

On the superficiality of Welsh agreement

Robert D. Borsley

Abstract A notable feature of Welsh is a number of agreement phenomena, all of which only occur with pronouns. Finite verbs agree with a following pronominal subject, prepositions agree with a following pronominal complement, and a particle which introduces non-finite clauses agrees with a following pronominal subject. Similarly, nouns have a preceding clitic agreeing with a following pronominal possessor, non-finite verbs have a preceding clitic agreeing with a following pronominal object, and what looks like the non-finite form of *bod* “be”, which introduces certain subordinate clauses, has a preceding clitic agreeing with a following pronominal subject. There seems to be a single phenomenon here. Approaches that involve an abstract level face problems and there is no evidence that the phenomenon involves an abstract level. It seems quite plausible to suggest that superficial linear order is crucial, and this idea can be implemented in a straightforward way within linearization-based Head-driven Phrase Structure Grammar (HPSG). The properties of agreement also entail that unexpressed noun phrases of various kinds must be represented in the superficial constituent structure and not just at some abstract level, contrary to the view of some frameworks.

Keywords Agreement · Linear order · Superficial phenomena · Welsh

1 Introduction

Agreement is a complex phenomenon, which has been discussed quite extensively; see e.g. Corbett (2006). Among the questions that arise about it is: What level of representation is relevant? It is widely assumed that agreement normally involves abstract levels. For Minimalism, it is one manifestation of the operation Agree. This applies in the course of the bottom-up construction of a sentence and later processes may modify the structures to which it applies in important ways. For Lexical Functional Grammar (LFG) agreement is the result of constraints on the f(unctional)-structure of a sentence, which may differ quite significantly from the superficial c(onstituent)-structure. For much work in Head-driven Phrase Structure Grammar (HPSG), agreement must refer to the lists that are the value of the feature ARG-ST (ARGUMENT-STRUCTURE). This is because the noun phrase that triggers agreement is only represented in such lists in certain kinds of sentence, e.g. null-subject sentences. (I use the term ‘noun phrase’ here and subsequently to remain neutral on whether the relevant expressions are NPs or DPs.)

It is fairly clear that abstract levels of structure are relevant for agreement in many languages. In English, for example, an important feature of agreement is that a verb agrees in the same way with a preceding subject and a following subject, as illustrated in (1) and (2).¹

¹ There are two sorts of exception to this generalization, exemplified by the following:

- (i) a. *I aren't clever.
b. Aren't I clever?
- (ii) a. A cow and a sheep were/*was in the field.

- (1) a. They are/*is here.
 b. Are/*is they here?
- (2) a. They have/*has been here.
 b. Have/*has they been here.

Within a transformational framework such as Principles and Parameters theory (P&P) or Minimalism, the obvious conclusion is that agreement applies before certain movement processes. Within a non-transformational approach such HPSG or LFG, the obvious conclusion is that agreement refers to a notion of subject, which is realized in more than one way. Another important feature of English agreement, illustrated in (3), is that a verb agrees in the same way with an in-situ subject and an ‘extracted’ subject.

- (3) a. The men know/*knows the answer.
 b. Which men do you think know/*knows the answer?

Again, this suggests within P&P/Minimalism that agreement applies before certain movement processes, and in LFG it suggests again that agreement refers to a notion of subject, which is realized in more than one way. In HPSG, it is assumed that subject in (3b) is an empty category or an unrealized element (a member of the ARG-ST list with no counterpart in constituent structure), which shares most features with *which men*.

In this paper, I will show that agreement in Welsh is rather different. I will argue that it involves superficial aspects of structure and not some kind of abstract structure. It is plausible to suggest that it involves linear order at a superficial level, and I will develop an analysis within HPSG embodying this assumption. I will also show that there is evidence here that unexpressed noun phrases of various kinds must be represented at a superficial level and not just at some abstract level, a position which is generally accepted within Principles and Parameters theory but which has been rejected in much work in HPSG and LFG.

The paper is organized as follows. In section 2, I introduce six kinds of agreement in Welsh with very similar properties. In section 3, I develop a more precise description of the data, and in section 4, I consider the implications of the data for various approaches. In section 5, I look at the implications of coordination, and in section 6, I discuss A'-movement sentences. In section 7, I show how the idea that Welsh agreement involves superficial linear order can be implemented within linearization-based HPSG. In section 8, I consider the implications of the analysis for the treatment of unexpressed noun phrases. Finally, in section 9, I summarize the paper.

b. There was/*were a cow and a sheep in the field.

It is clear, however, that these are exceptions.

2 The basic data

We turn now to the Welsh data.² Much of this has been discussed before, but there is more to be said. I will show that Welsh has six types of agreement with very similar properties. This suggests that they are in fact different manifestations of a single agreement phenomenon.

Welsh, like the other Celtic languages, is a VSO language with either a verb or an auxiliary preceding the subject in finite clauses. Finite verbs and auxiliaries agree with a following pronominal subject.³ The examples in (4) illustrate this for verbs.

- (4) a. Gwelodd e/hi ddraig.
see.PAST.3SG he/she dragon
 “He saw a dragon.”
 b. Gwelon nhw ddraig.
see.PAST.3PL they dragon
 They saw a dragon.

With a non-pronominal subject, singular or plural, the third person singular form of the verb appears. Thus, we have the following data:

- (5) Gwelodd y bachgen/bechgyn ddraig.
see.PAST.3SG the boy boys dragon
 The boy/boys saw a dragon.
 (6) *Gwelon y bechgyn ddraig.
see.PAST.3PL the boys dragon

There are two views one take of this datum. One might assume that a special mechanism ensures that a finite verb with a non-pronominal subject has the same agreement features as a finite verb with a third person singular pronominal subject. Alternatively one might assume that finite verbs have a form with no agreement features with a non-pronominal subject but that this is identical in form to the third person singular form. This approach would make finite verbs like other heads considered below which have a form with no agreement features. Thus, finite verbs show agreement with a pronoun but either a special kind of agreement or no agreement at all with a non-pronominal noun phrase.

We turn now to prepositions. Most prepositions agree with a following pronominal object, as is illustrated in (7).⁴

² As discussed in Borsley, Tallerman, and Willis (2007: chapter 1), Welsh shows a complex stylistic continuum, ranging from a very literary style to very colloquial styles varying from region to region. Following Borsley, Tallerman, and Willis, I focus on neutral syntactic patterns and phonological and morphological forms, which are not confined to speech and writing or to specific geographic areas.

³ For full paradigms for each of the agreement phenomena discussed here see Borsley, Tallerman, and Willis (2007: chapter 6).

⁴ As is implicitly indicated here, Welsh has some prepositions which do not inflect, e.g. *gyda* ‘with’. One might try to ensure that such prepositions have no agreement features. Alternatively, one might assume that they have agreement features just like inflecting prepositions but have the same phonological form whatever their agreement features. I will not try to choose between these two approaches.

- (7) a. arno fo
 on.3SGM he
 on him
 b. arni hi
 on.3SGF she
 on her
 c. arnyn nhw
 on.3PL they
 on them

Notice that, unlike with finite verbs, inflecting prepositions have separate masculine and feminine third person singular forms. The basic uninflected form of the preposition *ar* appears with a non-pronominal object. Thus, we have (8) and not (9).

- (8) ar y bachgen/yr eneth/y bechgyn
 on the boy the girl the boys
 on the boy/the girl/the boys

- (9) a. *arno 'r bachgen
 on.3SGM the boy
 b. *arni 'r eneth
 on.3SGF the girl
 c. *arnyn y bechgyn
 on.3PL the boys

Here, then, we have agreement with a pronoun but no agreement with a non-pronominal noun phrase.

A similar pattern of agreement is found in a very common type of subordinate clause which resembles an English *for-to* clause. This is introduced by what looks like the preposition *i* “to, for”, analyzed as a prepositional complementizer in Borsley (1986). This element agrees with a following pronominal subject:⁵

- (10) a. Disgwyliodd Emrys [iddo fo fynd i Fangor].
 expect.PAST.3SG Emrys to.3SGM he go to Bangor
 Emrys expected him to go to Bangor.
 b. Disgwyliodd Emrys [iddi hi fynd i Fangor].
 expect.PAST.3S Emrys to.3SGF she go to Bangor
 Emrys expected her to go to Bangor.
 c. Disgwyliodd Emrys [iddyn nhw fynd i Fangor].
 expect.PAST.3S Emrys to.3PL they go to Bangor
 Emrys expected them to go to Bangor

There is no agreement with a following non-pronominal subject, as (11) and (12) show:

⁵ Clause-initial *i* only shows visible agreement with a third person pronouns. In this it is like the homophonous preposition

- (11) Diswyliodd Emrys i 'r bachgen/eneth/bechgyn fynd i Fangor.
expect.PAST.3SG Emrys to the boy girl boys go to Bangor
 Emrys expected the boy/girl/boys to go to Bangor.
- (12) a. *Diswyliodd Emrys iddo 'r bachgen fynd i Fangor.
expect.PAST.3SG Emrys to.3SGM the boy go to Bangor
 b. *Diswyliodd Emrys iddi 'r eneth fynd i Fangor.
expect.PAST.3SG Emrys to.3SGF the girl go to Bangor
 c. *Diswyliodd Emrys iddyn y bechgyn fynd i Fangor.
expect.PAST.3SG Emrys to.3PL the boys go to Bangor

Again, then, there is agreement with pronouns but not with a non-pronominal noun phrase.

Consider next nouns. When a noun is followed by a pronominal possessor, it is preceded by an agreeing clitic:

- (13) a. ei dad o
3SGM father he
 his father
 b. ei thad hi
3SGF father she
 her father
 c. eu tad nhw
3PL father they
 their father

The third person singular masculine and feminine are homophonous but differ in that the former triggers soft mutation while the latter triggers aspirate mutation. There is no clitic with a non-pronominal possessor:

- (14) tad y bachgen/bechgyn
father the boy boys
 the boy's/boys' father

- (15) a. *ei dad y bachgen
3SGM father the boy
 b. *ei thad yr eneth
3SGF father the girl
 c. *eu tad y bechgyn
3PL father the boys

Once more, then, we see agreement with pronouns but not with non-pronominal noun phrases.

Non-finite verbs behave in much the same way as nouns. They are preceded by the same clitics when there is a pronominal object:⁶

⁶ Non-finite verbs are traditionally known as verb-nouns. It is this similarity that is the main reason for this label. See below and Borsley (1993, 1997) and Borsley Tallerman and Willis (2007: 3.1) for relevant discussion.

- (16) a. Gwnaeth Emrys ei weld o.
do.PAST.3SG Emrys 3SGM see he
 Emrys saw him.
- b. Gwnaeth Emrys ei gweld hi.
do.PAST.3SG Emrys 3SGF see she
 Emrys saw her.
- c. Gwnaeth Emrys eu gweld nhw.
do.PAST.3SG Emrys 3PL see they
 Emrys saw them.

There is no clitic with a non-pronominal object:

- (17) Gwnaeth Emrys weld y bachgen/bechgyn.
do.PAST.3SG Emrys see the boy boys
 Emrys saw the boy/boys.
- (18) a. *Gwnaeth Emrys ei weld y bachgen.
do.PAST.3SG Emrys 3SGM see the boy
- b. *Gwnaeth Emrys eu gweld y bechgyn.
do.PAST.3SG Emrys 3PL see the boys

Again, then, there is agreement with pronouns but not with non-pronominal noun phrases.

Finally we consider what look like non-finite subordinate clauses introduced by the verb *bod* ‘be’. Tallerman (1998) and Borsley, Tallerman, and Willis (2007: 3.3) show that there is evidence that these clauses are probably finite, but this is not particularly important in the present context.⁷ In these clauses, *bod* is preceded by an agreeing clitic when it has a following pronominal subject:

- (19) a. Dywedodd Gwyn ei fod o yn ddiog.
see.PAST.3SG Gwyn 3SGM be he PRED lazy
 Gwyn said he was lazy.

⁷ It seems that *bod* in these examples is a special present or imperfect form. The normal present forms do not appear in affirmative complement clauses, and for many speakers this is also true of the normal imperfect forms. Thus, (i) is ungrammatical for most speakers and (ii) is ungrammatical for many.

- (i) *Mae Aled yn credu [y mae Elen yn darllen y llyfr].
be.PRES.3SG Aled PROG believe PRT be.PRES.3SG Elen PROG read the book
 Aled believes that Elen is reading the book.
- (ii) %Mae Aled yn credu [roedd Elen yn darllen y llyfr].
be.PRES.3SG Aled PROG believe be.IMPF.3SG Elen PROG read the book
 Aled believes that Elen was reading the book.

The normal forms do appear in negative complements, as the following illustrate:

- (iii) Mae Aled yn credu [nad ydy Elen yn darllen y llyfr].
be.PRES.3SG Aled PROG believe NEG be.PRES.3SG Elen PROG read the book
 Aled believes that Elen is/was reading the book.

Notice, however, that the copula has different forms in a negative clause.

- b. Dywedodd Gwyn ei bod hi yn ddiog.
see.PAST.3SG Gwyn 3SGF be she PRED lazy
 Gwyn said she was lazy.
- c. Dywedodd Gwyn eu bod nhw yn ddiog.
see.PAST.3SG Gwyn 3PL be they PRED lazy
 Gwyn said they were lazy.

There is no clitic with a non-pronominal subject:

- (20) Dywedodd Gwyn bod y bachgen/yr eneth/y bechgyn yn ddiog.
see.PAST.3SG Gwyn be the boy the girl the boys PRED lazy
 Gwyn said the boy/the girl/the boys was/were lazy.
- (21) a. *Dywedodd Gwyn ei fod y bachgen yn ddiog.
see.PAST.3SG Gwyn 3SGM be. INF the boy PRED lazy
- b. *Dywedodd Gwyn ei bod yr eneth yn ddiog.
see.PAST.3SG Gwyn 3SGF be the girl PRED lazy
- c. *Dywedodd Gwyn eu bod y bechgyn yn ddiog.
see.PAST.3SG Gwyn 3PL be the boys PRED lazy

Yet again, there is agreement with pronouns but not with non-pronominal noun phrases.

There is one further situation in which a clitic may appear, illustrated by the following:

- (22) Fe 'u cerais i nhw.
PRT 3PL love.PAST.1SG I they
 I loved them.

Here an enclitic attached to a pre-verbal particle is associated with the object of a finite verb. The enclitics that are used here are similar but not identical to those which appear with nouns and non-finite verbs, discussed in the next section. Unlike the other phenomena that we have looked at, this phenomenon is confined to a very literary variety of Welsh. It also differs formally from the other phenomena in two ways. Firstly, it is optional. Literary Welsh also allows the following, with no clitic:

- (23) Fe gerais i nhw.
PRT love.PAST.1SG I they
 I loved them.

Secondly, the clitic is not associated with the nearest following noun phrase, which is the subject *i* and not the object *nhw*. For these reasons I assume that this is one of a number of separate agreement phenomena.

With the exception of clitics associated with objects of finite verbs, the agreement phenomena outlined above are clearly very similar to each other. The similarities are listed in (24).

- (24) a. All involve agreement with pronouns, but not full noun phrases.
 b. All are obligatory.
 c. In all cases the pronoun follows the realization of agreement.

The obvious conclusion is that they are all manifestations of a single phenomenon. This was first noted by McCloskey and Hale (1984: 520)⁸, and a number of other researchers have come to the same conclusion, including Sadler (1988: 104), who remarks that agreement morphemes and clitics are “essentially the same phenomenon”, Roberts and Shlonsky (1996: 184), who conclude that Welsh has a “single agreement system”, Pollard and Sag (1994: 9.3), and Rouveret (1994). Obviously if this is right, a unified account is desirable.

3 Towards a description of the data

Having argued in the last section that the phenomena surveyed there are different aspects of a single phenomenon, I want now to consider how exactly the facts should be described. This will depend on how inflections and clitics are analyzed and also on the analysis of attributive adjectives.

3.1 Inflections and clitics

Within P&P both inflections and clitics would be analyzed as the realization of a functional head. This position is developed in Rouveret (1994: chapter 2), Roberts and Shlonsky (1996), and Roberts (2005). For Rouveret, finite verb and preposition inflections are the realization of agreement heads, while the clitics are realizations of D before a noun and a non-finite verb other than *bod* “be”, and a realization of C before *bod*. However, the idea that clitics before a non-finite verb are a realization of D is quite problematic. On standard assumptions, it entails that non-finite VPs are embedded in a DP but, as noted in Borsley (1993), DPs and non-finite VPs behave rather differently. For example, they differ with respect to binding. In (25), *Emrys* and *o* can refer to the same person, but this is not possible in (26).

(25) Gwelodd Emrys ei wraig o.
see.PAST.3SG Emrys 3SGM wife he
 Emrys saw his wife.

(26) Ceisiodd Emrys ei weld o.
try.PAST.3SG Emrys 3SGM see he
 Emrys tried to see him.

The natural conclusion from these examples is that *Emrys* and *o* are separated by a DP boundary in (25) but not in (26). DPs and non-finite VPs also differ with respect to *wh*-movement. When the possessor in a DP is a *wh*-element, the whole DP is fronted. Thus, (27) is fully acceptable while (28) is very marginal.

(27) Gwraig pw y welaist ti?
wife who see. PAST.2SG you.SG
 Whose wife did you see?

⁸ McCloskey and Hale’s main concern is Irish, but they also look briefly at Welsh (see section 6.5 below).

(28) ??Pwy welaist ti ei wraig?
who see.PAST.2SG you.SG 3SGM wife

In contrast, when the object in a non-finite VP is a *wh*-element, only the object is fronted, as in (29), and not the whole VP, as in (30).

(29) Pwy geisiaist ti ei weld?
who try.PAST.2SG you.SG 3SGM see
 Who did you try to see?

(30) ?*Gweld pwy geisiaist ti?
see who try.PAST.2SG you.SG

Again this suggests that there is no DP in the case of a non-finite VP. Thus, the idea that pre-verbal clitics are a realization of D is dubious. The obvious alternative is that these elements are the realization of some other functional element, and in fact Roberts (2005: chapter 3) proposes that they are realizations of an Agr head.⁹

Although there are questions as to what functional categories are relevant, it is fairly clear that agreement will involve the following configuration within P&P:

(31) Functional Head [Lexical Head ... Pronoun ...]

The pronoun will be the first noun phrase following the lexical head. Hence, we have the following generalization about agreement:

(32) An agreeing element agrees with the first following noun phrase if and only if the latter is a pronoun.

What about frameworks which assume less abstract structures, such as LFG and HPSG? Here it is natural to assume that inflections are the realization of features on the heads to which they are attached, and that they involve the following configuration, where the head is a finite verb, a preposition, or clause-initial *i*:

(33) Head ... Pronoun ...

In these frameworks, as in P&P, one might assume that clitics are the realization of functional elements. Alternatively, one might assume that they are just morphological elements which realize agreement on nouns, non-finite verbs and *bod*. This position is developed in Pollard and Sag (1994: 357), who propose that clitics are prefixes on these elements (McCloskey and Hale (1984: 512) make the same assumption). It is fairly clear, however, that clitics are not ordinary prefixes since in the nominal case they can be separated from the associated noun by a numeral or an adjective, as the following illustrate:

⁹ For both Rouveret and Roberts the clitics are assigned to a number of different categories. On the face of it, this position makes it rather surprising that they have the same form in all cases. This looks like a weakness in these approaches.

- (34) a. ei dri llyfr (o)
 3SGM three book he
 his three books
 b. ei hen lyfr (o)
 3SGM old book he
 his old book

Moreover, if the preceding word ends in a vowel, they may be replaced by enclitics attached to this word, as in the following:¹⁰

- (35) i 'w dy (o)
 to 3SGM house he
 to his house
- (36) i 'w weld (o)
 to 3SGM see he
 to see him

One might propose, however, that the clitics are what Anderson (1992) calls phrasal affixes, affixes which are attached not to a word but to a phrase, giving a structure of the following form for (34a):

- (37)
-
- ```

graph TD
 A[] --- B[ei dri llyfr]
 A --- C[o]
 B --- D[ei]
 B --- E[dri llyfr]
 style A fill:none,stroke:none
 style B fill:none,stroke:none
 style C fill:none,stroke:none
 style D fill:none,stroke:none
 style E fill:none,stroke:none

```

On this view, agreement involves the configuration in (33) in all cases, and the head may be a finite verb, a preposition, clause-initial *i*, a noun (possibly combined with a numeral or a pre-nominal adjective), a non-finite verb, or *bod*.

### 3.2 Attributive adjectives

If clitics are morphological elements realizing agreement on nouns, non-finite verbs and *bod*, the agreeing head and the pronoun with which it agrees are generally adjacent. This is the case with finite verbs, prepositions, *i*, non-finite verbs and *bod*. The one apparent exception is nouns, where examples like the following occur:

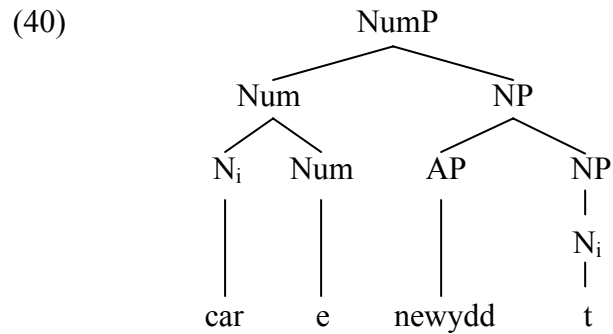
- (38) ei gar newydd o  
       3SGM car new he  
       his new car

Attributive adjectives normally follow the noun and precede a possessor. Hence, when the possessor is pronominal, as here, we apparently have agreement with a pronoun which does not immediately follow the agreeing element.

<sup>10</sup> First and second person singular enclitics only appear with certain lexical items and belong to a rather literary style. Other enclitics are more generally available. See Borsley, Tallerman, and Willis (2007: 5.2.2).

This is the right conclusion if we accept the proposal of Rouveret (1994: chapter 3) and Roberts (2005: chapter 3) that adjectives are left-adjoined to an NP whose head is extracted. On this approach, (39) has the structure in (40).<sup>11</sup>

- (39) *car newydd*  
*car new*  
 a new car



However, the evidence for such an analysis is not strong (see Willis 2006 and Borsley, Tallerman, and Willis 2007, chapter 5). Rouveret (1994) argues that it explains the order of adjectives in an example like (41).

- (41) *cup mawr gwyrdd Sieineaidd*  
*cup big green Chinese*  
 a big green Chinese cup

Here the order of the adjectives is the same as in English, where it is plausible to assume that adjectives are adjoined to a following NP. However, the examples in (42)–(43) show that the order of adjectives is not always the same as in English.

- (42) *athro ifanc hoffus*  
*teacher young likeable*  
 a likeable young teacher

- (43) *bardd ifanc addawol*  
*poet young promising*  
 a promising young poet

In the absence of good evidence for a left-adjunction analysis, it is reasonable to consider an alternative in which adjectives are adjoined to a preceding noun, forming a complex nominal constituent.<sup>12</sup>

Some evidence for such an analysis comes from what is known as mutation, systems of word-initial consonant alternations, which are a prominent feature of

<sup>11</sup> It is not literally possible to have an extracted head within LFG and HPSG, but mechanisms have been proposed in both frameworks which allow a head to appear outside the associated maximal projection. See Bresnan (2001: 7.1) for LFG and Borsley (1989b) for HPSG.

<sup>12</sup> An analysis of this kind is proposed in Samvelliian (2007) for Persian attributive adjectives, which also appear between the noun and a following possessor.

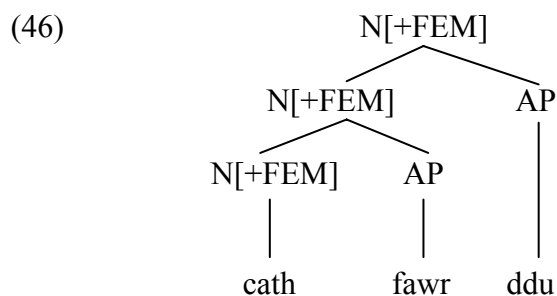
Welsh. Three types of mutation are traditionally identified in Welsh: soft mutation, aspirate mutation and nasal mutation. The first is by far the most common and occurs in many contexts.<sup>13</sup> What is important here is that an adjective undergoes soft mutation after a feminine singular noun. This is illustrated in (44), where the mutated adjective is in bold and the basic form in parentheses.

- (44) cath **fawr** (mawr)  
*cat big*  
 a big cat

A second adjective is also mutated.

- (45) cath **fawr ddu** (mawr, du)  
*cat big black*  
 big black cat

This is not surprising if adjectives are right-adjoined, giving structures like the following:



Here, the second adjective follows a feminine singular nominal element just as much as the first, and so the mutation is only to be expected. Thus, there is some evidence that a noun and a following adjective or adjectives form a constituent.

If this analysis is right, we can assume that examples like (38) involve agreement between a complex nominal element and an immediately following pronoun. This will mean that all agreement involves the following configuration, where the head if nominal may be complex:

- (47) Head Pronoun ...

This gives us the following generalization:

- (48) An agreeing element agrees with an immediately following noun phrase if and only if the latter is a pronoun.

<sup>13</sup> Soft mutation involves the following alternations:

- |            |             |                         |
|------------|-------------|-------------------------|
| (i) p > b  | b > f([v])  | m > f([v])              |
| t > d      | d > dd([ð]) | ll [H] > l              |
| c([k]) > g | g > ∅       | rh[r <sup>h</sup> ] > r |

### 3.2 Summary

It seems, then, that either (32) or (48) is a correct generalization about Welsh agreement. (32) is correct if inflections and clitics are the realization of functional heads. However, if they are the realization of features on lexical heads and attributive adjectives are adjoined to a preceding noun, (48) may be correct. We will consider some further data in sections 5 and 6, but as we will see, whichever is the correct generalization holds without exception. It is not just the norm.

## 4. Preliminary conclusions

Having identified two possible generalizations about the data, we can now draw some preliminary conclusions. A satisfactory analysis does not necessarily need to encode one of the generalizations directly, but in the absence of any reason to think that the generalizations are spurious, it should provide a single generalization covering the same ground. I will argue here that there are problems for approaches involving certain kinds of abstract structure.

### 4.1 Lexical Functional Grammar

We can begin with an approach to agreement based on grammatical function (GFs) of the kind that is assumed in LFG. Here, we have the situation in Table 1 if we assume that possessors are subjects, which is plausible in at least some cases. We would have a more complex picture if some possessors are not subjects.<sup>14</sup> As we see, some heads agree with a subject and others with an object. Essentially we have two sorts of agreement, and it is accidental that they have various properties in common. On the face of it, the facts would be no more complex if finite verbs agreed with an object or non-finite verbs with a subject. We would just have two rather different sorts of agreement. There seems to be no possibility of formulating a single generalization here. It has often been argued that approaches which do not allow reference to grammatical functions miss generalizations, for example about passive sentences. In the present area, it seems that an approach referring to grammatical functions misses a generalization. In the absence of any evidence that the generalization is spurious, this is a serious problem.

**Table 1** Agreement and grammatical functions

| Agreeing head           | GF of relevant noun phrase |
|-------------------------|----------------------------|
| Finite verb             | Subject                    |
| Preposition             | Object                     |
| Clause-initial <i>i</i> | Subject                    |
| Noun                    | Subject                    |
| Non-finite verb         | Object                     |
| <i>Bod</i>              | Subject                    |

<sup>14</sup> As noted below, Borsley (1999) and Borsley and Jones (2005: chapter 8) argue that *i* takes the following subject and predicate as two separate complements. If this analysis is adopted, *i* is another element which agrees with an object.

An anonymous referee has suggested to me that one might propose within LFG that agreement is with “the highest available GF”. The idea is that agreement is with a subject except where none is ‘available’, in which case it is with an object. The question, of course, is what it means for a GF to be available. A non-finite verb has a subject in f-structure just like a finite verb. Hence, a characterization of availability must refer to c-structure. One might suggest that a GF is available if and only if it is realized in c-structure. However, this is not very promising. Consider first (10a), repeated here as (49).

- (49) Disgwylodd Emrys [iddo fo fynd i Fangor].  
*expect.PAST.3SG Emrys to.3SGM he go to Bangor*  
 Emrys expected him to go to Bangor.

Here, *fo* is the subject of the non-finite verb *fynd*. Hence, *fynd* has a realized subject, but it does not agree with it. Consider also (16a), repeated here as (50).

- (50) Gwnaeth Emrys ei weld o.  
*do.PAST.3SG Emrys 3SGM see he*  
 Emrys saw him.

Within some frameworks it might be correct to say that the subject of the non-finite verb *weld* is unrealized. However, within LFG, this would be a case of functional control, in which a single expression is the subject of two verbs. On this approach *Emrys* is the subject of both the finite auxiliary *gwnaeth* and the non-finite verb *weld*. Thus, *weld* has a realized subject, and it is problematic that it does not agree with it. Within LFG, there are two sorts of examples in which a subject is unrealized: cases of control which LFG classifies as anaphoric control, where an unexpressed subject is coindexed with a controller, and null-subject sentences. A non-finite verb does not agree with its subject in any kind of control sentence. However, as we will see in section 8, the unrealized subject of a null-subject sentence must be visible to agreement. Thus, it is doubtful whether there is a viable notion of availability. It is worth adding that any attempt to develop such a notion accepts that the superficial structure of sentences is crucial for Welsh agreement, which is a central theme of the present paper.

#### 4.2 Head-driven Phrase Structure Grammar

A rather different GF-based approach to Welsh agreement is outlined in Pollard and Sag (1994: chapter 9) within HPSG. Central to HPSG are the valence features SUBJ and COMPS. The former normally indicates what kind of subject a head requires and the latter normally indicates what complements a head takes. Pollard and Sag propose that Welsh agreement is with the first member of a COMPS list. Objects of non-finite verbs and prepositions are realizations of the first member of a COMPS list. Following Borsley (1989a), Pollard and Sag propose that subjects of finite clauses, and *bod*-clauses, and possessors within noun phrases are the realization not of the single member of the SUBJ list but of the first member of a COMPS list.<sup>15</sup> Thus, finite verbs and *bod* have categories of the form in (51), where the NP is the subject,

<sup>15</sup> Borsley (1989) argues for this position partly on the basis of agreement, and partly on the grounds that it automatically accounts for the fact that these elements follow the associated head.

and possessed nouns have categories of the form in (52), where the NP is the possessor. (I adopt the standard HPSG assumption that noun phrases are NPs.)

(51)

$$\left[ \begin{array}{l} \text{HEAD } \textit{verb} \\ \text{SUBJ } \langle \rangle \\ \text{COMPS } \langle \text{NP}, \dots \rangle \end{array} \right]$$

(52)

$$\left[ \begin{array}{l} \text{HEAD } \textit{noun} \\ \text{SUBJ } \langle \rangle \\ \text{COMPS } \langle \text{NP}, \dots \rangle \end{array} \right]$$

Given these analyses, the approach extends to agreement with subjects of finite clauses and *bod*-clauses and possessors. It also extends to agreement between clause-initial *i* and the following subject given the assumption, defended in Borsley (1999) and Borsley and Jones (2005), that *i* has the category in (53) and takes the following subject and predicate as two separate complements.

(53)

$$\left[ \begin{array}{l} \text{HEAD } \textit{prep - comp} \\ \text{SUBJ } \langle \rangle \\ \text{COMPS } \langle \text{NP}, \text{VP} \rangle \end{array} \right]$$

The absence of agreement between a non-finite verb and a preceding subject is expected within this approach. The subject of a clause introduced by *i* is not a member of the COMPS list of the following verb. Hence, one would not expect the verb to agree with it. It seems, then, that this approach can handle the basic data quite satisfactorily. Notice, however, that within this approach there is close correspondence between COMPS lists and superficial structure. Essentially, the COMPS list of a head encodes its following dependents. Thus, this approach is very similar to an approach referring to superficial linear order.

### 4.3 Transformational analyses

An approach involving c-command of the kind that is assumed within transformational work also seems quite plausible. There are two analyses to consider here, those of Rouveret (1994: chapter 2) and Roberts (2005: chapter 2). Rouveret proposes that pronouns are NumPs containing Num and an NP complement and that agreement involves the incorporation of Num to a higher head leaving the NP behind, as in (54), where H is the head.<sup>16</sup>

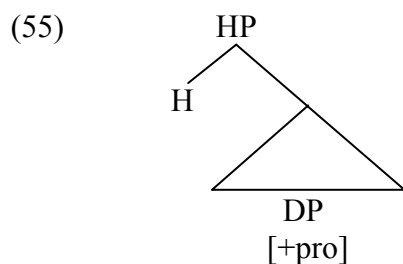
<sup>16</sup> Rouveret also assumes that NP moves to Spec NumP. Thus, strictly speaking he has structures of the following form:

(i) ... Num<sub>i</sub> X ... [NumP NP<sub>j</sub> t<sub>i</sub> t<sub>j</sub>]

However, this is of no importance in the present context.

(54) [H [... [<sub>NumP</sub> Num NP] ...]] ⇒ [Num<sub>i</sub> H [... [<sub>NumP</sub> t<sub>i</sub> NP] ...]]

Incorporation is a case of head-movement, which is standardly seen as movement of a head to the nearest c-commanding head. Rouveret assumes that it is subject to Relativized Minimality. Hence, it is not necessarily movement to the nearest head, but it is movement to the nearest head of the appropriate type. Roberts proposes that Welsh agreement involves the minimalist operation Agree. In early formulations, e.g. Chomsky (2001: 122), this is an operation involving a head and the nearest appropriate constituent which it c-commands. In more recent formulations, e.g. Chomsky (2004: 13), it is only restricted by the division of the derivation into phases and a probe may have more than one goal. In the present situation, Agree will involve a structure of the following form:



Thus, for both these approaches, Welsh agreement involves a head and a constituent which it c-commands. Such an approach seems quite plausible.

A question arises, however, about the fact highlighted above that a non-finite verb does not agree with the preceding subject in clauses introduced by *i*. A finite verb agrees with the following subject, as does *bod* in subordinate clauses which it introduces. There is no obvious reason why these processes should not apply in just the same way in clauses introduced by *i*. One might propose, however, that Welsh unlike English has no mechanism to transfer agreement features from T to an in-situ V. Thus, the absence of agreement between verb and subject in clauses introduced by *i* is probably not a problem although it is a complication.<sup>17</sup>

A more important question arises about the fact that there are no object clitics with a finite verb except optionally in a very literary variety of Welsh. If object clitics did appear with finite verbs, they would provide an exception to the generalizations formulated in section 4 since finite verbs are not adjacent to their objects. However, as we noted in section 4, these generalizations are exceptionless. If object clitics are the result of either head-movement or Agree applying within VP, it is not obvious why it should not apply with a finite verb which is subsequently raised to T, just as it applies with non-finite verbs which remain in situ. The Welsh situation contrasts with the situation in Arabic, which has examples like the (56).<sup>18</sup>

---

<sup>17</sup> It has been suggested to me that this is much like the fact that they do not agree with their subjects in English. It seems to me, however, that it is rather different. While English non-finite verbs do not show any agreement, Welsh non-finite verbs do show agreement but not with a preceding subject.

<sup>18</sup> I am grateful to Nushour Alduaij for providing the Arabic examples.

- (56) a. *al modaris ra?aa-h(u)*  
           *the teacher saw-him*  
           The teacher saw him.
- b. *ra?a-hu el modaris*  
           *saw-him the teacher*

The sentence in (56a) shows that finite clauses with SVO order may contain an object clitic, and (56b) shows that the same is true of finite clauses with VSO order. The latter suggests that the mechanism that is responsible for clitics applies before verb-movement. If object clitics appeared with finite verbs in Welsh, it would be evidence within a transformational approach that the mechanism which is responsible applies before verb-movement. Since they do not, there is no evidence that the responsible mechanism applies before verb-movement.

In fact, we can go further than this. The absence of clitics with finite verbs suggests that agreement applies after verb-movement. If it applied before verb-movement, some stipulation would be necessary to ensure that object clitics do not appear with finite verbs. One possibility would be to stipulate that the functional category which is realized as an object clitic is absent when T is finite. However, if agreement applies after verb-movement, the absence of object clitics can be attributed to a general locality restriction. We can simply say that there are no object clitics with finite verbs because a finite verb is not close enough to the object on the surface. Thus, we seem to have some evidence that agreement is a superficial matter.

#### 4.4 Summary

We have seen in this section that the basic data, introduced in section 2, are problematic for a GF-based approach to agreement of the kind that is assumed in LFG, but not for the COMPS list-based approach developed within HPSG by Pollard and Sag (1994). However, the latter, is very similar to an approach referring to superficial linear order. A c-command-based approach of the kind that has been advanced within transformational work also seems quite promising. However, the absence of object clitics with finite verbs suggests that an approach involving superficial c-command would be preferable to one involving c-command at an abstract level.

### 5 Coordination and its implications

I want now to consider the interaction of agreement and coordination. This appears to provide further support for the idea that Welsh agreement is a superficial matter.

#### 5.1 Data

In all the data we have considered so far we have a simple, non-coordinate noun phrase in the position associated with agreement. Naturally it is also possible to have a coordinate noun phrase in these positions. As noted by Morris-Jones (1931: 84), Rouveret (1994: chapter 5.1) and Sadler (1999), when a coordinate noun phrase appears in a position associated with agreement, the agreement is apparently with the first conjunct if it is a pronoun. Consider, for example, the following data:

- (57) Gweles [i a Megan] ddafad.  
*see.PAST.ISG I and Megan sheep*  
 Megan and I saw sheep.
- (58) arnaf [i a Megan]  
*on.ISG I and Megan*  
 on me and Megan
- (59) fy nhad [i a Megan]  
*ISG father I and Megan*  
 my and Megan's brother
- (60) Gwaeth Emrys fy ngweld [i a Megan].  
*did.3SG Emrys ISG see I and Megan*  
 Emrys saw me and Megan.

Peterson (2004) suggests that single-conjunct agreement is an extragrammatical phenomenon, a 'strategy' that speakers resort to 'patch up' gaps left by the grammar. This may be in the case in some languages, but it is not what we have in Welsh. Agreement with the first conjunct is a robust phenomenon, and it seems clear that it is an ordinary part of the grammar.<sup>19</sup>

First conjunct agreement would pose no problem if we could assume that coordinate structures have the person, number and gender features of the first conjunct so that agreement is really with the coordinate structure, and this is essentially what Johannessen (1998) proposes for data like (58)–(60) in a variety of languages.<sup>20</sup> However, there is evidence from anaphora that coordinate structures have their own person, number and gender features, distinct from those of the first conjunct. Consider the following examples:

- (61) a. Gwelais [i a Megan] ein hunain.  
*see.PAST.ISG I and Megan 1PL self*  
 I and Megan saw ourselves.
- b. Gwelest [ti a Megan] eich hunain.  
*see.PAST.2SG you.SG and Megan 2PL self*  
 You and Megan saw yourselves.
- c. Gwelodd [e a Megan] eu hunain.  
*see.PAST.3SG he and Megan 3PL self*  
 He and Megan saw themselves.

In each of these examples the verb appears to agree with the first conjunct of the following coordinate subject but the reflexive agrees with the whole coordinate subject.

There are two conclusions one might draw from these data. One is that coordinate structures must have two sorts of person, number and gender features relevant to different kinds of agreement. This position is developed within HPSG in connection

<sup>19</sup> Some evidence for this comes from the fact that it is mentioned without special comment in Morris-Jones (1931), which is concerned with formal, 'correct' Welsh.

<sup>20</sup> For some critical discussion of Johannessen's ideas see Borsley (2005).

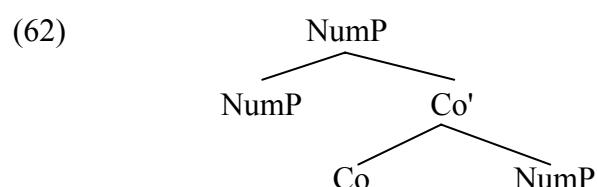
with Portuguese data in Villavicencio, Sadler, and Arnold (2005).<sup>21</sup> They propose that constituents have three sets of agreement features. These are identical in non-coordinate constituents, but in coordinate structures, one reflects the first conjunct, one the last, and one reflects the whole structure. Given these three sets of features, one can require agreement under certain circumstances with the first or last conjunct of a coordinate structure or the structure as a whole. In case of Welsh, it would be easy to require a head to agree with the first conjunct when a coordinate structure appears in the relevant position. The problem is that it would be just as easy to require a head to agree with the final conjunct. Thus, while it is possible to describe the facts within such an approach, the facts are not explained.

The alternative conclusion is that what looks like agreement between a head and one conjunct is just that. It has been developed by researchers in a variety of frameworks—see, for example, McCloskey (1986), Doron (2000), Citko (2004), Marten (2005), and van Koppen (2005). If the agreement is really with the first conjunct, it may be possible to explain the facts. I will assume from now on that this is the right approach. One point that is worth noting here is that this assumption is no problem for the generalizations in (32) and (48) above; it just means that the first following or the immediately following noun phrase may be a conjunct.

## 5.2 Implications

What follows if we accept that what looks like agreement with a conjunct is just that? For LFG it entails that agreement is not necessarily with a subject or object but may be with part of a subject or object. For HPSG it entails that agreement is not necessarily with the first member of a COMPS list but may be with part of the first member of a COMPS list. Thus, the position that agreement really is with the conjunct is quite problematic for both an LFG approach and the HPSG approach of Pollard and Sag (1994: chapter 9). It looks, then, as if the coordination data provide further evidence against an LFG approach and also show that one HPSG approach is dubious.

What are the implications for transformational approaches? Consider first Rouveret's incorporation analysis. On this analysis examples like those in (61) appear to violate the Coordinate Structure Constraint, as McCloskey and Hale (1984: 527) note for similar data in Modern Irish. Rouveret (1994: 306) suggests in effect that the crucial constituents are not real coordinate structures; rather they have the following structure:



Here the conjunction (Co) and the second NumP form a constituent (headed by the conjunction), which is adjoined to the first NumP. Given such a structure it should be possible for the head of the first NumP to move to some higher head. However, it is unclear how the whole structure could have agreement features reflecting both of the NumPs that it contains, something which the examples in (61) suggests is necessary.

<sup>21</sup> A broadly similar position is developed in Yatabe (2004).

It is also unclear why standard examples of movement should not be able to apply to what looks like the first conjunct of a coordinate structure. For example, we might expect it to be possible to derive a *wh*-interrogative like (63).

- (63) \*Pwy welodd a Megan ddafad?  
*who see.PAST.3SG and Megan sheep*  
 Who and Megan saw a sheep?

Thus, the coordination data seem quite problematic for an incorporation analysis.<sup>22</sup>

What, then, of an Agree-based approach? Agree-based approaches to somewhat similar data in various languages are developed in Doron (2000), Citko (2004), and van Koppen (2005). Particularly interesting is van Koppen (2005), who is concerned with certain dialects of Dutch. She considers data like the following from Tegelen Dutch:

- (64) Ich dink de-s doow en ich ôs kenne treffe.  
*I think that-2SG you.SG and I each other can.PL meet*  
 I think that you and I can meet.

Here the complementizer *de* shows agreement in the form of a suffix with the first conjunct of the following coordinate subject. The following verb agrees with the whole coordinate subject. Van Koppen suggests that “the syntactic component provides the configuration in which two Goals are available, and morphology chooses which one of these Goals eventually defines the agreement morphology of the probe” (2005: 4–5). More precisely, she proposes that both coordinate structures and their first conjuncts are accessible to Agree and that which actually undergoes Agree is determined by morphology. She proposes that Agree applies in a way that gives the most specific agreement morphology—it applies to the first conjunct in (64) and not to the coordinate structure because this allows the only distinctive form of the Tegelen Dutch complementizer, the second person singular form, to appear.

Could this approach be applied to the Welsh data? If the first conjunct of the coordinate subject in (64) is accessible to Agree, then presumably the first conjuncts of the Welsh examples in (57)–(60) will be too. Will morphological considerations ensure that Agree actually applies to the first conjunct in each case? On the face of it, they could if a coordinate noun phrase is always non-pronominal even when both its conjuncts are pronominal. Consider (65):

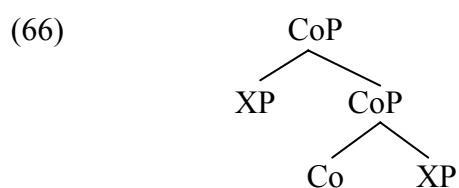
- (65) Gwelodd e a hi ddafad.  
*see.PAST.3SG he and she sheep*  
 He and she saw sheep.

Here the verb is third person singular, in agreement with the first conjunct, *e*. The coordinate subject must be third person plural. If it were pronominal, there is no obvious reason why the verb should not agree with it. In fact, one might expect it to since third person plural morphology is arguably more specific than third person singular morphology given that the latter, as noted in section 2, is the form which appears with a non-pronominal subject. I am not aware of any independent evidence

<sup>22</sup> Essentially the same point is made in Sobin (2004) in connection with English examples like *There is a frog and some fish in the pond*. Sobin points out that it is not plausible to suggest that movement is responsible for the agreement since it would violate the Coordinate Structure Constraint.

that a coordinate noun phrase is non-pronominal when both its conjuncts are pronominal. Equally, however, I am not aware of any difficulties which might arise from this assumption. It looks, then, as if it might be possible to apply van Koppen's approach to the Welsh data. However, before we conclude that the Welsh data are unproblematic for Agree we need to look more closely at van Koppen's analysis.

Van Koppen assumes the earlier version of Agree, in which it involves a head and the nearest appropriate c-commanded constituent. She makes a number of assumptions which entail that coordinate structures and their first conjuncts are equally near to a higher head. Firstly, she assumes the Conjunction Phrase (CoP) analysis of coordinate structures, where the first conjunct is the specifier, the conjunction the head, and the second conjunct the complement. Secondly, following Kayne (1994), she assumes that specifiers are adjoined constituents. Thus, coordinate structures have the following structure:



Finally, she assumes that adjunction gives a single category with two segments. These assumptions entail that a coordinate structure and its first conjunct are dominated by the same set of nodes and hence are equally close to any c-commanding head.

Questions arise about all of these assumptions. The CoP analysis of coordinate structures has been widely assumed but, as Borsley (2005) shows, it faces a number of problems. Most obviously, it is not clear how it can accommodate the fact that the distribution of coordinate structures depends on the nature of the conjuncts, the fact, that is, that coordinations of DPs appear in DP positions, coordinations of VPs appear in VP positions, and so on. On the face of it, CoPs must somehow acquire features from their specifier and complement, and hence cease to be just CoPs, but it is not clear how this should be done.<sup>23</sup> The assumption that specifiers are adjoined constituents also seems problematic given that it is explicitly rejected in standard Minimalism (Chomsky 1995, 2004). Finally, the idea that adjunction gives a single category with two segments seems to have been lost in Minimalism—Chomsky (1995: 248) proposes that adjunction of  $\alpha$  to  $\beta$  gives a structure whose label is the ordered pair  $\langle \beta, \beta \rangle$ . It seems, then, that the adjunction structure and its head have different labels. Thus, questions arise about all of the assumptions which van Koppen uses to derive the conclusion that both coordinate structures and their first conjuncts are equally close to any c-commanding head.

Coordinate structures and their first conjuncts must be equally close to any c-commanding head if agreement is the product of the Agree mechanism and Agree involves a head and the nearest c-commanded constituent. Recall, however, that in more recent versions of Agree the operation is only restricted by the division of the derivation into phases, and a probe may have more than one goal. Assuming such a version of Agree, both a coordinate structure and its first conjunct will be accessible

<sup>23</sup> Doron (2000: 87) suggests that “conjunctions lack any formal features, from which it follows that the category they project is that of conjuncts themselves”. It is not clear how conjunctions could lack any formal features since they clearly have certain syntactic properties which must be encoded in some way. It is also not clear why it would follow that they project the category of the conjuncts.

to Agree. On the face of it, however, its second conjunct will also be accessible. It is not clear, then, whether there is a plausible set of assumptions which entail that both a coordinate structure and its first conjunct but not its second conjunct are accessible to Agree.

What about the idea that it is morphology that determines whether agreement is with a coordinate structure or its first conjunct? English appears to provide a problem for this idea. It is quite well known that English has agreement with the first conjunct in *there*-insertion sentences such as (67).

(67) There is/\*are a cat and a dog in the garden.

This is expected given that *is* is a more specific form than *are*. Consider, however, the following:

(68) Are/??Is the cat and the dog in the garden?

(69) Where are/??is the cat and the dog?

Here *are* seems to be strongly preferred to *is*. On standard assumptions, both the auxiliary and subject occupy different positions in (68) and (69) from (67). However, (68) and (69) are more like (64) than (67) is. In (68) and (69), as in (64), the agreeing element is in C and the subject is Spec TP. It is not at all clear, then, why agreement should be with the whole coordinate structure in (68) and (69).

Perhaps the most important point to make about van Koppen's analysis is that Agree applies at the interface to PF, and it must do so if it is to be constrained by morphological factors. This assumption is also required by van Koppen's account of why a verb does not agree with the first conjunct of a preceding construction, illustrated by (64). If Agree applied before movement, T could agree with the first conjunct of a coordinate subject and the result would be that the verb showed agreement with this conjunct. Van Koppen proposes that the internal structure of copies left by movement is invisible. Hence, Agree can only apply to the whole copy of the moved coordinate subject. This means that van Koppen's Agree is quite different from Agree in standard Minimalism, which applies during the syntactic derivation and identifies constituents which undergo Move. Within van Koppen's approach, some other mechanism must perform this function.<sup>24</sup>

While there are a variety of questions about van Koppen's approach, it may be possible to develop an analysis of first conjunct agreement involving c-command restricted in some way so that it does not also allow agreement with a second conjunct. It is likely, however, that any such approach will involve c-command at a superficial level.

### 5.3 Summary

If agreement is really with the first conjunct, just as it appears to be, the data considered in this section provide further evidence against a GF-based approach to agreement of the kind that is assumed in LFG and also evidence against the COMPS list-based approach developed within HPSG by Pollard and Sag (1994). A c-command approach of the kind developed within P&P/Minimalism seems more

---

<sup>24</sup> For quite different arguments that agreement must be separate from the basic operations of 'narrow syntax' see Bobaljik (2006).

promising although it is not without its problems. However, it is likely that any such approach should refer to a superficial level of structure. As noted in section 4, this is a conclusion which is also suggested by the absence of object clitics with finite verbs.

## 6 A'-movement

We noted in section 4 that there is no evidence that agreement applies before verb-movement and some evidence that it applies after. Clearly we should ask about other kinds of movement, in particular A'-movement in *wh*-interrogatives and related constructions.

In section 1, we noted that examples like (3b), repeated here as (70), are traditionally seen as evidence that agreement applies before A'-movement in English:

(70) Which men do you think know/\*knows the answer?

In this section we will see that there is an agreement process which appears to apply before A'-movement, but there is evidence that this is separate from the agreement which is the focus of the paper. We will also see that there appear to be no examples which could show that the agreement we are focusing on applies before A'-movement. Hence, the facts are probably compatible with the assumption that this agreement applies before A'-movement, but they provide no evidence for this assumption.

Before we proceed, it is worth noting that it is not really clear that examples like (70) provide evidence that agreement applies before A'-movement in English. Given the copy theory of movement, in which movement operations leave behind a copy of the moved constituent, agreement could on the face of it apply just as easily after A'-movement.

### 6.1 *Wh*-interrogatives

There are two types of sentence which might be seen as evidence for an agreement process applying before A'-movement. The first is exemplified by *wh*-interrogatives like (71).

(71) Pwy soniodd Gwyn amdano?  
*who talk.PAST.3SG Gwyn about.3SGM*  
 Who did Gwyn talk about?

Here it appears that the preposition *amdano* agrees with the fronted noun phrase *pwy*. There is evidence, however, that *pwy* is non-pronominal. Welsh allows pied piping of prepositions. Hence, we could have the following instead of (71):

(72) Am bwy soniodd Gwyn?  
*about who talk.PAST.3SG Gwyn*  
 About whom did Gwyn talk?

Here, *pwy*, which appears as the mutated form *bwy*, is object of the preposition *am*, which is in its basic, uninflected form. This suggests that *pwy* is non-pronominal. It seems to follow that the agreement in (71) is unexpected. In fact, however, we can

argue that (71) is a standard case of agreement. It is generally accepted that such examples involve a phonologically empty resumptive pronoun following the preposition (see Borsley, Tallerman, and Willis 2007, section 4.1.7 for discussion). An overt resumptive pronoun is possible in this position, as (73) illustrates.

- (73) Pwy soniodd Gwyn amdano fo?  
*who talk.PAST.3SG Gwyn about.3SGM he*  
 Who did Gwyn talk about?

It looks, then, as if we can assume that we actually have agreement with an empty resumptive pronoun in (71) and hence not with the fronted noun phrase.

The second type of sentence that is relevant here is exemplified by (74).

- (74) Pwy wnest ti ei weld?  
*who do.PAST.2SG you.SG 3SGM see*  
 Who did you see?

Here the non-finite verb is preceded by a clitic agreeing with the fronted noun phrase *pwy*. It is sometimes assumed that sentences like this have an empty resumptive pronoun in object position (Rouveret 1994: chapter 5). On this view, (74) involves agreement with an empty resumptive pronoun just like (71). However, as Willis (2000: 545) notes, an overt resumptive pronoun is not possible here:

- (75) \*Pwy wnest ti ei weld o?  
*who do.PAST.2SG you.SG 3SGM see he*  
 Who did you see?

Thus, it is unlikely that sentences like (74) contain an empty resumptive pronoun. Here, then, the idea that the agreement is the result of a mechanism applying before A'-movement seems quite plausible. Notice, however, that the fronted noun phrase is again non-pronominal. It looks, then, as if the clitic in (74) must be the manifestation of some special mechanism distinct from whatever mechanism is responsible for the agreement that is the main focus of this paper.

We have the expected situation when a subject is fronted. Consider, for example, the following:

- (76) Pa ddynion welodd / \*welon ddafad?  
*which men see.PAST.3SG see.PAST.3PL sheep*  
 Which men saw a sheep.

Here the fronted *wh*-phrase is plural but the following verb is singular. This is what we expect if *wh*-phrases are non-pronominal. This reinforces the idea that the agreement in an example like (74) involves a special mechanism.

## 6.2 Cleft sentences

If *wh*-expressions are non-pronominal and the agreement that is the focus of this paper involves pronouns, *wh*-interrogatives could not provide any evidence that this agreement applies before A'-movement. What we need is a construction in which

pronouns undergo A'-movement. Cleft sentences appear to be such a construction. I will argue, however, that appearances are misleading.

Welsh cleft sentences consist of a focused constituent followed by a clause containing a gap or a resumptive pronoun and look rather like *wh*-interrogatives. The following is a typical example:

- (77) Y dynion welodd ddafad.  
*the men see.PAST.3SG sheep*  
 It's the men that saw a sheep.

Here the focused constituent is associated with a gap in subject position. Since it is non-pronominal, the following singular verb is what we would expect. Consider now a similar example, where the focused constituent is a pronoun.

- (78) Nhw welodd ddafad.  
*they see.PAST.3SG sheep*  
 It's they that saw a sheep.

Again, the verb is singular. On both of the transformational analyses of agreement introduced in section 4, one would expect the verb to agree with a pronominal subject before it undergoes A'-movement. In other words, one would expect the following to be acceptable, which it is not:

- (79) \*Nhw welon ddafad.  
*they see.PAST.3PL sheep*  
 It's they that saw a sheep.

The fact that we have (78) and not (79) seems problematic for the assumption that agreement applies before A'-movement. In fact this seems problematic even if agreement applies after A'-movement. As noted earlier, within Minimalism, movement leaves not a trace but a copy of the moved constituent within Minimalism. Obviously, a copy of a pronoun is a pronoun. Hence, on the face of it, one would expect agreement here even if agreement applies after A'-movement.

The fact that we have (78) and not (79) also looks problematic for LFG and HPSG. In both frameworks we expect to have (79) and not (78) if the initial constituent is a filler. It looks, then, as if the data are problematic for a number of approaches.

In fact, however, there is evidence that there is no problem here. It seems that a focused constituent is not a filler and hence not the result of movement on a transformational approach. Consider the following:

- (80) a. Fi mae Gwyn wedi 'i ddewis/\*fy newis.  
*I be.PRES.3SG Gwyn PERF 3SGM choose 1SG choose*  
 It's me that Gwyn has chosen.  
 b. Ti mae Gwyn wedi 'i ddewis/\*dy ddewis.  
*you.SG be.PRES.3SG Gwyn PERF 3SGM choose 2SG choose*  
 It's you that Gwyn has chosen.

In these examples the clitic is third person although the focused constituents are first and second person, respectively. Assuming that the clitics agree with the first post-

verbal gaps, it follows that the focused constituents differ in person from the associated gaps. This suggests rather strongly that these constituents are not fillers.<sup>25</sup>

If the initial constituents are not fillers, there is no reason to assume that (78) has a pronominal subject. In fact, it is probably easy in all three frameworks to ensure that the gap in a cleft sentence is non-pronominal so that if it is a subject the verb is singular even if the gap is plural. In the case of P&P/Minimalism one could assume that what is fronted in clefts is an empty operator which is non-pronominal. There are no empty operators in LFG and HPSG but there are various ways in which one might ensure that the gap in a cleft sentence is pronominal. We return to this matter in section 7. It seems, then, that the fact that we have (78) and not (79) is not a problem for standard views of agreement.

If we do not have a pronoun undergoing A'-movement in an example like (78), it is probable that there are no such cases. In other words, there are probably no examples which could show that the agreement that is the focus of this paper applies before A'-movement.

### 6.3 Summary

We have seen in this section that Welsh has an agreement process applying before A'-movement on standard transformational assumptions. However, this is separate from the agreement that is the focus of the paper because it is not restricted to pronouns.<sup>26</sup> We have also seen that there are probably no cases where a pronoun undergoes A'-movement and hence no examples which could show that this agreement applies before A'-movement. It follows that the facts are compatible with the assumption that agreement applies before A'-movement. However, they provide no evidence for this assumption. We saw in section 4 that there are reasons for thinking that it applies at a superficial level. There seem to be no reasons not to accept this position.

There is one further point to note about the data considered in this section. If there were cases where a pronoun triggers agreement and then undergoes movement,

<sup>25</sup> In Borsley (2008), I argue that the initial constituent of a Welsh cleft sentence and the following clause are two terms of a hidden identity predication and that the examples in (80) are similar to the following quasi pseudo-clefts:

- (i) a. *Fi ydy 'r un mae Gwyn wedi 'i ddewis.*  
*I be.PRES.3SG the one be.PRES.3SG Gwyn PERF 3SGM choose*  
 The one that Gwyn has chosen is me.
- b. *Ti ydy 'r un mae Gwyn wedi 'I ddewis.*  
*you.SG be.PRES.3SG the one be.PRES.3SG Gwyn PERF 3SGM choose*  
 The one that Gwyn has chosen is you.

<sup>26</sup> There is one type of example which might be seen in a transformational approach as evidence for an agreement process applying before A-movement. This is passives such as (i).

- (i) *Mi gafodd Gwyn ei daro gan Emrys.*  
*PRT get.PAST.3SG Gwyn3 SGM hit with Emrys*  
 Gwyn was hit by Emrys.

Here the non-finite verb is preceded by a clitic agreeing with the subject. On transformational assumptions, the subject *Gwyn* would originate in a position following the non-finite verb. Hence, one might suppose that the clitic is the result of some agreement process applying before A-movement. Notice, however, that the subject is non-pronominal. It seems, then, that the clitic must be the manifestation of some special mechanism, rather like the clitic in (74).

the result would be cases in which a pronoun precedes the associated agreement. Since there are no cases where a pronoun triggers agreement and then undergoes movement, a pronoun never precedes the associated agreement. It follows that the agreement that we are concerned with here always involves a following pronoun, not just normally. Thus, whichever of the generalizations formulated in section 3 is correct applies without exception.

## 7 A linear approach

In the preceding sections we have seen that there is evidence against the idea that Welsh agreement involves abstract levels of various kinds and no evidence that such a level is relevant. We have also seen that whichever of the generalizations formulated in section 3 is correct holds without exception. Thus, the evidence suggests that it is not some abstract structure but a fairly superficial structure that is relevant for agreement. One possibility is an approach like that proposed by van Koppen involving c-command at a superficial level. Another possibility is an approach involving linear order at a superficial level, an analysis incorporating one of the generalizations formulated in section 4 fairly directly. I will develop an approach of the second kind here.

The idea that what looks like agreement may involve linear order at a superficial level is explicitly advanced by Ackema and Neeleman (2004: chapter 6). They propose that some apparent examples of agreement are really a case of context-sensitive spell-out, where the form of some element is influenced by an adjacent element. They apply this approach to finite verbs in Irish, suggesting that the verb combines with a following pronoun, which is realized as a suffix, and they also suggest that it should be applied to Welsh finite verbs. In Irish, verbal morphology and overt subjects do not co-occur, but in Welsh, verbal morphology can co-occur with an overt pronominal subject. This leads them to suggest, like Rouveret (1994), that pronouns involve two elements. The problem for this approach is that while verbal and prepositional agreement morphology is adjacent to the associated pronoun, this is obviously not true of clitics. Thus, unless the approach is extended in some way it cannot accommodate the Welsh data.

Ackema and Neeleman (2004: chapter 7) argue for a rather different approach to agreement between a complementizer and the following subject in some Dutch dialects. They assume that this is real agreement, but agreement which is established between adjacent elements at PF.<sup>27</sup> Van Koppen (2005: 175–181) seeks to show that a c-command approach is preferable to a linear approach in Dutch. However, his discussion is rather inconclusive. She concedes (p.179) that his approach is “not as straightforwardly equipped” to account for certain data as Ackema and Neeleman’s approach. Van Koppen pays particular attention to certain cases where some element intervenes between the complementizer and the following subject. In some cases, there is no agreement, but in others agreement does occur, and she suggests that these

---

<sup>27</sup> A related view is applied to Scots Gaelic in Adger (2000). Working within a version of Minimalism, Adger proposes that verb–subject agreement in Scots Gaelic is the product of a morphological process involving adjacent elements. Linear order is involved in agreement in a more complex way in McCloskey’s (1986) analysis of Irish. McCloskey proposes that agreement in Irish involves a head and a pronoun which it governs, but (in part) to allow agreement with a first conjunct he proposes a language-specific definition of government referring to linear order.

cases are problematic for Ackema and Neeleman's approach. On quite plausible assumptions the elements involved in agreement are always adjacent in Welsh. Hence, whatever may be the case in Dutch, a linear approach to Welsh looks unproblematic. There is also a general argument for a linear approach. It is articulated by Culicover and Jackendoff (2005: 52), who comment that:

Given the epistemological priority of linear order – it is immediately available to the learner in a way that structure is not – it seems to us that the natural approach would be to see how much explanatory mileage one could get out of linear order.

In other words, a linear account is preferable other things being equal to a structural account. I will seek to develop a linear approach here, but I will depart from Ackema and Neeleman's approach in an important way. In their approach, agreement applies in PF and refers to prosodic phrases. They assume that agreement is with the first conjunct in some dialects because coordinate subjects are parsed into two prosodic phrases, but they do not provide any independent evidence for this proposal. I will assume that Welsh agreement involves a superficial level of structure which is closely related to phonology but is a syntactic level. This is the level of order domains proposed in linearization-based HPSG. I will first introduce this framework and then show how it can accommodate the Welsh data.

### 7.1 Linearization-based HPSG

In much HPSG work, e.g. Pollard and Sag (1994), order is a reflection of constituent structure, but for linearization-based HPSG, developed in Pollard, Kasper, and Levine (1993), Reape (1994) and especially Kathol (2000), it is defined in terms of a separate system of order domains. For linearization-based HPSG, phrasal constituents have both a list of daughters and a list of domain elements. Within Kathol's framework, the former are signs, linguistic expressions with syntactic, semantic and phonological properties, and, if phrasal, their own internal structure, while the latter include syntactic, semantic and phonological information, but do not include information about internal structure. The domain elements of a constituent may be 'compacted' to form a single element in the order domain of the mother or they may just become elements in the mother's order domain. In the latter case, the mother has more domain elements than daughters, and some members of the order domain are not sisters.

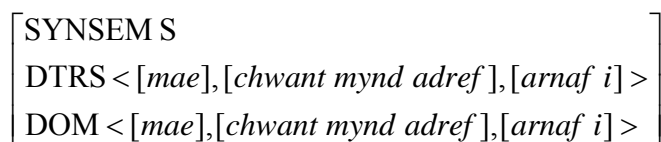
Linearization-based HPSG has been applied to word order phenomena in German (Kathol 2000), Breton (Borsley and Kathol 2000), Serbo-Croatian (Penn 1999) and Japanese (Yatabe 2001), among other languages. Among other things, the distinction between daughters and domain elements permits an analysis of certain extraposition phenomena. Consider, for example, the following Welsh examples:

- (81) a. Mae            chwant mynd adref arna' i.  
           *be.PRES.3SG desire go home on.1SG I*  
           I desire to go home.  
       b. Mae            chwant arna' i fynd adref.  
           *be.PRES.3SG desire on.1SG I go home*  
           I desire to go home.

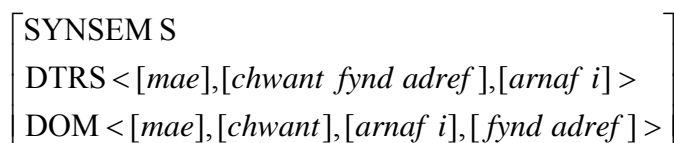
The sentence in (81a) has a subject containing an abstract noun *chwant* and an infinitival complement *mynd adref*, while in (81b) the complement is extraposed.

Within linearization-based HPSG, we can propose that these sentences have the same three daughters but that the second has one more domain element. For HPSG, the syntactic and semantic properties of an expression are encoded as the value of a feature SYNSEM and the daughters and domain elements are encoded as the value of the features DTRS and DOM, respectively. Using bracketed orthography to represent both daughters and domain elements, we might propose the following schematic analyses:

(82) a.

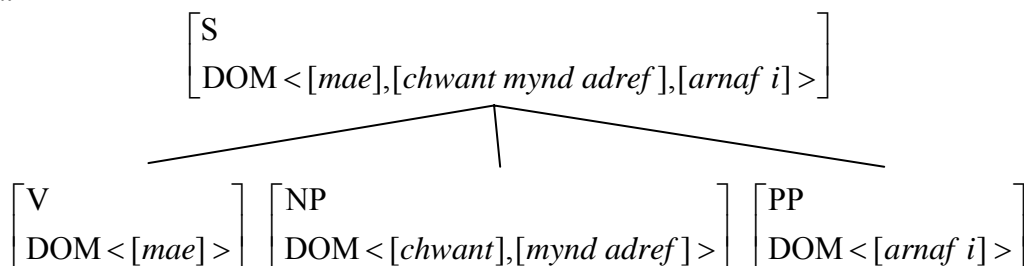


b.

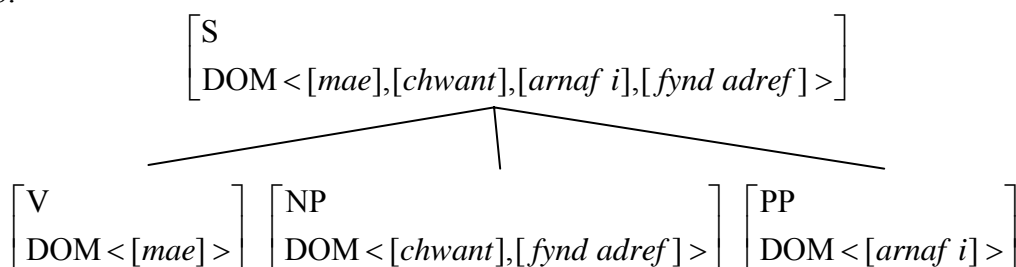


Alternatively, using the standard tree format to represent constituent structure, we might give slightly fuller analyses, as in (83). (As before, I adopt the standard HPSG assumption that noun phrases are NPs.)

(83) a.



b.



I will use this latter form of representation in the following discussion.

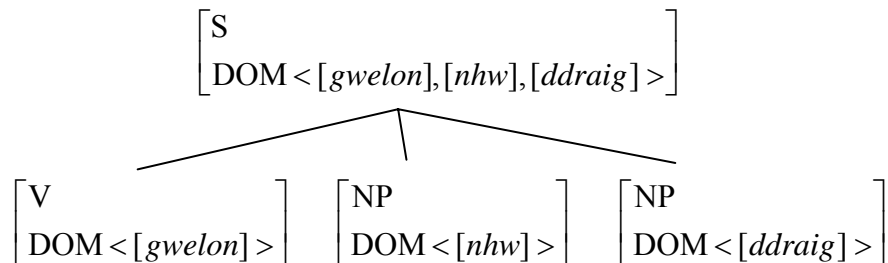
## 7.2 The analysis

I return now to the agreement data. The basic data are a fairly simple matter. They involve analyses in which there is a one-to-one correspondence between daughters and domain elements. The example in (4b), repeated here as (84), will have the

schematic analysis in (85).

- (84) Gwelon      nhw    ddraig.  
*see.PAST.3PL they dragon*  
 They saw a dragon.

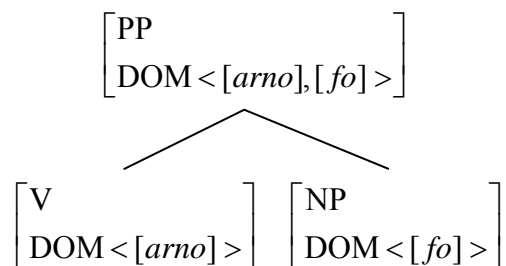
(85)



The PP in (7a), repeated as (86), will have the analysis in (87).

- (86) arno      fo  
*on.3SGM he*  
 on him

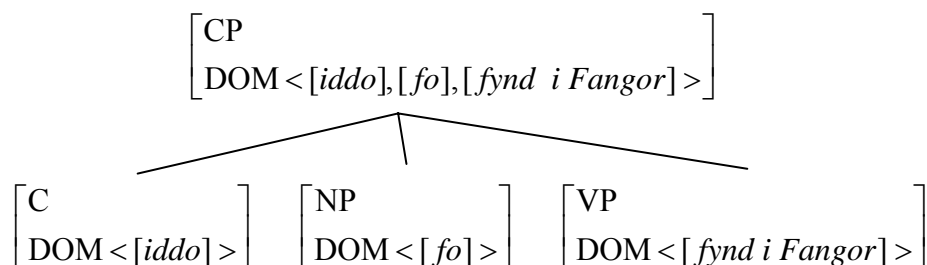
(87)



Similarly, the clause in (10a), repeated as (88), will have the analysis in (89).

- (88) iddo      fo    fynd i Fangor  
*to.3SGM him go to Bangor*  
 him to go to Bangor

(89)

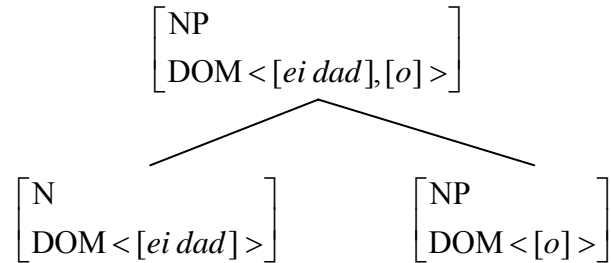


In all these cases, a head agrees with an element which immediately follows it in the topmost order domain.

What about examples where agreement takes the form of a clitic? If clitics are phrasal affixes, (13a), repeated here as (90), will have the schematic analysis in (91).

- (90) ei dad o  
 3SGM father he  
 his father

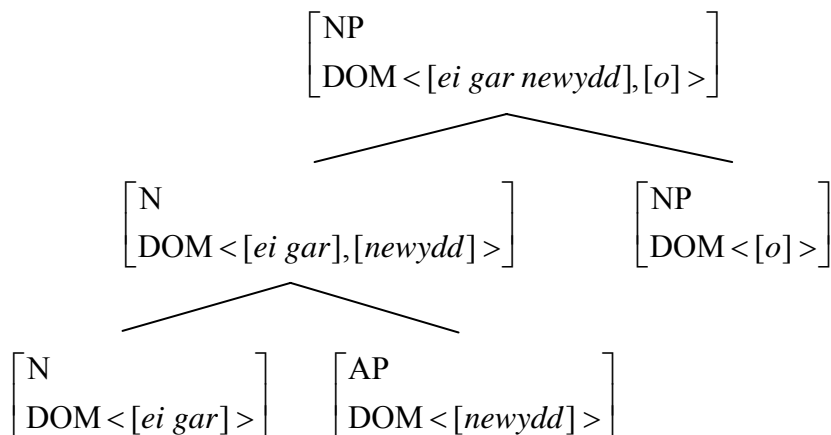
(91)



If attributive adjectives are right-adjoined to a preceding nominal element, the more complex example in (38), repeated here as (92), will have the analysis in (93).

- (92) ei gar newydd o  
 3SGM car new he  
 his new car

(93)

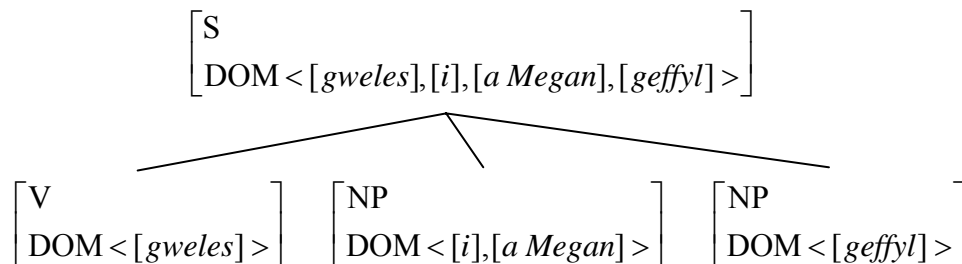


In both cases agreement is with an immediately following element within the order domain of NP. These examples would have somewhat different structures and order domains if clitics were analyzed as functional heads, but I won't explore this possibility.

We can turn now to the coordination data (cf. section 5 above). Here the distinction between daughters and domain elements is crucial. I will assume that Welsh coordinate structures are not compacted. This means that a typical two-conjunct coordinate structure will give two domain elements in the order domain of its mother and that coordinate structures are invisible at the level that is relevant to agreement. Given this assumption, (57), repeated here as (94), will have the analysis in (95).

- (94) Gweles i a Megan geffyl.  
*see.PAST.ISG I and Megan horse*  
 Megan and I saw a horse.

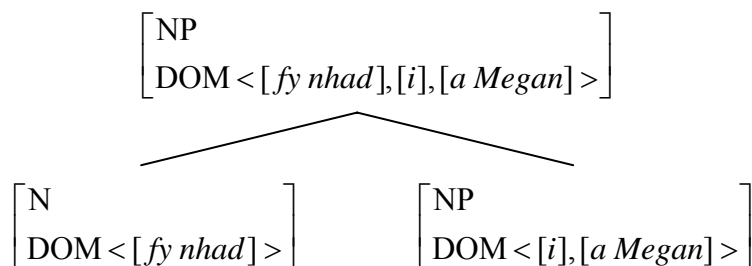
(95)



Here S has three daughters but four domain elements. As in (85), the verb agrees with the element which immediately follows it within the order domain of S. For the NP in (59), repeated here as (96), we will have the analysis in (97).

(96) *fy nhad i a Megan*  
*1SG father I and Megan*  
 my and Megan's brother

(97)



Here, the higher NP has two daughters but three domain elements. As in (91), the noun agrees with the element which immediately follows it within the order domain of NP. We will have similar analyses for the other examples in which agreement is with the first conjunct of a coordinate structure.

Although coordinate structures are *ex hypothesi* not compacted, their domain elements must remain adjacent. Thus, as the translation makes clear, (98) can only be an example with a coordinate NP in object position. It is not an alternative form of (94).

(98) *Gweles i ddraig a Megan.*  
*saw.1SG I dragon and Megan*  
 I saw a dragon and Megan.

It seems, then, that we need to allow not only constraints requiring one domain element to precede another but also constraints requiring one domain element to immediately precede another.<sup>28</sup>

We can now consider what constraints are appropriate here. I will assume, as is standard in HPSG, that the person, number and gender features of a nominal constituent are properties of its index. I will also assume, following Kathol (1999),

<sup>28</sup> For independent suggestions that it is necessary to assume immediate precedence rules, see Zwicky and Nevis (1986) and Ojeda (1988).

that agreement, whatever form it takes, is the realization of a feature AGR, whose value is person, number and gender features or *none* where there is no agreement. Given these assumptions, we need to ensure that where a head is immediately followed by a pronoun the value of the head's AGR feature is the person, number and gender features of the pronoun's index. I am assuming that noun phrases contain a possibly complex head. If numerals and pre-nominal adjectives, as in (34), are part of this head, then the head will always be domain-initial. This suggests that we can propose the following constraint:

(99) [DOM <[AGR [1]], NP: *ppro*[2], ...>] → [1] = [2]

Following standard practice, I use 'X:Y' to stand for a category X whose semantic content is Y. Thus, NP:*ppro*[n] stands for an NP whose semantic content is that of a personal pronoun whose person, number and gender features are n. If attributive adjectives could not be analyzed as forming a constituent and a domain element with the preceding noun, and if clitics were analyzed as the realization of separate heads, we would need a constraint requiring agreement with the first following NP if it is a pronoun. Normally, the value of AGR will be *none*. I assume that this is a result of the constraint in (100), where '/' indicates that it is a default constraint:

(100) [DOM <[AGR [1]], ...>] → /[1] = *none*

This will be overridden by (99). It will also be overridden by a constraint applying to non-finite verbs whose object is an unbounded dependency gap and by a constraint applying to non-finite verbs in passives such as (i) in footnote 25.<sup>29</sup>

An important point to note about the analysis outlined here is that it necessitates the assumption that VSO clauses have a flat structure and not an NP VP structure.<sup>30</sup> Consider the following example:

(101) Gwelodd Megan nhw.  
*see.PAST.3SG Megan they*  
 Megan saw them.

If we assumed an NP VP structure for VSO clauses, this would have the schematic analysis in (102).

(102)

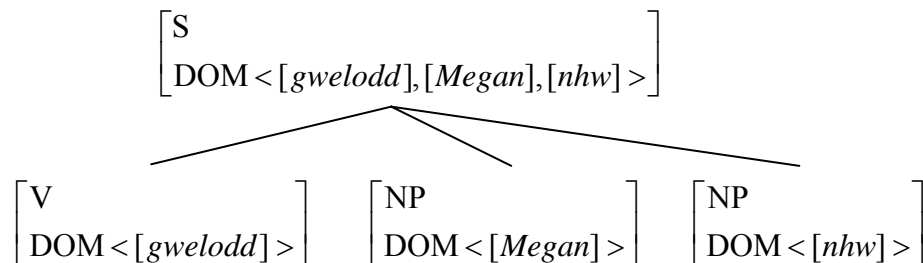
$$\begin{array}{c}
 \left[ \begin{array}{l} \text{S} \\ \text{DOM} < [\textit{gwelodd}], [\textit{Megan}], [\textit{nhw}] > \end{array} \right] \\
 \swarrow \quad \searrow \\
 \left[ \begin{array}{l} \text{NP} \\ \text{DOM} < [\textit{Megan}] > \end{array} \right] \quad \left[ \begin{array}{l} \text{VP} \\ \text{DOM} < [\textit{gwelodd}, \textit{nhw}] > \end{array} \right]
 \end{array}$$

<sup>29</sup> If what look like third person singular verbs with a non-pronominal subject are really third person singular verbs and not verbs with no agreement, they will require another constraint overriding (100).

<sup>30</sup> I am grateful to Andreas Kathol for drawing this point to my attention.

Here, within the order domain of the VP, the verb *welodd* is immediately followed by the object pronoun *nhw*. Thus, we would expect to have agreement. However, finite verbs do not agree with their objects. As was noted in section 2, a very literary variety of Welsh optionally allows object clitics, but it is fairly clear that this is a separate phenomenon. The absence of agreement in (101) is unproblematic if we assume a flat structure for VSO clauses. This will give us the following schematic analysis:

(103)



Here *welodd* is only followed by other domain elements within the order domain of the S and the first NP that follows it is the proper name *Megan*. Hence, we do not expect agreement. Borsley (2006) argues that there are no compelling arguments for an NP VP analysis of Welsh VSO clauses and that such an analysis faces certain problems. I think, then, that it is safe to assume that the right analysis is (103) and not (102), and hence that there is no problem here.

There is one further matter that we should consider: the fact that cleft sentences such as (77) and (78), repeated here as (104) and (105), have a singular verb:

(104) *Y dynion welodd ddafad.*  
*the men see.PAST.3SG sheep*  
 It's the men that saw a sheep.

(105) *Nhw welodd ddafad.*  
*they see.PAST.3SG sheep*  
 It's they that saw a sheep.

We argued that the initial constituent in such sentences is not a filler (cf. section 6.2 above). Thus, it seems that gaps are non-pronominal when there is no filler just as much as when there is a non-pronominal filler. How this should be accounted for depends on what approach is taken to gaps. In much work in HPSG, as in pre-Minimalist P&P, it is assumed that unbounded dependency gaps are filled by an empty category. This view is rejected in some HPSG work, e.g. Ginzburg and Sag (2000) and Bouma, Malouf, and Sag (2001). However, the view that unbounded dependency gaps do not contain an empty category is incompatible with the approach to agreement that we have developed here. Consider the following:

(106) *Pwy welodd ti?*  
*who saw.3SG you.SG*  
 Who saw you?

Here the filler *pwy* is associated with a subject gap. Given the analysis we are developing, there must be an empty category in subject position since otherwise the verb would be immediately followed by the object *ti* and would agree with it, giving

(107).

(107) Pwy welest ti?  
*who saw.2SG you.SG*  
 Who did you see?

This is perfectly grammatical, but it can only have an object gap, as the translation makes clear. It is clear, then, that we must analyze unbounded dependency gaps as empty categories.<sup>31</sup> Pollard and Sag (1994: chapter 4) assume an empty category of the following form:

(108)

$$\left[ \begin{array}{l} \text{LOCAL}[1] \\ \text{SLASH} \{[1]\} \end{array} \right]$$

Here the value of the LOCAL feature, which encodes the main syntactic and semantic properties of an expression, is the single member of the set which is the value of the SLASH feature. Interacting with certain constraints, this ensures that information about the main syntactic and semantic properties of the gap is available higher in the tree. Hence, where there is a filler higher in the tree it will have these syntactic and semantic properties. In the present context, what we need to ensure is that a nominal gap is non-pronominal. We can do this with the following constraint:

(109)

$$\left[ \begin{array}{l} \text{LOCAL}[1] \left[ \begin{array}{l} \text{HEAD } \textit{noun} \\ \text{CONTENT} [2] \end{array} \right] \\ \text{SLASH} \{[1]\} \end{array} \right] \rightarrow [2] = \textit{npro}$$

This says that if a gap has a nominal LOCAL value (and hence a nominal value in SLASH) its CONTENT value is non-pronominal.<sup>32</sup>

### 7.3 Summary

I have now provided an account of the main Welsh agreement facts within linearization-based HPSG. As noted at the outset, it may be possible to provide a similar analysis within other frameworks. The important point is that there are grounds for thinking that Welsh agreement, unlike agreement in many other languages, involves a superficial level of structure and that an account involving superficial linear order is quite plausible.

<sup>31</sup> For independent arguments that unbounded dependency gaps should generally be analyzed as empty categories, see Levine and Hukari (2006).

<sup>32</sup> We will also need a further constraint to exclude resumptive pronouns from the topmost subject position. I will not try to decide what form this should take.

## 8 Unexpressed noun phrases

In this section I want to consider an important fact about the Welsh agreement phenomena which I have ignored so far. In colloquial Welsh, the pronouns that are associated with agreement are usually present, but they are often absent in more literary varieties. In these varieties, all the agreement contexts allow an unexpressed noun phrase instead of the pronoun. Since Chomsky (1982), it has been assumed within P&P that such unexpressed noun phrases are instances of a phonologically empty pronoun *pro*, and it is natural to assume that this is what we have here.<sup>33</sup> However, such phonologically empty pronouns are rejected in much work in HPSG and within other frameworks such as LFG. I will argue in this section that we should indeed assume a phonologically empty pronoun in the Welsh examples when there is no overt pronoun present.

Given phonologically empty pronouns, the analysis extends without any modification to cases where there is no overt pronoun. An agreeing head is immediately followed by a pronoun in such examples just as it is in examples with an overt pronoun. On the other hand it is easy to show that this analysis cannot be maintained without empty pronouns. Examples like the following are crucial:

- (110) Welest fi?  
*saw.2SG I*  
 Did you see me?

Here, if we do not assume an empty pronominal subject, we have a second person singular verb immediately followed by first person singular pronoun. Notice that this argument is similar to the argument in section 7 for the view that unbounded dependency gaps are filled by an empty category.

If unexpressed noun phrases are only represented at some abstract level, then agreement must be characterized at that level. Consider what this implies. In the case of LFG, unexpressed noun phrases are only represented in f-structure. It follows that agreement must be characterized in terms of f-structure. We saw, however, in section 4, that an f-structure approach to agreement cannot provide a unified account of the Welsh data. In an LFG discussion of some of the data that we have been concerned with here, Sadler (1997: 7) suggests that the assumption that clitics are associated with an empty category “requires us to extend the use of null categories in a wholly unmotivated manner”. This would be true if agreement could be characterized in terms of f-structure. However, as we have seen, there is good evidence against this.

We have a similar situation in much HPSG work, where it is assumed that unexpressed noun phrases are only represented in the lists that are the value of the feature ARG-ST (ARGUMENT-STRUCTURE); see, for example, Manning and Sag (1999). In the case of verbs, the subject is the first member of the ARG-ST list, and an object is the second member. It follows that agreement is with the first member of the ARG-ST list in the case of finite verbs and *bod*, and with the second member of the ARG-ST list in the case of non-finite verbs. Thus, an ARG-ST list approach cannot provide a single generalization about agreement. In fact such an approach is less satisfactory than the COMPS list-based approach discussed in section 4. As noted, the latter handles the basic data quite satisfactorily and only has problems with the

---

<sup>33</sup> This position was applied to Irish, where an unexpressed NP is not just possible but required in the sorts of context discussed here, in McCloskey and Hale (1984).

coordination data. An ARG-ST list approach has essentially the same problems as an LFG GF-based approach.

It seems, then, that unexpressed noun phrases must be analyzed as phonologically empty pronouns if we want to maintain the approach to agreement developed in section 7. Thus, Welsh provides important evidence that unexpressed noun phrases must be represented at a superficial level of structure, as has generally been assumed in Principles and Parameters theory, but contrary to the standard position in LFG and a widely accepted view in HPSG.<sup>34</sup>

## 9 Concluding remarks

The main focus of this paper has been the nature of agreement in Welsh. I have looked at six kinds of agreement and argued, following a number of earlier researchers, that they are aspects of a single phenomenon. I have argued that the data provide evidence against the idea that Welsh agreement involves abstract levels of various kinds and that there is no evidence that an abstract level is relevant. I have suggested that the data involve linear order at a superficial level, and I have developed an analysis within HPSG embodying this assumption. I have also argued that this analysis entails that unexpressed noun phrases of certain kinds must be represented in the superficial constituent structure and not just at some abstract level as assumed in LFG and in much HPSG work.

Naturally, the proposals developed here raise various questions. One is whether there is any connection between the strongly head-initial nature of Welsh and the type of agreement that it displays. We have very similar data in Irish, which is also strongly head-initial. It would obviously make sense to look at other VSO languages.

Questions also arise about languages with mixed systems of agreement in which superficial order seems crucial in some cases but not others. English is a case in point. As noted in the introduction, superficial order is generally irrelevant to agreement in English. However, as noted in section 5, English has agreement with the first conjunct in *there*-insertion sentences. A rather different mixed system is found in Polish. Here, when a coordinate subject follows the verb, the verb may agree either with the whole subject or the first conjunct, as the following illustrate:

---

<sup>34</sup> Further evidence for this contention comes from mutation. Borsley (1999) argues that soft mutation is triggered *inter alia* by an immediately preceding complement and that the relevant constraint applies to order domains. On the assumption that post-verbal subjects are complements, this accounts among other things for the mutation of the object of a finite verb illustrated in (i).

- (i) Gwelais i **gi**. (ci)  
*see.1SG I dog*  
 I saw a dog.

The same mutation occurs in a null-subject sentence, as (ii) shows:

- (ii) Gwelais **gi**. (ci)  
*see.1SG dog*  
 I saw a dog.

It follows that null-subjects must be represented in order domains.

- (111) a. Do pokoju weszli jeden facet i dwaj chłopcy.  
*to room came.MPL one man and two boys*  
 b. Do pokoju wszedł jeden facet i dwaj chłopcy.  
*to room came.MSG one man and two boys*  
 One man and two boys came into the room.

When a coordinate subject precedes the verb, the verb may only agree with the subject as a whole, as the following show:

- (112) a. Jeden facet i dwaj chłopcy weszli do pokoju.  
*one man and two boys came.MPL to room*  
 b. \*Jeden facet i dwaj chłopcy wszedł do pokoju.  
*one man and two boys came.MSG to room*  
 One man and two boys came into the room.

It looks as if agreement in both languages must involve both a mechanism referring to abstract levels of structure of the kind assumed in most syntactic frameworks and a superficial mechanism of the kind proposed here for Welsh. How exactly this should be worked out is not immediately obvious.

Thus, there are various questions that arise here. I hope, however, that I have established some fundamental properties of Welsh agreement and shown how they can be accommodated within one framework.

## ACKNOWLEDGEMENTS

This paper is a distant descendant of a paper called ‘Agreement, mutation and missing NPs in Welsh’ presented at the Ninth International Conference on Head-driven Phrase Structure Grammar, Seoul, Korea, 8–9 August 2002. An earlier version of the present paper was presented at the Eleventh Welsh Syntax Seminar at Gregynog, Wales, 12–13 July 2004. Some of the ideas presented here appear in chapter 6 in Borsley, Tallerman and Willis (2006). At various times I have benefited from the comments of David Adger, Emily Bender, Andreas Kathol, Ad Neeleman, Adam Przepiórkowski, Andrew Radford, Louisa Sadler, Ivan Sag, Nick Sobin, Andrew Spencer, and especially Maggie Tallerman and David Willis. I have also benefited from the comments of Peter Culicover, John Moore, and a number of anonymous *NLLT* referees. I am grateful to Ewa Jaworska for help with the preparation of the final draft of the paper. Any bad bits are my responsibility.

**References**

- Ackema, Peter and Ad Neeleman. 2004. *Beyond morphology: Interface conditions on word formation*. Oxford: Oxford University Press.
- Adger, David. 2000. Feature checking under adjacency and VSO clause structure. In *The nature and function of syntactic categories*, ed. Robert D. Borsley, 79–100. New York: Academic Press.
- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
- Bobaljik, Jonathan D. 2006. Where’s  $\Phi$ ? Agreement as a post-syntactic operation. In *Phi-Theory: Phi features across interfaces and modules*, eds. David Adger, Susana Béjar, and Daniel Harbour, 295–328. Oxford: Oxford University Press.
- Borsley, Robert D. 1986. Prepositional complementizers in Welsh. *Journal of Linguistics* 20: 277–302.
- Borsley, Robert D. 1989a. An HPSG approach to Welsh. *Journal of Linguistics* 25: 333–354.
- Borsley, Robert D. 1989b. Phrase structure grammar and the Barriers conception of clause structure. *Linguistics* 27: 843–863.
- Borsley, Robert D. 1993. On so-called ‘verb-nouns’ in Welsh. *Journal of Celtic Linguistics* 2: 35–64.
- Borsley, Robert D. 1997. On a nominal analysis of Welsh verb-nouns. In *Essays in memory of Conn R. Ó Cleirigh*, eds. Anders Ahlquist and Vera Čapková, 39–47. Dublin: The Linguistics Institute of Ireland.
- Borsley, Robert D. 1999. Mutation and constituent structure in Welsh. *Lingua* 109: 263–300.
- Borsley, Robert D. 2005. Against ConjP. *Lingua* 115: 461–482.
- Borsley, Robert D. 2006. On the nature of Welsh VSO clauses. *Lingua* 116: 462–490.
- Borsley, Robert D. 2008. On some Welsh unbounded dependency constructions. *Essex Research Reports in Linguistics* 57.4: 1–21.  
<http://www.essex.ac.uk/linguistics/errl/errl57-4.pdf>. Accessed 4 November 2008.
- Borsley, Robert D. and Bob Morris Jones. 2005. *Welsh negation and grammatical theory*. Cardiff: University of Wales Press.
- Borsley, Robert D. and Andreas Kathol. 2000. Breton as a V2 language. *Linguistics* 38: 665–710.
- Borsley, Robert D., Maggie Tallerman, and David [W. E.] Willis. 2007. *The syntax of*

- Welsh*. Cambridge: Cambridge University Press.
- Bouma, Gosse, Robert Malouf, and 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.
- Chomsky, Noam. 1982. *Some concepts and consequences of the theory of Government and Binding*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond – The cartography of syntactic structure*, vol. 3, ed. Adriana Belletti, 104–131. Oxford: Oxford University Press.
- Citko, Barbara. 2004. Agreement asymmetries in coordinate structures. In *Formal approaches to Slavic linguistics 12: The Ottawa Meeting 2003*, eds. Olga Arnaudova, Wayles Browne, María-Luisa Rivero, and Danijela Stojanović, 91–109. Ann Arbor, MI: Michigan Slavic Publications.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Culicover, Peter W. and Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.
- Doron, Edit. 2000. VSO and left-conjunct agreement: Biblical Hebrew vs. Modern Hebrew. In *The syntax of verb-initial languages*, eds. Andrew Carnie and Eithne Guilfoyle, 75–95. Oxford: Oxford University Press.
- Ginzburg, Jonathan and Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford, CA: CSLI Publications.
- Johannessen, Janne Bondi. 1998. *Coordination*. Oxford: Oxford University Press.
- Kathol, Andreas. 1999. Agreement and the syntax–morphology interface in HPSG. In *Studies in contemporary phrase structure grammar*, eds. Robert D. Levine and Georgia Green, 223–274. Cambridge: Cambridge University Press.
- Kathol, Andreas. 2000. *Linear syntax*. Oxford: Oxford University Press.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Levine, Robert D. and Thomas E. Hukari. 2006. *The unity of unbounded dependency constructions*. Stanford, CA: CSLI Publications.
- Manning, Christopher and Ivan A. Sag. 1999. Dissociations between argument structure and grammatical relations. In *Lexical and constructional aspects of linguistic explanation*, eds. Andreas Kathol, Jean-Pierre Koenig, and Gert Webelhuth, 63–78. Stanford, CA: CSLI Publications.
- Marten, Lutz. 2005. The dynamics of agreement and conjunction. *Lingua* 115: 527–547.
- McCloskey, James. 1986. Inflection and conjunction in Modern Irish. *Natural Language & Linguistic Theory* 4: 245–82.
- McCloskey, James and Kenneth Hale. 1984. On the syntax of person–number inflection in Modern Irish. *Natural Language & Linguistic Theory* 1: 487–533.
- Morris-Jones, John. 1931. *Welsh syntax: An unfinished draft*. Cardiff: University of Wales Press.
- [http://kimkat.org/amryw/1\\_gramadeg/gramadeg\\_4\\_1931\\_jmj\\_welsh\\_syntax\\_01\\_sganiadau\\_amrwd\\_2640e.htm](http://kimkat.org/amryw/1_gramadeg/gramadeg_4_1931_jmj_welsh_syntax_01_sganiadau_amrwd_2640e.htm). Accessed 4 November 2008.
- Ojeda, Almerindo E. 1988. A linear precedence account of cross serial dependencies. *Linguistics and Philosophy* 11: 457–492.

- Penn, Gerald. 1999. Linearization and *wh*-extraction in HPSG: Evidence from Serbo-Croatian. In *Slavic in Head-driven phrase structure*, eds. Robert D. Borsley and Adam Przepiórkowski, 149–182. Stanford, CA: CSLI Publications.
- Peterson, Peter G. 2004. Coordination: Consequences of a lexical-functional account. *Natural Language & Linguistic Theory* 22: 643–679.
- Pollard, Carl J., Robert Kasper, and Robert D. Levine. 1993. Studies in constituent ordering: Towards a theory of linearization in Head-driven Phrase Structure Grammar. Research Proposal to the National Science Foundation, Ohio State University.
- Pollard, Carl J. and Ivan A. Sag. 1994. *Head-driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press.
- Reape, Michael. 1994. Domain union and word order variation in German. In *German in Head-driven Phrase Structure Grammar*, eds. John Nerbonne, Klaus Netter, and Carl J. Pollard, 151–98. Stanford, CA: CSLI Publications.
- Roberts, Ian. 2005. *Principles and parameters in a VSO language: A case study in Welsh*. Oxford: Oxford University Press.
- Roberts, Ian and Ur Shlonsky. 1996. Pronominal enclisis in VSO languages. In *The syntax of the Celtic languages*, eds. Robert D. Borsley and Ian Roberts, 171–199. Cambridge: Cambridge University Press.
- Rouveret, Alain. 1994. *Syntaxe du gallois: Principes généraux et typologie*. Paris: Editions du CNRS.
- Sadler, Louisa. 1988. *Welsh Syntax: A Government-Binding approach*. London: Croom Helm.
- Sadler, Louisa. 1997. Clitics and the structure–function mapping. In *LFG97 Conference*, eds. Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/2/sadler-lfg97.pdf>. Accessed 4 November 2008.
- Sadler, Louisa. 1999. Non-distributive features and coordination in Welsh. *LFG99 Conference*, eds. Miriam Butt and Tracy Holloway King. Stanford, CA: CSLI Publications. <http://csli-publications.stanford.edu/LFG/4/sadler/lfg99-sadler.html>. Accessed 4 November 2008.
- Samvellián, Pollet. 2007. A (phrasal) affix analysis of the Persian Ezafe. *Journal of Linguistics* 43: 605–645.
- Sobin, Nicolas. 2004. Expletive constructions are not “lower right corner” movement constructions. *Linguistic Inquiry* 35: 503–508.
- Tallerman, Maggie. 1998. The uniform case-licensing of subjects in Welsh. *The Linguistic Review* 15: 69–133.
- Van Koppen, Marjo. 2005. *One probe – two goals: Aspects of agreement in Dutch dialects*. Ph.D. thesis, Leiden University. [LOT-publications 105.]
- Villavicencio, Aline, Louisa Sadler, and Doug Arnold. 2005. An HPSG account of closest conjunct agreement in NP coordination in Portuguese. In *12th International Conference on Head-driven Phrase Structure Grammar*, ed. Stefan Müller, 427–447. 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.
- Willis, David W. E. 2006. Against N-raising and NP-raising analyses of Welsh noun phrases. *Lingua* 116: 1807–1839.
- Yatabe, Shuichi. 2001. The syntax and semantics of left-node raising in Japanese. In *7th International Conference on Head-Driven Phrase Structure Grammar*, eds. Dan Flickinger and Andreas Kathol, 325–344. Stanford, CA: CSLI Publications.

- Yatabe, Shuichi. 2004. A comprehensive theory of coordination of unlikes. In *11th International Conference on Head-driven Phrase Structure Grammar*, ed. Stefan Müller, 335–355. Stanford, CA: CSLI Publications.
- Zwicky, Arnold M. and Joel A. Nevis. 1986. Immediate precedence in GPSG. In *Interfaces: Papers by Arnold M. Zwicky* (Working Papers in Linguistics 32), 133–138. Columbus, OH: Ohio State University.