

Confensiynau trawsgrifio CorCenCC

CorCenCC's transcription conventions¹



Corpws Cenedlaethol Cymraeg Cyfoes
National Corpus of Contemporary Welsh

Crynodeb | Précis version

1 Trawsgrifio / Transcribing

1.1 Troeon Siaradwyr / Speaker Turns

Yn ystod sgysiaiu, bydd pobl yn cymryd tro i siarad. Gelwir y rhain yn '**droeon siaradwyr**'. Dechreuwch bob tro gyda'r cod siaradwr cywir. Gwasgwch **Enter** yn *Transcriber* ar ddiwedd bob tro.

*During conversations, people take it in turns to talk. We call these '**speaker turns**'. Start each turn with the correct speaker code. Press **Enter** in *Transcriber* at the end of each turn.*

1.2 Atalnodi / Punctuation

1.2.1 Marciau Atalnodi / Punctuation Marks

- Dechreuwch bob tro gyda'r **cod siaradwr** cywir
- Dechreuwch y gair nesaf gyda **phrif lythyren**
- Defnyddiwch **farc cwestiwn** ar ôl pob cwestiwn, neu pan fydd y siaradwr yn troi gosodiad yn gwestiwn (e.e. *Y car hwn?*).
- Defnyddiwch **atalnod llawn** ar ddiwedd bob tro arall, heblaw pan fydd lleferydd yn gorgyffwrdd

Ni ddylid defnyddio marciau atalnodi eraill – e.e. comas, ebychnodau. Dylid defnyddio collnodau pan fydd llythyrennau ar goll.

- *Start each turn with the correct **speaker code***
- *Start the following word with a **capital letter***
- *Use a **question mark** after every question, or when a speaker is turning a statement*

¹ Datblygwyd y confensiynau trawsgrifio hyn ar brosiect Corpws Cenedlaethol Cymraeg Cyfoes (The National Corpus of Contemporary Welsh): Dull cymunedol o greu corpws ieithyddol (Rhif Grant ES/M011348/1) a gyllidwyd gan Gyngor Ymchwil Economaidd a Chymdeithasol y DU (ESRC) a Chyngor Ymchwil i Celfyddydau a'r Dyniaethau (AHRC). Am fwy o wybodaeth, ewch i www.corcencc.cymru
These transcription conventions were developed on the UK Economic and Social Research Council (ESRC) and Arts and Humanities Research Council (AHRC) funded Corpws Cenedlaethol Cymraeg Cyfoes (The National Corpus of Contemporary Welsh): A community driven approach to linguistic corpus construction project (Grant Number ES/M011348/1). For more information, go to www.corcencc.org

Wrth ddyfynnu confensiynau CorCenCC gofynnir i chi gyfeirio at y fersiwn lawn a gyhoeddir yn: **Knight, D., Morris, S., Arman, L., Needs, J. a Rees, M.** (2021, ar waith). *Blueprints for minoritised language corpus design: a focus on CorCenCC*. Llundain: Palgrave.

When citing CorCenCC's conventions please refer to the full version published in:

Knight, D., Morris, S., Arman, L., Needs, J. and Rees, M. (2021, in prep.). *Blueprints for minoritised language corpus design: a focus on CorCenCC*. London: Palgrave.

into a question

- Use a **full stop** in all other instances, unless there is overlapping speech

Do not use any other punctuation marks, e.g. commas, exclamation marks. Apostrophes should be used when letters are missing.

1.2.2 Unedau / Units

Fel arfer bydd un tro wedi ei ffurfio o un uned (h.y. 'brawddeg' ar lafar), e.e. <S1> *Rydw i'n hoffi caws.* Ambell waith bydd siaradwr yn defnyddio mwy nag un uned mewn tro.

Defnyddiwch **atalnod llawn** (neu **farc cwestiwn**) ar ddiwedd bob uned, hyd yn oed os yw'n digwydd yng nghanol y tro.

*A speaker turn usually consists of one unit (i.e. a spoken 'sentence'), e.g. <S1> I like cheese. Sometimes, however, a speaker will use several units in their turn. Use a **full stop** (or **question mark**) at the end of each unit, even if this occurs mid-turn.*

Enghreifftiau / Example:

<S1> *Pa fath o dad yw e? Cwestiwn da. Weithiau mae'n anodd gwybod pwy yw'r oedolyn a phwy yw'r plentyn.*

1.2.3 Gofodau / Spaces

Teipiwch un bwlch yn unig rhwng geiriau, bracedi ac unedau ac eithrio lle bo lleferydd yn gorgyffwrdd (+) neu heb ei orffen (-).

Type a single space between words, brackets and 'units', unless there is speech overlap (+) or unfinished speech (-).

1.2.4 Enwau priod / Proper Nouns

Dylid defnyddio priflythrennau ar ddechrau pob gair wrth deipio enwau pobl, lleoedd, pynciau, adrannau, ffilmiau, sefydliadau, cyrsiau, ac ati. Cofiwch gyfeirio at y Confensiynau Anonymeiddio wrth ddelio ag enwau.

Names of people, places, subjects, departments, institutions, courses, films, books, etc. should be typed with initial capitals. Remember to check the Anonymisation Conventions when dealing with names.

1.2.5 Dyfyniadau, Darllen ac Adrodd / Quotations, Reading and Reciting

Dylid defnyddio marciau dyfynnu dwbl (" nid ') lle bo'n amlwg bod y siaradwr yn dyfynnu rhywun, yn darllen yn uchel, neu'n adrodd.

Double quotation marks (" not ') should be used when it is obvious that a speaker is quoting someone, reading aloud, or reciting.

Enghreifftiau / Example:

<S4> *Dyna oedd fy nhaid yn wastad yn dweud "dyfal donc a dyr y garreg".*

1.2.6 Acronymau / Acronyms

Dylid defnyddio priflythrennau ar gyfer acronymau. Ni ddylid defnyddio atalnodau llawn na bylchau. Lle bo acronym wedi ei ynganu fel petai'n air, dylid defnyddio priflythrennau ar ddechrau'r gair yn unig:

Acronyms should be typed in capitals. No spacing or full stops should be used. When an acronym is pronounced as a word, it should be typed with an initial capital only:

Enghreifftiau / Example:

<S1> Mae'n gweithio i'r BBC erbyn hyn.

<S1> Mae hi'n eistedd ei harholiadau TGAU eleni.

1.2.7 Byrfoddau a Ffurfiâu Byrion / Abbreviations and Short Forms

Ni ddylid defnyddio atalnodau llawn ar ôl teitlau neu fyrfoddau.

Full stops should not be used following abbreviations or titles.

Enghreifftiau / Example:

<S1> Mae Dr Jones yn barod i'ch gweld nawr

1.3 Tagiau Trawsgrifio / Transcription Tags

1.3.1 Codau Siaradwyr / Speaker Codes	
<S1> <S2> <S3>	Defnyddiwch god newydd ar gyfer bob siaradwr(aig) ond cofiwch ail-dddefnyddio'r un cod pob tro mae'r person yna yn siarad. <i>Use a new code for each speaker but remember to use the same code each time that person speaks.</i>
<SB> <SG> <S?>	Lle nad yw'n bosib adnabod y siaradwr (SG: "gwrywaidd"; SB: "benywaidd"; S?: methu ag adnabod y rhyw). <i>Where it's not possible to identify the speaker (SG: male; SB: female; S?: unable to identify the gender)</i>
<S1?>	Lle nad yw'n bosib adnabod y siaradwr, ond lle bo'r trawsgrifydd yn gymharol ffyddiog o ran pwy sy'n siarad. <i>When it's not possible to identify the speaker, but the transcriber is reasonably confident of who's speaking.</i>
<SS>	Pan fo nifer o bobl yn siarad ar yr un pryd. <i>When a number of speakers talk at once.</i> <S1> Bore da blant. <SS> Bore da Syr.
1.3.2 Lleferydd yn gorgyffwrdd / Overlapping speech	
+	Lle bo lleferydd un siaradwr yn gorgyffwrdd â lleferydd siaradwr arall. Wrth ddefnyddio Transcriber, dylid gwasgu Shift a Return gyda'i gilydd i wahanu'r siaradwyr lle bo lleferydd yn gorgyffwrdd. <i>When one speaker's speech overlaps another's. When using Transcriber, press the Shift key and the return key together to separate speakers where there is overlapping speech:</i> <S1> Roedd hyn yn ddarlith radio+ <S2> Ie <S1> +hynod bwysig.
1.3.3 Lleferydd sydd heb ei orffen, ymyriadau ac ail-ddechrau / Unfinished speech, interrupted speech and false starts	
<->	Lle bo ymyriad neu lle bo tro siaradwr heb ei orffen. Defnyddir hefyd pan fydd siaradwr yn ymyrryd ar ei l/leferydd ei hun

	<p><i>Where a turn is unfinished or interrupted. Also used when the speaker interrupts him/herself and changes direction (a false start)</i></p> <p><S1> <i>Es i am dro unwaith a <->.</i></p> <p><S2> <i>Am dro? I ble?</i></p>
-	<p>Defnyddiwrch gysylltnod a bwlch lle bo geiriau yn dameidiog. (Defnyddiwrch gysylltnod heb fwlch lle bo'r cysylltnod yn rhan o sillafiad y gair – e.e. cyd-destun)</p> <p><i>Use a hyphen plus a space for word fragments. (Use a hyphen without a space when the hyphen forms part of the word's spelling, e.g. sugar-free)</i></p> <p><S1> <i>Lle brafydy Ab- Aberystwyth</i></p>
1.3.4 Ailadrodd / Repetition	
<=>	<p>Pan fydd siaradwr yn ailadrodd gair neu eiriau llawn. Hefyd pan fydd siaradwr yn cywiro ei hun wrth ailadrodd.</p> <p><i>When a speaker repeats a whole word or words. Also when a speaker corrects themselves when repeating.</i></p> <p><S4> <i>Ble <=> mae mae mae </=> mae Daniel?</i></p> <p><S6> <i>Dw i'n dod o <=> Caerdydd </=> Gaerdydd.</i></p>
1.3.5 Seibiau / Pauses	
<saib>	<p>Ystyrier saib fel bwlch yn y lleferydd sy'n hwy na'r arfer. Gall seibiau ddigwydd yng nghanol uned ("brawddeg") neu rhwng unedau.</p> <p><i>A pause is a gap in speech that's longer than usual. Pauses can occur mid-unit (or "sentence"), or between units.</i></p> <p><S5> <i>Mae hwn yn adlewyrchu'n <saib> ddrwg arnom.</i></p> <p><S4> <i>Sut mae plannu tatws? <saib> Fel hyn.</i></p>
1.3.6 Nodiadau trawsgrifydd / Transcriber notes	
<N> </N>	<p>I farcio unrhyw nodiadau angenrheidiol neu wybodaeth gyd-destunol.</p> <p><i>To record any necessary notes or contextual information.</i></p> <p><N> <i>Rhywun yn baglu dros gadair yn y cefndir </N>.</i></p> <p>Wrth ddefnyddio <i>Transcriber</i>, dylid gwasgu Shift a Return gyda'i gilydd i wahanu troeon y siaradwyr a nodiadau'r trawsgrifydd.</p> <p><i>In Transcriber, press Shift and Return to separate speaker turns and add a transcriber note</i></p>
1.3.7 Dyfaliadau / Guesses	
<D> </D>	<p>Lle nad oes sicrwydd o ran beth mae'r siaradwr wedi ei ddweud.</p> <p><i>When the transcriber is unsure of what's being said.</i></p> <p><S5> <i>Wel y brif <D> broblem </D> yw eu bod nhw'n methu newid.</i></p>
1.3.8 Aneglurder / Unintelligible	
<aneglur>	<p>Gair aneglur unigol.</p> <p><i>A single unintelligible word.</i></p> <p><S3> <i>Dwedodd Jac fod <aneglur> ganddo.</i></p>
<aneglur2>	<p>Lle bo mwy nag un gair aneglur (defnyddiwrch rif i ddynodi sawl gair sydd ar goll).</p>

	<i>Where more than one word is unintelligible (use the number to indicate how many words are missing).</i> <S2> Dw i'n credu bod <aneglur4>.
<aneglur?>	Lle nad yw'n bosib gwybod sawl gair sy'n aneglur. <i>Where it's not possible to know how many words are missing.</i>
1.3.9 Sillafu geiriau allan / Spelling out words	
<sillafu> </sillafu>	Lle bo siaradwr yn sillafu gair fesul llythyren. Gadewch bwlch rhwng bob llythyren. <i>Where a speaker spells out a word letter by letter. Leave a space between each letter.</i> <S1> Na Lowri ydw i <sillafu> l o w r i </sillafu>
1.3.10 Synau / Sounds	
<[enw'r sŵn/name of sound]>	Gweler yr Atodiad am restr gynhwysfawr o synau posib. <i>See the Appendix for a comprehensive list of possible sounds.</i>

2 Confensiynau Sillafu / Spelling Conventions

2.1 Egwyddor Gyffredinol / General Principle

Nid oes angen safoni ffurfiau rhanbarthol. **Ni** ddylid cywiro camgymeriadau gramadegol (e.e. cam-dreiglo) - teipiwrch yr hyn a glywir.

There is no need to 'standardise' regional pronunciations. Do not correct grammatical mistakes (e.g. incorrect mutations) - type what you hear.

Enghreifftiau / Examples:

adag neu adeg,

llywodr'eth neu llywodraeth

oes, o's neu wes

pethau, pethe neu pethau

Eithriadau / Exceptions:

Peidiwch â thynnu sylw at namau lleferydd nag anawsterau ynganu (e.e. pobl yn dweud -th- yn lle -s-, neu chl-/thl- yn lle -ll-).

Do not draw attention to speech impediments or pronunciation difficulties (e.g. people saying -th- instead of -s-, or -chl-/thl- instead of -ll-).

2.2 Llythrennau sydd wedi'u hepgor / Omitted Letters

Defnyddiwch gollnod i ddangos bod un neu fwy o lythrennau wedi'i/u hepgor. Cofiwch fod yn gyson trwy'r trawsgrifiad. Gweler [Adran 2.4](#) am gyfarwyddiadau ar sut i drin rhagenwau, e.e. dw i.

Use an apostrophe to show that one or more letters have been omitted. Remember to be consistent throughout the transcription. See [Section 2.4](#) for instructions on how to handle pronouns, e.g. dw i.

Enghreifftiau / Examples:

ar ddechrau gair / at the start of a word:

'i ("ei")
'lly ("felly")
'na ("yna", "dyna", "hwnna")
'di ("wedi")
'dy ("ydy")
'to ("eto")

ar ddiwedd gair / at the end of a word:

myn' i ffwr'
be' dach chi'n feddwl
yn yr ha'

yng nghanol gair / in the middle of a word:

ca'l
lla'th
gw'ith'o
er'ill

wrth gyfuno geiriau / when combining words:

(doe)s + (y)na + (ddi)m = 's'na'm
dw i + (y)n = dw i'n
dw i + (ddi)m = dw i'm
(dy)dy o + (ddi)m = 'dy o'm
(ddi)m (yn) + (gwy)bo(d) =
'm'bo'
gyda + (e)i = gyda'i
sy(dd) + (y)n = sy'n
t(i'n) + (g)wel(d) = t'wel

Eithriadau:

- **Nid** oes angen collnod ar ddechrau geiriau pan fo'r llythyren g- wedi'i cholli drwy dreiglo – e.e. *yr ardd*, *dwyr afwr*, *welest ti ddim*.
- **Nid** oes angen collnod ar ddiwedd geiriau lluosog megis *geiria*, *petha*, *papura*, *'sgidia*.
- **Exceptions:**
- **No** apostrophe is required at the start of words when the letter g- has been lost due to mutation – e.g. *yr ardd*, *dwyr afwr*, *welest ti ddim*.
- **No** apostrophe is required at the end of plural words such as *geiria*, *petha*, *papura*, *'sgidia*.

2.3 Homoffonau ac Amwysedd / Homophones and Ambiguity

Weithiau, mae dwy sain yn swnio'r un fath (h.y. homoffonau ydyn nhw), ond yn aml bydd hi'n amlwg o'r cyd-destun pa eiriau oedd gyda'r siaradwr mewn golwg.

Sometimes, two words sound the same (i.e. they are homophones), but often it will be obvious from context which word was meant.

Enghreifftiau / Examples:

Ma' fe'n dod.
(*ma'* = "mae")
Y petha 'ma.
(*'ma* = "yma")

Weithiau bydd achosion amwys. Mewn achosion o'r fath, dylid amgylchynu'r sain gyda thagiâu <ynganu> ac </ynganu>.

Sometimes there will be ambiguous cases. In such cases, the sound should be marked between <ynganu> and </ynganu>.

Enghreifftiau / Examples:

Ma fe.

(Dyma fe? Mae fe?)

<S1><Ynganu> Ma </ynganu> fe.

2.4 Bylchau o Flaen Rhagenwau / Spaces Between Pronouns

Rhowch fwllch rhwng berfau neu arddodiaid wedi'u rhedeg a'r rhagenw sy'n dilyn.

Put a space between conjugated verbs or prepositions and the following pronoun.

Enghreifftiau / Examples:

dw i (~~dwi~~)

Ga' i fynd? (~~Gai fynd?~~)

arna' i (~~arnai~~)

dach chi (~~dachi~~)

2.5 Cyfnewid Cod / Code-switching

Lle bo enghreifftiau o siaradwr yn defnyddio iaith arall am gyfnod (geiriau / ymadroddion / brawddegau) dylid amgylchynu'r lleferydd di-Gymraeg gyda thagiau sy'n dynodi'r iaith.

Dylid defnyddio codau dwy lythyren ISO 639-1 ar gyfer ieithoedd eraill. Er enghraifft, defnyddiwch y tagiau <fr> a </fr> ar gyfer cyfnewid cod i'r Ffrangeg.

Where a speaker uses words, phrases or sentences in another language, the non-Welsh speech should be enclosed in tags which indicate the language used. Refer to this list of two-letter codes in these scenarios (column headed "639-1"). For example, use the tags <fr> and </fr> when code-switching into French.

Enghreifftiau / Examples:

<S4> Rwy'n dy garu di o hyd <fr> Je ne sais pas pourquoi </fr>.

Lle bo'r lleferydd di-Gymraeg yn parhau ar draws sawl tro siaradwr, dylid marcio'r iaith amgen ar bob tro.

Where the non-Welsh speech continues across several speaker turns, the non-Welsh speech should be marked for every turn.

Enghreifftiau / Examples:

<S4> Rwy'n dy garu di o hyd <fr> je ne sais pas pourquoi </fr>.

<S1> <fr> Mais qui </fr>.

<S3> <fr> Je ne parle pas Français </fr>.

<S4> Paid â gofidio.

Lle nad oes modd adnabod yr iaith a siaredir, dylid nodi hyn gyda'r tag <lleferydd di-Gymraeg>.

Where it is not possible to identify the language, use <lleferydd di-Gymraeg> ("non-Welsh speech").

Enghreifftiau / Examples:

<S4> Rwy'n dy garu <lleferydd di-Gymraeg>.

2.6 Geiriau benthg / Loanwords

Mae “gair benthg” yn golygu gair sydd wedi’i fabwysiadu o un iaith a’i ymgorffori mewn iaith arall. Er enghraifft, mae’r geiriau *wel, jest/jyst, lot, off, so, reit, siampion/tsiampion* bellach yn ymddangos yng Ngeiriadur Prifysgol Cymru, felly **nid** oes angen defnyddio’r tagiau lleferydd di-Gymraeg yn yr achosion hyn. Yn hytrach, defnyddiwch sillafiadau Cymraeg y geiriau hyn.

*A “loanword” is a word which has been adopted from one language and incorporated into another. Many English words have now been incorporated into Welsh, for example, wel, jest/jyst, lot, off, so, reit, siampion/tsiampion. These now appear in the Geiriadur Prifysgol Cymru, so there is **no need** to put these words in non-Welsh tags. Instead, use the Welsh spellings of these words.*

Cymharwch / Compare:

<S1> *Es i yno ond <en> apparently </en> dyw e ddim yn byw ’na rhagor.*

Siaradwr yn cyfnewid cod i’r Saesneg (Adran 2.5)

Speaker code-switches to English (Section 2.5)

<S1> *So ro’n i’n meddwl wel mae hynny jyst yn grêt.*

Siaradwr yn defnyddio gair benthg - grêt - sydd bellach wedi’i ymgorffori i’r iaith Gymraeg

Speaker uses a loanword - grêt - which has already been incorporated into the Welsh language.

2.7 Geiriau benthg sydd wedi’u haddasu / Adapted Loanwords

Ar gyfer geiriau di-Gymraeg sydd wedi’u haddasu i system ramadeg y Gymraeg (e.e. drwy ychwanegu *-io* neu *-o* at ferf a/neu drwy dreiglo), dylid trawsgrifio’r rhain yn ôl sillafiadau Cymraeg a’u tagio yn ôl yr iaith fenthg. Yn ogystal, dylai’r tagiau nodi’r sillafiadau gwreiddiol.

For non-Welsh words adapted into the Welsh grammar system (e.g. by adding -io or -o to a borrowed verb and/or by mutating), these should be transcribed according to Welsh spellings and tagged according to loan language. In addition, the tags should note the original spellings.

Enghreifftiau / Examples:

<S2> *Dw i heb <en gair=“apply”> apleio </en> eto.*

2.8 Rhifolion / Numerals

Teipiwch rifau yn llawn fel geiriau.

Type out numbers in full as words.

Enghreifftiau / Examples:

cant ac ugain

un deg tri

deunaw

un deg wyth

2.9 Seiniau nad ydynt yn eiriol / Non-verbal sounds

Dylid sillafu'r canlynol yn gyson:

Use consistent spellings for the following:

ââ (“ <i>ahh</i> ”)	pan fo'r ceiniog yn syrthio / i ddynodi dealltwriaeth / ar ddarganfod rhywbeth <i>when the penny drops / to denote understanding / on discovering something</i>
ah (“ <i>aargh</i> ”)	i ddynodi dicter, poen, rhwystredigaeth <i>to denote anger, pain, frustration</i>
ahem	pan fo rhywun yn ceisio dal sylw rhywun arall <i>when someone is trying to get someone else's attention</i>
aw (“ <i>ow</i> ”) / awtsh (“ <i>ouch</i> ”)	pan fo rhywbeth yn rhoi loes/dolur <i>when something hurts</i>
doh	pan fo rhywbeth wedi mynd o chwith / pan fo rhywun wedi gwneud rhywbeth dwl / pan fo rhywun yn teimlo'n rhwystredig <i>when something has gone wrong / when someone has done something foolish / when someone is feeling frustrated</i>
dyh (“ <i>duh</i> ”)	pan fo rhywun wedi gwneud neu ddweud rhywbeth dwl <i>when someone has done something stupid</i>
eh?	pan fo rhywun heb ddeall neu'n cwestiynu rhywbeth <i>when someone doesn't understand or is questioning something</i>
err / erm / hmm / mmm	pan fo rhywun heb ddeall neu yn ystyried neu'n meddwl am rywbeth <i>when someone doesn't understand or is thinking about something</i>
grr	pan fo rhywun yn grac <i>when someone is angry</i>
haha	i ddynodi ffug-chwerthin <i>when someone is fake-laughing</i>
meh	i ddynodi difaterwch <i>to denote apathy</i>
m-hm / yh-hyh (“ <i>uh-huh</i> ”)	pan fo rhywun yn cytuno â rhywun arall neu'n dangos eu bod nhw'n deall neu wedi clywed <i>to show agreement or understanding</i>
mmm / iy (“ <i>yum</i> ”)	pan fo rhywbeth yn flasus <i>when something is tasty</i>
ôô (“ <i>aww</i> ”)	pan fo rhywbeth yn giwt / pan fo rhywun yn tosturio dros rywun neu rywbeth <i>when something is cute /when someone is pitying someone else</i>
shh or shwsh (“ <i>shush</i> ”)	pan fo rhywun yn ceisio tawelu rhywun <i>when someone is trying to silence someone</i>
whiw (“ <i>phew</i> ”)	i ddynodi rhyddhad <i>to denote relief</i>
wps (“ <i>oops</i> ”)	i ddynodi camsyniad <i>to denote a mistake</i>
ww (“ <i>ooh</i> ”)	pan fo rhywbeth yn ddiddorol neu'n annisgwyl / pan fo rhywbeth yn swnio'n ddeniadol

	<i>when something is interesting or unexpected / when something sounds appealing</i>
ych (“ugh”) / iww (“ew”)	i ddynodi ffeidd-dod <i>to denote disgust</i>
yh-oh (“uh-oh”)	pan fo rhywbeth wedi mynd yn anghywir / pan fo rhywun wedi gwneud rhywbeth dwl <i>when something goes wrong / when someone’s done something stupid</i>
yh-yh (“uh-uh”)	pan fo rhywun yn gwrthod gwneud rhywbeth <i>when someone’s refusing to do something</i>

3 Anonymeiddio / Anonymisation

3.1 Data Sensitif / Sensitive Data

Mae angen bod yn dra gofalus pan fo data yn sensitif. Gellir ystyried data’n sensitif pan:

- fo’n bosib adnabod y siaradwr neu unrhyw berson maent yn eu trafod wrth ddarllen y trawsgrifiad
- fo’n enllib neu’n athrod posibl
- allai beri embaras i rywun (e.e. datgelu gwybodaeth bersonol)
- allai achosi colled ariannol (e.e. datgelu manylion banc)
- allai beryglu enw da rhywun (e.e. trafodaeth o gamymddwyn yn y gwaith neu drafodaeth o weithgareddau anghyfreithlon)
- allai achosi niwed mewn unrhyw ffordd arall (e.e. datgelu cyfrinair)

Special care needs to be taken when data is sensitive. Data can be considered sensitive if it:

- is possible to identify the speaker or anybody they’re talking about while reading the transcript
- is a potential libel or slander
- could cause embarrassment (e.g. revealing personal information)
- could cause someone financial loss (e.g. divulging bank details)
- could put someone in a compromising position (e.g. discussion of misconduct at work or illegal activities)
- could cause someone harm in any other way (e.g. divulging passwords)

3.2 Egwyddor Gyffredinol / General Principle

Dylid cofnodi pob achos o ddata sensitif, gan nodi y credir bod y data yn sensitif. Cofnodwch bob achos o anonymeiddio yn nogfen ‘Templed.docx’ pan ydych yn dychwelyd gwaith. Yn y ffurflen hon, ceir lle i nodi:

- unrhyw enghreifftiau o **ddata sensitif**
- yr **enwau sydd wedi eu hanonymeiddio** ynghyd â’r **codau a ddefnyddiwyd** yn eu lle
- **unrhyw ymholiadau neu gwestiynau** sy’n codi wrth anonymeiddio
- bod yna enghreifftiau yn y trawsgrifiad o **iaith sy’n anaddas i blant** (e.e. iaith sarhaus, pynciau 18+)

Ar ôl cwblhau’r ffurflen hon, dylid ei chadw, gan enwi’r ffeil yn ôl enw ffeil y trawsgrifiad.

All instances of sensitive and/or personal data should be anonymised and logged. Record each case of anonymisation in the 'Templed.docx' document that you fill out and return with your completed transcription file. In this form, there are spaces to note:

- *any instances of **sensitive data***
- ***names anonymised** along with the **code used** to anonymise them*
- *any **comments or questions** that arise during anonymisation*
- *that the transcription includes instances of **language which is unsuitable for children** (e.g. offensive language, 'adult' topics).*

Having completed this form, it should be saved using the same name as the transcription file name.

3.3 Enwau Pobl / People's Names

Oni bai bod caniatâd penodol wedi ei roi gan berson, dylid anonymeiddio enwau fel a ganlyn:
Unless specific consent is given by the person, names should be anonymised as follows:

3.3.1 Enwau Cyntaf / First Names

<code><anon> enwg1 </anon></code>	ar gyfer yr enw gwrywaidd cyntaf / <i>for the first male name</i>
<code><anon> enwb1 </anon></code>	ar gyfer yr enw benywaidd cyntaf / <i>for the first female name</i>
<code><anon> enwg2 </anon></code>	ar gyfer yr ail enw gwrywaidd / <i>for the second male name</i>
<code><anon> enwb2 </anon></code>	ar gyfer yr ail enw benywaidd, ac ati / <i>for the second female name, and so on</i>

Pan fydd enw sy'n gallu perthyn i ddyn neu fenyw yn cael ei ddefnyddio (e.e. Charlie, Bobi, Pat), a phan nad oes modd gwybod p'un ai ddyn neu fenyw yw'r person, dylid defnyddio `<anon> enwb1 </anon>`, `<anon> enwb2 </anon>` ac yn y blaen.

When a name which could belong to a man or a woman is used (e.g. Charlie, Bobbie, Pat), and when there is no way of knowing whether the person is male or female, use `<anon> enwb1 </anon>`, `<anon> enwb2 </anon>` and so on.

PWYSIG / IMPORTANT:

Dylid defnyddio yr un cod ar gyfer yr un enw pob tro (o fewn y trawsgrifiad cyfredol – dechreuwch restr newydd ar gyfer pob trawsgrifiad newydd).

Use the same code for the same name each time (within that specific transcription – start a new list for each new transcription you undertake).

Data gwreiddiol / Original data:

`<S1>` Aeth Angharad i'r dafarn gyda Paul wythnos ddiwethaf.

`<S2>` Gyda Paul?

`<S1>` Ie, gyda Paul.

Data wedi ei anonymeiddio / Anonymised data:

`<S1>` Aeth `<anon> enwb1 </anon>` i'r dafarn gyda `<anon> enwg1 </anon>` wythnos ddiwethaf.

`<S2>` Gyda `<anon> enwg1 </anon>`?

<S1> Ie, gyda <anon> enwgl </anon>.

3.3.2 Cyfenwau / Surnames

Dylid anonymeiddio cyfenwau fel a ganlyn:

Anonymise surnames as follows:

<anon> cyfenw1 </anon>

<anon> cyfenw2 </anon>, ac yn y blaen / and so on

Data gwreiddiol / Original data:

<S1> Aeth Angharad i'r dafarn gyda Paul wythnos ddiwethaf.

<S2> Gyda Paul Williams?

<S1> Nage, gyda Paul Evans.

Data wedi ei anonymeiddio / Anonymised data:

<S1> Aeth <anon> enwb1 </anon> i'r dafarn gyda <anon> enwgl </anon> wythnos ddiwethaf.

<S2> Gyda <anon> enwgl </anon> <anon> cyfenw1 </anon>?

<S1> Nage, gyda <anon> enwgl </anon> <anon> cyfenw2 </anon>.

3.3.3 Ffigurau Cyhoeddus / Public Figures

Pan fydd siaradwyr yn trafod ffigurau cyhoeddus neu enwogion, ni ddylid anonymeiddio enwau'r ffigurau cyhoeddus. Pan fydd y siaradwr ei hun yn ffigwr cyhoeddus sy'n cael sgwrs breifat, dylid anonymeiddio enw'r siaradwr yn unol ag adran **ENWAU POBL**.

*When speakers discuss public figures or celebrities, the names of the public figures should not be anonymised. If the speaker is a public figure having a private conversation, names should be anonymised as per **PEOPLE'S NAMES** section.*

3.4 Data eraill / Other data

Anonymeiddiwrch unrhyw ddata all roi modd o adnabod y siaradwyr neu'r bobl a drafodir gan y siaradwyr. Os nad oes risg o hyn, h.y. mae'r data yn gyffredinol, does dim angen anonymeiddio felly. Os nad ydych yn siŵr, anonymeiddiwrch y data a gwnewch nodyn yn y ddogfen 'Templed' i esbonio eich rhesymau.

Anonymise any data that could potentially identify a speaker or individuals discussed. If there is no risk of this, i.e. the data is general, then there is no need to anonymise. If unsure, anonymise the data and make a note in the Templd document as to your reasons.

	Data gwreiddiol / Original data:	Wedi'i anonymeiddio / Anonymised:
Dyddiadau / Dates	Ebrill 12 fed 1957; y 15fed	<anon> dyddiad </anon>
Rhifau Ffôn / Phone numbers	02920 217751	<anon> Rhif ffôn </anon>

Cyfeiriadau e-bost ac “IPs” / Emails and IPs	<i>corcenc@caerdydd.ac.uk; 80.36.158.1</i>	<i><anon> cyfeiriad e-bost </anon>; <anon> cyfeiriad IP </anon></i>
Gweithleoedd a sefydliadau / Workplaces and institutions	<i>Siop Bapur Archie’s; Tesco</i>	<i><anon> enw sefydliad </anon></i>
Gwefannau / Websites	<i>www.corcenc.org</i>	<i><anon> cyfeiriad gwefan </anon></i>
Teitlau swyddi neu alwedigaethau / Job titles or occupations	<i>Swyddog hyrwyddo’r Gymraeg</i>	<i><anon> teitl swydd </anon></i>
Mannau cyhoeddus, adeiladau, neu dirnodau / Public places, buildings, or landmarks	<i>Prifysgol Caerdydd Eglwys y Santes Fair</i>	<i><anon> prifysgol </anon> <anon> eglwys </anon></i>
Cyfeiriadau / Addresses	<i>101 Cammock Drive, Waunarlwydd, Abertawe, SA4 3DL</i>	<i><anon> cyfeiriad </anon></i>
Lleoliadau daearyddol / Geographical locations	<i>Waunarlwydd; Llanelli</i>	<i><anon> lleoliad </anon>; <anon> lleoliad </anon></i>

4 Atodiad / Appendix

4.1 Rhestr Synau / List of Sounds

Dyma restr drwyadl o synau. Ni ddylid defnyddio cod gwahanol neu ddyfeisio codau newydd.

This is an exhaustive list of sounds. Do not use alternative tags or try to create your own.

<i><anadlu i mewn yn sydyn> (“sharp intake of breath”)</i>	<i><hisian> (“hiss”)</i>
<i><clecian gwefusau> (“lip smacking”)</i>	<i><hymian> (“humming”)</i>
<i><clirio gwddf> (“clear throat”)</i>	<i><iapian> (“yelp”)</i>
<i><cnoi> (“munch”)</i>	<i><igian> (“hiccup”)</i>
<i><usanu> (“kiss”)</i>	<i><llefain> (“cry”)</i>
<i><chwerthin> (“laugh”)</i>	<i><ocheneidio> (“sigh”)</i>
<i><chwibanu> (“whistle”)</i>	<i><pesychu> (“cough”)</i>
<i><chwyrnu> (“snore”)</i>	<i><sgrechian> (“scream”)</i>
<i><chwythu allan yn sydyn> (“sharp exhalation”)</i>	<i><slochian> (“slurp”)</i>
<i><chwythu trwyn> (“blow nose”)</i>	<i><sniffian> (“sniff”)</i>
<i><dyheu> (“pant”)</i>	<i><swnian> (“whine”)</i>
<i><dylyfu gêm> (“yawn”)</i>	<i><tagu> (“choke”)</i>
<i><ebychu> (“gasp”)</i>	<i><tisian> (“sneeze”)</i>
<i><giglan> (“giggle”)</i>	<i><torri gwynt> (“break wind, e.g. fart/belch”)</i>
<i><griddfan> (“groan”)</i>	<i><traflyncu> (“gulp”)</i>
<i><gwichian> (“wheeze”)</i>	<i><wffio> (“raspberry”)</i>

