

## POS tagset for ChronHib

These tags are for the ‘Part\_Of\_Speech’ and ‘Classification’ fields.

POS	Classification	examples
adjective	(stem)	
adjective_numeral		2, 3, 4
adjective_quantifier		unstressed cach, nech, etc. and uile
adjective_pronominal		aile, alaile, ala, nach 1
adverb		
complementiser		co 3, co 4, ara, día 1
conjunction		
definite_article		
interjection		
noun	(stem)	
noun_numeral	(stem)	20, 30, etc.
numeral		
numeral_adjective_noun		óen
particle_augment		con 6
particle_anaphoric		
particle_comparative		oldaas
particle_demonstrative_proximate		
particle_demonstrative_distal		
particle_discourse		
particle_focus		acht 1
particle_interrogative		
particle_negative_main		
particle_negative_subordinate		ná 2
particle_numerative		a
particle_prefix		
particle_preverb		
particle_pronominal		nota augens
particle_relative		
particle_vocative		a
preposition	(cases governed)	
pronoun_anaphoric		side
pronoun_demonstrative_distal		
pronoun_demonstrative_proximate		
pronoun_indefinite		sechi
pronoun_independent		
pronoun_infixed		
pronoun_interrogative		
pronoun_possessive		
pronoun_propword		í ‘that one’
pronoun_quantifier		cách, ní, etc.
pronoun_emphatic		
pronoun_relative		
pronoun_suffixed		
proper_noun		
verb	(stem class)	
verbal_noun		

verbal_of_necessity		
verbal_participle		

proper\_noun:

at the moment only names of persons and places, languages and spiritual entities.

verbal\_of\_necessity

including those lemmata that do not have an underlying verb.

### Classification:

multiple values are separated by '/', e.g. *eclais* 'ā/i/ī'

dubious value is followed by '?'

nominal stems:

o	ā	o/ā	io	iā	io/iā	i	ī	u	k	g	t
	d	nt	n	r	s	irreg.	indecl.	unk.			

the tag irreg. is for nominal whose paradigm does not follow the regular patterns of the other nominal stems in Old Irish, e.g. *bó* 'cow'.

the tag indecl. is for nominal which is indeclinable, e.g. *Ísu* 'Jesus'

unknown declension is marked by unk.

verbal stem is named after McCone's system:

W1	W2a	W2b	S1a	S1b	S1c	S2	S3a	S3b	H1	H2	H3
----	-----	-----	-----	-----	-----	----	-----	-----	----	----	----

plus: copula, defective

Cases govern by preposition:

acc. dat. gen. acc./dat.

### Gender:

masc. neut. fem.

uncertain genders: in the order of the list above, e.g. masc./neut./fem., whichever genders are possible.

### Language

When the morph is not Irish, fill this field with the following values:

Latin	Greek	Hebrew	British	Pictish
	Old_English	Old_Norse	Old_High_German	Old_Breton
	Old_Welsh	Frankish		

Discuss with the team if other tags are to be added.

### Etymology:

???Pre-Irish etymology (DS please define)? e.g. \*to-ber-; \*kejl̥io/ā-

### Onomastic Complex

This field is for marking the whole string that makes up a personal name (excluding genealogical information), place name, epithet, hydronym or ethnonym, in case that the complex is not visible under an individual token, e.g. under the tokens with morph *glenn*, *dá* and *locha*, one marks in each token “*Glenn dá Locha*”, to which the tokens belong; usw. *Cú cen Máthair*; *Uí Fidgenti*; *Glúin Duib*; *Mac an Bhaird* etc. Leave blank if there is no onomastic complex but a single word, e.g. *Breg*; *Díarmait*; *Lífe*; *Dergfholt*

### Onomastic Usage:

When a Onomastic Complex field is populated, choose from the following tags:

anthroponym    surname                  epithet                  ethnonym                  hydronym                  toponym

### Morphological Tagset for Chronhib annotation

this is the information in the ‘Analysis’ field in the newest database template.

tags should be concatenated without blank space but with dashes, e.g. *nom.sg.neut.*; *augm.3sg.pres.ind.pass.rel.*

#### 1. noun, verbal noun, numerals 2.4:

attribute	(for numerals)	case	number
value	stress.	nom.	sg.
		gen.	du.
		dat.	pl.
		acc.	
		voc.	
		compos.	
		abl.	
		loc.	

uncertain case/number/genders: in the order of the list above, e.g. *masc./neut./fem.*, whichever genders are possible.

‘compos.’ are for the compositional forms in a nominal composition, e.g. *ríg* in *rígdonnae*

abl. and loc. tags are saved for Latinised proper nouns governed by Latin prepositions.

#### 2. adjective, participle, definite article:

attribute	degree	case	number	gender
value	Ø (positive)	nom.	sg.	masc.
		gen.	du.	neut.
		dat.	pl.	fem.
		acc.		
	equ.			

	comp.
	sup.

### 3. verbs:

cf. Stifter in *The Celtic Languages* (2<sup>nd</sup> ed.), pp. 85.88.

attribute	augment (morph.)	pers. & num.	tense & mood	voice	relativity (morph.)
values	augm./Ø	1sg.	pres.ind.	pass./Ø	rel./Ø
		2sg.	hab.		
		3sg.	impf.		
		1pl.	pret.		
		2pl.	fut.		
		3pl.	pres.subj.		
			past.subj.		
			impv.		
			cond.		

All augmented preterite verbal forms (or ‘perfect’) should be tagged with the ‘augm.’ tag, including forms that are not augmented by augment particles such as *do-ucc*. For forms that are augmented by augment particles, such as *do-cuaid* or *do-rigni*, the ‘augment’ field in the ‘morphology’ table should be filled as ‘Yes’.

### 4. preposition

put in the case that it governs in the context: acc., dat., gen.

conjugated preposition:

attribute	person	case
value	1sg.	gen.
	2sg.	dat.
	3sg.masc./neut.	acc.
	3sg.fem.	
	1pl.	
	2pl.	
	3pl.	

e.g. *linn* 1pl.acc.; *for* 3sg.masc./neut.dat.

### 5. preverb:

the etymology of the whole compound verb, e.g. \*to-are-ber-, is put in the ‘Comment’ column.

## 6. pronoun:

( 'headwords' already contain information such as: 3sg.masc.indp.pron., 1pl.inf.pron., etc., since they are etymologically distinct) where applicable:

attribute	class	case	gender (pronoun_propword <i>í</i> )
value	A	nom.	masc.
	B	acc.	neut.
	C	gen.	fem.
	emph.	dat.	

stressed forms *aí*, *moí*, *toí* have the lemmata '3sg.neut.indp.pron.', '1sg.indp.pron.' etc, and they are tagged as gen .

## 7. adverb:

pure adverbs do not have morphological tags, e.g. *beus* 'yet, still', *sís* 'below'.

for adverbs derived from adjectives, the morph's POS is still 'adjective' and they probably have the following tags in Analysis:

dat.sg.neut., acc.sg.neut.

for the *-aid*, *-aith* adverbial forms of *-de* adjectives in the Glosses, they have been given separate lemmata, e.g. *immfólngidid* 'effectively', POS = adverb, analysis = NULL

## 8. verbal boxes:

1) deponentiality: Ø = non.deponent verb, Yes = deponent ending, No = active ending for verb otherwise attested as deponent.

2) dependency: absolute, conjunct, deutorotonic, prototonic.

3) transitivity: transitive, intransitive, passive, unclear

## 9. Mutation

tags: + Nasalization; - Nasalization; + Lenition; - Lenition; + Gemination; - Gemination.

These tags are given in the following contexts:

1) when the initial consonant of a morph is mutated, or *n-* or *h-* is added before an initial vowel of a morph, and these mutations are shown in the spelling of the morph, the + tags are given in this field.

e.g. *chíall* '+ Lenition'; *mmolad* '+ Nasalization' *h-ech* '+ Gemination'

note: in *indfír* 'the men', *fír* '+ Lenition', because the lenited *f* is indicated by the *-d* in the def.art.

note: prepositions prefixed with *h* are not treated as being affected by germination, e.g. *hi* 'in', *hiter* 'between'.

2) when morpho-syntactic rules demand that the initial consonant or vowel of a morph be mutated (or added *n-* or *h-*), and standardised Old Irish orthography (as adopted for the Lemmata) allows these mutations to be indicated, but such mutations are not shown in the spelling of the morph, the - tags are given.

Mutations allowed by the standardised orthography include:

nasalization: *n-V*, *mb*, *nd*, *ng*, *ll*, *rr*

germination: *hV*, *mm*, *nn*, *pp*, *bb*, *ss*, *ll*, *gg*, *rr*

lenition: *ch*, *ph*, *th*.

e.g. *la amrae* ‘with a miracle’, *amrae* ‘- Gemination’

*no·guidim* ‘I beseech him’, *·guidim* ‘- Nasalization’

*a carpat* ‘his chariot’, *carpat* ‘- Lenition’

*anro·chésasat* ‘that which they had suffered’, *ro·* ‘- Nasalization’

3) when mutations are indicated in a way that the standardised Old Irish orthography does not allow, e.g. in later manuscripts or with idiosyncratic spelling, the + tags are used, but it should be indicated in the ‘comment’ that this is not the standardised Old Irish spelling. e.g.

*dí gheinntib* ‘by the heathens’, *gheinntib* ‘+ Lenition’, comment: ‘non-standard spelling’

*a śarugud* ‘violating him’, *śarugud* ‘+ Lenition’, comment: ‘non-standard spelling’

*i bpein* ‘in pain’, *bpein* ‘+ Nasalization’, comment: ‘non-standard spelling’

*con·dánicc* ‘until he came’, *·dánicc* ‘+ Nasalization’, comment: ‘non-standard spelling’

4) When mutations are indicated in the morph which are against the morpho-syntactic rules, they are marked with the + tags, but it should be indicated in the ‘comment’ that these mutations are not grammatical/expected. e.g.

*co chéill* ‘with sense’, *chéill* ‘+ Lenition’, comment: ‘unexpected lenition’