WELSH CLAUSAL *I* AND THE HIERARCHICAL LEXICON

Robert D. Borsley, University of Essex & Bangor University, rborsley@essex.ac.uk

At least three types of Welsh subordinate clause are introduced by what looks like the preposition *i* 'to', 'for': subjectless infinitives in control sentences, subjectless infinitives in raising sentences, and full clauses with an overt subject, reminiscent of English *for-to* clauses (henceforth *i*-clauses). The following illustrate:

(1) Mae Heledd yn awyddus [i weld Rhiannon].
be.PRS.3SG Heledd PROG eager to see.INF Rhiannon
'Heledd is eager to see Rhiannon.'

(2) Mae Heledd yn mynd [i weld Rhiannon].
be.PRS.3SG Heledd PROG go.INF to see.INF Rhiannon
'Heledd is going to see Rhiannon.'

(3) Disgwyliodd Heledd [i Sioned weld Rhiannon].
expect.PAST.3SG Heledd to Sioned see.INF Rhiannon
'Heledd expected Sioned to see Rhiannon.'

An obvious question here is: how many *i* lexemes are there? Everyone assumes that English has the same *to* lexeme in control and raising sentences. English has a different lexeme *for* introducing a full clause with an overt subject in examples like *I arranged* [*for Kim to see Lee*]. But Welsh could be different from English. It could be that there is a single *i* lexeme here.

Borsley (1986) argued that the Welsh data are problematic for Government Binding Theory because they seem to entail that *i* can be followed by PRO, a trace, or a lexical NP, contrary to central assumptions of the framework. However, this is only true if it is the same *i* lexeme in all three cases. Tallerman (1998) argued against this position. There is no reason to think that control and raising complements involve different *i* lexemes, but Tallerman showed that predicates which can take both a full clause introduced by *i* and a subjectless infinitive do not necessarily have *i* with the subjectless infinitive. Instead, they may be introduced by zero or an element homophonous with the preposition *o* 'from':

(4) a. Disgwyliodd Heledd [i Sioned weld Rhiannon].
expect.PAST.3SG Heledd to Sioned see.INF Rhiannon
'Heledd expected Sioned to see Rhiannon.'

b. Disgwyliodd Heledd [weld Rhiannon].
expect.PAST.3SG Heledd see.INF Rhiannon
'Heledd expected to see Rhiannon.'

(5) a. Roedd hi 'n siŵr [iddi hi glywed y gwcw].
be.IMPF.3SG she PRED sure to.3SGF she hear.INF the cuckoo
'She was sure she heard the cuckoo.'

b. Roedd hi 'n siŵr [o gyrraedd yn hwyr].
be.IMPF.3SG she PRED sure from arrive.INF PRED late
'She was sure to arrive late.'

This suggests that the *i* of subjectless infinitives and the *i* of *i*-clauses are distinct lexemes.

Tallerman (1998) also showed that while some *i*-clauses are non-finite clauses, rather like English *for-to* clauses, others are finite. *I*-clauses with *disgwylio* 'expect' and many other verbs are clearly non-finite. They are negated by the negative verb *peidio* like subjectless infinitives. (It is mutated as *beidio* in both cases.)

(6) Disgwyliodd Heledd [i Sioned beidio â gweld Rhiannon].
expect.PAST.3SG Heledd to Sioned NEG with see.INF Rhiannon
'Heledd expected Sioned not to see Rhiannon.'

(7) Disgwyliodd Heledd [beidio â gweld Rhiannon].
expect.PAST.3SG Heledd NEG with see.INF Rhiannon
'Heledd expected not to see Rhiannon.'

But other *i*-clauses appear with verbs which normally take a finite clause such as *meddwl* 'think'. Past tense forms of Welsh verbs are generally not acceptable in positive complement clauses:

(8) %Meddyliodd Heledd [aeth Sioned adre'].
think.PAST.3SG Heledd go.PAST.3SG Sioned home
'Heledd thought that Sioned had gone home.'

In colloquial Welsh, a perfect clause involving *bod* 'be' and the particle *wedi* appears instead (Jones 2010: 172):

(9) Meddyliodd Heledd [bod Sioned wedi mynd adre'].
think.PAST.3SG Heledd be.INF Sioned PERF go.INF home
'Heledd thought that Sioned had gone home.'

Despite appearances, this is a type of finite clause, as Tallerman (1998) and Bonami, Borsley & Tallerman (2016) show. In literary Welsh, an *i*-clause appears:

(10) Meddyliodd Heledd [i Sioned fynd adre'].
think.PAST.3SG Heledd to Sioned go.INF home
'Heledd thought that Sioned had gone home.'

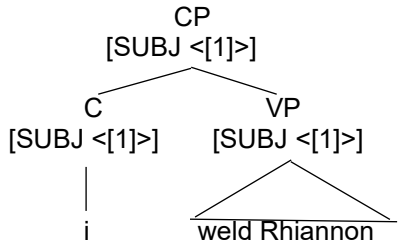
The interpretation suggests that this clause is actually finite, and so does the fact that it is in a context where a finite clause is expected. The fact that a negative counterpart of this clause is the ordinary finite clause in (11) points to the same conclusion:

(11) Meddyliodd Heledd [aeth Sioned ddim adre'].
think.PAST.3SG Heledd go.PAST.3SG Sioned NEG home
'Heledd thought that Sioned had not gone home.'

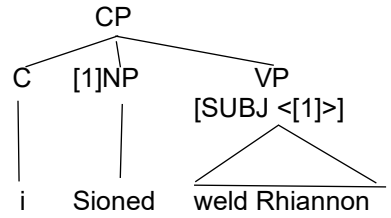
Tallerman shows that coordination and binding data also point to this conclusion. It seems, then, that there are non-finite *i*-clauses and finite *i*-clauses, and hence two *i* lexemes: one non-finite, and one finite and past tense. Thus, there are three *i* lexemes altogether.

Following earlier work, I assume these lexemes are all complementizers, one taking a VP complement, and the other two, like English *for* as analysed in Sag (997), taking an NP and a VP complement. Tallerman (1998), assumes an orthodox Chomskyan view of clause structure, in which there is a distinction between C(omplementizer) and I(nflection), and proposes that the *i* of *i*-clauses is in the I position. But the argument for this analysis is quite weak even within Chomskyan assumptions. Outside those assumptions there is no reason to think that *i* occupies a different position in *i*-clauses and subjectless infinitives. Thus, I assume that the complements in (1) and (2) have the form in (12), while the complement in (3) has the form in (13), and the complement in (10) has a similar structure.

(12)



(13)



I assume here that the *i* of subjectless infinitives is a raising predicate, whose subject is the subject of its complement, and that the *i* of *i*-clauses is a raising to object predicate, whose first complement is the subject of its second complement.

It is not unusual for a language to have homophonous lexemes. But these lexemes have properties in common other than just their phonological form, and they all share properties with the preposition *i*. All four lexemes trigger soft mutation on the following constituent. This is an NP in (14), (16) and (17), and a VP in (15). In each case, the mutated word is shown in bold and the basic unmutated form appears in brackets.

(14) i **Fangor** (Bangor)
to Bangor

(15) Mae Heledd yn awyddus [i **weld** Rhiannon]. (gweld)
be.PRS.3SG Heledd PROG eager to see.INF Rhiannon
'Heledd is eager to see Rhiannon.'

(16) Disgwyliodd Heledd [i **ddau** dyn weld Rhiannon]. (dau)
expect.PAST.3SG Heledd to two man see.INF Rhiannon
'Heledd expected two men to see Rhiannon.'

(17) Meddyliodd Heledd [i **ddau** dyn fynd adre']. (dau)
think.PAST.3SG Heledd to two man go.INF home
'Heledd thought that two men had gone home.'

(The mutation of *weld* in (16) is triggered not by *i*, but by the preceding subject *ddau dyn*.)

The preposition *i* and the *i* of non-finite and finite *i*-clauses also agree with a following third person pronoun:

(18) iddo fo / iddi hi / iddyn nhw
to.3SGM he to.3SGF she to.3PL they
'to he/she/them'

(19) Disgwyliodd Heledd [iddo fo / iddi hi / iddyn
expect.PAST.3SG Heledd to.3SGM he to.3SGF she to.3PL
nhw weld Rhiannon].
they see.INF Rhiannon
'Heledd expected he/she/them to see Rhiannon.'

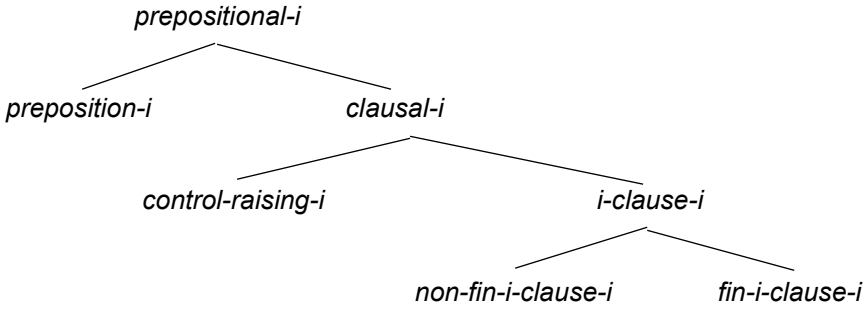
(20) Meddyliodd Heledd [iddo fo / iddi hi / iddyn
think.PAST.3SG Heledd to.3SGM he to.3SGF she to.3PL
nhw fynd adre'].
they go.INF home
'Heledd thought that he/she/they had gone home.'

The *i* of subjectless infinitives does not show agreement, but it does not have the opportunity because it is never immediately followed by a pronoun. Thus, we can say that all four lexemes have the same agreement potential, and more generally that they have the same morphological properties.

The four lexemes have different syntactic properties, but the three complementizers have in common the fact that they are complementizers, and the two *i*-clause complementizers have the same complement selection properties.

How should the various similarities be captured? Standard HPSG assumptions about the lexicon allow a simple solution. They allow the four lexemes to be analysed as four realisations of a 'super-lexeme' and all the shared properties to be specified just once. We can propose the type hierarchy in (21) for this part of the lexicon. Note that *prepositional-i* and *preposition-i* are quite different types, and that I am using the type *control-raising-i* for the *i* of subjectless infinitives.

(21)



We have seen that all four lexemes have the same phonological and morphological properties. I assume that the phonological properties are a reflection of the morphological properties, and that the morphological properties reflect two features. First, following Borsley (2009), I assume that agreement in Welsh is the realization of a feature AGR, whose value is the index of a following pronoun with its PERSON, NUMBER, and GENDER features, or *none* when there is no following pronoun. (Agreement is only triggered by pronouns.) I assume that the mutation-triggering property of a lexeme reflects a feature MUT(ATION)-TR(IGGER) with the values *soft*, *aspirate*, *nasal* for the three kinds of mutation that occur in Welsh, or *none*. (Only the first is relevant here.) With these assumptions, we can attribute the phonological and morphological properties of the four lexemes to the following constraint on

prepositional-i (where the MARKING feature allows heads to select a constituent headed by one of these lexemes):

(22) *prepositional-i* $\Rightarrow$ $\begin{bmatrix}\text{MARKING } i \\ \text{AGR } index \vee none \\ \text{MUT} - \text{TR } soft\end{bmatrix}$

For the two immediate subtypes of *prepositional-i*, we just need the following simple constraints:

(23) *preposition-i* $\Rightarrow$ $\begin{bmatrix}\text{HEAD } prep \\ \text{COMPS} < \text{NP} >\end{bmatrix}$

(24) *clausal-i* $\Rightarrow$ $[\text{HEAD } comp]$

For the two immediate subtypes of *clausal-i*, we need the following, slightly more complex, constraints:

(25) *control-raising-i* $\Rightarrow$ $\begin{bmatrix}\text{HEAD } [\text{VFORM } inf] \\ \text{SUBJ} < [1] > \\ \text{COMPS} < \text{VP}[inf, \text{SUBJ} < [1] >] >\end{bmatrix}$

(26) *i-clause-i* $\Rightarrow$ $\begin{bmatrix}\text{SUBJ} <> \\ \text{COMPS} < [1]\text{NP}, \text{VP}[inf, \text{ SUBJ} < [1] >] >\end{bmatrix}$

For the two subtypes of *i-clause-i*, we can propose the following:

(27) *non-fin-i-clause-i* $\Rightarrow$ $[\text{HEAD } [\text{VFORM } inf]]$

(28) *fin-i-clause-i* $\Rightarrow$ $[\text{HEAD } [\text{VFORM } fin, \text{TENSE } past]]$

The [TENSE *past*] specification in (27) ensures that finite *i*-clauses have the sort of interpretation that one would expect to be expressed by a complement clause with a past tense verb. But what about the fact that a positive past tense verb is generally ungrammatical in a complement clause? One possibility is an analysis of the kind outlined in Bonami, Borsley & Tallerman (2016), in which finite *i* is literally a positive past tense form of the associated verb. However, as noted above, finite *i* is generally confined to the literary language. In more colloquial Welsh a perfect clause involving *bod* 'be' and the particle *wedi* appears instead. I will assume, then, that there is a constraint ruling out a past tense verb in a positive complement clause, and that different varieties have different ways of expressing the meanings which cannot be expressed by a past tense verb, finite *i* fulfilling this role in the literary language.

There are some other cases in Welsh of homophonous lexemes which should probably be analysed as alternative realizations of a single super lexeme. I assume the element *o* in (5b) is another complementizer homophonous with a preposition. This element triggers soft mutation (the unmutated form of the following verb is *cyrraedd*). In this, it just like the preposition:

(30) Dw i wedi dôd o **Gaernarfon**. (Caernarfon)
be.PRES.1SG I PERF come.INF from Caernarfon
'I have come from Caernarfon.'

This suggests that we have two realizations of a single super lexeme.

There are at least two other cases for which an analysis of this kind seems appropriate. Welsh has a number of aspectual particles which are homophonous with prepositions. The most common, homophonous with the prepositions *yn* 'in' and *wedi* 'after', share no other properties with the prepositions. But two others, homophonous with the prepositions *ar* 'on' and *heb* 'without', share with the prepositions the property of assigning soft mutation to their complement. Here are examples with *ar*:

(31) a. Mae 'na wylan ar **gar** Heledd. (car)
be.PRES.3SG there seagull on car Heledd
'There is a seagull on Heledd's car.'

b. Mae o ar **ganu**. (canu)
be.PRES.3SG he on sing.INF
'He's about to sing.'

Here are examples with *heb*:

(32) a. Dw i heb **gar** yr wythnos 'ma. (car)
be.PRES.1SG I without car the week here
'I'm without a car this week.'
b. Maen nhw heb **gyrraedd** eto. (cyraedd)
be.PRES.3PL they without arrive.INF yet
'They haven't arrived in Bangor yet.'

A super lexeme treatment seems appropriate for both cases.

REFERENCES

Bonami, Olivier, Robert D. Borsley & Maggie O. Tallerman. 2016. On pseudo-non-finite clauses in Welsh. *Proceedings of the Joint 2016 Conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar*, 104–124.

Borsley, Robert D. 1986. Prepositional complementizers in Welsh. *Journal of Linguistics* 22, 67–84.

Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language & Linguistic Theory* 27, 225–265.

Jones, Bob Morris. 2010. *Tense and Aspect in Informal Welsh*. Berlin & New York: Mouton de Gruyter.

Sag, Ivan A. 1997. English relative clause constructions. *Journal of Linguistics* 33, 431–484.

Tallerman, Maggie O. 1998. The uniform case-licensing of subjects in Welsh. *The Linguistic Review* 15, 69–133.