Utilities enable the identification of a source language document's terminology prior to its translation. This allows the key terms contained within the document to be identified and standardized in advance, which ensures increased consistency and efficiency during the translation process. It also enables the identified terms to be used with the TEnT's inbuilt glossary or termbase functionality, so as to increase the translator's productivity and consistency during the translation process itself.

Bilingual Terminology Extraction Utilities

Rather than processing monolingual texts, Bilingual Terminology Extraction Utilities analyse parallel bilingual texts in an attempt to identify candidate terms in the source language along with the corresponding terms in the target language. This process automatically creates bilingual terminology lists which are consistent with the texts inputted during the extraction process, and is therefore extremely useful for translators needing to retain consistency with previously translated work. Bilingual Terminology Extraction can also highlight internal inconsistencies found within the inputted documents, allowing a standardized term to be agreed for future use. Should the texts used during the extraction process include texts from various organizations, inconsistencies in the terminology used by different organizations can be identified.

Issues and Improvements

For highly inflected languages such as Welsh, frequency or statistically based methods of term extraction are less effective than for less inflected languages such as English. Unless frequency or statistic based term extraction utilities also possess language dependent features such as part of speech taggers and lemmatizers especially tailored to the needs of specific languages, the results of those utilities will be poor compared to those for similar systems designed to be used with less inflected languages. Part of speech taggers and lemmatizers need to be part of a Terminology Extraction Utility for heavily inflected languages. Most commercial TEnT feature some sort of basic Terminology Extraction Utility. However these are usually language independent, and if they are closed source, may not be developed without permission. Open source possibilities include the Okapi Framework's Text Extraction Utility, Tim Craven's ExtPhr utility for Java, and Poterminology from Translate Toolkit.

4.4 Collaborative Terminology Standardization Tools

The LTU has developed a collaborative tool designed to enable the consensus based standardization of terms for standardized terminology dictionaries. Known as Maes-T, this web-based tool allows a dispersed team of subject specialist to collaborate with linguists on the development of terminology dictionaries over the internet. The software enables a process, including the assigning of varying degrees of edit rights.

Maes-T enables standardized terminology dictionaries, once completed, to be exported in the following formats:

- Printed dictionaries
- Electronic dictionaries distributed on CD-Rom
- Java based dictionary format for mobile phones
- Searchable online dictionary

It is based on the bespoke OLEW implementation of the OLIF lexical interchange international standard. Maes-T has been adapted to serve as an environment for the digitization and editing of the main English-Welsh dictionary, *Geiriadur yr Academi*. All of the Language Technologies Unit's content resources, where appropriate, are currently

being consolidated within Maes-T to enable the improved cross referencing, interchange and export of all data.

Maes-T enables public bodies to have control over their terminology as well as access to the expertise of linguists. It could potentially serve as a hub for consensus-based terminology standardization between public bodies in Wales. Maes-T may be adapted to enable the export of terminological resources into formats such as TBX and TMX which are usable by translators within their translation environments and workflows. Maes-T may also be developed as a web server to serve the required terminology on request to translator's TEnT systems, doing away with the need to manage and revise local terminology lists.

5 Thesauri

Thesauri are useful tools for translators, as having a number of synonyms at hand greatly aids the translator in choosing the most appropriate translation.

The LTU has a thesaurus application found within its Cysill software program. Currently Cysill's thesaurus content and Cysgeir's dictionary content exist as separate entities. Further development of the LTU's lexical resources should include the integration and cross-referencing of both sets of data within the Maes-T platform. The integration should be matched at the interface level, where thesaurus entries should be accessible as synonyms directly from the Cysgeir dictionary user interface. The platform should enable adaptations for other languages as needed.

Further enhancements to the thesaurus program could include a graphical interface for displaying the semantic relationships between words.

6 Machine translation (MT)

Machine translation is a thorny issue for the translation industry. Some see it as a threat to their livelihoods, thinking it will make human translators obsolete, while others criticize the low quality of translation obtained by automatic means. Customers of the translation industry however see it as a way of bringing down the cost and the time delay that translation involves.

The quality of results gained from commercial machine translation such as Google Translate when translating between major languages has improved greatly in recent years, due in part to the increased processing power and storage facilities now available to run statistical based translation. Commercial MT applications are available for some language pairs and some are far superior to those freely available on the web. However, the results are far from perfect and machine translated texts remain unsuitable for publication without human input.

A full discussion of MT is beyond the scope of this study, but some relevant points are covered here.

6.1 Gist translation

MT may be used to access texts in other languages where the reader needs a rough idea of the content, without expecting the translation to be correct in terms of either style or content. It is often used in this context by web users needing to look at web pages in languages they are not familiar with. It is also used by some international companies for internal communications with a multilingual workforce, where errors of style and inelegancies of translation may be tolerated.

English to Welsh MT has recently had very bad publicity due to a spate of awful translations caused by a poor machine translation engine marketed by TransXP being made freely available online. This was used to translate public signage in Wales, with the highly-visible errors causing great embarrassment.

A Welsh to English gist translator would however be useful to Wales where non-Welsh speakers are not able to access Welsh language content. It may enable private businesses for example to understand the gist of a customer email quickly, and judge whether a human translation needs to be commissioned. Gist translation would not be appropriate in the opposite direction however, where the customer would likely to be very upset to receive a reply in poorly formulated Welsh. A gist translator may be developed for Welsh however from the mouse-over translator which forms part of the Cysill program.

6.2 MT for less resourced languages

Minority languages are often less suited to statistical methods of machine translation, as these methods rely on access to a large corpus of internet text. The amount of text to be found on the internet in minority languages may be significantly lower than is available for major languages. Some minority languages such as Welsh are more morphologically complex and less standardized than some of the major languages such as English, necessitating a larger corpus to be used in order to deliver statistic based machine translation of a comparable level.

MT less resourced languages is not widely available, the high development costs being another impeding factor. The Welsh Language Board commissioned a report on MT for Welsh which reported in 2004 (see <http://www.byig-wlb.org.uk/english/publications/pages/publicationitem.aspx?puburl=/english/publications/publications/2302.doc>). However, despite a tender being used subsequently, no further work was commissioned, due to the high cost of developing such systems.

Open source systems provide a way forward in that may be freely adapted to specific language pairs offer low cost solutions, provided that adequate electronic data for those language pairs are available. These include

1. Moses, a statistical based approach developed in the international EuroMatrix project (see <http://www.statmt.org/moses/>). A pilot project at the LTU trialled the use of Moses for Welsh/English translation, using a limited domain bilingual corpus derived from the Welsh Assembly Government Record. This pilot achieved a high degree of accuracy, but quality increasingly deteriorated with distance from the domain used. The process was also very slow, with twelve sentences taking two hours to be run through the system. More information can be found within the SALT Cymru Feasibility study at www.saltcymru.org.
2. Apertium, a language-independent machine translation engine developed by University of Alicante (see <http://www.apertium.org/>). Originally intended for closely related pairs, it is now extended to cater for other languages. It also includes tools to manage the linguistic data necessary to build a machine translation system for a given language pair. It needs linguistic data from relevant language pairs in order to function. A Welsh to English facility is currently being investigated under this system.
3. Avenue, a hybrid (or „omnivorous“) MT system that combines different approaches, developed at Carnegie Mellon University, Pittsburgh (see <http://www-2.cs.cmu.edu/~avenue/>). It has been especially designed to work with less-resourced languages, and is adaptable to the resources available. So far it has worked with indigenous languages in South America, but the CMU team are interested in working with other languages such as Welsh.

6.3 Combining MT with a TEnT

Increasingly, simple machine translation functionality is appearing within TEnT software, working in conjunction with the translation memory systems and terminology databases which are a part of such systems. However, these features are language dependent, and are not usually to be found for languages other than the most resourced.

They may be used however to ‘fill in the gaps’ where the TM supplies already translated segments, and the MT attempts to translate where there are no matches or where the match is low. A further refinement of this is to use the parallel text corpus already in the TM as a resource to create example-based MT (the TM being used to supply the examples). An extremely large TM is needed for this approach to be fruitful at the moment, but this may be a way forward.

7 Concordance Software

Concordance software enables translators to quickly search existing digital texts to find examples of the use of a word or phrase in its various possible contexts. Following a search, the results are displayed in such a way that they can be easily compared. Often, this includes displaying multiple examples onscreen at once, with the subject of the search centrally aligned and highlighted. In addition to monolingual texts, bilingual or multilingual texts or corpora can also be searched, including Translation Memories. These are especially useful for the translator as they can allow access to translation strategies that have been previously employed.

7.1 Examples of concordance software:

- Wordsmith
- Antconc
- Sketch Engine

Many high-end commercial translation environment software suites include concordance functions. However, these may lack some of the more specialized features found in dedicated concordance software.

Although designed primarily as a TM editor, the free, open source Okapi Olifant can be used as a concordancing software for TMX translation memory files.

Due to its open source nature, Okapi Olifant could be potentially further enhanced to address the particular needs of the Translation Industry in Wales.

7.2 Issues with current concordance software

Inflexion

Inflexion is the phenomenon where words, under certain circumstances dictated by the rules of the language, change their form. In English, for example, the word *mouse* changes in the plural to *mice*. This creates a problem when searching for examples referring to mice in general, as searching for *mouse* or *mice* alone would not produce a complete result.

English features a comparatively low degree of inflexion and this problem is more acute in many other languages. For concordance software to be useful in heavily inflected languages such as Welsh, the main language of translation in Wales, language dependent morphological functions which can identify the basic form of a word (known as a lemma) are required.

8 File Format Filters

In recent years the format of files that translators are required to translate have become increasingly complex.

These include:

- Graphic Design Formats
 - .eps
 - .psd
 - .png

- Web Formats
 - .html
 - .php
 - .cfm
 - .aspx

- Software File Formats
 - .ini
 - .properties
 - .po
 - .fla
 - .xml

Freelance translators and small translation agencies currently find it difficult to work with such formats due to:

- the cost of such software
- the unfamiliarity of such software
- the lack of translation aids in such software

In response to these issues, file filters have been developed that allow translators to translate the language content of various files from within their translating environment, whatever the file formats, without having access to the specialist software used to create the documents. This allows the translator to work in a familiar environment, with access to translation aids such as translation memories and terminology databases. However, specific filters must be created for each individual file formats. Some of the more common file formats such as PDF have filters available for download, or are included with TEnT software.

Many of the required file filters are not available for all translation environment systems. This is especially true for software language files, where many various file formats are employed. For example, Wordfast 5.5, the most commonly used software amongst freelancers in Wales, is only suitable for translating documents in formats which are compatible with Microsoft Word. Some high-end commercial TEnT software suites contain functionality that enables the translator to create filters for file formats where no file filters currently exists. Owing to the complex nature

of file formats and the need for technical expertise, this is often beyond the capability of most translators, even when software wizards exist to aid their creation.

In this respect large, well-supported translation agencies have the advantage over small enterprises. They are able to employ dedicated staff to deal with complex or specialist formats, and can extract text for translation, as part of the general pre-translation process. However, small enterprises in Wales could share resources in this area, and be supported by a dedicated translation technology centre.

9 Translation Management Systems

Translation Management Systems are software systems specifically designed to facilitate the management of multiple translators, clients and translation projects. These systems manage the availability, specialization, and language skills of their translators, and the details and requirements of their clients. They often also aid the management and exchange of translation memories between the translator and the client/agency pre and post-translation.

These systems may be unnecessarily complicated for individual translators working as freelancers, but may be suitable for freelancers looking to expand from translation alone to running a translation agency managing the work of other translators. Currently, the main market for such software is private translation companies that employ numerous translators. The commercial system LTC Worx has recently attracted much interest in Wales.

An open source Translation Management System was launched in January 2009 by Welocalize. Known as Globalsight, it allows the management of multiple translators and translation tasks, facilitates interaction with clients, and aims to improve the commissioning and delivery systems of translation service providers. Owing to its open source nature, Globalsight may offer possibilities for creating a stripped down version to aid freelancers in managing their Translation Memories. Freelancers could then benefit from simple but effective ways of managing interaction with customers and suppliers, including:

- generating quotes and invoices
- documenting jobs
- assigning translators
- prioritizing projects
- measuring work rate
- managing holidays and general availability

10 Source Language Writing Aids

Translation costs can be cut significantly by introducing improvements to the way in which the original documents are written. Writing aids in the form of software programs are available for the authors of original documents. They are therefore of interest to customers wishing to cut down on their translation costs. In addition to helping to improve the clarity

and style of the author's writing, these writing aids often help the author produce more concise texts.

One such piece of software, Stylewriter, demonstrates the ability to reduce a text from 149 words to 86 in the process of increasing its readability. As translation is usually priced according to the number of words translated, the use of such software on a text intended for translation would result in an immediate 42% reduction in the cost of translation.

The cost of translation of a 20,000 word report into Welsh at the current standard rate of £80 per thousand words would be £1600. A 42% reduction in the word count of the report to 11600 words would result in a reduction in translation cost from £1600 to £928 and a corresponding cost saving of £672.

A case study involving Hewlett-Packard's use of an authoring software called Author-it also demonstrated considerable savings in translation costs through changes made to the source language:

"In a pilot study with a large U.S. company, a word-count reduction of 32.5% for new translations was found, resulting in direct savings of more than \$75,000 in external costs for just one language."

(see <http://www.author-it.com/index.php?page=hpiss>)

In addition to these explicit cost savings, a clearer, more readable document would also be easier, quicker and more satisfying for the translator to translate. Reducing the length of sentences also increases the likelihood of matches being found in the translation memory, further reducing the time and cost of translation.

Author-it also features an authoring memory, similar to a translation memory, but intended for the source language rather than the target language. Sentences that have previously been typed by the author reappear when similar sentences begin to be typed at a later point, allowing their re-use, and resulting in an increased internal consistency. This increased consistency produces more repetitions at the translation stage, cutting costs further.

11 Conclusions

All the software resources outlined above have the potential to help the translation industry and their customers in Wales. Demonstrating the range of software solutions already in existence is the first step towards improving translation technology take-up amongst translators. Some items of software are also relevant for use by customers of the translation industry, and all lead to saving in costs and time, with some also leading to improved quality control.

Collating dispersed software resources, or inputting relevant language content into them and making them available for use, is also achievable within the short to medium term. Longer term solutions include the creation of new and improved language tools specifically to answer the needs of Wales and Welsh industry. Increased productivity and saving generated

within the Translation Industry, coupled with enhanced economic benefits for the industry's customers, will ensure the sustainability of the knowledge base, as the true value of translation technology becomes known and appreciated.

Appendix 2: Scoping Discussions Report: the ‘toolkit’ needs of the Translation Industry and their customers

Scoping Discussion 1: Freelance translators

Scoping Discussion 2: Public sector translation services

Scoping Discussion 3: Private Translation Bureaux

Scoping Discussion 4: Disabled translators

Scoping Discussion 5: Customers of the translation industry

Scoping Discussion 1: Freelance translators

Profile

The largest group of translators in Wales for all languages combinations work as sole traders or freelancers. 39% of participants in CCC's 2008 conference focus groups were freelancers. As such they form the backbone of the industry, and provide a flexible and responsive service to their customers. Typically, freelancers work full-time, some of them working very long hours to keep up with their workload and meet tight deadlines. Others however, mainly in the older age group, work part-time as translators, having retired from other professions. These tend to be highly experienced language professionals, but may have lower computing skills because they were introduced to computers fairly late in their working careers. A significant number of translators who work in the public sector reported that they undertook additional work as freelancers, working in the evening and at weekends. This was reported by some as something they needed to do to top up low wages, while others reported that they had begun to take on extra translation work in their own time due to urgent calls for their services because of lack of capacity in the industry.

The reported that demand for their services continually outstripped their ability to supply. "I only take on work from regular customers" was a common comment, "and I can't even keep up with their demands". Many reported that there was a huge unmet demand for translators: "I would need to treble my capacity in order to answer all the demand for my work".

The translation supply chain

As part of the translation supply chain freelancers provide direct services to both industry and public bodies and also act as subcontractors taking in translation work from larger translation agencies and from the translation services of various public and voluntary bodies.

Many of the private translation companies and agencies are based in Wales, but increasingly freelance translators are reporting requests from translation agencies based outside Wales for their services. Freelancers perceive this as a threat to their ability to deal directly with end customers, citing the very low rates offered by some of the agencies, and the way the middleman appropriates a large proportion of the profit.

Issues of quality control were also raised, as many of the outside agencies did not seem to operate any additional editing or proofreading measures to translations provided to them, and often possessed no in-house expertise in the languages being served and therefore of the linguistic context in which the translation would eventually appear. The agencies' ability to explain the differences between linguistic contexts sympathetically is also limited as a result. This was especially noticeable in the case of Welsh where most of the linguistic expertise resides within Wales and is rarely found outside her borders. Because demand outstripped supply there was a perception that outside agencies used unqualified translators. "Some people think that because they can speak two languages that they can then work as professional translators. Qualified translators wouldn't work for such low wages, but some agencies don't ask too many questions, and lots of people see it as a way into work".

Use of technology and resources

Freelancers also reported increased demand from translation agencies for them to be in possession of specific translation tools, especially one of the leading commercial translation memory systems. Typically, if this was done, the translation given to them would already have been pre-edited, and processed through the agency's TM system before being passed on, so that only new translation segments would need to be dealt with. This can in itself raise problems of quality control and stylistic consistency. Sometimes, freelancers are also given specialist terminology resources either from the agency or the end-customer to aid the translation task. End-customers are usually more generous in sharing their terminology resources, presumably because they wish to see consistency in the use of their in-house terminology across all languages. Agencies are sometimes reluctant to share their own terminology dictionaries, seeing them as valuable commercial assets and the freelancers they employ as potential competitors in the translation market.

Translation offices in public bodies who employ freelance translators either do so routinely, keeping only a small in-house team of translators for urgent or core translation tasks, and farming out large amounts of translations to freelancers. Alternatively, they may do most of their translation work in-house, but still employ freelancers occasionally, in especially busy periods or when they face staff shortages. Freelancers report that public bodies are much less likely than private agencies to ask them to use TM systems and to pre-edit translation work for them, although there are some exceptions, and there are signs that the public sector is also moving in this direction as they see that significant savings to their translation bills could be made in this way.

Freelancers also report that public and voluntary bodies are far more generous in sharing their terminology resources with them in the interest of consistency and quality control. A notable example of this is the TermCymru database provided on-line by the Welsh Assembly Government. This contains the terms used by WAG's own translation service, extracted from their own in-house translation memory.

Apart from other on-line resources developed at the LTU at Bangor University, the other resource much valued by freelancers is the welsh-termau-cymareg e-mail discussion list, run by the LTU at Bangor University (see <https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=welsh-termau-cymraeg>).

This was reported as the first line of call by freelancers who wanted to know if another translator had come across a difficult to translate term or phrase, and they much appreciated the instant replies received from other translators with relevant information. The list has over 230 members, and on a typical working day can generate between 10 and 20 answers. All past answers are archived and available on-line so that translators can revisit previous discussions.

Internet and IT support

Freelance translators typically work from home, and internet connectivity is still an issue for some. This is now the preferred method of receiving and sending documents for translation, as well as for billing and general communication with customers. It is also used extensively for

accessing on-line resources such as terminology dictionaries and relevant documents. While very few freelance translators reported that they did not use the internet at all in their work, 14% of freelancers questioned did not have access to broadband, and therefore had to put up with a slow and sometimes undependable internet connection.

In general, freelancers use standard computer workstations, and seem reluctant to invest in computer hardware and software. The idea that improving their working environment could result in improved productivity was greeted by surprise by many, who expressed themselves as satisfied with what they had. Many freelancers rely heavily on a local computer retailer or IT specialist to advise them on computer matters, and the support given varies widely, with some having a close relationship with a company that provide advice, upgrade and repair services, to others who have no such support and are at a loss of where to turn to for such help.

Simultaneous translation and interpretation

Although most freelancers are engaged in written translation work, oral translation, usually in the form of simultaneous translation at meetings, is also undertaken by freelancers. This may be undertaken in addition to written translation work, but some freelancers concentrate almost exclusively on oral translation or interpretation.

Interpreters who do a great deal of simultaneous translation tend to invest in their own portable translation equipment, which they take with them to events where they need to translate. Alternatively, they borrow or hire equipment from a private or public agency. The body asking for their interpretation services may sometimes provide equipment for them, especially if they make extensive use of simultaneous translation services. In some circumstances there is built-in translation equipment in the meeting or conference rooms that the freelance interpreter is expected to use.

Further Needs

Impartial advice was cited as one of the main unmet needs of freelance translators. “I don’t know where to start” was a comment often heard, as well as the suspicion that retailers were only interested in selling their own products. This was especially true of dedicated translation technology software. A recurrent theme was “it seems a lot of money to spend when I’m not sure that it will suit my needs”. Freelancers were reluctant to approach individual TM vendors as they felt that the pressure would then be on for them to buy that particular system. Many felt that they would also be pressured to keep on spending on their systems, either on helpline advice, upgrades or new features. The lack of a centre that they could visit and have different software solutions explained to them, with the ability to try out various products without pressure to buy was mentioned by many as one of the main provisions that would help them as freelancers. In the words of one freelance translator at a focus group in Caernarfon:

“I would like to come face to face with someone (at Bangor) who could introduce a package of suitable programmes for me (as one who has not yet used any programme such as Cysill on my computer)”.

The other urgent call was for “joined-up” resources, where tools and resources could be readily accessed from one portal:

“I spend a lot of time just finding out what’s out there. It would be much better if it were all grouped together.”

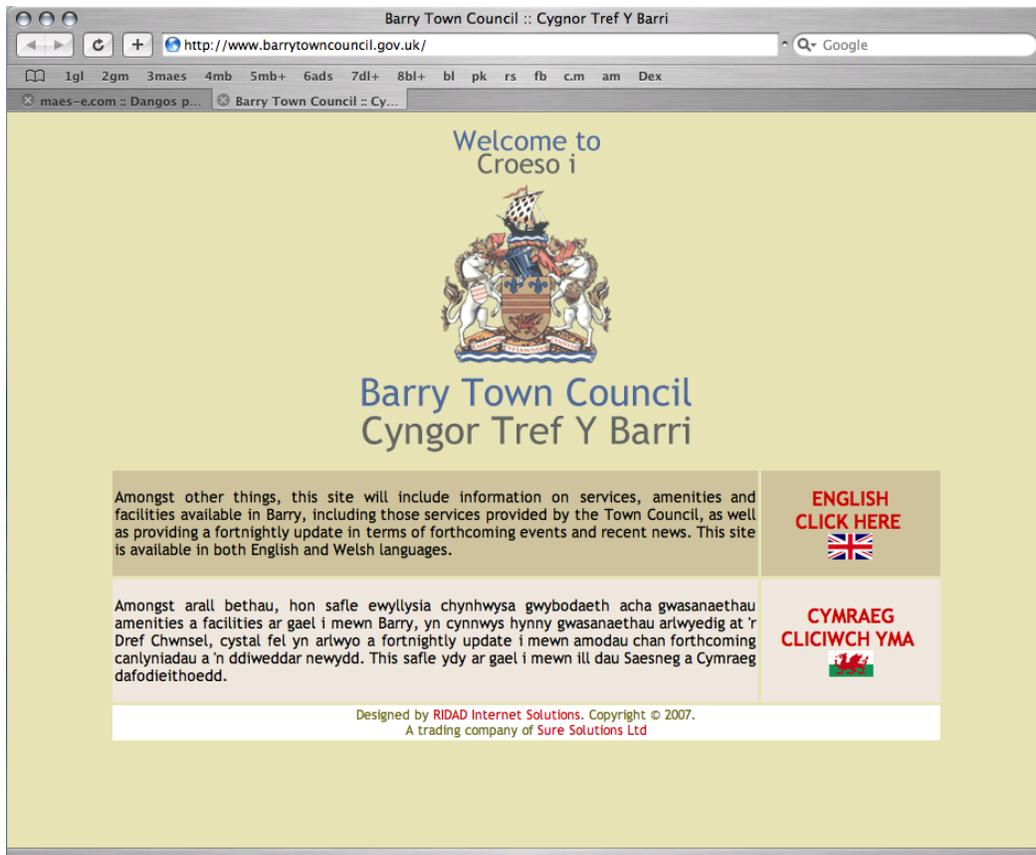
Although the isolation of freelancers working on their own was rarely mentioned, the project team felt that the seminars and focus groups provided by the project and the activities of CCC provided valuable opportunities for freelancers to meet and exchange information relevant to their situation as freelance translators. Translators often trusted the advice and experience of their peers (sometimes obtained via the discussion group *welsh-termau-cymraeg*) far more than that of commercial vendors.

Scoping Discussion 2: Public sector translation services

Profile

Provision of services in both Welsh and English is a statutory requirement established for the public sector in Wales by the Welsh Language Act of 1993. This requirement led to a huge increase in demand for translation services in Wales, leading to the expansion of the private translation sector as well as the establishment or growth of translation offices within public bodies in Wales. Next to freelance translators, the largest numbers of translators in Wales are employed by the public sector. 38% of participants at the focus groups in CCC's 2008 conference were employed by the public sector.

Public bodies with a statutory obligation to provide services in both languages include all local authorities in Wales, the Welsh Assembly Government, and public services such as health, police, law courts etc. The provision extends to other public agencies such as environmental and heritage agencies, and the education system may also be considered under this heading. Some smaller public bodies with no in-house expertise have tried to minimize their translation costs by using free machine translation on the web. This has caused some embarrassing mistranslations, especially where they appear on highly visible public signage or website. One such famous case was that of Barry Town Council's homepage on the web, which reads as pure gibberish in Welsh (see below).



“This is the sort of thing that gives translators a bad name” said one translator in a local government office, although fortunately in this case the translation was so obviously the work of very poor quality machine

translation that no human translator could possibly be blamed for it.

Public bodies do usually have their own in-house translation office to serve their internal translation needs. Alternatively they may use external translators, either translation agencies or freelance translators. Often, public bodies use a mixture of both, keeping in-house staff for core translation tasks, supplementing it with the use of external translators as needed. Public bodies may share translation services, as in the case of the City and County of Swansea and the Neath Port Talbot County Borough Council. Bodies in closely related disciplines, such as the Countryside Council for Wales and the Environment Agency, may work closely together, sharing terminology and other resources. Some public agencies, such as North Wales Police, the Court Justice System, and the Fire and Ambulance services may hold joint meetings to discuss language issues that are relevant to their translation provision and the provision of bilingual services in general.

Language Officers

The post of Language Officer within public bodies in Wales is a comparatively new one. It seems however to be a role that is expanding quickly. Five percent of the participants at CCC's 2008 conference were employed as Language Officers by the public sector. The creation of such a role is often seen by public bodies as a way of providing a 'one-stop shop' for dealing with all the organisation's linguistic issues. The officer may be in charge of a translation unit, or may be the gate-keeper who deals with translation requests and administers the use of external translation companies. The Officer may be called upon to undertake some translation work themselves, or to edit, proofread and quality assure work done by external translators and agencies.

Because of their key role in administering the translation process, Language Officers can be very influential in the choice of computer hardware and software used by translators, and the provision of TMs, terminology and other relevant resources within their organisation. They also have their own networks of contacts e.g. amongst the local health boards and trusts of NHS Wales (see <http://www.wales.nhs.uk/sites3/page.cfm?orgid=415&pid=33967>) which are used to cooperate on translation and related matters.

Multilingual issues in the public sector

Although translation services in the public sector in Wales were initially set up to provide for translations between Welsh and English, increased globalization and the mobility of populations has led to public bodies in Wales also needing translation services in other languages. These services are often offered under equal rights and race equality schemes, and could mainstream the provision of Welsh and English in a multilingual environment. Some local authorities reported the extension of use of simultaneous translation equipment traditionally employed for Welsh and English to other languages such as Polish. However, some stakeholders expressed concern that providing for Welsh alongside other languages or under the umbrella of general equality rights issues could dilute Welsh language provision.

Computers and Translation aids

Public bodies usually have designated IT support services to help staff with their computer hardware and software needs. Computers are usually networked in larger organisations, with programs and documents shared on the organisation's servers rather than on the translators' local workstation. Despite having a much greater level of IT support from that available to the freelance translators, many translators in the public sector expressed frustration at the lack of understanding their organisations had of translators technical needs, and the low priority afforded to translators by their IT departments. "We're always the last to get new computers in this place" complained one translator, adding that all other offices had been moved to the new XP Windows operating system while the translators were stuck with Windows 98. "This meant that we couldn't run many of the newer software programs, and had problems with converting files received in the newer formats". Others complained about not being able to access the welsh-termau-cymraeg e-mail list since they were blocked from receiving external emails for security reasons. Even accessing external web-based resources such as terminology resources were problematic for some: "they think we're looking to book ourselves a foreign holiday" one translator said if they asked to be able to access external web-sites.

However, IT support staff also complained of lack of communication with their translation service. "I wish they'd ask us before buying a piece of software" was one comment. "They don't check before hand that it will run on our system or that they're buying the right version". Since translation technology tools can be specialist software and involve wider issues of support on organisation network and company intranets, consultation and agreement between translators and IT personnel is essential before any new software is introduced. IT staff have reported problems with licensing issues where a network's set-up clashes with the licence servers of commercial software.

Many translators thought that opportunities for improving productivity and quality through making better use of translation technology and their own IT services were being lost. "We still haven't set up a central shared translation memory on our network", said one head of translation services. "If we did that, we could share our resources better, and if one translator was ill, or left, we'd still be able to access the valuable translations she'd already done, especially if someone had to take over a task that she'd previously done".

Others reported that they still did not have basic translation technology resources such as any form of translation memory systems, but were however expanding the size of their translation units. Even sharing common resources such as an in-house terminology bank were not universal. "Each translator keeps her personal little dictionary of difficult words, but we've never looked at putting them all together and sharing them between us".

However, some translators could see huge benefits from networking and sharing translation memories, even across institutions. Local government and higher education institutions were seen as two areas where there was much overlap in translation work and where content could be gainfully shared.

Issues of confidentiality

Some translators however said that their institutions could never provide content for TM systems, since some of it was confidential in nature. “We deal with cases of child abuse and everything” said one translator. “We could never allow that to go on any system.” Care has to be taken therefore that no confidential material ever gets into a TM system. Some material is so sensitive that it should never be included in such a scheme. However, technologies exist that could be adapted to anonymize content by taking out personal details, named entities etc. and this should be used as a matter of course with any such technology.

Administrating the translation process

The project identified a wide variety of administrative practices for tracking translation work from the initial request for translation through to the delivery of the final translation. These included steps such as the taking of contact details, dates for delivery, special requests, in addition to assigning roles such as translators and reviser, as well as ensuring quality control checks were carried out and that the completed translation was returned to the relevant contact person. Some offices relied on traditional secretarial help to accomplish these tasks. In some cases, details such as the number of words in each translation task, date of receipt and date of return, details of translator used etc. were not kept. In these offices, it would be therefore difficult to achieve an accurate audit of the translation work undertaken by the unit. Translators and others commented that it could sometimes be difficult to trace a translation within the system, and find out who had a particular piece of work and when it was likely to be ready, without recourse to phoning around individual translators. Some units admitted that work had on occasion been translated twice due to administrative errors, although where translation memories were employed such duplications were immediately apparent as the entire documents were contained in the translation memory and would therefore in effect translate themselves.

Offices who gave out work to external translators had additional requirements in terms calculating translation costs, billing and payments. While the larger translation offices were seen to be making progress towards the use of software to assist them in all these translation administration tasks, many other offices seem unaware of the existence of such software or were reluctant to spend on them. Others had created their own internal systems, based on shared databases on their intranets, in line with their administration needs. “It’s so much better” said one translator. “We can now see at a glance how much work we’ve done that month, and how many translation jobs are in the queue waiting to be dealt with”.

Further needs

Any engagement with translators in the public sector needs to include their IT support staff and the translators’ line managers. Translation services often come under the heading of communications and public relations, and their senior managers may be unaware of translation technology and the resources and aids that can improve productivity and keep down costs. Introducing computer aided translation tools where they do not yet exist would be a huge step forward. Sharing resources such as personal dictionaries and federating translation memories both within and between institutions would be another significant advance.

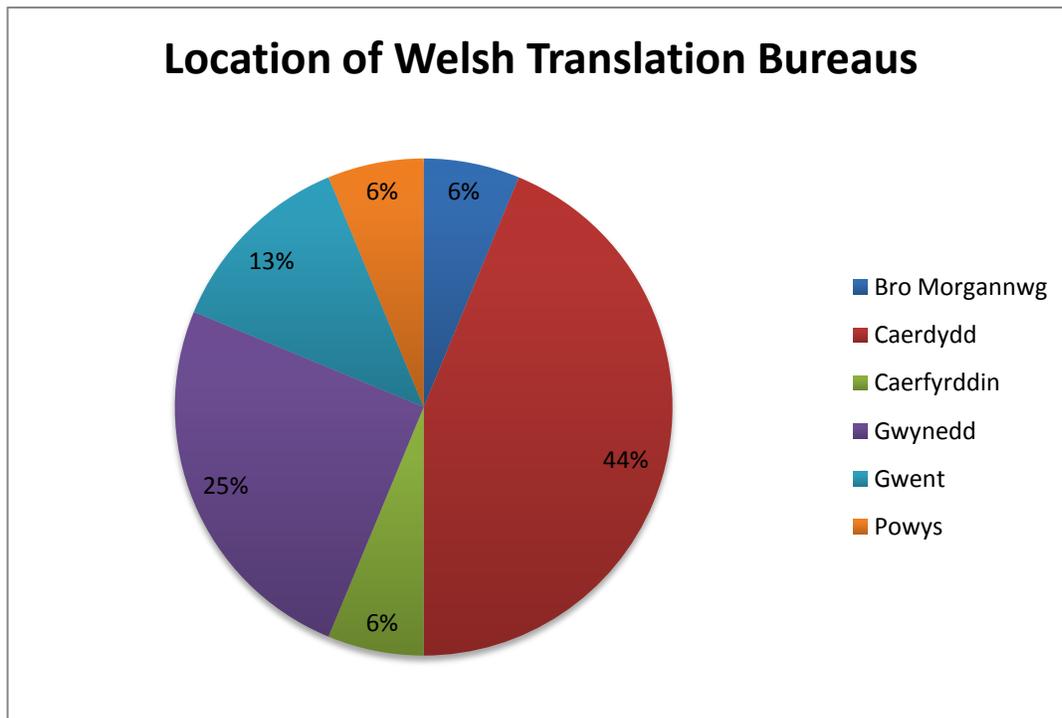
Public translation offices are not divorced from the rest of the translation industry in Wales. Personnel move freely between them, and they meet each other in common seminar and training events. The use of freelancers and private translation agencies by the public sector also means that they are closely associated. Public bodies are important customers of translators working in the private sector. Translators in general commented favourably on the generosity of the public sector in sharing terminology resources and expertise with them. Any solution such as a demonstrator centre, terminology portal and shared translation memories should therefore include the public sector in its remit.

Scoping Discussion 3: Private Translation Bureaux

Profile

Seventeen Private Translation Bureaux employing two or more translators on-site were identified in Wales. Eleven of the seventeen (65%) specialize almost exclusively in translation to and from Welsh.

The majority of the private translation bureaux are based in Cardiff and Gwynedd respectively. The numbers based in Cardiff reflect the activity surrounding the number of organizations based in the capital and requiring translation services, whilst the number of translation bureaux based in Gwynedd can be explained by the demands of a predominantly Welsh speaking area for translations into both Welsh and English, the local capability for translation, the policy of Gwynedd County Council to operate through the medium of Welsh.



Three bureaux are owned by companies based outside of Wales, including one bureau almost exclusively involved in translation work involving Welsh. This reflects a growing outside interest in the Welsh translation market that could potentially lead to the loss of Welsh investment to the Welsh economy.

Of the bureaux who work with languages other than Welsh, most said they would accept translation work for almost any language. This willingness to take on translation work for languages in which there is no in-house expertise stems from the widespread use of freelancers and of subcontracting work to them. In fact some translation service providers now work exclusively as middlemen, passing on translation commissions to freelancers as necessary. This

has allowed these companies to focus on managing and administrating the translation process, and to do this effectively, they have invested in the latest translation technology, including workflow managers and multilingual document management systems, as well as translation memories and tools that translators themselves use.

One translator service provider said: “One advantage we offer is that we can give the customer a service in a number of different languages. Therefore, if someone wants their web-site in half a dozen different languages we can do it all for them, so that they don’t have to find different translators themselves”. It was also noted that with the ease of modern global communications, such companies can employ translators based in any country, and have native standard translations done quickly, and often cheaply, if salaries and costs in the translators’ home countries are cheaper than in the UK.

Many translation service providers expressed the desire to expand. “There is so much unmet need out there. We aim to increase our capacity 300% in the next five years, and could probably exceed that if current trend continue”. Growth was anticipated by companies expanding the range of languages they offered, and outsourcing more work to freelance translators, using translation technology tools to pre-prepare translations, and to streamline administrative tasks.

Within the UK however, the cost of living in Wales is comparatively low, and given the lack of Welsh language expertise outside Wales, private companies based in Wales lead the Welsh-language translation market in competitiveness. Some worries were expressed however that Wales-based translation agencies were losing out to outside companies even for Welsh language translations because of better use of technology: “They’re able to undercut us, but they don’t really care for quality and go for the cheapest freelancers who are usually ones who are least experienced or don’t translate that well”.

The loss of business to outside companies could also be interpreted as reflecting an inability within Wales to put into place the changes necessary to adopt advanced translator technologies, despite their promise of increased translator efficiency and improved administration methods. One small company based in rural north-west Wales said:

“We still haven’t networked our own computers yet, and we don’t really have an IT person we can turn to when we need help. We could expand, and we know the standard of our translation work is excellent, but we need outside help on the technology front”.

However, some companies who now deal exclusively with Welsh and English translation work saw the possibility of expanding into other languages if they employed the right translation tools. “If we make this investment for our current needs it would be a pity to waste the opportunity and not take in work for other languages as well. Certainly using a proper workflow manager would allow us to use more freelancers and keep a track of everything. They wouldn’t even need to be based in Wales, and that would help us use good native speakers of other languages”.

Customer base and their requirements

Most of the translation bureaux identified specialized in Welsh-language translation, and most identified the public sector as their main clients. However, some noted an increase in work from private companies operating in Wales.

It was reported that translation bureaux tended to deal with larger contracts than those which could be dealt with by freelancers. These included requirements to be available at all times, and undertake a specific quantity of work within a specific timeframe. Such contracts are awarded by the UK Government or by public sector organizations based in England who are required to produce Welsh language documentation, either alone or as part of wider multilingual provision. The need for translations into multiple languages was another reason given for the need to use translation bureaux rather than deal directly with freelancers. A third reason was that the format of the translation work received was becoming increasingly sophisticated. More and more documents are produced for the web, often using content management systems, and software translation is a service increasingly requested by clients. Some companies did not possess the technical capability to deal with work which was received from customers in complicated software formats. Translation agencies from outside Wales with appropriate software were said to win Welsh translation contracts, convert the file format into a simpler translation-friendly format for subcontracted translators in Wales, and re-convert it upon receiving back the final translation.

Use of Technology

Private Translation Companies on the whole were found to make significantly greater use of translation environment (TeNT) technologies than their freelance and public sector counterparts.

It was noted that the benefits in terms of efficiency and consistency from the adoption of TeNTs were greater for companies employing numerous translators than for single freelance translators. For example, the ability to share translation memories on a company's networked drives multiplies the benefits gained from the use of translation memory systems by the number of translators employed, as each company translator benefits from the translations of their fellow translators. The higher volume of translation processed in general by private translation companies also increase the benefits of using such technology, as higher translation volumes result in larger TMs and consequently more matches when translating. The administrative needs of keeping track of a number of translators, including often outside subcontractors, and of accounting and invoicing customers, have also been driving the move towards use of technology in this sector.

It is this sector that is most sensitive to the economic savings possible through appropriate use of technology. Despite the difficulties of finding impartial IT advice and support, and the reluctance of some translators to embrace new technology, company owners have had to focus on increasing efficiency and competitiveness. Companies that can afford to invest in staff training also reported that they occasionally sent staff to seminars and training sessions where new technology was being demonstrated, in order to bring new knowledge back into the company.

Larger companies are also able to employ support staff to help translators and take administrative and IT issues from them. “We found it cheaper to employ office staff to deal with customers rather than waste translators’ time dealing with requests” was one comment. “Buying in outside IT help has been really worth it” was another comment, although some companies expressed dissatisfaction with the lack of understanding of specific translation technologies from the IT support companies they had access to. “My local IT supplier had never heard of translation memory systems” was one comment.

Translation Technicians

International translation agencies often employ dedicated translation technicians to facilitate the work of their translators. This includes converting complicated file formats into files which translators can easily translate using their TEnT software. This allows the agencies to outsource translation to translators whatever their technical ability. One company said that they did not have enough work to employ a dedicated translation technician themselves “but maybe we could use one if they offered technical support to freelancers as well”.

However, trained translation technicians are currently rare. An increased emphasis on training students in these technologies within Wales would help to satisfy the requirements of the domestic translation industry.

Lack of Awareness of Standards

Only two private translation companies in Wales were found to be advertising their adherence to *BSI EN 15 0 38 Translation services – services requirements*, and little knowledge of the standard was found amongst translators in Wales.

Very little awareness of technology standards relevant to translation, such as TMX, TBX, and ISO standards was found amongst private translation companies, reflecting a lack of technical knowledge about translation technology.

Unless customers asked specifically to adhere to standards, there was very little pressure on companies to advertise adherence to them. In general, customers did not seem to be aware of translation standards, and were focused instead on price and delivery time of translation services.

Further needs

Amongst those companies who already use some form of translation technology, the general finding was that their use of such technology could be improved in nearly all instances. A common comment was that although training sessions might have been provided when a new piece of software was first introduced to a company, staff turnover often resulted in the relevant personnel having left, and new members of staff were left to find their way on their own.

Upkeep and upgrading of resources were also named as issues. In-house terminology banks could grow unwieldy and outdated over time, but not enough attention was

being paid (with some notable exceptions) to pruning and reorganising them. In-house translation memories were often left to grow to the point where they became slow and cumbersome and in need of an overhaul. “I know we should split it into separate TMs for different clients and subjects” said one company, “but we haven’t got round to it yet”.

Tools for filtering TMs and organizing translation units by relevance could enable companies to run smaller multiple TMs without sacrificing the number of matches found. In addition, alignment software would help them reuse legacy translations and gaining a head start in populating new TMs.

Many companies however had a clear idea of what they wanted, and expressed their frustration at the lack of available tools to work with Welsh language texts. “A style checker for Welsh would be really useful, because some of translators have really different styles”.

Two large companies were found to be having issues relating to the use of Welsh spelling and grammar checkers (despite using different brands of TEnT software), and were engaging in time consuming workarounds to proofread their documents and ensure that only segments that had been proofread were added to their TMs. One of these workarounds included the manual realignment of TMs, a laborious and time-consuming task. Help in the integration of spelling and grammar tools and terminology dictionaries into their TEnT products was the most consistent appeal heard. This is was primarily for use with Welsh, but it was also pointed out that if an adaptable platform could be developed, this could then be reused for other languages, especially the long tail of less-resourced languages used by many emerging economies.

Scoping Discussion 4: Disabled translators

Profile

It is difficult to estimate the number of disabled translators in Wales as no comprehensive survey has been undertaken. The numbers would anyway change according to the definition of disability used, and so these discussions took the broader view, encompassing all translators and their employers who wished to discuss disability issues.

Translation can be an attractive profession for people with disabilities that impair their movement and/or sight. Translation tasks are capable of being performed remotely, with flexible working hours, and technology aids help disabled translators process and prepare text documents. However, prolonged use of keyboard and computer workstations have led some translators to become disabled, as they become vulnerable to musculo-skeletal disorders, commonly known as Repetitive Strain Injuries or RSI.

This scoping discussion examines the case of three disabled translators contacted during the course of this study. For reasons of confidentiality, their names and personal details are not disclosed.

Translator 1

This translator had recently developed a motor neuron disease. Motor neuron diseases destroy the cells that control voluntary muscle activity, gradually preventing sufferers from controlling their bodies, initially making general movement such as walking, typing or holding a biro difficult, but leading to an inability to speak, breathe and swallow.

During the time span of this study the translator came to require the use of a wheelchair and could no longer type with ease. The translator works in the public sector fulfilling the organization's need to operate a bilingual policy. The translator's daughter was employed by the translator's public sector employer as a typist. This has allowed the translator to continue in employment, although the extra cost of a typist is significant and would perhaps be prohibitive for a freelance translator. Unfortunately, as the translator's main work involves translating from English into Welsh, the use of speech recognition products was not open to him, as Welsh is not supported by the commercial speech recognition software currently available.

Although Welsh speech recognition resources are currently being developed at the Language Technologies Unit at Bangor University, research is at an early stage, and professional quality applications are not yet fully ripe. Investment to continue this pioneering work is currently being sought.

One strategy open to the translator and his employer would be encourage the organization to produce more documents through the medium of Welsh, and for the translator to utilize existing English speech recognition software such as Dragon Naturally Speaking to translate those document into English. However this is no substitute for the development of commercial quality speech recognition for currently unsupported languages such as Welsh.

Translator 2

Another translator contacted during the course of this disability study had developed an upper limb musculo-skeletal disorder (commonly known as RSI or Repetitive Strain Injury) after years of typing at her keyboard. Upper limb disorders include conditions such as tenosynovitis, carpal tunnel, and tendinitis. These arise from repetitive or forceful tasks, including the use of devices such as keyboards and mice that allow an individual to interface with computers. Bad posture when seated is also cited as a cause of such conditions, with stress also believed by some to be a contributing factor.

The translator could no longer continue to type using a keyboard due to the pain. As she, like the first translator, translates from English into Welsh, no appropriate speech recognition software was available to enable her to bypass using the keyboard. When she contacted her local *Menter Iaith* (language support association) she was advised to dictate and record her translations and employ a typist to transcribe them. At a typical cost of £10 per hour this was an expense she felt that she could not afford as a freelance translator.

The temptation for this translator and others like her is to continue to struggle to type, and to aggravate their medical condition. Several workarounds are available before the condition becomes too serious, such as the use of shortcut keys and predictive software, allied to translation memory systems, to cut down on the need for typing. However, ideally these should be used as preventative, rather than palliative measures, and speech recognition is the best solution for those with established RSI symptoms.

Translator 3

This translator is visually-impaired. He works as a bilingual administrative officer undertaking general office tasks who is occasionally called upon to translate Welsh documents as part of his duties rather than being employed solely on translation duties. He needs to magnify text on his computer monitor as much as possible in order to see the text. He prefers to use the text-to-speech capability of his computer to read the words aloud to him. Since the advent of Welsh language text-to-speech he has been able to use a screen reader to have Welsh read aloud to him on screen. He is then able to touch-type the English translation or he conveys the translation orally to a colleague, if that is what is required.

Greater attention to accessibility standards, especially for high screen contrast, large icons, and the provision for magnifying text within dictionary applications, would improve his access translation to resources on the computer screen and on-line.

Further needs

As there are many kinds of disabilities, there are many kinds of needs for accessibility aids for translators. Many of these can be catered for though the provision of general aids to allow access to computers and the internet. The disabled translators in the above studies felt excluded because of their need for Welsh language functionality, rather than because of their impairments. Speech recognition for Welsh, and

adherence to standards on accessibility for the visually impaired within Welsh language resources were cited as priorities by these users.

Scoping Discussion 5: Customers of the translation industry

Profile

Customers and potential customers of the translation industry in Wales are many and varied. Both freelancers and private translation bureaux reported that the majority of their customers in Wales were public bodies. This is in line with the statutory obligations for public bodies in Wales to operate Welsh language policies. Although there is no statutory obligation on them to do so, many private companies are also increasing their bilingual provision within Wales. Public-facing companies such as retailers, financial service providers, and leisure and tourism facilities are at the forefront of this change.

Customers needing translations into languages other than Welsh fall into two categories: those who are trying to export their products to other countries, and those who are trying to attract visitors from other countries to visit Wales. Very few large international companies have their headquarters in Wales. In many instances translation work is commissioned and dealt with outside Wales. “We send all our translation requests to head office” was one common reply. Smaller companies tend to deal with translation needs on an ad-hoc basis: “it depends what we need” was one reply received. The need for translation into Asian languages for newly emerging economies was also emphasised, as India and China were seen as markets with great potential for Welsh exporters. Translation support for SMEs wishing to export was seen as non-existent, with companies having to look outside Wales for advice and services.

Creative Industries Sector

Companies in the creative industries sector reported the need to translate not only marketing and administrative material, but also entire film scripts. “We see a huge potential in selling programmes outside Wales. To do this effectively we need to be able to translate both content and promotional material into other languages. This is a new area for us, but we need to be proactive to get our products out there. The returns are far greater than what we can get at home”.

Testun Cyf, a translation company based in Cardiff who specialize in subtitling television programmes is one company who foresee a potential in selling translated programmes to international markets. Their current Knowledge Transfer Programme (KTP) project with the LTU at Bangor University hopes to deliver automated translation aids to their Welsh and English language activities in the first instance, but then to adapt the technology to meet the anticipated demand from media companies to localize their programs to a wide range of other languages for European and world-wide markets.

Tourism

Companies working within the tourism sector were also aware of the benefits of multilingual marketing, but even here, the use of translation service providers was patchy. “We could do far more to attract visitors if we had a multilingual website” was one comment. “As it is, we expect everyone to understand English”.

An increasing number of tourism websites were found to be providing bilingual English and Welsh websites. Websites such as those for the Celtic Manor Resort and Oakwood Theme Park were found to possess Welsh mini-websites, but made no provision for international visitors in languages such as French and German. This is surprising considering the large numbers of international visitors to such facilities, and anticipated major international sporting events due to visit Wales. As there is no real difference in technical provision between a bilingual and multilingual architecture for a website, translating key web pages into other popular languages need not be expensive or involve a large redesign of such websites.

The general impression therefore was of a huge unmet opportunity for industry in Wales to improve their marketing and their penetration of export markets through improved use of multilingual provision. Multilingual websites were seen as one area for immediate improvement, and an area where translation technology tools could make a huge impact.

Guidelines, demonstrations, and case studies aimed at encouraging the internationalization of Welsh tourism websites could provide a significant marketing boost to the tourism industry, especially at a time when the fall in the value of sterling makes Wales a more attractive destination than ever for travellers from abroad.

Welsh-language Translation in the Private Sector

The commissioning of Welsh translation by the private sector is increasing as the demand from Welsh-speakers for Welsh speaking services is increasingly recognised.

Large multinationals such as Microsoft have commissioned and delivered Welsh translations of their products, and Facebook, one of the best able to identify the demands of their users, chose Welsh as one of their first non top 4 languages to be translated using their crowdsourcing translation mechanism.

In the retail sector, large supermarkets chains such as Tesco, Sainsbury's, Morrisons and M&S now routinely display bilingual English and Welsh signage in their stores.

Given that the sector is new to the commissioning of translations, this development has not always proceeded smoothly, and quality control has been a recurring problem, especially on signage. Some of the errors cause Welsh-speakers to smile, such as a "Parking for parents with children" sign rendered as *Parcio ar gyfer rhieni a phla* ("Parking for parents with plague"), the result of the last two letters being truncated due to lack of space. Such mistranslations have spawned a group on the Flickr website (see <http://www.flickr.com/groups/scymraeg>) dedicated to showing photographs of badly translated signs from all over Wales.

Not all these errors are linguistic in nature, and include other issues such as quality control with sign writers. This highlights the fact that private sector customers of the translation industry require support and resources that are often different to those which exist in Wales for the public sector, and their needs should be included in the remit of a translation technology centre.

Welsh Branding

Food and drink manufacturers such as Braces, Penderyn Whisky and Murrroughs Welsh Brew Tea (marketed as *Paned Gymreig*) have used bilingual Welsh and English packaging on their products to strengthen their branding as authentic Welsh products. This is an increasing phenomenon with unique requirements for Welsh translators, but where extending such branding to export markets can increase the appeal of products from Wales. Asian markets such as Japan may be particularly attracted to such branding, and they increase the distinctiveness of Welsh products.

Multilingual Packaging

Multilingual packaging is the norm in Europe, where packaging often features two, three or more languages. This means that the more aware potential exporters are of the translation commissioning process, the better placed such companies are to access European export markets.

For example, it is important for the exporters of foodstuffs to use translators with access to resources such as standardized multilingual terminology databases for the names of food ingredients, or risk paying for additional work and receiving nonstandard translations.

Encouraging Exports

Discussions with Dr Christian Galinski, director of UNESCO's Infoterm service, emphasized the strategic opportunities for globalization and the key role multilingualism has to play in allowing SMEs to reach new export markets. Conformity to international guidelines and standards is necessary if translation resources are not to be wasted, and companies able to respond imaginatively to the challenges of multilingualism are likely to score ahead of their competitors.

In his book *Hidden Champions* (Leeds Press 2008), Hermann Simon proves that it is possible for SMEs to compete in global markets and presents a sample of small and medium enterprises who achieve more than half their sales in export markets. Although language can be a barrier to globalization, it is in no way insurmountable given the services available to would-be exporters from translation service providers.

The true impediment to export according to Hermann is a mental one, a point emphasised by Clemens C. Steiner in his valuable guide to selling in the global marketplace *SMEs Go Global!* (Wienn 2003). He found that many companies who were identified as not having internationalized their businesses were too focused on optimizing their home markets, and had in many cases been doing so for decades. These companies he categorized as *hesitators*.

Steiner saw this as evidence that:

“Hesitators have not understood the Pareto skimming principle, hence spend 80% of their energy to capture the remaining 20% in the existing market rather than grasping 320% in new other markets, spending the same time and money” (p120)

As a result, Welsh SMEs who have gained a significant share of their home market should be encouraged to channel their energies into seeking new export markets rather than attempt to gain further market share at home.

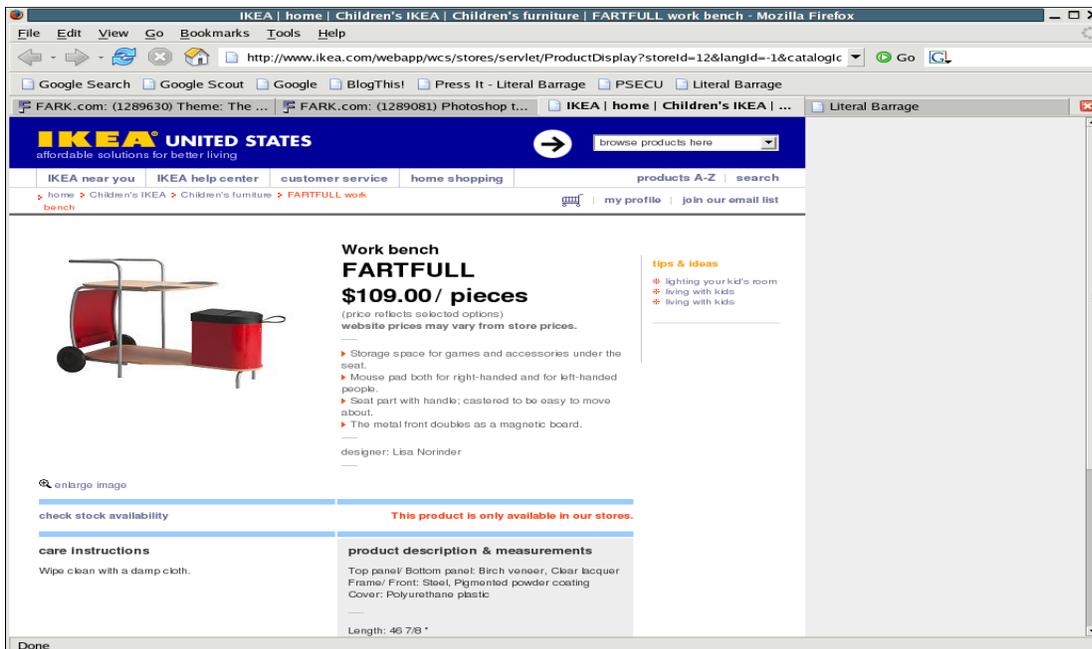
Another error highlighted by Steiner was that SMEs who did consider branching out into new export markets often automatically looked to neighbouring markets, rather than taking the global view.

“Hesitators are not aware that transactional costs for goods and services nowadays are not dominated any longer by geographical distance as the main cost driver – as a matter of fact they would need to widen their horizon urgently to understand that usually the richest markets are not necessarily next door.”

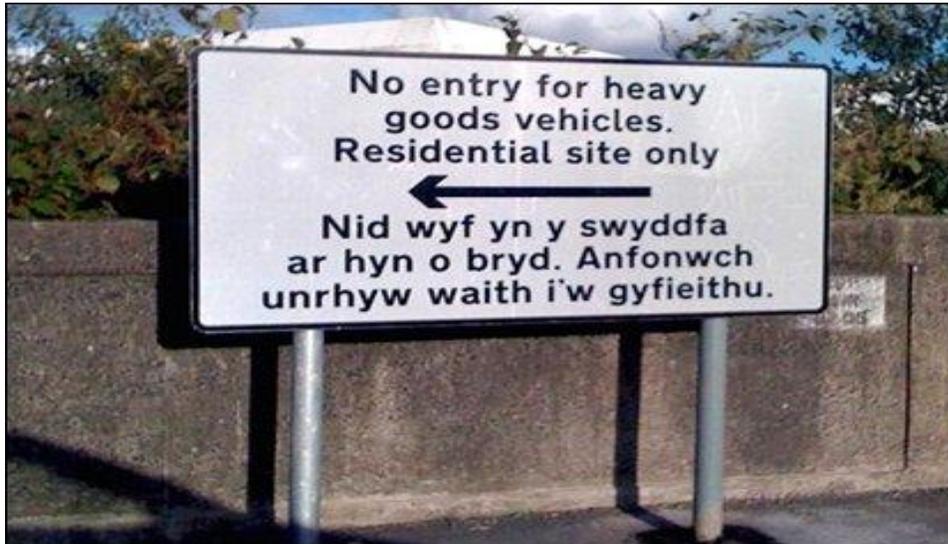
Just as these markets may not be next door, they may not share a common language with the SME’s home market. In today’s globalized world the richest market may be found in areas speaking Hindi, Russian or Chinese rather than the more traditional major European languages. However, as with distance, translation and localization services are often not substantial cost drivers in comparison with the gains to be made from the successful exploitation of new markets.

Improving the commissioning process

Lack of understanding of the translation process is not a specifically Welsh problem. Many international misunderstandings and unintentionally funny mistranslations happen when the customer has no quality control mechanism for catching rogue translations. However, these can cause embarrassment and loss of revenue in a commercial situation. Even the naming of a product for a foreign market needs to be carefully checked in case the chosen name means something unfortunate in the target market. One such an example occurred in the *Ikea* catalogue, where a workbench was marketed under the name FARTFULL:



Translation errors on visible signs are particularly embarrassing for commerce. They can occur because of lack of quality assurance on the commissioning side. A recent sign in a Welsh supermarket which gained world-wide notoriety was when an automated out of office message from a translator was taken for the translation of the message itself:



The Welsh reads: “I am not in the office at present. Please send any work to be translated”.

Another area where it is vital that no translation errors are made is in legal documents, such as contracts with suppliers of goods and services. In such circumstances, specialist legal translators should be used, and technical tools such as a dedicated legal TM, and good quality legal terminology resources used to help ensure quality and consistency.

Need for commissioning guidelines

The potential pitfalls of badly translated materials could be discouraging to industry. The need for guidelines was emphasised, and the possibility of providing industry in Wales with simple to follow steps towards establishing and managing translation process, such as:

- deciding on the scope of the project
- drawing up specifications for translation needs
- choosing a translation service provider
- adherence to translation industry standards
- pre-editing source text to be easily and unambiguously translated
- costings, deadlines and other administrative issues
- arrangements for upkeep and revisions
- proof-reading and quality control

Many of these issues are addressed by the BSI's BS EN 15038 *Translation services – service requirements* standard. Although these requirements are approached primarily

from the standpoint of translation service providers rather than that of their customers, valuable insights can be gained by clients of translation service as to the services available and the scope and level of service they should expect. Chris Cox's *Translation Service Provider's Guide to BS EN 15038* (BSI, 2006) provides an informative and clearly written guide to the standard. The publishing of a similar guide for *customers* of TSPs could prove highly beneficial in improving the translation industry's use of translation services.

Other issues to be covered could be the appropriate use of machine translation, especially in the form of gist machine (automatic) translation for internal needs, the keeping of in-house glossaries to ensure terminology consistency, the keeping of an in-house translation memory system to cut down on external translation costs, and all issues of ensuring a fruitful dialogue with the translation industry in Wales.

Further needs

Although authoring tools for better preparation of source text before sending it to be translated was not mentioned by customers of the translation industry in this study, this technology has the potential to save money and improve quality. This is especially true where text has to be translated into multiple languages. Effort spent in ensuring clear, concise sentences, no use of ambiguous terms, and consistency of expression, is amply repaid each time the text is converted into another language. Substantial savings were reported by Hewlett-Packard in a case study (see <http://www.author-it.com/index.php?page=hpiss>) through using such authoring tools. Providing translation industry customers and potential customers with access to the same demonstration facilities as those enjoyed by translators themselves is seen as crucial to enabling industry in Wales make full use of multilingual resources in their marketing and exporting activities. A better understanding of the translation process and the benefits of multilingual provisions will benefit both the translation industry and their customers.

Although the main focus of translation technology is on providing tools that translators can use, some tools may be of benefit to commissioners of translation. These include tools to manage their in-house terminology, their own translation memory systems, and authoring tools for writing original text. Many of these issues are relevant to customers in both the public and the private sectors, and to bilingual Welsh/English and broad multilingual translation projects.

At the moment, customers seem to be the poor relations in the translation supply chain. However, without customers the translation industry would not exist. Helping one part of the supply chain helps all, and the largest economic benefits could be those reaped from enabling Welsh industry to compete better in on the global, multilingual stage.

Appendix 3: Planning for the Future – A Discussion Paper

This project began with the simple idea of developing a toolkit of resources and software that translators in Wales could use to improve their translation processes. These tools would, as far as possible, be generic in nature, and adaptable to any language pair or combination of multiple languages. The provision of these tools would encourage the uptake of translation technology within Wales, increase capacity within the industry, improve quality and keep a lid on translation costs. It would help the public sector deliver on its statutory obligations to service a bilingual Wales, and would help Welsh industry to compete with its exports in a multilingual, global environment.

The project tried to identify mechanism of knowledge transfer, enabling the knowledge base to share its expertise with industry, and provide for sustainable growth within the Welsh economy. It has seen the translation needs of Wales, not as an economic burden, but as an opportunity to build a vibrant translation sector, using experience gained in providing for domestic needs to compete in the wider world. In this way its bilingual experiences at home will feed through into increasing capacity within the international market.

Academic researchers are often accused of chasing after “blue sky” ideas, and of being too far removed from the immediate needs of industry. In this project the focus has been on practical, achievable solutions, ensuring value for money and the targeting of effort and funding in order to gain maximum impact on the industry sector and the Welsh economy.

Encouraging uptake of existing technology and providing guidelines for industry customers were therefore identified as priorities. Collating existing tools and packaging them for ease of use was also seen as a necessary prerequisite before moving on to develop more sophisticated solutions. However, in the longer time frame, innovation and development of new technical solutions are seen as gaining greater prominence. The dialogue between the translation industry, their customers and the academic knowledge base will generate new ideas, and the increased capacity of the knowledge base, through a dedicated translation technology centre at Bangor University, will ensure that such ideas are explored in specific research projects, leading to the development of new products and services.

Two other issues came into focus in this project, namely Welsh language needs and long term economic impact and sustainability.

In the past the needs of Welsh language translators have been treated in isolation from the needs of the rest of the translation industry in Wales. This has led to lack of investment as Welsh is often not seen to be commercially interesting. This perception needs to be challenged, as the size of the Welsh language translation market is by now significant in itself. There has also been a lack of focus on the need of translators in Wales working with languages other than Welsh. Here, there has been very little support at a Wales level. Welsh industry have thereby suffered, and as customers and potential customers of this sector, they have not been able realize the potential of multilingual marketing and communication in their export activities.

The mainstreaming of Welsh language provision, as part of the wider translation industry in Wales is therefore one strong message from this project. Designing translation tools such as terminology portals and translation memory systems to be adaptable for any language combination will deliver support for all translators in Wales, irrespective of the languages they work with. The establishment of a translation technology centre, accessible and welcoming to all translators in Wales, as well as to their customers and potential customers, is seen as a value for money solution to the translation and multilingual needs of Welsh industry.

The long-term economic impact of such provision is likely to be significant. The figure of 19% business loss to the Welsh economy through lack of language skills has been quoted as an indicator of opportunities missed. It is argued that in view of the current economic downturn, providing Welsh industry with support and solutions becomes even more urgent. Selling to multiple export markets also help diversify the customer base, thereby lessening reliance on any one market. This may involve commissioning translations into multiple languages, and companies need to resist the urge to stay in familiar territory with a small number of languages that they have grown comfortable dealing with. Rather than economizing on translation costs, companies need to invest more in improved translation services to compete in a tougher international market, where competitors with superior multilingual provision may gain the advantage.

Providing long-term support to Welsh industry means that the long-term sustainability of the translation technology centre has to be safeguarded. Initial investment is seen as a pump primer leading to the centre and its activities becoming self sustaining. The academic knowledge base may be strengthened through development of new research, including grant applications to HE Funding Councils. Joint ventures with Welsh industry will be sought, such as those offered under the new KESS programme, where private companies will co-fund PhDs and MRes research together with Convergence funding. The KTP programme is another opportunity for Welsh industry to co-fund research and development in various aspects of translation technology. These are both schemes which are eminently suitable for small-scale investment by micro SMEs who make up the bulk of the translation industry in Wales and will encourage them to invest where there has not been a tradition of funding R&D.

Inward investment is also anticipated. In future UNESCO's invitation for cooperation with the LTU at Bangor University, and the opening of a UK/Ireland office there will increase the centre's international profile. Extending the centre of excellence into a cross-border facility with Ireland with the aid of Interreg IV funding is another strategic aim.

Licensing new products will be another means of attracting revenue. Commercially viable translation tools may be developed, both for inclusion in other products commercialized by industry (e.g. spellcheckers in TM programmes) and in products sold directly to the end user. This will capitalize on university-owned IP owned, and will be used to fund further research and development.

All these initiatives entail the continued support and encouragement of all relevant stakeholders in Wales. The support of the Welsh Assembly Government and the A4B programme which funded the original investigation has been much appreciated, and will continue to be vital to

continued success. The support of Cymdeithas Cyfieithwyr Cymru, the Association of Welsh Translators and Interpreters, is also a key element in developing further support to the translation industry in Wales. Reaching other relevant stakeholders, in the translation industry, the public sector, and Welsh business and industry in general will also allow us to plan ahead, and ensure that translation technology will play its role in creating a vibrant and outward-looking Welsh economy.