

On pseudo-non-finite clauses in Welsh

Olivier Bonami Robert D. Borsley Maggie Tallerman
Université Paris Diderot University of Essex Newcastle University

1 Introduction

Languages differ in how they employ finite and non-finite clauses. Welsh finite and non-finite clauses have a similar distribution to their counterparts in English. However, it doesn't look like this because Welsh has certain finite clauses which look rather like non-finite clauses (Tallerman, 1998; Borsley et al., 2007). One type just involves the verb *bod* 'be'. The other involves all verbs.

2 Data

Standard Welsh verbs have three synthetic paradigms: future, past, and conditional. *Bod* 'be' has two additional subparadigms, present and imperfect, illustrated with one type of complement in (1). These forms are fine in main clauses and clauses involving an unbounded dependency such as relative clauses (2), but present forms (3) and for some speakers imperfect forms too (4) are ungrammatical in complement clauses:

- (1) Mae / Roedd Elen yn darllen y llyfr.
be.PRS.3SG be.IMPF.3SG Elen PROG read the book
'Elen is reading the book.'
- (2) y llyfr [mae / roedd Elen yn ei darllen]
the book be.PRS.3SG be.IMPF.3SG Elen PROG 3SG.M read
'the book that Elen is/was reading'
- (3) *Mae Aled yn credu [mae Elen yn darllen y llyfr].
be.PRS.3SG Aled PROG believe be.PRS.3SG Elen PROG read the book
'Aled believes that Elen is reading the book.'
- (4) %Mae Aled yn credu [roedd Elen yn darllen y llyfr].
be.PRS.3SG Aled PROG believe be.IMPF.3SG Elen PROG read the book
'Aled believes that Elen was reading the book.'

Instead of present forms of *bod* and for some speakers imperfect forms as well, what looks like the non-finite form *bod* appears:

- (5) Mae Aled yn credu [bod Elen yn darllen y llyfr].
be.PRS.3SG Aled PROG believe be Elen PROG read the book
'Aled believes that Elen is/was reading the book.'

We will call the complement clause here a *bod*-clause. Table 1a summarizes the distribution.

Past tense forms of Welsh verbs appear in main clauses and in negative complement clauses, but in literary varieties they are unacceptable in positive complement clauses:

- (6) Aeth Mair adre.
go.PAST.3SG Mair home
'Mair went home.'
- (7) Meddylodd Aled [aeth Mair ddim adre].
think.PAST.3SG Aled go.PAST.3SG Mair NEG home
'Aled thought that Mair had not gone home.'
- (8) %Meddylodd Aled [aeth Mair adre].
think.PAST.3SG Aled go.PAST.3SG Mair home
'Aled thought that Mair had gone home.'

Instead of a positive complement clause with a past tense verb, what looks rather like an English *for-to* clause appears:

- (9) Meddylodd Aled [i Mair fynd adre].
think.PAST.3SG Aled to Mair go home
'Aled thought that Mair had gone home.'

There is evidence of various kinds that these clauses are finite, and we will call them finite *i*-clauses, but elsewhere we have *i*-clauses that are clearly non-finite e.g. the following:

- (10) Disgwylodd Aled [i Elen ddarllen y llyfr].
 expect.PAST.3SG Aled to Elen read the book
 ‘Aled expected Elen to read the book.’

Both *bod*-clauses and finite *i*-clauses express finite meanings and appear instead of certain finite clauses which are unexpectedly ungrammatical. They can also coordinate with ordinary finite clauses, and like finite clauses but unlike uncontroversial non-finite clauses they do not allow a reflexive in subject position with an antecedent in the superordinate clause. *Bod*-clauses also have the internal structure of finite clauses. They show verb-subject order, which is found in finite clauses, and not subject-verb order, which occurs in uncontroversial non-finite clauses such as that in (10). They are also like finite clauses and unlike non-finite clauses in showing negation with the post-subject negative adverb *ddim* (11) and allowing the expletive subject *yna* ‘there’ (12). The only two properties that make *bod* similar to a nonfinite verb are its phonological form and the fact that agreement is expressed through a clitic preceding the verb rather than as suffix (13).

- (11) Mae Aled yn dweud [bod Mair ddim yn barod].
 be.PRS.3SG Aled PROG say be Mair NEG PRED ready
 ‘Aled says that Mair isn’t ready.’
- (12) Mae Gwyn yn meddwl [bod yna ddafad yn yr ardd]. yr ardd
 be.PRS.3SG Gwyn PROG think be there sheep in the garden
 ‘Gwyn thinks that there is a sheep in the garden.’
- (13) Mae Aled yn dweud [ei bod hi yn barod].
 be.PRS.3SG Aled PROG say 3SG.F be 3SG.F PRED ready
 ‘Aled says that she is ready.’

In contrast, finite *i*-clauses have essentially the same internal structure as non-finite *i*-clauses. They differ from non-finite *i*-clauses only in that they cannot be negated.

3 Towards an analysis

In this paper we attempt to account for two key properties of *bod* clauses and finite *i*-clauses: the fact that both appear to fill a gap in a paradigm of otherwise finite constructions, as illustrated in Tables 1a and 1b; and the fact that both contain what looks like a non-finite verb. In addition, we aim to account for the fact that *bod*-clauses have the internal structure of finite clauses while finite *i*-clauses have the same internal structure as non-finite *i*-clauses.

Our account relies on two mechanisms familiar from recent work in realisational morphology. First, the statement of the morphology-syntax interface necessitates a distinction between two sets of morphosyntactic features (Sadler and Spencer, 2001; Stump, 2006; Bonami, 2015) corresponding to the syntactic and the morphological view of the inflectional paradigm. Here we locate these concurrent morphosyntactic feature bundles respectively inside HEAD and INFL. INFL is a feature carried only by words (not by phrases) and serves as the input to inflectional morphology (Bonami, 2015). Second, the morphology-syntax interface is governed by an information-based principle of morphosyntactic blocking (Andrews, 1990; Koenig, 1999). Specifically, we assume with Stump (2006, 2015) that a paradigm is licensed by a set of conditional statements such that if the antecedents of two statements stand in an asymmetric entailment relation, only the more specific statement may apply. This may be implemented in HPSG using the general strategy advocated by Bonami and Crysmann (2013) of assuming a compilation step where the antecedent of each statement is strengthened with the conjunction of the negations of all more specific statements.

In the case at hand, we assume the set of interface statements in (14). (14a) states the canonical situation, where HEAD and INFL values are constrained to be identical. Finite *bod* in the present is licensed by the statement in (14b), which allows a mismatch between *bod*’s HEAD|VFORM and INFL|VFORM values.¹ The feature STATUS is a replacement for ROOT allowing appropriately for a more fine-grained distinction between complement ([STATUS *subord*]) and relative and wh-interrogative ([STATUS *udc*]) clauses. Independent constraints ensure that only nouns and [INFL|VFORM *inf*] verbs may take an agreement clitic as a specifier. Figure 1 illustrates the analysis.²

The analysis of *i*-clauses depends on setting up appropriate lexical entries for *i*. We assume that finite and non-finite *i* are both instances of the lexical type *i-clausal* in (15a): complementizers raising the subject of their infinitival complements, and realized by what is morphologically a preposition, as per the [INFL|LID *i*] specification. This ensures that finite *i*, nonfinite *i* and the preposition

¹Varieties where *bod* is also used in the imperfect use a slightly different statement.

²We follow Borsley (1989) in assuming that the subject of a Welsh finite verb is the first member of the COMPS list.

i share the same irregular morphological paradigm, where the 1SG, 1PL, 2SG and 2PL cells are unexpectedly syncretic with the absolute form of the preposition (compare the paradigms in Table 1). Nonfinite *i* is then simply subject to the further constraint that its VFORM is nonfinite (15b). Since nothing is stated to the contrary, clause (14a) applies, and nonfinite *i* is morphosyntactically canonical. Figure 3 illustrates the analysis.

The analysis of finite *i* involves both the specialized lexical entry in (15c) and the interface statement in (14c). The lexical entry ensures that finite *i* inherits the HEAD|LID value of its complement. Since the complement is a verb and no verb carries an [INFL|LID *i*] specification, (15c) can only license noncanonical words that are exceptions to (14a) and must be licensed by more specific interface statements. This is the role of (14c), which states that the finite subordinate positive past must be realized by a complementizer specified as [INFL|LID *i*]*i*—a constraint satisfied by finite *i* and no other word. Figure 3 illustrates the analysis.

(14) **Syntax-morphology interface statements**

$$\begin{array}{l}
 \text{a. } [] \rightarrow \left[\begin{array}{l} \text{HEAD} \quad [1] \\ \text{INFL} \quad [1] \end{array} \right] \\
 \\
 \text{b. } \left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l} \text{LID} \quad \textit{bod} \\ \text{VFORM} \quad \textit{fin} \\ \text{STATUS} \quad \textit{subord} \\ \text{TMA} \quad \textit{pres} \end{array} \right] \\ \text{SLASH} \quad \{ \} \end{array} \right] \rightarrow \left[\text{INFL} \quad \left[\begin{array}{l} \text{LID} \quad \textit{bod} \\ \text{VFORM} \quad \textit{inf} \end{array} \right] \right] \\
 \\
 \text{c. } \left[\begin{array}{l} \text{HEAD} \quad \left[\begin{array}{l} \text{VFORM} \quad \textit{fin} \\ \text{STATUS} \quad \textit{subord} \\ \text{TMA} \quad \textit{past} \\ \text{POL} \quad \textit{pos} \end{array} \right] \right] \rightarrow \left[\begin{array}{l} \text{HEAD} \quad \left[\textit{complementizer} \right] \\ \text{AGR} \quad [1] \\ \text{INFL} \quad \left[\begin{array}{l} \text{LID} \quad \textit{i} \\ \text{AGR} \quad [1] \end{array} \right] \end{array} \right]
 \end{array}$$

(15) **Lexical entries**

$$\begin{array}{l}
 \text{a. } \textit{i-clausal} \rightarrow \left[\begin{array}{l} \text{HEAD} \quad \textit{complementizer} \\ \text{INFL} \quad \left[\text{LID} \quad \textit{i} \right] \\ \text{ARG-ST} \quad \left\langle [1], \left[\begin{array}{l} \text{HEAD} \quad \left[\textit{verb} \right] \\ \text{VFORM} \quad \textit{inf} \end{array} \right] \right\rangle \right. \\ \left. \left[\text{SUBJ} \quad \langle [1] \rangle \right] \right] \\
 \\
 \text{b. } \textit{i-inf} \rightarrow \textit{i-clausal} \wedge \left[\text{VFORM} \quad \textit{inf} \right] \\
 \\
 \text{c. } \textit{i-fin} \rightarrow \textit{i-clausal} \wedge \left[\begin{array}{l} \text{HEAD} \quad \left[\text{LID} \quad [2] \right] \\ \text{ARG-ST} \quad \left\langle [], \left[\text{HEAD|LID} \quad [2] \right] \right\rangle \right]
 \end{array} \right]
 \end{array}$$

The analysis of finite *i*-clauses above constitutes a periphrastic analysis in the spirit of Bonami and Webelhuth (2013) and Bonami and Samvelian (2015): the head of the clause contributes to filling a cell in the paradigm of one lexeme (the verb) but is realized morphologically as another lexeme (the preposition *i*). Unlike those previous analyses, the present approach does not postulate valence-changing morphological rules and takes morphosyntactic blocking at face value. Unlike the approach of Bonami (2015), it does not rely on the hypothesis that a morphologically nonfinite verb fills a cell in a finite paradigm. Quite on the contrary, it implements rather directly Blevins’s (forthcoming) notion of ‘periphrasis as syntactic exponence’: the head value of the whole clause is the locus of evaluation of what constitutes an extended paradigm.

In addition, the present analysis appropriately captures the similarities and differences between the three constructions. Finite *bod* clauses and *i*-clauses both instantiate a morphosyntactic mismatch where a cell in a finite paradigm is realized by something else than a finite verb. Finite and nonfinite *i*-clauses have the same basic structure and are licensed by the same lexical type; they differ only in finiteness and morphosyntactic canonicity.

| (a) Sample 3sg forms of <i>bod</i> | | | | (b) Past 3sg constructions with the verb <i>mynd</i> 'go' | | |
|------------------------------------|----------------|------------|--------------|---|-------------------|--------------------------------------|
| | | present | past | imperfect | positive | negative |
| main clause | | <i>mae</i> | <i>buodd</i> | <i>roedd</i> | main clause | <i>aeth</i> <i>aeth ...ddim</i> |
| UDC clause | subject gap | <i>sy</i> | <i>buodd</i> | <i>roedd</i> | UDC clause | <i>aeth</i> <i>aeth ...ddim</i> |
| | nonsubject gap | <i>mae</i> | | | complement clause | <i>i ...mynd</i> <i>aeth ...ddim</i> |
| complement clause | | <i>bod</i> | <i>buodd</i> | <i>%roedd/bod</i> | | |

Table 1: Paradigmatic description of the distribution of finite *bod* and *i*

| (a) Regular P: <i>dan</i> 'under' | | | | (b) Irregular P: <i>i</i> 'to' | | | |
|-----------------------------------|-----|--------------|---------------|--------------------------------|-----|-------------|--------------|
| | ABS | <i>dan</i> | | | ABS | <i>i</i> | |
| | | SG | PL | | | SG | PL |
| AGREEING | 1 | <i>dana</i> | <i>danon</i> | AGREEING | 1 | <i>i</i> | <i>i</i> |
| | 2 | <i>danat</i> | <i>danoch</i> | | 2 | <i>i</i> | <i>i</i> |
| | 3.M | <i>dano</i> | <i>danyyn</i> | | 3.M | <i>iddo</i> | <i>iddyn</i> |
| | 3.F | <i>dani</i> | | | 3.F | <i>iddi</i> | |

Table 2: Sample prepositional paradigms

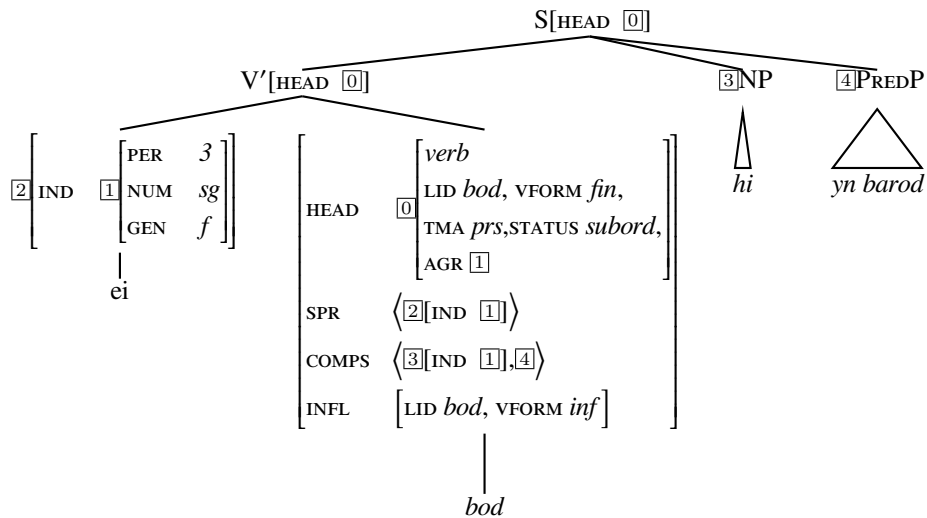


Figure 1: Analysis of a finite *bod* clause

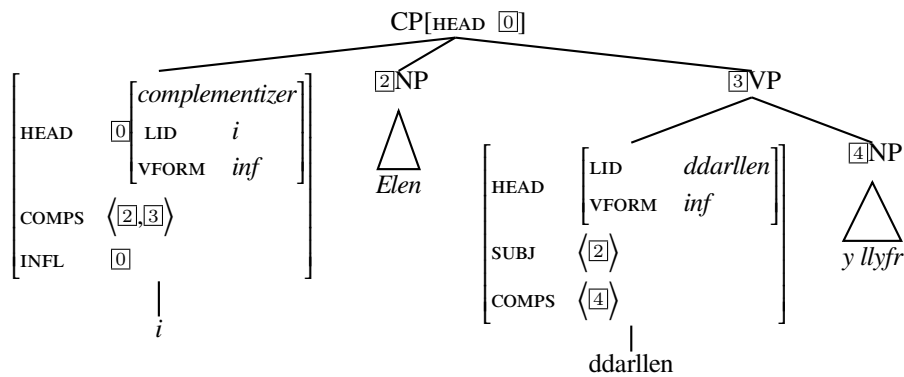


Figure 2: Analysis of a nonfinite *i* clause

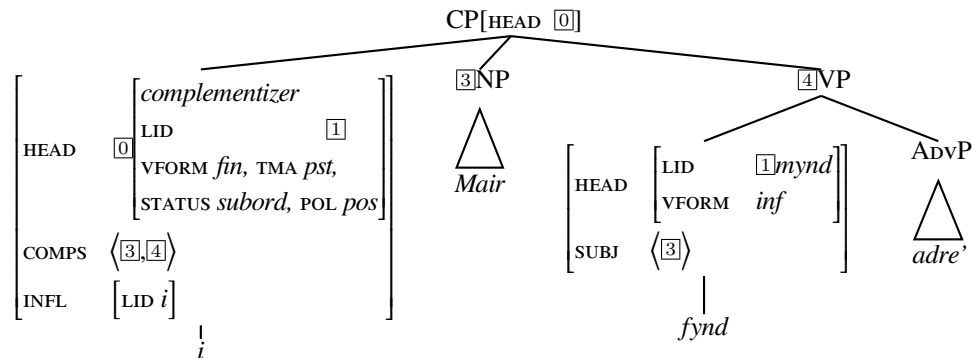


Figure 3: Analysis of a finite *i* clause

References

- Andrews, A. D. (1990). 'Unification and morphological blocking'. *Natural Language and Linguistic Theory*, 8:507–557.
- Blevins, J. P. (forthcoming). 'Periphrasis as syntactic exponence'. In F. Ackerman, J. P. Blevins, and G. T. Stump (eds.), *Paradigms and Periphrasis*. Stanford: CSLI Publications.
- Bonami, O. (2015). 'Periphrasis as collocation'. *Morphology*, 25:63–110.
- Bonami, O. and Crysmann, B. (2013). 'Morphotactics in an information-based model of realisational morphology'. In S. Müller (ed.), *Proceedings of HPSG 2013*. Stanford: CSLI Publications, 27–47.
- Bonami, O. and Samvelian, P. (2015). 'The diversity of inflectional periphrasis in Persian'. *Journal of Linguistics*, 51:327–382.
- Bonami, O. and Webelhuth, G. (2013). 'The phrase-structural diversity of periphrasis: a lexicalist account'. In M. Chumakina and G. G. Corbett (eds.), *Periphrasis: The role of syntax and morphology in paradigms*. Oxford: Oxford University Press, 141–167.
- Borsley, R. D. (1989). 'An HPSG approach to Welsh'. *Journal of Linguistics*, 25:333–354.
- (1999). 'Mutation and constituent structure in Welsh'. *Lingua*, 109:267–300.
- (2009). 'On the superficiality of Welsh agreement'. *Natural Language and Linguistic Theory*, 27:225–265.
- Borsley, R. D., Tallerman, M., and Willis, D. (2007). *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Koenig, J.-P. (1999). *Lexical relations*. Stanford: CSLI Publications.
- Sadler, L. and Spencer, A. (2001). 'Syntax as an exponent of morphological features'. In G. Booij and J. van Marle (eds.), *Yearbook of Morphology 2000*. Dordrecht: Kluwer, 71–96.
- Sag, I. A. (1997). 'English relative clause constructions'. *Journal of Linguistics*, 33:431–484.
- Stump, G. T. (2006). 'Heteroclis and paradigm linkage'. *Language*, 82:279–322.
- (2015). *Inflectional paradigms*. Cambridge: Cambridge University Press.
- Tallerman, M. (1998). 'The uniform case-licensing of subjects in Welsh'. *The Linguistic Review*, 15:69–133.