

## Handout: extended tagset for Old Irish

The following Old Irish inflected forms, based on a 'test set' of five verbs employed in the research, partly constitute the model for a proposed (preliminary) extended PAROLE<sup>1</sup> tagset for Old Irish verbs (see Table next page). Examples 4(c) and 4(d) are Middle and Modern Irish cognates of 4(b),<sup>2</sup> respectively; 5(b) is a Modern Irish cognate of 5(a). These later forms serve to illustrate morphological simplifications from the viewpoint of Old Irish, reflected in the assigned tags (see Table next page). On p. 3 the PROIEL tagset is given for comparison.

### 1. *as-rubart*

<b>as</b>	<b>ru</b>	<b>bart</b>
lexical preverb	= <i>ro</i> , augment <sup>3</sup>	carry_pret.3sg
'(s)he has said'		

### 2. *brisfid*

<b>bris.fid</b>
break.future3sg
'(s)he will break'

### 3. *(ní) imthiget*

<b>(ní)</b>	<b>im</b>	<b>thig.et</b>
(negative conjunct particle)	lexical preverb	go away_depend_presInd.3pl
'they do not go away'		

### 4. *in-fid | in-cuaid<sup>4</sup> -> ro indis -> d'in(n)is sé/sí*

(a)	in-		fid	
	lexical preverb		tell_preterite3sg	
	'(s)he told'			
(b)	in-	cu	(a)id	
	lexical preverb	= com, augment	tell_preterite3sg	
	'(s)he has told'			
(c)	ro		indis	
	augment ( > perfect verbal particle)		tell_preterite3sg	
	'(s)he has told'			
(d)	do		in(n)is	sé/sí
	past verbal particle		tell_past	persPron3sg_m/f
	'he/she told'			

### 5. *hore no-n-dob-molor-sa<sup>5</sup> -> óir molaim(-)se sibh*

(a)	<b>hore</b>	<b>no</b>	<b>n</b>	<b>dob</b>	<b>mol.or</b>	<b>sa</b>	
	for (conj.)	preverbal particle	relative marker	obj. pronoun 2pl	praise.presInd1sg	emphasising Pron1sg	
(b)	<b>óir</b>				<b>mol.aim</b>	<b>se</b>	<b>sibh</b>
	for (conj.)				praise.presInd1sg	emphasising Pron1sg	persPron2pl
	'for / praise you (pl.)'						

<sup>1</sup> LE-PAROLE stands for Language Engineering - Preparatory Action for Linguistic Resources Organisation for Language Engineering. The goal of this project was to create and harmonize text corpora and lexica for 14 European languages. These standardisation efforts included a common PAROLE tagset and morpho-syntactic annotation standards (Ruimy N., Corazzari O., Gola E., Spanu A., Calzolari N., Zampolli A. (1998) The European LE-PAROLE Project: The Italian Syntactic Lexicon. In Proceedings of the First International Conference on Language Resources and Evaluation, Granada. Vol. I. pp. 241-248.)

<sup>2</sup> The distinction preterite/perfect is lost in the (later) Modern Irish period, with the preterite becoming obsolete in the spoken language by the beginning of the 17th century (Williams 1994). Contemporary Irish past tense forms derive from perfect forms (which generally derive from augmented preterite formations). Furthermore, the modern verb form is analytic, i.e. we need a subject pronoun.

<sup>3</sup> *ro* creates an augmented preterite and functions as a perfective marker. The augment may also signify potentiality of the action communicated by the verb. Another example of an augment occurs in 4(b), i.e. *cu* (*com*). A discussion of augmentation in Early Irish can be found in McCone (1997).

<sup>4</sup> The underlying structure being *ind(e)-cum-féd*, with loss of *-mf-* (McCone 1997: 138).

<sup>5</sup> Glossing Latin *quia laudo vos*, Wb. 14<sup>c</sup>18 (Kavanagh and Wodtko 2001). The Old Irish temporal conjunction (*h*)*óre* 'for' enforces a relative form of the verb. For a discussion of relative constructions in Old Irish cf. Thurneysen (1946: 314-9).

Proposed PAROLE Morphosyntactic Tagset (verbs) for Old Irish, with additions to the modern tagset right of the dividing line (i.e. 10.-16.) as well as within 4. Tense ( $t$  = *preterite*). Note that the order of elements in the tag does not reflect the order of morphemes in the inflected word (the morphological features used during morphological analysis<sup>6</sup> are mapped to the Parole tagset<sup>7</sup> in a final step). Headwords for Old and Middle Irish forms are taken from the *electronic Dictionary of the Irish Language* while modern (standard) headwords are supplied from *Ó Dónaill (1977)*. The preverbal past particle ( $d'$  < *do*) and the subject pronoun (*se/si*) are not part of the Verb tag class in Modern Irish forms and are ignored in example 4(d).

2. VERB																		
Example	Head-word	1. POS	2. Type	3. Mood	4. Tense	5. Person	6. Number	7. Gender	*8. Dependency	*9. Contrast	10. Infl. type (1)	11. Infl. type (2)	12. Infl. type (3)	13. Pre-verb	14. Augment	15. Obj. pron. pers.	16. Obj. pron. numb.	Tag
		V	m = main	i = indic s = subj. m = imp. c = cond.	p = pres. s = past h = past hab. f = future g = future pres. hab. t = preterite	1 = first 2 = sec. 3 = third 0 = auto	s = sing. p = pl.	n/a	d = dependent r = relative n = negative	e = emphatic	a = active d = dependent	w = weak s = strong h = hiatus	s = simplex c = compound	l = lexical n = 'no'	r = resulative p = potential	1 = first 2 = sec. 3 = third	s = sing. p = pl.	
1. as-rubart	as-beir	V	m	i	t (bar)	3 (t)	s (t)	-	-	-	a	s	c	l (as)	r (ru)	-	-	Vmit3s---asclr---
2. brisfid	bris(s)id	V	m	i (bris)	f (f)	3 (id)	s (id)	-	-	-	a	w	s	-	-	-	-	Vmif3s---aws----
3. (n)l inthiget	imm-téit	V	m	i	p (tig)	3 (et)	p (et)	-	n	-	a	s	c	l (im)	-	-	-	Vmip3p-n-ascl---
4. (a) in-fid	in(d)-fét	V	m	i	t	3 (fid)	s (fid)	-	-	-	a	s	c	l (in)	-	-	-	Vmit3s---ascl---
4. (b) in-cuaid	in(d)-fét	V	m	i	t	3 (id)	s (id)	-	-	-	a	s	c	l (in)	r (cu)	-	-	Vmit3s---asclr---
4. (c) ro indis	in(d)isid	V	m	i (indis)	t (indis)	3 (indis)	s (indis)	-	-	-	a	w	s	-	r (ro)	-	-	Vmit3s---aws-r---
4. (d) (d') in(n)is	inis	V	m	i (inis)	s (inis)	-	-	-	-	-	-	-	-	-	-	-	-	Vmis-----
5. (a) no-n-dob-mol-sa	molaíd	V	m	i (mol)	p (mol)	1 (or)	s (or)	-	r (n)	e (sa)	d	w	s	n (no)	-	2 (dob)	p (dob)	Vmip1s-redwsn-2p
5. (b) molaim(-)se	mol	V	m	i (mol)	p (mol)	1 (aim)	s (aim)	-	-	e (se)	-	-	-	-	-	-	-	Vmip1s-e-----

<sup>6</sup> <https://www.scss.tcd.ie/SLP/morphtag.htm>

<sup>7</sup> <https://www.scss.tcd.ie/SLP/parole.htm>

# Eckhoff, Hanne, Kristin Bech, Gerlof Bouma, Kristine Eide, Dag Haug, Odd Einar Haugen, and Marius Jøhndal. "The PROIEL treebank family: a standard for early attestations of Indo-European languages." Language Resources and Evaluation (2017): 1-37.

Table 3 Morphological tags

1. Person	1. 2. 3. x (uncertain)
2. Number	s (singular), d (dual), p (plural), x (uncertain number)
3. Tense	p (present), i (imperfect), r (perfect), s (resultative, i.e. 1-form), a (aorist), u (past), l (pluperfect), f (future), t (future perfect), x (uncertain tense)
4. Mood (combined mood and finiteness)	i (indicative), s (subjunctive), m (imperative), o (optative), n (infinitive), p (participle), d (gerund), g (gerundive), u (supine), e (indicative or subjunctive), f (indicative or imperative), h (subjunctive or imperative), x (uncertain mood)
5. Voice	a (active), m (middle), p (passive), e (middle or passive)
6. Gender	m (masculine), f (feminine), n (neuter), p (masculine or feminine), o (masculine or neuter), r (feminine or neuter), q (masculine, feminine or neuter), x (uncertain gender)
7. Case	n (nominative), a (accusative), o (oblique), g (genitive), c (genitive or dative), d (dative), b (ablative), i (instrumental), l (locative), v (vocative), e (accusative or dative), x (uncertain case), z (no case)
8. Degree	p (positive), c (comparative), s (superlative), x (uncertain degree)
9. Strength <sup>a</sup>	w (weak, i.e. long form), s (strong, i.e. short form), t (weak or strong)
10. Inflection	n (non-inflecting), i (inflecting)