

Complementizer-Verb Interactions in Welsh

Frances Dowle
University of Oxford

1 Introduction

Welsh has a comprehensive inventory of complementisers, and several of which cannot occur with certain forms of the verb *bod* 'be' (hereafter referred to as (*Welsh*) *be*, to avoid using as the citation form the non-finite form *bod*, the distribution of which merits discussion in its own right). This paper will provide a diachronic analysis of complementizer-*be* interactions in Welsh, arguing that in some cases, the data supports a lexical sharing (Wescoat, 2002, 2005, 2007) analysis of certain forms of *be* in Welsh. This case study will allow for discussion about the type of evidence that supports the use of this analytic device - or any other that permits a single indivisible form to correspond to multiple syntactic nodes - in LFG.¹

2 Two types of incompatibility with *be*

There are two types of complementizer-*be* incompatibilities that will be central to this paper. The first of these concerns the modern spoken complementizers *mi* (North Wales) and *fe* (South Wales). The incompatibilities of these complementizers with *be* forms is summarised by Borsley et al. (2007):

1. "*Fe* occurs with any finite verb except present or imperfect forms of [*be*]"
2. "*mi* occurs with any finite verb except third-person present tense forms of [*be*]"

The following examples show these (in)compatibilities with *be* and with other auxiliaries:

- | | |
|--|--|
| (1) *Mi/*Fe mae Non yn siarad.
C.AFF be.PRS Non PROG talk.NF
'Non is talking.' | (2) Mi/*Fe (r)ydw i =n siarad..
C.AFF be.PRES.1SG 11SG PROG talk.NF
'Non was talking.' |
| (3) Mi/*Fe (r)oedd Non yn siarad.
C.AFF be.IMPF.3S Non PROG talk.NF
'Non was talking.' | (4) Mi/Fe fydd Non yn siarad.
C.AFF be.FUT.3S Non PROG talk.NF
'Non will be talking.' |
| (5) Mi/Fe wnaeth Non siarad.
C.AFF seido.PAST.3S Non speak.NF
'Non spoke.' | |

Another type of incompatibility of complementizers with *be* concerns affirmative complement clauses. In affirmative complement clauses (but not subordinate clauses generally), present-tense finite forms of *be*, and (in some dialects) imperfect-tense finite forms of *be*, do not occur. Instead, complement clauses with present/imperfect interpretation contain a morphologically non-finite form of *be*, that is, *bod*. The restriction against morphologically-finite forms of verbs in complement clauses with particular tense interpretations is general to all verbs and not specific to *be*. *Bod*, as all other morphologically non-finite verbs found clause-initially in affirmative complement clauses with present/imperfect interpretation, is incompatible with the complement clause complementizer.

- | | |
|--|--|
| (6) Dw =i =n meddwl...
be.PRS.1SG =1SG =PROG think.NF...
'I think...' | |
| (7) *...[mae Bethan yn= mynd i =r
...[be.PRS.3SG Bethan PROG= go.NF to =the
siopau heddiw].
shops today]
<i>Intended:</i> '... (that) Bethan is going to the shops to-
day' | (8) ...[bod Bethan yn= mynd i =r siopau
...[be.NF Bethan PROG= go.NF to =the shops
heddiw].
today]
'... (that) Bethan is going to the shops today' |
| (9) [y bydd Bethan yn= mynd i
[C.AFF.COMP be.FUT Bethan PROG= go.NF to
=r siopau yfory].
=the shops tomorrow]
'(that) Bethan will go to the shops tomorrow.' | (10) *...[y bod Bethan yn= mynd i
...[C.AFF.COMP be.NF Bethan PROG= go.NF to
=r siopau heddiw].
=the shops today]
<i>Intended:</i> '... (that) Bethan is going to the shops to-
day.' |

¹In addition to abbreviations found in the *Leipzig Glossing Rules*, the following abbreviations are used in this paper: NF = NONFINITE, PRS = PRESENT, C = COMPLEMENTISER. Broad phonological transcriptions are given in the IPA upon first use of Welsh word throughout the paper, except where the IPA transcription is identical to the Welsh orthography.

3 Stages of development: complementizers and *be*

In the period known as Early Modern Welsh (approx. 1500-1700CE), Welsh clauses can either be formed with a finite lexical verb in initial position ($V_{\text{LEX.FIN}}\text{SO}$) or with a finite auxiliary verb in initial position and a later non-finite lexical verb ($V_{\text{AUX.FIN}}\text{SV}_{\text{LEX.NF}}\text{O}$). Note that the former construction declines throughout the development of Welsh, with the result that the latter almost entirely dominates in the modern language in most tenses.

In early modern Welsh, affirmative main clauses that contained *be* in initial position (in either its auxiliary or existential use) were headed by a complementizer $y(r)$:²

- (11) ...ac yr oedd rhyw bendefig...
 ...and C.AFF be.IMP.3SG some ruler...
 ‘...and there was a certain ruler...’ (from William Morgan’s 1588 translation of the Bible)

Complement clauses (COMPs) are also headed by a complementizer $y(r)$, which we shall treat as a separate lexical item, at least for now. This complementizer behaves as outlined in the first section, i.e., it is incompatible with morphologically-non-finite verbs, which are required in affirmative complement clauses of particular types.

Negative clauses have a morphophonologically distinct complementizer depending on whether they are main clauses, in which case it takes the form *ni(d)*, or subordinate clauses, in which case it takes the form *na(d)*. The negative complement clause complementizers are compatible with all forms of *be*, except that in the 3SG present we find *ydy* (a kind of dependent 3SG form of *be*) rather *mae*.

The situation at this stage of the language concerning compatibilities of complementizers and *be* can be summarised in the following table. In general throughout this section, the distribution of the 1SG.PRES form of *be* is illustrative of the distribution of all non-3SG present tense forms:

	CLAUSE TYPE	3SG.PRES	1.SG.PRES	3SG.IMP	3SG.FUT
(12)	AFF MAIN	y=mae	yr=ydw	yr=oedd	y=bydd
	AFF COMP	bod	bod	(%yr=oedd)	y=bydd
	NEG MAIN	nid=ydy	nid=ydw	nid=oedd	ni=fydd
	NEG COMP	nad=ydy	nad=ydw	nad=oedd	na=fydd

Over time, the complementizers erode, giving us the situation found in later Modern Welsh (and still found in some registers today), where the affirmative main clause complementizer has come to be just a syllable onset at the beginning of *be* forms, and not present at all with the form *mae*. The same process has occurred with negative clauses. In negative and in complement clauses, the complementizer may be deleted entirely. This process is more advanced in negative clauses, although the soft mutation effects of the negative complementizer may remain. All negative forms are at this stage reinforced with the negative marker *ddim* /*ðim*/ later in the sentence.

	CLAUSE TYPE	3SG.PRES	1.SG.PRES	3SG.IMP	3SG.FUT
(13)	AFF MAIN	mae	r-ydw	r-oedd	bydd
	AFF COMP	bod	bod	(%r-oedd)	(y=)bydd
	NEG (MAIN, COMP)	dydy	d-ydw	d-oedd	fydd

During the period of early modern Welsh, a new complementizer system begins to develop Willis (2007). The end result of this process is that a dialect boundary emerges. On one side of the boundary, *mi* becomes the new affirmative main clause complementizer. On the other side of the boundary, *fe* is established. By this stage of development, many clauses in the spoken language use an auxiliary, but in the preterite this is a form of the verb *do* rather than of *be*. Both *mi* and *fe* can be used with the *do* auxiliary and with finite lexical verbs (where these are still found), but from the earliest stages of their development, they exhibit restrictions with *be* forms:

	CLAUSE TYPE	3SG.PRES	1.SG.PRES	3SG.IMP	3SG.FUT
(14)	AFF MAIN <i>mi</i> areas	(*mi=)mae	mi=(ry)dw	(*mi=(r)oedd	mi=fydd
	AFF MAIN <i>fe</i> areas	(*fe=)mae	(*fe=)dw	(*fe)(r)-oedd	fe=fydd
	AFF COMP all areas	bod	bod	(%r)oedd	bydd
	NEG (MAIN and COMP)	dydy	d-ydw	d-oedd	fydd

4 Analysis

There are clearly some parallels between the main clause and complement clause data presented above. In particular, *fe* is excluded in present and imperfect main clauses and subordinating $y(r)$ / $@(r)$ / is excluded in present and imperfect complement clauses in present-day Welsh, where *be* is almost always found as the clause-initial verb. However, a difference emerges in the origin and development of these restrictions. One of these originates from a general incompatibility between the complementizer and certain clause types (the same clause types which require morphologically-non-finite forms). At no point is the restriction unique to *be* forms, and in case, the complementizer is soon lost entirely. As such,

²As with other complementizer forms presented in this paper, the bracketed final consonant in $y(r)$ / $@(r)$ / surfaces only before vowels.

it is entirely possible for speakers to acquire a complement clause complementizer that has certain functional restrictions, preventing it from occurring in clauses of a certain type, and later, for speakers to require no complement clause complementizer at all. At the same time, certain types of finite verb also are restricted from these clauses: there is no direct incompatibility between *bod* and the complement clause complementizer required in the analysis. The situation is rather different with affirmative main clauses. Here, direct comparisons can be made, as restrictions are specific to *be* forms (cf. (5) above). As such, speakers acquire a complementizer form which is specifically incompatible with forms of the verb *be* that occur only in clause types which we would otherwise expect to be compatible with the relevant complementizer. In other words, speakers acquire the knowledge that *mi* cannot occur with *mae*, even though they know that it can occur with other 3sg forms (cf. 3) and with other present tense forms (cf. 2).

I argue that between the middle and latter stages presented above, speakers acquire a lexical-sharing analysis of certain *be* forms. Although the phonological evidence of the presence of the affirmative main clause complementizer $y(r)$ may be slight or non-existence, speakers still perceive the forms in (13) as consisting of both a complementizer and a finite verb form. We can thus see several stages of development, from the stage in which the main clause complementizer and finite verb are two separate lexical items (as in 15), to their development as a lexical sharing form in (16). This analysis parallels some of the diachronic analysis of English possessives put forward by Lowe (2015, 2016), and the formal notation used for lexical sharing forms follows Lowe's rather than Wescoat's presentation.

(15)	$y(r) : \hat{C}$ $\neg (\text{COMP}\uparrow)$								
	$oedd : I^0$ $(\uparrow \text{TENSE}) = \text{IMP}$ $(\uparrow \text{SUBJ PERS}) = 3$ $(\uparrow \text{SUBJ NUM}) = \text{SG}$								
(16)	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;">$r-oedd : \hat{C}$</td> <td style="width: 50%;">I^0</td> </tr> <tr> <td>$\neg (\text{COMP}\uparrow)$</td> <td>$(\uparrow \text{TENSE}) = \text{IMP}$</td> </tr> <tr> <td>$(\uparrow \text{SUBJ PERS}) = 3$</td> <td></td> </tr> <tr> <td>$(\uparrow \text{SUBJ NUM}) = \text{SG}$</td> <td></td> </tr> </table>	$r-oedd : \hat{C}$	I^0	$\neg (\text{COMP}\uparrow)$	$(\uparrow \text{TENSE}) = \text{IMP}$	$(\uparrow \text{SUBJ PERS}) = 3$		$(\uparrow \text{SUBJ NUM}) = \text{SG}$	
$r-oedd : \hat{C}$	I^0								
$\neg (\text{COMP}\uparrow)$	$(\uparrow \text{TENSE}) = \text{IMP}$								
$(\uparrow \text{SUBJ PERS}) = 3$									
$(\uparrow \text{SUBJ NUM}) = \text{SG}$									

If speakers have acquired a lexical sharing analysis of the form *r-oedd*, or even a form like *mae*, it is then straightforward to explain why they should not later allow the new complementizers *mi* or *fe* to occur with these forms. This incompatibility maintains the possibility of acquiring a lexical sharing analysis of *be* forms. This analysis on behalf of speakers acquiring the language is informed by and maintains the impossibility of the co-occurrence of these *be* forms with *mi* or *fe*. This co-occurrence is blocked at c-structure, because the lexical sharing *be* form already projects to a C node, and so another C item cannot also be present. In southern varieties, there are many more lexical-sharing forms of *be* than in northern varieties, hence the less-restricted distribution of *mi*. Crucial to the acquisition of any of the lexical sharing forms is the power of analogy and contrast to show that *mi* and *fe* are restricted with specific verb forms, not specific clause types. The contrast between affirmative main clause *be* forms and the forms found elsewhere is also of import.

In summary, this paper will set out the stages of analysis and reanalysis that lie behind both the rise and restrictions of *fe* and *mi*, and the loss of complementizers in subordinate clauses. In so doing, it will show the kind of evidence that allows speakers to acquire a lexical sharing analysis (or any analysis which involves a single word corresponding to multiple syntactic nodes). In turn, this paves the way for discussion about the kind of linguistic evidence (if any) that licenses the use of this analytic device within LFG.

References

- Borsley, Robert D., Tallerman, Maggie and Willis, David. 2007. *The syntax of Welsh*. Cambridge University Press.
- Lowe, John J. 2015. Degrees of degrammaticalization: A lexical sharing approach to the English possessive. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG'15 Conference*, pages 208–228, CSLI Publications.
- Lowe, John J. 2016. English possessive 's: clitic and affix. *Natural Language & Linguistic Theory* 34(1), 157–195.
- Wescoat, Michael T. 2002. *On lexical sharing*. Ph. D.thesis, Stanford University.
- Wescoat, Michael T. 2005. English nonsyllabic auxiliary contractions: An analysis in LFG with lexical sharing. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG'05 Conference*, pages 468–486, CSLI Publications.
- Wescoat, Michael T. 2007. Preposition-determiner contractions: An analysis in Optimality-Theoretic Lexical-Functional Grammar with Lexical Sharing. *Proceedings of the LFG'07 Conference* page 439–459.
- Willis, David. 2007. Specifier-to-head reanalyses in the complementizer domain: Evidence from Welsh. *Transactions of the Philological Society* 105(3), 432–480.