

Robert D. Borsley*

Apparent filler–gap mismatches in Welsh

DOI 10.1515/ling-2015-0024

Abstract: Filler–gap dependencies involving a clause-initial filler constituent of some kind followed by a matching gap are an important feature of human languages. There are also certain cases where what looks like a filler differs in some way from the following gap. In the case of Welsh there is a mismatch between apparent filler and gap in some nominal cleft sentences. It can be argued, however, that the initial constituent is not a filler but one term of a hidden identity predication. There are various other complexities in this area. There is one word, the identity copula, which only allows a complement that is a gap. There are two cases where a deletion process conceals the identity of the initial constituent in a cleft sentence, making a Progressive Phrase look like a Verb Phrase and a Predicative Phrase look like an Adjective Phrase or a Noun Phrase. Finally, there are three cases where a verb with a gap as a dependent has a special form, two cases involving the predicational copula and one involving all transitive verbs. Thus, a number of mechanisms are required to deal with the full set of facts.

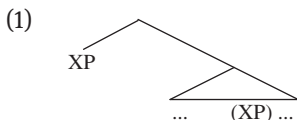
Keywords: filler–gap dependencies, fillers, gaps, clefts, copula

1 Introduction¹

A central feature of natural languages is what are often known as filler–gap dependencies, where there is an extra clause-initial constituent of some kind and a gap somewhere later in the clause (possibly in an embedded clause). The basic situation is as follows, where I use an underscore to indicate the gap followed by a bracketed category to indicate its type:

¹ Earlier versions of this paper were presented at the Seventeenth Welsh Syntax Seminar at Gregynog, Wales, July 5–6, 2010, the Fourteenth International Celtic Congress at Maynooth, Ireland, August 1–5, 2011, and the HPSG Workshop, at Frankfurt University, May 10–11, 2012.

*Corresponding author: Robert D. Borsley, Department of Language and Linguistics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK, E-mail: rborsley@essex.ac.uk



The following *wh*-interrogatives provide a simple illustration:

- (2) a. [_{NP} *Who*] *did Kim talk to* ___ (NP)?
 b. [_{PP} *To whom*] *did Kim talk* ___ (PP)?
 c. [_{AP} *How long*] *is a piece of string* ___ (AP)?
 d. [_{AdvP} *How quickly*] *did you do it* ___ (AdvP)?

In each case the filler and the gap are of the same category. They typically match in other respects as well. For example, if they are nominal, they match in number, as the following illustrate:

- (3) a. [_{NP[SG]} *Which student*] *do you think* ___ (NP[SG]) knows the answer?
 b. [_{NP[PL]} *Which students*] *do you think* ___ (NP[PL]) know the answer?

In languages with grammatical gender or morphological case, they also share these properties.

Filler–gap dependencies have had a great deal of attention, and from time to time attention has been drawn to what can be called filler–gap mismatches, where there is apparently a filler–gap dependency but where what looks like a filler differs from the gap in some way. There are essentially two sorts of example. On the one hand, there are examples in which only some other kind of overt constituent is possible in the position of the gap. A well-known example of this kind, discussed in Bouma, Malouf and Sag (2001), Bresnan (2001), and Webelhuth (2012), is illustrated by the following:

- (4) *That he might be wrong, he didn't think of* ___.

Here, the apparent filler is a clause, but as the following shows, only an overt NP and not an overt clause is possible in the position of the gap.

- (5) *He didn't think of* $\left. \begin{array}{l} \textit{that matter} \\ \textit{*that he might be wrong} \end{array} \right\}$.

On the other hand, there are examples where no overt constituent can appear in the position of the gap, e.g. (6), from Kayne (1980):

(6) *This candidate, they assured me ___ to be reliable.*

The following shows that no overt constituent is possible in the position of the gap:

(7) **They assured me this candidate to be reliable.*

For Minimalism and earlier transformational approaches, filler-gap dependencies are the result of movement of the filler constituent from the position of the gap. Hence, within such approaches one expects filler and gap to have all the same properties. For Head-driven Phrase Structure Grammar (HPSG), as developed, for example, in Ginzburg and Sag (2000), filler-gap dependencies involve the SLASH feature, which makes information about a gap available higher in the structure. Its value is a set of local feature structures (normally a singleton set), and constraints ensure that a filler and the associated gap have the same local feature structure. A local feature structure encodes most of the syntactic and semantic properties of an expression. It does not include the WH feature which identifies interrogative *wh*-elements, but it includes categorial features and, in the case of noun phrases, person, number, gender and case features. It follows that a gap associated with a filler that is an interrogative *wh*-element does not have the WH feature, but filler and gap have the same categorial features, and the same person, number, gender and case features if the filler is a noun phrase. Thus, an example like (4) above poses a challenge for both transformational approaches and HPSG. Examples like (6) are no problem for HPSG since, as Bouma et al. (2001: 25) point out, it is not difficult to stipulate that only a gap is possible in some position. Whether they are problematic for Minimalism is not clear to me.

As emphasized in Webelhuth (2008), there are a number of types of filler-gap mismatch in English, and a number of approaches that might be taken to them within HPSG. In the case of examples like (5), Webelhuth (2012) argues that what looks like a filler is not really a filler although it is coindexed with the gap.² If this is the right approach to take here, it does not necessarily follow that it is

² Chomsky (1977: 91) proposed that clause-initial topics are base-generated and hence not fillers, but he did not discuss apparent filler gap mismatches like (4).

the right approach to take to other filler–gap mismatches. It may well be that different cases require different sorts of analyses.

In this paper I will investigate a number of apparent filler–gap mismatches in Welsh and develop analyses within HPSG. I will argue that there are reasons for employing a number of different mechanisms, including one not envisaged in Webelhuth’s discussion of the English phenomena. The analyses will be quite complicated but that is because the facts are complicated. I doubt whether simpler analyses are possible.

The paper is organized as follows. In Section 2 I look at what I will call nominal cleft sentences, which provide an interesting example of an apparent filler–gap mismatch. Then in Sections 3 and 4, I consider a variety of cleft sentences involving *bod* ‘be’. In Section 5 I consider some further apparent mismatches, which arise not just with clefts but also with *wh*-interrogatives. Finally, in Section 6 I provide some concluding remarks.

2 Nominal cleft sentences

Nominal cleft sentences, which I discussed in Borsley (2008), provide a notable Welsh example of an apparent filler–gap mismatch. Here are some typical examples:³

- (8) a. *Emrys (a) brynodd ___ lyfr.*
 Emrys PART buy.PAST.3SG book
 ‘It was Emrys that bought a book.’
 b. *Llyfr (a) brynodd Emrys ___.*
 book PART buy.PAST.3SG Emrys
 ‘It was a book that Emrys bought.’

In (8a) the gap is in subject position, which is post-verbal because Welsh is a VSO language, and in (8b) it is in object position. Tallerman (1996) assumed that clefts are a filler–gap construction and proposed a transformational analysis in which the initial constituent is the result of movement to Spec CP. However, the initial constituent may differ from the associated gap in certain ways. This makes it unlike a typical filler constituent.

³ Both the verb *brynodd* ‘buy’ and the noun *lyfr* ‘book’ in (8a) are mutated. The basic unmutated forms are *prynodd* and *llyfr*. In the present context mutation is generally not important, and I will pass over most instances without comment.

An important feature of Welsh is that a verb agrees with a pronoun but not with a non-pronominal NP.⁴ The examples in (9) illustrate agreement with a following pronominal subject:

- (9) a. *Gwelodd* *o*.
 see.PAST.3SG he
 ‘He saw.’
 b. *Gwelon* *nhw*.
 see.PAST.3PL they
 ‘They saw.’

With a following non-pronominal subject, singular or plural, the third person singular form appears:

- (10) a. *Gwelodd* *y bachgen*.
 see.PAST.3SG the boy
 ‘The boy saw.’
 b. *Gwelodd* *y bechgyn*.
 see.PAST.3SG the boys
 ‘The boys saw.’
 c. **Gwelon* *y bechgyn*.
 see.PAST.3PL the boys
 ‘The boys saw.’

In a cleft sentence with an initial constituent associated with a subject gap, the finite verb is third person singular, whether the initial constituent is pronominal, as in (11), or non-pronominal, as in (12):

- (11) a. *Nhw welodd* ___ *ddraig*.
 they see.PAST.3SG dragon
 ‘It was they that saw a dragon.’
 b. **Nhw welon* ___ *ddraig*.
 they see.PAST.3PL dragon
- (12) a. *Y bechgyn welodd* ___ *ddraig*.
 the boys see.PAST.3SG dragon
 ‘It was the boys that saw a dragon.’

⁴ For detailed discussion see Borsley (2009).

- b. *Y *bechgyn welon* — *ddraig*.
 the boys see.PAST.3PL dragon

This suggests that the gap is non-pronominal whatever the nature of the associated initial constituent.

A second type of contrast between initial constituent and gap is highlighted by the examples in (13):

- (13) a. *Fi mae Gwyn wedi* $\left\{ \begin{array}{l} 'i \quad ddeuwis \\ *fy \quad newis \end{array} \right\}$ —.
 I be.PRES.3SG Gwyn PERF $\left\{ \begin{array}{l} 3SGM \quad choose.INF \\ 1SG \quad choose.INF \end{array} \right\}$
 'It's me that Gwyn has chosen.'
 b. *Ti mae Gwyn wedi* $\left\{ \begin{array}{l} 'i \quad ddeuwis \\ *dy \quad newis \end{array} \right\}$ —.
 you.SG be.PRES.3SG Gwyn PERF $\left\{ \begin{array}{l} 3SGM \quad choose.INF \\ 1SG \quad choose.INF \end{array} \right\}$
 'It's you that Gwyn has chosen.'

In these examples the gap is object of a nonfinite verb. In this situation, the nonfinite verb is preceded by a clitic agreeing with the gap. The clitic is third person singular masculine, and so we presumably have a third person singular masculine gap, but the initial constituent is first person singular in (13a) and second person singular in (13b). The clitic cannot be first person in (13a) or second person in (13b). Thus, initial constituent and gap differ in person.⁵

Thus, it seems that the initial constituent in a cleft sentence and the associated gap differ in two important respects. One might try to account for the absence of agreement in (11a) by stipulating that agreement is only with pronominal elements which are not gaps, but this would do nothing to account for the appearance of a third person clitic in (13a) and (13b). Saying that cleft sentences have a non-pronominal gap provides a simple account of both sets of data. There is no agreement in (11a) because agreement is only with pronouns, and a third person clitic appears in (13a) and (13b) because a non-pronominal gap like an overt non-pronominal NP is third person.

⁵ I present an analysis of clitics associated with a gap that is object of a nonfinite verb in Section 5 below.

If clefts involve a non-pronominal gap even when the initial constituent is a pronoun, this suggests that the initial constituent is not a filler. This was the conclusion I came to in Borsley (2008). I proposed there that the initial constituent in a cleft sentence is one term of an identity predication. On this view, the cleft sentences in (8a), (11a), and (13a) above are similar to examples like the following, which we might call quasi-pseudoclefts:

- (14) a. *Emrys ydy 'r un (a) brynodd ___ lyfr.*
 Emrys be.PRES.3SG the one PART buy.PAST.3SG book
 'The one that bought a book is Emrys.'
- b. *Nhw ydy 'r rhai (a) welodd ___ ddraig.*
 they be.PRES.3SG the ones PART see.PAST.3SG dragon
 'The ones that saw a dragon were them.'
- c. *Fi ydy 'r un mae Gwyn wedi 'i ddewis ___.*
 I be.PRES.3SG the one be.PRES.3SG Gwyn PERF 3SGM choose.INF
 'The one that Gwyn has chosen is me.'

They are also similar to English examples like the following, from Akmajian (1970: 150):

- (15) *It's me who ___ is responsible.*

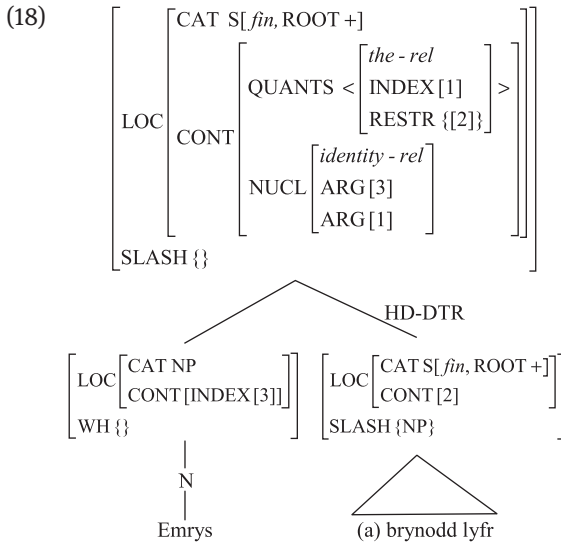
Both types of example contain an overt identity predication. In the Welsh examples there is no requirement of pronominality and person identity between the two terms of the identity relation. Similarly in the English example, there is no requirement for the focused constituent and the following relative clause to have the same person. Thus, the contrasts between the initial constituent and the gap that we have seen are only to be expected on this approach.

I suggested in Borsley (2008) that negative sentences like the following provide evidence for the hidden identity predication:

- (16) a. *Nid/dim Emrys (a) brynodd ___ lyfr.*
 NEG Emrys PART buy.PAST.3SG book
 'It wasn't Emrys that bought a book.'
- b. *Nid/dim llyfr (a) brynodd Emrys ___.*
 NEG book PART buy.PAST.3SG Emrys
 'It wasn't a book that Emrys bought.'

CAT, and CONT, respectively, and I will use the traditional tree format to represent constituent structure.

Assuming that clefts involve a hidden identity predication, (7a) above will have an analysis that can be represented as follows:

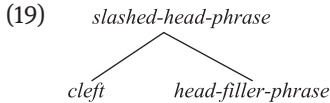


Here, the second daughter is the head but the first daughter is not a filler and its LOC value is not identified with the local feature structure in the value of SLASH in the second daughter. The value of SLASH in the mother is the empty set {} because the head daughter is the top of the dependency. The CONT value of the mother makes it clear that the second daughter is interpreted as a definite description and identified with the first daughter.⁷

How should structures like (18) be licensed? Following Borsley (2008), I will assume a type *slashed-head-phrase* with subtypes *cleft* and *head-filler-phrase* as follows:⁸

⁷ The CONT value is a simplified version of the type of value assumed in Ginzburg and Sag (2000). For Ginzburg and Sag, QUANTS and NUCLEUS appear in the value of a feature SOA (STATE-OF-AFFAIRS), which appears in the value of the feature PROP (PROPOSITION). The extra complexity is not important in the present context.

⁸ In Borsley (2008) I treated *slashed-head-phrase* as a subtype of *slashed-daughter-phrase*, the latter having two daughters where one has a single local feature structure within the value of SLASH and neither is identified as the head. This was to provide an account of free relatives, which I argued involve a slashed daughter that is not a head. I ignore this matter here.



This allows us to capture both the similarities and the differences between clefts and head-filler-phrases. Slashed-head-phrases will be subject to the following constraint:

(20)

$$sl-hd-ph \Rightarrow \left[\begin{array}{l} SS[SLASH \{\}] \\ DTRS < [phrase], [1] S[SLASH \{\{\}] > \\ HD - DTR [1] \end{array} \right]$$

This says that a slashed-head-phrase is SLASH {} and has one daughter which is a phrase and another which is a head, and a clause with a single local feature structure within the value of SLASH. Clefts will be subject to the constraint in (21).

(21)

$$cleft \Rightarrow \left[\begin{array}{l} \left[\begin{array}{l} CAT S[fin, ROOT +] \\ SS | LOC \left[\begin{array}{l} CONT \left[\begin{array}{l} QUANTS < \left[\begin{array}{l} the - rel \\ INDEX [1] \\ RESTR \{ [2] \} \end{array} \right] > \oplus L \\ NUCL \left[\begin{array}{l} identity - rel \\ ARG [3] \\ ARG [1] \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \\ DTRS < \left[\begin{array}{l} SS \left[\begin{array}{l} LOC | CONT [INDEX [3]] \\ WH \{ \} \end{array} \right] \end{array} \right], [SS | LOC | CONT [2]] > \end{array} \right]$$

This says that a cleft is a finite root clause whose first daughter is not an interrogative *wh*-phrase and the two daughters are interpreted as the two terms of an identity predication.

To complete this analysis, we need to say something about gaps. Bouma et al. (2001) propose that the *synsem* objects that encode the syntactic and semantic properties of linguistic expressions have a number of subtypes. In particular, there are *canonical-synsem* objects, which are realized as ordinary constituents, and *gap-synsem* objects, which are realized as gaps. I assume that nominal gaps are required to be non-pronominal and third person by the following constraints:

- (22) a. $\left[\begin{array}{l} gap \\ LOC | CAT [HEAD noun] \end{array} \right] \Rightarrow [LOC | CONT npro]$
 b. $[CONT npro] \Rightarrow [PERS third]$

The first requires a nominal gap to be non-pronominal. The second requires all non-pronominal NPs to be third person. Given these constraints, there will be a mismatch between the initial constituent and the gap whenever the initial constituent is not third person or is pronominal.

Here, then, we have an analysis which seems to capture the central properties of nominal clefts. Crucially it claims that the initial constituent is not a filler. Hence, we only have apparent filler–gap mismatches here. There is, however, more to be said about Welsh clefts. There are a number of additional types of apparent filler–gap mismatches, mainly involving *bod* ‘be’, which require a number of different mechanisms. There are situations where only a gap is possible. Then there are situations where the initial constituent undergoes a deletion process that makes it look like a different category. Finally, there are situations where a verb with a gap as its complement has a distinctive form.

3 Identity *bod*

I will look first look at what Borsley, Tallerman and Willis (2007: Section 4.4) call the identity copular construction. What we called quasi-pseudoclefts in the last section are an example of this construction. There are simpler examples, such as the following:

- (23) *Y meddyg ydy Sioned* ____.
 the doctor be.PRES.3SG Sioned
 ‘Sioned is the doctor.’

In the examples in (14) above, the post-verbal constituent is a complex NP containing a relative clause. In (23) it is a simple NP. I have indicated in (23) that the post-verbal constituent is a subject and that there is a gap in complement position. The agreement in an example like the following, from Zaring (1996: 130), provides evidence for this:

- (24) *Y tîm arall ydyn nhw* ____.
 the team other be.PRES.3PL they
 ‘They are the other team.’

Thus, what we have here are cleft sentences where the verb has an identity interpretation and there is a gap in complement position. Notice that given the analysis that I am proposing for cleft sentences there are in fact two identity

predications in these examples. In effect, the meaning of (23) is ‘It is the doctor that Sioned is’.

The initial constituent in this construction is most often a definite NP, but can also be an indefinite NP, a PP, or a VP, as the following show:

- (25) a. *Rhaff ydy 'r ateb* ____.
 rope be.PRES.3S the answer
 ‘The answer is a rope.’

(Jones and Thomas 1977: 49)

- b. *Yn yr ardd ydy 'r lle i fod ar ddiwrnod braf* ____.
 In the garden be.PRES.3SG the place to be on day fine
 ‘In the garden is the place to be on a fine day.’

- c. *Gweithio ydy beth mae Siôn yn*
 work be.PRES.3SG what be.PRES.3SG Siôn PROG
ei wneud ____.
 3SGM do.INF

‘What Siôn is doing is working.’ (‘Siôn is working.’)

(Zaring 1996: 134)

The (b) and (c) examples show that the construction is not limited to equating two individuals but can also equate two locations or two actions.⁹

As emphasized by Zaring (1996), the construction has two surprising properties. First, the copula has a surprising form. In ordinary declarative sentences the present tense of *bod* is *mae* with third person singular pronominal subjects and non-pronominal subjects singular or plural, and *maen* with third person plural pronominal subjects. In interrogative and conditional clauses, *ydy* appears with third person singular pronominal subjects and *ydyn* with third person plural pronominal subjects.¹⁰ With non-pronominal subjects, *ydy* appears if the subject is definite while *oes* appears if it is indefinite. Among other things, this means that we have contrasts like that between (26) and (27).

- (26) *Mae Sioned yn aros.*
 be.PRES.3S Sioned PROG stay.INF
 ‘Sioned is staying.’

⁹ Note that (25b) and (25c) are pseudo-clefts. These are often said to be “specificational” clauses, and they are sometimes seen as different from “equative” clauses such as (23), (24) and (25a) (see e.g. Mikkelsen 2011). In Welsh it seems clear that they are two examples of single type of clause. Heycock and Kroch (1999) argue that this is generally true.

¹⁰ There are differences with other persons, but the differences are clearest in the third person. See Thorne (1993: 248–251), King (1993: 145–146) for further information about the various forms of the copula.

- (27) a. *Ydy Sioned yn aros?*
 be.PRES.3S Sioned PROG stay.INF
 ‘Is Sioned staying?’
 b. *os ydy Sioned yn aros*
 if be.PRES.3S Sioned PROG stay.INF
 ‘if Sioned is staying’

Secondly, there is no possibility of an identity interpretation with a verb-initial clause with any form of the copula:

- (28) **Mae/Ydy Sioned y meddyg.*
 be.PRES.3SG Sioned the doctor
 ‘Sioned is the doctor.’

It seems, then, that the identity copular construction has some surprising properties. However, it is not too difficult to propose an analysis. Following Zaring (1996), I assume that Welsh has two different copula lexemes, identity *bod*, and predicational *bod*. This position is motivated by the very different syntactic properties that are associated with the two meanings.¹¹ We will be concerned with the properties of predicational *bod* in the next section. For identity *bod* we need to ensure that it only has a gap and not an overt constituent as its complement and that it has the right form. Ginzburg and Sag (2000) and Bouma et al. (2001) assume that *gap-synsem* objects appear in the ARG-ST (ARGUMENT-STRUCTURE) lists of words, which encode their basic combinatorial potential, but not in their COMPS lists, which indicate what complements they actually combine with. However, there is evidence from mutation (Borsley 1999) and agreement (Borsley 2009) that gaps in Welsh should be analyzed as empty categories. I assume, therefore, that *gap-synsem* objects appear in both ARG-ST lists and COMPS lists. Their special property is that signs that are *gap-synsem* are subject to the following constraint, which requires that they have no phonology:

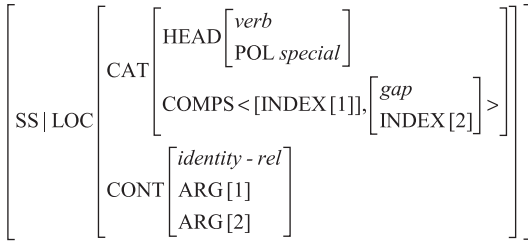
- (29) [SS[*gap*]] ⇒ [PHON <>]

I also assume, following Borsley (1989), that the post-verbal subjects of Welsh finite verbs are realizations of an extra member of the COMPS list. Finally, I

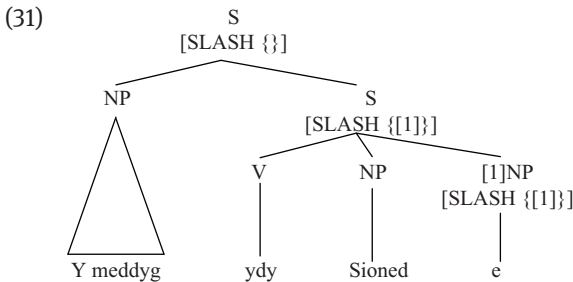
¹¹ Some have argued that English has a single copula, which can express both predicational, and identity meanings (see e.g. Partee 1986). Whatever may be the case in English, this position seems untenable in Welsh.

assume, that forms like *ydy* have the value *special* for a POL(ARITY) feature. The idea is that these are special forms distinct from the ordinary positive forms such as *mae* and certain negative forms.¹² Given these assumptions, we can specify the syntactic and semantic properties of identity *bod* as follows:¹³

(30) *Identity bod 'be'*



This ensures that identity *bod* has the appropriate form and takes a subject and a complement that is a gap. Given the properties in (30), (23) will have the following structure:



As noted above, there are two identity predications here on the analysis that I am proposing. However, there is just a single identity predication in the related *wh*-interrogative in (32).

(32) *Pwy ydy Sioned ___?*
 who be.PRES.3SG Sioned
 'Who is Sioned?'

¹² I derive this approach from Borsley and Jones (2005: Chapter 8), but they call the relevant value *int(errogative)-cond(itional)*. As an anonymous referee points out, this is potentially confusing given that these forms are not confined to interrogatives and conditionals.

¹³ I simplify here by writing just INDEX and not LOC|CONT|INDEX within the synsem objects in the value of COMPS.

It seems, then, that it is not difficult to accommodate the rather surprising properties of the identity copular construction. We just need an appropriate set of properties for identity *bod*.¹⁴ In the next two sections I will be concerned with some surprising properties of predicational *bod*.

4 Predicational *bod*

Predicational *bod* allows a number of types of complement. We can have a PP, as in (33).

- (33) *Mae Gwyn yn yr ardd.*
 be.PRES.3SG Gwyn in the garden
 ‘Gwyn is in the garden.’

We can also have what I will call a Perfect Phrase (PerfP), consisting of the perfective particle *wedi* and a VP, and what I will call a Progressive Phrase (ProgP), consisting of the progressive particle *yn* and a VP.¹⁵ The following illustrate:

14 One might wonder whether it is possible to have a nonfinite identity predication. Strictly speaking, the answer is no. All one can do is use predicational *bod* and rely on the context to convey an identity meaning. The following are relevant examples:

- (i) *Dw i 'n gobeithio bod yn gapten.*
 be.PRES.1SG I PROG hope.INF be.INF PRED captain
 ‘I hope to be a captain.’/‘I hope to be the captain.’
- (ii) *Dw i 'n disgwyl i Gwyn fod yn gapten.*
 be.PRES.1SG I PROG expect.INF for Gwyn be.INF PRED captain
 ‘I expect Gwyn to be a captain.’/ ‘I expect Gwyn to be the captain.’

15 In addition to these aspectual particles, Welsh has a number of aspectual particles that are homophonous with prepositions. In the following, we have what look like the prepositions *ar* ‘on’, *heb* ‘without’, and *am* ‘about’ (hence the glosses):

- (i) a. *Mae Rhiannon ar adael.*
 be.PRES.3S Rhiannon on leave.INF
 ‘Rhiannon is about to leave.’
- b. *Mae Rhiannon heb adael.*
 be.PRES.3S Rhiannon without leave.INF
 ‘Rhiannon has not left.’

- (34) *Mae Gwyn wedi darllen y llyfr.*
 be.PRES.3SG Gwyn PERF read.INF the book
 ‘Gwyn has read the book.’
- (35) *Mae Gwyn yn darllen y llyfr.*
 be.PRES.3SG Gwyn PROG read.INF the book
 ‘Gwyn is reading the book.’

Finally, we have what I will call a Predicative Phrase (PredP), consisting of the predicative particle *yn* and an AP or NP, as in the following:

- (36) *Mae Gwyn yn glyfar.*
 be.PRES.3SG Gwyn PRED clever
 ‘Gwyn is clever.’
- (37) *Mae Gwyn yn feddyg.*
 be.PRES.3SG Gwyn PRED doctor
 ‘Gwyn is a doctor.’

Predicative *yn* differs from progressive *yn* in triggering soft mutation. Thus, *glyfar* in (36) is the mutated form of *clyfar*, and *feddyg* in (37) is the mutated form of *meddyg*.

Predicational *bod* is rather like English *be* in its complement selection properties. The latter is assumed, e.g. by Warner (2000), to take complements which are [PRED+]. It seems reasonable to propose a similar analysis for predicational *bod*. I will assume then that the perfective, progressive, and predicative particles are heads which are [PRED+] and that certain prepositions are too. This will entail that all the phrases that can appear as complements of

-
- c. *Mae Rhiannon am adael.*
 be.PRES.3S Rhiannon for leave.INF
 ‘Rhiannon wants to leave.’

There is also one aspectual particle that looks like an adjective. In the following, we have what looks like the adjective *newydd* ‘new’:

- (ii) *Mae Rhiannon newydd adael.*
 be.PRES.3S Rhiannon new leave.INF
 ‘Rhiannon has just left.’

See Jones (2010: Chapter 9) for discussion.

predicational *bod* are [PRED +]. I assume that nonfinite verbs, adjectives, and nouns are all [PRED -]. This will ensure that the phrases they head cannot be complements of predicational *bod*, ruling out examples like the following:¹⁶

- (38) a. **Mae* *Gwyn ddarllen y llyfr.*
 be.PRES.3SG Gwyn read.INF the book
 b. **Mae* *Gwyn glyfar.*
 be.PRES.3SG Gwyn clever
 c. **Mae* *Gwyn feddyg.*
 be.PRES.3SG Gwyn doctor

This looks like a promising approach. However, before we accept it, there is an alternative that might be proposed within HPSG. Various elements rather like these particles have been analyzed as a type of non-head called a marker. For example, English complementizers are analyzed as markers in Pollard and Sag (1994: Section 1.6). Markers combine with a phrase to form a larger phrase that is identical in its feature makeup to the smaller phrase except for the value of the MARKING feature. The smaller phrase is [MARKING *none*], and the larger phrase has some other value which comes from the marker. Thus, markers appear in structures of the following form:

- (39)
- $$\begin{array}{c}
 \text{XP} \\
 \text{[MARKING [1]]} \\
 \diagdown \quad \diagup \\
 \text{[MARKING [1]]} \quad \text{XP} \\
 \quad \quad \quad \text{[MARKING none]}
 \end{array}$$

Applying this view to PerfPs, we would have structures of the following form:

- (40)
- $$\begin{array}{c}
 \text{VP} \\
 \text{[MARKING wedi]} \\
 \diagdown \quad \diagup \\
 \text{[MARKING wedi]} \quad \text{VP} \\
 \quad \quad \quad \text{[MARKING none]}
 \end{array}$$

However, there is a problem with this approach. If PerfPs and ProgPs are [PRED +], the VPs that they contain will also be [PRED +]. Similarly, if PredPs are [PRED +], the APs and NPs that they contain will also be [PRED +]. Hence, it will not be possible to rule out the examples in (38) by stipulating that

¹⁶ The complements in these examples are mutated with *ddarllen*, *glyfar*, and *feddyg* instead of *darllen*, *cllyfar*, and *meddyg*. Post-subject constituents are normally mutated. See Borsley et al. (2007: Chapter 7).

predicational *bod* takes a [PRED +] complement. It would be possible to get the right results by stipulating that predicational *bod* takes a complement which is either a PP or a VP, AP or NP with a certain MARKING value, but of course this is more complex than stipulating that it takes a complement which is [PRED +].

I think, then, that there are reasons for favoring an analysis in which the three particles are [PRED +] heads taking a [PRED -] complement. Of course, this doesn't say what exactly they are. One possibility is that they are nonstandard prepositions. On such an analysis they would need to be distinguished from ordinary prepositions and from each other by some features. Another possibility is that they are members of special categories each with a single member. This requires the feature HEAD, which encodes part of speech information, to have *perf*, *prog* and *pred* among its possible values. I will adopt this position in subsequent discussion. However, the analyses that I will outline would not be very different if the particles were nonstandard prepositions or heads of some other kind.

Assuming that the various complements of predicational *bod* are all [PRED +], it seems reasonable to suggest that finite forms of *bod* have the following COMPS feature:

$$(41) \left[\text{COMPS} < \text{NP}, \left[\begin{array}{l} \text{HEAD}[\text{PRED } +] \\ \text{SUBJ} < \square > \\ \text{COMPS} < > \end{array} \right] > \right]$$

Given the assumptions we are making about PRED, this will allow PerfPs, ProgPs, PredPs and PPs, but not VPs, APs and NPs.

Complications arise when predicational *bod* appears in a cleft sentence. We have examples with an initial PP or PerfP, but it seems no examples with an initial ProgP:

(42) *Yn yr ardd mae Gwyn ____.*
 in the garden be.PRES.3SG Gwyn
 'Gwyn is IN THE GARDEN.'

(43) *Wedi darllen y llyfr mae Gwyn ____.*
 PERF read.INF the book be.PRES.3SG Gwyn
 'Gwyn has READ THE BOOK.'

(44) **Yn darllen y llyfr mae Gwyn ____.*
 PROG read.INF the book be.PRES.3SG Gwyn
 'Gwyn is READING THE BOOK.'

It also seems that we can have a VP-initial constituent although a VP complement is not possible.

(45) *Darllen y llyfr mae Gwyn ____.*
 read.INF the book be.PRES.3SG Gwyn
 ‘Gwyn is READING THE BOOK.’

(46) **Mae Gwyn ddarllen y llyfr.*
 be.PRES.3SG Gwyn read.INF the book
 ‘Gwyn is reading the book.’

Finally, it seems that a PredP-initial constituent is not possible.

(47) **Yn glyfar mae Gwyn ____.*
 PRED clever be.PRES.3SG Gwyn
 ‘Gwyn is clever.’

(48) **Yn feddyg mae Gwyn ____.*
 PRED doctor be.PRES.3SG Gwyn
 ‘Gwyn is a doctor.’

What should we make of these facts? If we assume that the initial constituent in a cleft sentence must have the same category as the gap, the facts suggest that we have a number of contrasts between overt constituents and gaps as complements summarized in Table 1.¹⁷

Table 1: Complements of predicational *bod*: first version.

Complement	Overt constituent	Gap
PP	Yes	Yes
PerfP	Yes	Yes
ProgP	Yes	No
PredP	Yes	No
VP	No	Yes

¹⁷ That the initial constituent in a cleft sentence must have the same category as the gap might follow from the nature of identity predications. Alternatively it could be the result of a further stipulation on cleft sentences.

The COMPS feature in (41) will not allow a VP gap, and will allow a ProgP and a PredP gap. It looks, then, as if we require something more complex. The obvious suggestion is that we need the following COMPS feature:

$$(49) \left[\text{COMPS} < \text{NP}, \begin{bmatrix} \textit{canon} \\ \text{HEAD}[\text{PRED } +] \\ \text{SUBJ} < \square > \\ \text{COMPS} < > \end{bmatrix} \vee \begin{bmatrix} \textit{gap} \\ \text{HEAD } \textit{prep} \vee \textit{pref} \vee \textit{verb} \\ \text{SUBJ} < \square > \\ \text{COMPS} < > \end{bmatrix} > \right]$$

However, further data provides evidence for a somewhat different treatment.

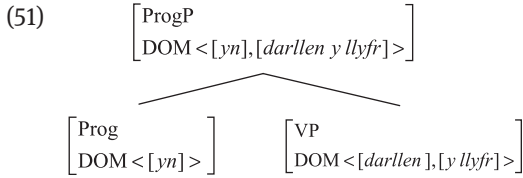
There are in fact certain examples with predicational *bod* and a ProgP-initial constituent. Consider, for example, the following:

- (50) a. *Wrthi yn golchi 'r car mae Mair* ____.
 at.3SGF PROG wash.INF the car be.PRES.3SG Mair
 'Mair is in the process of washing the car.'
 b. **Wrthi golchi 'r car mae Mair* ____.
 at.3SGF wash.INF the car be.PRES.3SG Mair

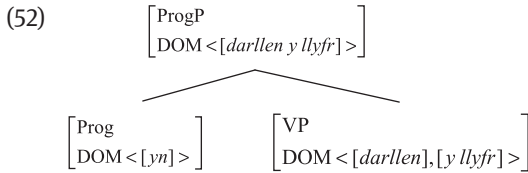
In (50a) the initial constituent contains progressive *yn*, and (50b) shows that this is obligatory. This might suggest that we have a VP-initial constituent under some circumstances and a ProgP-initial constituent under others. I want to propose, however, that we always have a ProgP-initial constituent but that under certain conditions the progressive particle *yn* is suppressed or deleted.

It is not difficult within HPSG to require the head of a phrase to be deleted under certain circumstances. Since Kathol (2000), much HPSG work has assumed that expressions have an order domain, which provides a basis for an account of word order facts among other things.¹⁸ Normally the domain elements of a constituent become elements in the order domain of the mother or are “compacted” to form a single element in the mother’s order domain. The latter is the norm in languages with a fixed word order. Order domains are encoded as the value of a feature DOM(AIN). If we use bracketed orthography to represent domain elements, we can give the following schematic analysis for the ProgP complement in (35):

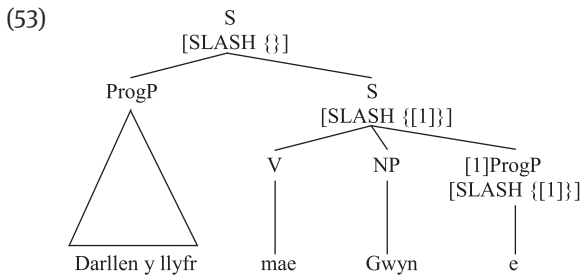
¹⁸ In Borsley (1999) I propose that Welsh mutation involves constraints on order domains, and in Borsley (2009) I argue that the same is true of Welsh agreement.



Notice that the two elements in the order domain of VP are compacted to form a single element in the order domain of ProgP. As discussed in Crysmann (2003), Beavers and Sag (2004) and Chaves (2006), deletion can be analyzed within this framework as a situation in which an element in the order domain of some expression is neither an element nor part of an element in the mother’s order domain. Adopting this approach, we can assign the following representation to the initial constituent in (45):



This sentence will then have the following structure:



Requiring representations like (52) under appropriate circumstances will account for the general apparent absence of ProgP-initial constituents with *bod* and the apparent appearance of VP-initial constituents.¹⁹

¹⁹ It may be that deletion within a filler is involved in English examples like the following, highlighted by Pullum (2009):

(i) *Good linguist though he is* ____, ...

A related example with an in-situ complement contains the indefinite article:

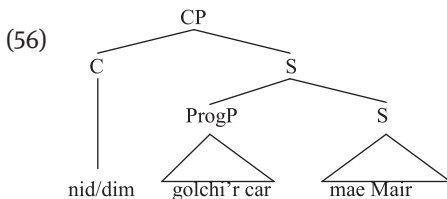
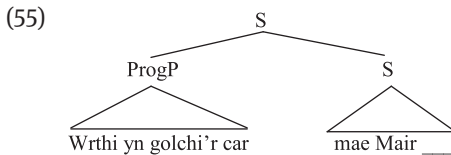
(ii) *He is a good linguist.*

Thus, there is an apparent filler-gap mismatch in (i). It may well be that deletion of *a* is responsible for this.

Of course, a full analysis needs to specify exactly when deletion occurs. This is quite complex. The contrast between (44) and (50a) might lead one to propose that *yn* is deleted in sentence-initial position. However, the following shows that it is also deleted after the negative particle *nid/dim*:

- (54) *Nid/dim* (**yn*) *golchi* 'r *car* *mae* *Mair* ____.
 NEG PROG wash.INF the car be.PRES.3SG Mair
 'It's not washing the car that Mair is doing.'

Why do *wrthi* and *nid/dim* differ in this way? One possibility is that the contrast reflects a structural difference. Tallerman (1996) suggests that *nid/dim* is a kind of complementizer. On this view it is separate from the initial phrase of the cleft. *Wrthi* is presumably part of the initial phrase. It looks, then, as if we might have two rather different structures as follows:



If these structures were viable, there would be an obvious account of the contrast between (50a) and (54). We could say that progressive *yn* is deleted just in case it is sentence-initial. However, the analysis in (56) seems rather dubious. It complicates the description of the word *mai* seen in the following examples:

- (57) a. *Dywedd* *Gwyn* *mai* (*nid/dim*) *Megan* *welish* *i* ____.
 say.PAST.3SG Gwyn COMP NEG Megan see.PAST.1SG I
 'Gwyn saw that it was (not) Megan that I saw.'
 b. *Credodd* *Emrys* *mai* (*nid/dim*) *draig* *welodd* *o* ____
 believe.PAST.3SG Emrys COMP NEG dragon see.PAST.1SG he
 'Emrys believed that it was (not) a dragon that he saw.'

I assume that *mai* is a complementizer (hence the gloss). It would be natural to say that it is a complementizer that introduces embedded cleft sentences. However, given the structure in (56), it would be necessary to say that it introduces either a cleft sentence or a CP containing a cleft sentence. This seems undesirable. It is likely, then, that (54) will have a structure rather like that of (50a). If this is right, it is not just in sentence-initial position that progressive *yn* is deleted.

However examples like (54) are analyzed, there is evidence that progressive *yn* is sometimes deleted when it is not in sentence-initial position. Consider first the following example, drawn to my attention by David Willis:

- (58) *Nid/dim draig ond unicorn welish i ____.*
 NEG dragon but unicorn see.PAST.1SG I
 ‘It was not a dragon but a unicorn that I saw.’

Here, it seems that *nid/dim* is part of the initial constituent of the cleft since the following material cannot appear without *nid/dim*:

- (59) **Draig ond unicorn welish i ____.*
 dragon but unicorn see.PAST.1SG I

In this sort of example, as in (45), we have what looks like a VP where a ProgP is expected. The following illustrates:

- (60) *Nid/dim sgrifennu ond darllen mae Mair ____.*
 NEG write.INF but read.INF be.PRES.3SG Mair
 ‘It is not writing but reading that Mair is doing.’

Notice that we actually have two instances of *yn* deletion in a non-sentence-initial position here.

Structures of the form *nid/dim X ond Y* are probably coordinate structures. It is natural to ask about more ordinary instances of coordination. Consider, then, the following example:

- (61) *Darllen a sgrifennu mae Gwyn ____.*
 read and write be.PRES.3SG Gwyn
 ‘Gwyn is reading and writing.’

In (61), as in (60), there are two instances of *yn* deletion. Here, however, only the second is not sentence-initial.

It seems, then, that progressive *yn* is deleted when the phrase it heads is in clause-initial position or is a conjunct of a coordinate structure in clause-initial position and is not preceded by certain adverbial elements (*wrthi* but *nid/dim*). It seems that we need some distinguishing feature here. If we call this feature F, we can say that the head of a phrase is deleted when it has the following description:

$$(62) \left[\text{HEAD} \left[\begin{array}{c} \textit{prog} \\ \text{F+} \end{array} \right] \right]$$

To ensure the correct distribution of [F+], we need firstly to ensure that clause-initial focused constituents but not other constituents are [F+]. Then we need to ensure that conjuncts have the same value for F as the coordinate structure. Finally we need to distinguish two situations where certain elements appear in initial position in an [F+] constituent. With *nid*, *dim*, *ond*, and *a* the constituent with which it combines must be [F+], whereas with *wrthi* (and some other elements that we will consider later) it must be [F-]. I won't try to decide exactly how exactly these things should be done since this would involve taking a stance on a variety of debatable issues. As far as I can see, however, there are no particular problems here.

If apparent VP-initial constituents with *bod* are really ProgP constituents, then rather than the situation summarized in Table 1 above we have the following somewhat simpler situation:

Table 2: Complements of predicational *bod*: second version.

Complement	Overt constituent	Gap
PP	Yes	Yes
PerfP	Yes	Yes
ProgP	Yes	Yes
PredP	Yes	No

However, there is evidence that this is still more complex than necessary.

We noted earlier that we do not seem to have cleft sentences with *bod* with a PredP-initial constituent. In other words, we do not seem to have cleft sentences related to (36) and (37) above. But we do have the following cleft sentences:

- (63) *Clyfar ydy Gwyn* ____.
 clever be.PRES.3SG Gwyn
 'Gwyn is clever.'

- (64) *Meddyg ydy Gwyn* ____.
 doctor be.PRES.3SG Gwyn
 ‘Gwyn is a doctor.’

These examples are surprising in two ways. First, they seem to have an AP- and an NP-initial constituent and not the PredP-initial constituent that we would expect. Secondly, they have the form of the copula that appears in interrogatives and conditionals and identity statements. I will consider how they should be analyzed in the remainder of this section.

The first point to note is that just as we find some examples with a ProgP-initial constituent, so we find some examples with a PredP-initial constituent. Consider, for example, the following:

- (65) a. *Bron yn barod ydy Mair* ____.
 almost PRED ready be.PRES.3SG Mair
 ‘Mair is almost ready.’
 b. **Bron parod ydy Mair* ____.
 almost ready be.PRES.3SG Mair
- (66) a. *Braidd yn siomedig ydy hi* ____.
 rather PRED disappointed be.PRES.3SG she
 ‘She is rather disappointed.’
 b. **Braidd siomedig ydy hi* ____.
 rather disappointed be.PRES.3SG she
- (67) a. *Bron yn fradychwr ydy o* ____.
 almost PRED traitor be.PRES.3SG he
 ‘He is almost a traitor.’
 b. **Bron bradychwr ydy o* ____.
 almost traitor be.PRES.3SG he

The (a) examples contain predicative *yn* and the (b) examples show that it is obligatory. I suggest, then, that examples like (63) and (64) involve a PredP-initial constituent where the predicative particle *yn* is deleted. As with progressive *yn*, the deletion applies not just in sentence-initial position but also after the negative particle *nid/dim*:

- (68) a. *Nid/dim parod ydy Mair* ____.
 NEG ready be.PRES.3SG Mair
 ‘Mair is not READY.’
 b. **Nid/dim yn barod ydy Mair* ____.
 NEG PRED ready be.PRES.3SG Mair

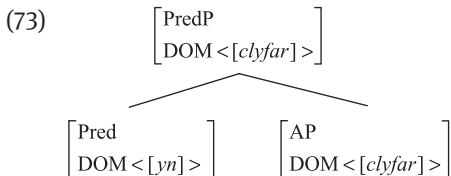
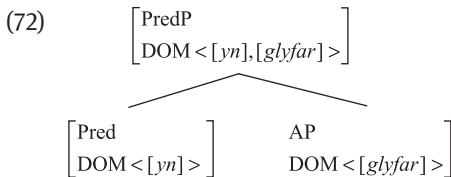
- (69) a. *Nid/dim siomedig ydy hi.*
 NEG disappointed be.PRES.3SG she
 ‘She is not DISAPPOINTED.’
- b. **Nid/dim yn siomedig ydy hi ____.*
 NEG PRED disappointed be.PRES.3SG she
- (70) a. *Nid/dim bradychwr ydy o ____.*
 NEG traitor be.PRES.3SG he
 ‘He is almost a traitor.’
- b. **Nid/dim yn fradychwr ydy o ____.*
 NEG PRED traitor be.PRES.3SG he

Also as with progressive *yn*, we have the deletion in both conjuncts of a coordinate structure:

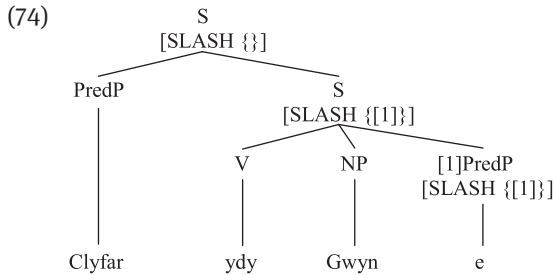
- (71) *Clyfar a chryf ydy Gwyn ____.*
 clever and strong be.PRES.3SG Gwyn
 ‘Gwyn is clever and strong.’

It looks, then, as if predicative *yn* is deleted under essentially the same conditions as progressive *yn*.

If we assume a deletion approach, while the PredP complement in (36) will have the schematic analysis in (72), the initial constituent in (63) will have the schematic analysis in (73).



(63) will then have the following structure:



Thus, (63) and (64) are another case where an initial constituent undergoes deletion, making it look like a different category.

One point to note about this analysis is that it correctly predicts the absence of mutation when predicative *yn* is deleted. In (72), *yn* and the adjective appear in the same order domain. Hence, we expect mutation. In (73), *yn* does not appear in the same order domain as the adjective. Hence, no mutation is expected.

Since predicative *yn* is deleted under essentially the same conditions as progressive *yn*, it is natural to propose essentially the same approach. It is natural, that is, to say that the head of a phrase is deleted when it has the following description:

(75) $\left[\text{HEAD} \begin{bmatrix} \textit{pred} \\ \text{F+} \end{bmatrix} \right]$

However, there is more to be said here.

As an anonymous referee has pointed out to me, there are examples where an initial PredP is understood as something other than the complement of predicational *bod*. Consider the following:

- (76) a. *Penodwyd ef yn arlywydd.*
 appoint.PAST.IMPERS him PRED president
 ‘He was appointed president.’
- b. *Yn arlywydd y penodwyd ef.*
 PRED president PRT appoint.PAST.IMPERS him
 ‘He was appointed PRESIDENT.’
- c. **Arlywydd y penodwyd ef.*
 president PRT appoint.PAST.IMPERS him
 ‘He was appointed PRESIDENT.’

- (77) a. *Daeth John adre 'n wlyb.*
 come.PAST.3SG John home PRED wet
 'John came home wet.'
- b. *Yn wlyb y daeth John adref.*
 PRED wet PRT come.PAST.3SG John home
 'John came home WET.'
- c. **Gwlyb y daeth John adref.*
 wet PRT come.PAST.3SG John home
 'John came home WET.'

In (76) a PredP is complement of the verb *penodi* 'appoint', and in (77) it is a secondary predicate. The (b) examples show that the PredP can appear in initial position in both cases, and the (c) examples show predicative *yn* cannot be deleted in these examples. Thus, we need some way to distinguish these examples from examples like (63) and (64). I will assume that a PredP is [PRED +] when it is the complement of predicational *bod*, but [PRED -] when it is the complement of some other verb or a secondary predicate. Then, if we assume that the initial constituent in a cleft sentence has the same value for PRED as the gap, we can account for the facts by saying that the head of a phrase is deleted when it has the following description:

$$(78) \left[\text{HEAD} \left[\begin{array}{l} \textit{pred} \\ \text{PRED+} \\ \text{F+} \end{array} \right] \right]$$

One might suppose that there would be examples with an initial ProgP that is understood as something other than the complement of predicational *bod*. If there were, and if they also did not allow the deletion of their head, we could extend this approach to progressive *yn*. However, there do not seem to be any relevant examples. Therefore, there is no need to add [PRED +] to the description in (62).

What, then, about the fact that we have *ydy* and not *mae* in (63) and (64)? We have already noted that *ydy* and not *mae* appears in interrogatives, conditionals, and identity statements. We just need to ensure that we have the same situation with predicational *bod* when it has a PredP gap as its complement. To do this we can propose the following constraint:

$$(79) \left[\begin{array}{l} \textit{predicational - bod} \\ \text{SS|LOC|CAT} \left[\begin{array}{l} \text{HEAD}[\text{TENSE } \textit{present}] \\ \text{COMPS} < [], \left[\begin{array}{l} \textit{gap} \\ \text{HEAD } \textit{pred} \end{array} \right] > \end{array} \right] \end{array} \right] \Rightarrow [\text{POL } \textit{special}]$$

This makes predicational *bod* with a PredP gap as its complement look like identity *bod*. Thus, (64) looks very similar to (23) above. However, on the analyses developed here, they are rather different. Whereas (23) has an initial NP, (64) has an initial PredP with a deleted head. As one might expect, it is possible to find examples that are ambiguous between an identity and a predicational interpretation. Consider, for example, the following, from Zaring (1996: 134):

- (80) *Anarferol ydy beth ydy Siôn* ____.
 unusual be.3SG what be.3SG Siôn
 ‘What Siôn is is unusual.’

This is ambiguous in the same way as the English translation. It may mean that the property that Siôn has is the property of being unusual (the identity interpretation) or that the property that Siôn has is unusual (the predicational interpretation).²⁰

We have now provided a fairly full account of apparent filler–gap mismatches with predicational *bod*. It involves three distinct mechanisms: (i) a deletion process affecting progressive *yn*, (ii) a deletion process affecting predicative *yn*, and (iii) the constraint in (79), which makes predicational *bod* with a PredP gap as its complement look like identity *bod*. This is quite complex, but the complexity seems justified. One point to note is that we now have the following very simple pattern of complement selection:

Table 3: Complements of predicational *bod*: final version.

Complement	Overt constituent	Gap
PP	Yes	Yes
PerfP	Yes	Yes
ProgP	Yes	Yes
PredP	Yes	Yes

This means that the simple COMPS feature in (41) is satisfactory after all.

²⁰ Note that the first of these interpretations is equivalent to the predicational interpretation ‘Siôn is unusual’.

5 Further phenomena

I will now consider two further examples of apparent filler–gap mismatches that arise not just with clefts but also with *wh*-interrogatives and relatives. Both are cases where a verb has a special form when a dependent is a gap. The first involves predicational *bod*. The second involves all transitive verbs.

The first of these phenomena is illustrated by the following examples:

- (81) a. *Gwyn sydd* ___ *yn canu*.
 Gwyn be.PRES.3SG PROG sing.INF
 ‘It’s Gwyn who is singing.’
 b. **Sydd* *Gwyn yn canu*.
 be.PRES.3SG Gwyn PROG sing.INF
 c. **Gwyn mae* ___ *yn canu*.
 Gwyn be.PRES.3SG PROG sing.INF

(81a) is a cleft sentence with a gap in subject position and the present tense of *bod* takes the form *sydd*. (81b) shows that this form cannot appear with an overt subject, and (81c) shows that the normal third person singular form *mae* cannot appear when there is a gap in subject position. The pattern that we see in (81) is also found in *wh*-interrogatives and relatives, as the following show:

- (82) a. *Pwy sydd* ___ *yn canu*?
 who be.PRES.3SG PROG sing.INF
 ‘Who is singing?’
 b. **Pwy mae* ___ *yn canu*?
 who bePRES.3SG PROG sing.INF
- (83) a. *y dyn sydd* ___ *yn canu*
 the man be.PRES.3SG PROG sing.INF
 ‘the man who is singing’
 b. **y dyn mae* ___ *yn canu*
 the man be.PRES.3SG PROG sing.INF

It is a fairly simple matter to account for this phenomenon. With identity *bod* and predicational *bod* with a PredP gap as its complement there is a set of special forms that appears. Here we are just concerned with a single special form. Hence we just need a constraint specifying the phonology of a present tense form of predicational *bod* when it has a gap as its subject. In other words, we need something like the following:

$$(84) \left[\begin{array}{l} \textit{predicational - bod} \\ \text{SS|LOC|CAT} \left[\begin{array}{l} \text{HEAD[TENSE } \textit{present}] \\ \text{COMPS} \langle [\textit{gap}], [] \rangle \end{array} \right] \end{array} \right] \Rightarrow [\text{PHON } \textit{sydd}]$$

Assuming that the constraint that is responsible for the normal realization of the present tense of predicational *bod* is a default constraint that can be overridden, this will handle the facts.

The other apparent filler-gap mismatch is more complex. It is illustrated by the following:

- (85) a. *Beth mae o 'n ei wneud ___?*
 what be.PRES.3SG he PROG 3SGM do.INF
 'What is he doing?'
- b. **Mae o 'n ei wneud rhywbeth.*
 be.PRES.3SG he PROG 3SGM do.INF something
 'He is doing something.'
- c. *Mae o 'n gwneud rhywbeth.*
 be.PRES.3SG he PROG do.INF something
 'He is doing something.'

The example in (85a) illustrates the fact, mentioned in Section 2 above, that a nonfinite verb with a gap as its object is preceded by a clitic agreeing with the gap. The examples in (85b) and (85c) show that a nonfinite verb with a non-pronominal NP as its object is not preceded by a clitic. One might suppose that the gap in (85a) is not a true gap but a phonologically null resumptive pronoun, and this is the conclusion that a number of researchers have reached (see Awbery 1977, Sadler 1988, and Rouveret 2002: 124). There are, however, reasons for rejecting this view. First, as emphasized in Willis (2000: 545), an overt resumptive pronoun is not possible in this position:

- (86) **Beth mae o 'n ei wneud o?*
 what be.PRES.3SG he PROG 3SGM do.INF he
 'What is he doing?'

Second, as noted in Borsley et al. (2007: 114), colloquial Welsh allows a third person singular masculine clitic to appear when the *wh*-phrase is feminine or plural. Thus, instead of the examples in (87), those in (88) may occur. (The third person singular masculine and feminine clitics are identical in form, but the former triggers soft mutation while the latter triggers aspirate mutation.)

- (87) a. *Pa gath ydych chi 'n ei phrynu ___?*
 which cat be.PRES.2PL you PROG 3SGF buy.INF
 'Which cat are you buying?'
 b. *Pa lyfre ydych chi 'n eu prynu ___?*
 which books be.PRES.2PL you PROG 3PL buy.INF
 'Which books are you buying?'
- (88) a. *Pa gath ydych chi 'n ei brynu ___?*
 which cat be.PRES.2PL you PROG 3SGM buy.INF
 b. *Pa lyfre ydych chi 'n ei brynu ___?*
 which books be.PRES.2PL you PROG 3SGM buy.INF

It seems clear, then, that examples like (85a) contain not a null resumptive pronoun but a true gap.²¹

We noted earlier that there is evidence that nominal gaps are non-pronominal. We have seen that a nonfinite verb does not have a clitic when its object is an overt non-pronominal NP. Thus, the appearance of a clitic in (85a) is surprising. In Borsley (2009), I propose that both agreement suffixes and clitics are realizations of an AGR(EEMENT) feature and that lexical heads are by default [AGR *none*]. Certain constraints override this and ensure that an agreement suffix or a clitic appears under certain conditions. In Borsley (2009), I propose that the main cases of agreement are the result of a constraint on order domains, and it seems reasonable to propose such a constraint here, as follows:

- (89)
$$\left[\begin{array}{l} \text{DOM} < \left[\begin{array}{l} \text{SS} | \text{LOC} | \text{CAT} | \text{HEAD} \left[\begin{array}{l} \textit{verb} \\ \text{VFORM } \textit{inf} \\ \text{AGR} [1] \end{array} \right] \end{array} \right], \\ \left[\begin{array}{l} \text{SS} \left[\begin{array}{l} \textit{gap} \\ \text{LOC} \left[\begin{array}{l} \text{CAT NP} \\ \text{CONT} [\text{INDEX} [2]] \end{array} \right] \end{array} \right] \end{array} \right], \dots > \end{array} \right]$$

$\Rightarrow [1] = [[2] \vee [3rd, \textit{sing}, \textit{masc}]]$

This requires the value of AGR on a nonfinite verb followed by a nominal gap to be either the person, number, and gender features of the gap's index or third person singular masculine. It is rather more complex than the constraint in (84), but that is because the facts are more complex.

²¹ There are, of course, questions about how resumptive pronouns should be analyzed. See Borsley (2010, 2013) for discussion.

6 Concluding remarks

I have now investigated a number of examples of apparent filler–gap mismatches and considered how they might be accounted for within HPSG. My main focus has been cleft sentences, where I have argued that the initial constituent is not a filler but one term of a hidden identity predication. As we have seen, however, there is much more to be said. We have one case where a word (identity *bod*) only allows a complement that is a gap. We have two cases where a deletion process conceals the identity of the initial constituent in a cleft sentence, making a ProgP look like a VP and a PredP look like an AP or an NP. Finally, there are three cases where a verb with a gap as a dependent has a special form, two cases involving *bod* and one involving all transitive verbs. The facts require a variety of constraints, but it is not difficult to accommodate them within HPSG.

What about other frameworks? It seems to me that it might well be possible to provide analyses within a transformational approach. It would presumably be possible to analyze cleft sentences as involving movement of an empty operator that is required to have the same category and, in the case of nominals, the same number and gender but not person as the clause-initial phrase. Identity *bod* would be no problem if one can stipulate that certain complements obligatorily undergo A'-movement. With predicational *bod* it would be necessary to require deletion to apply to certain constituents in Spec CP, which is presumably possible in a transformational approach. It would also be necessary to ensure that the present tense of predicational *bod* has a special form when a PredP complement is fronted. This is presumably not a problem. It would probably also be possible to handle the facts considered in the last section.

It looks, then, as if the Welsh data may be unproblematic for a transformational approach. However, it does look problematic for the Principles and Parameters view of language, at least if that is the position that grammatical systems are the result of setting a relatively small number of parameters.²² It seems most unlikely that the phenomena we have been concerned with here could be the product of setting parameters that have effects elsewhere. Rather, they look like the sort of idiosyncratic phenomena that Culicover (1999) calls “syntactic nuts”, which suggest that there must be more to the grammars of natural languages than parameter setting.

²² An anonymous referee asks if anyone still subscribes to this view, citing Baker (2001). It does appear that estimates of the numbers of parameters have been rising. Adger (2003: 16) asserted, “there are only a few parameters”, but Roberts and Holmberg (2006) suggest that “the correct figure” is “in the region of 50–100”. However, it seems unlikely that any plausible parameters could account for data discussed here.

Acknowledgments: I am grateful to Bob Morris Jones for valuable discussion and help with the data, and to Emyr Davies for help with the data. I have also benefited from discussion with Bob Levine and the members of the East Anglian Welsh Syntax Circle: Louisa Sadler, Ian Roberts, and especially David Willis, and from the comments of two anonymous referees. I am also grateful to Ewa Jaworska for editorial assistance. Any bad bits are my responsibility.

References

- Adger, David. 2003. *Core syntax: A minimalist approach*. Oxford: Oxford University Press.
- Akmajian, Adrian. 1970. On deriving cleft sentences from pseudo-cleft sentences. *Linguistic Inquiry* 1. 149–168.
- Awbery, Gwenllian. 1977. A transformational view of Welsh relative clauses. *Bulletin of the Board of Celtic Studies* 27. 155–206.
- Baker, Mark C. 2001. *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.
- Beavers, John & Ivan A. Sag. 2004. Coordinate ellipsis and apparent non-constituent coordination. In Stefan Müller (ed.), *Proceedings of the 11th International Conference on Head-driven Phrase Structure Grammar*, 48–69. Stanford, CA: CSLI Publications.
- Borsley, Robert D. 1989. An HPSG approach to Welsh. *Journal of Linguistics* 25. 333–354.
- Borsley, Robert D. 1999. Mutation and constituent structure in Welsh. *Lingua* 109. 263–300.
- Borsley, Robert D. 2008. On some Welsh unbounded dependency constructions. *Essex Research Reports in Linguistics* 57(4). 1–21.
- Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural Language & Linguistic Theory* 27. 225–265.
- Borsley, Robert D. 2010. An HPSG approach to Welsh unbounded dependencies. In Stefan Müller (ed.), *Proceedings of the 17th International Conference on Head-driven Phrase Structure Grammar*, 47–67. Stanford, CA: CSLI Publications.
- Borsley, Robert D. 2013. On the nature of Welsh unbounded dependencies. *Lingua* 133. 1–29.
- Borsley, Robert D. & Bob M. Jones. 2005. *Welsh negation and grammatical theory*. Cardiff: University of Wales Press.
- Borsley, Robert D., Maggie Tallerman & David Willis. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press.
- Bouma, Gosse, Robert Malouf & Ivan A. Sag. 2001. Satisfying constraints on extraction and adjunction. *Natural Language & Linguistic Theory* 19. 1–65.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Oxford: Blackwell.
- Chaves, Rui P. 2006. Coordination of unlikes without unlike categories. In Stefan Müller (ed.), *Proceedings of the 13th International Conference on Head-driven Phrase Structure Grammar*, 102–122. Stanford, CA: CSLI Publications.
- Chomsky, Noam. 1977. On *wh*-movement. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.
- Crysmann, Berthold. 2003. An asymmetric theory of peripheral sharing in HPSG: Conjunction reduction and coordination of unlikes. In Gerhard Jaeger, Paola Monachesi, Gerald Penn &

- Shuly Wintner (eds.), *Proceedings of FGVienna: The 8th Conference on Formal Grammar*, 45–64. Stanford, CA: CSLI Publications.
- Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory and language acquisition*. Oxford: Oxford University Press.
- Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford, CA: CSLI Publications.
- Heycock, Caroline & Anthony Kroch. 1999. Pseudocleft connectedness: Implications for the LF interface level. *Linguistic Inquiry* 30. 365–397.
- Jones, Bob M. 2010. *Tense and aspect in informal Welsh*. Berlin & New York: Mouton de Gruyter.
- Jones, Bob M. & Alan R. Thomas. 1977. *The Welsh language: Studies in its syntax and semantics*. Cardiff: University of Wales Press.
- Kathol, Andreas. 2000. *Linear syntax*. Oxford: Oxford University Press.
- Kayne, Richard. 1980. Extensions of binding and case-marking. *Linguistic Inquiry* 11. 55–91.
- King, Gareth. 1993. *Modern Welsh: A comprehensive grammar*. London: Routledge.
- Mikkelsen, Line. 2011. Copular clauses. In Claudia Maienborn, Klaus von Stechow & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 2, 1805–1829. Berlin & New York: De Gruyter Mouton.
- Partee, Barbara. 1986. Ambiguous pseudoclefts with unambiguous *be*. In Stephen Berman, Jae-Wong Choe & Joyce McDonough (eds.), *Proceedings of the North Eastern Linguistic Society 16*, 354–366. Amherst, MA: GLSA.
- Pollard, Carl & Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Pullum, Geoffrey K. 2009. Model-theoretic syntax and the implausibility of movement. Paper presented at the University of Essex, December.
- Roberts, Ian & Anders Holmberg. 2006. On the role of parameters in Universal Grammar: A reply to Newmeyer. In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz & Jan Koster (ed.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*, 538–553. Berlin & New York: Mouton de Gruyter.
- Rouveret, Alain. 2002. How are resumptive pronouns linked to the periphery? In Paul Pica & Johan Rooryck (eds.), *Linguistic variation yearbook 2*, 123–184. Amsterdam & Philadelphia: John Benjamins.
- Sadler, Louisa. 1988. *Welsh syntax: A government-binding approach*. London: Croom Helm.
- Tallerman, Maggie. 1996. Fronting constructions in Welsh. In Robert D. Borsley & Ian Roberts (eds.), *The syntax of the Celtic languages*, 97–124. Cambridge: Cambridge University Press.
- Thorne, David A. 1993. *A comprehensive Welsh grammar*. Oxford: Blackwell.
- Warner, Anthony R. 2000. English auxiliaries without lexical rules. In Robert D. Borsley (ed.), *The nature and function of syntactic categories*, 167–220. New York: Academic Press.
- Webelhuth, Gert. 2008. A lexical-constructional approach to ‘movement mismatches’. Paper presented at The Fifth International Conference on Construction Grammar. The University of Texas at Austin, 26–28 September 2008.
- Webelhuth, Gert. 2012. The distribution of *that*-clauses in English: An SBCG account. In Hans Boas & Ivan A. Sag (eds.), *Sign-Based Construction Grammar*, 203–227. Stanford, CA: CSLI Publications.
- Willis, David W. E. 2000. On the distribution of resumptive pronouns and *wh*-trace in Welsh. *Journal of Linguistics* 36. 531–573.
- Zaring, Laurie. 1996. Two ‘be’ or not two ‘be’: Identity, predication and the Welsh copula. *Linguistics and Philosophy* 19. 103–142.