

# Clitics and the Structure-Function Mapping

Louisa Sadler

University of Essex

**Proceedings of the LFG97 Conference**

University of California, San Diego

Miriam Butt and Tracy Holloway King (Editors)

1997

CSLI Publications

<http://www-csli.stanford.edu/publications/>

### Abstract

Bresnan (1997) proposes a model of constituent structure revolving around two universal modes of organisation - endocentricity and lexocentricity. For endocentric structures, a set of universal principles of structure-function association provides a canonical mapping from nodes to f-structures. This model proposes a highly constrained view of constituent structure, but does not take into account syntactic cliticisation.

We offer an account of Welsh pronominal cliticisation within this framework. We argue that the pronominal clitics in question are syntactically (rather than morphologically) positioned and provide a structure for them. We extend the universal principles of structure-function association to permit argument functions to be associated with clitic positions. We then turn to distributional questions concerning forms of pronominal expression in Welsh, demonstrating the relevance of the notions of mutual exclusion and pre-emption to cliticisation. The LFG account captures mutual exclusion phenomenon in a simple and straightforward way. Finally we show that the proposed structure goes some way to explaining the pre-emptive quality of cliticisation.<sup>1</sup>

## 1 Introduction

Bresnan (1997) accounts for the observed variability of phrasal or external syntactic structures by postulating two modes of phrasal organisation - endocentricity and lexocentricity - as the unmarked options made available universally. Endocentricity is modelled by a version of  $X'$  syntax in which the categorial signature includes a set of functional categories. Variability follows, *inter alia*, because languages may exhibit a mixed endocentric/exocentric organisation: in particular the subject-predicate distinction may be configurationally encoded in the exocentric construction  $S \rightarrow DP, XP$ .

The principles of structure-function correspondence for endocentric constructions are:

- (1) (a) complements of lexical categories are non-discourse arguments functions
- (b) complements of functional categories are co-heads
- (c) specifiers of lexical categories are adjuncts
- (d) specifiers of functional categories are discourse functions
- (e) c-structure heads are f-structure heads

This version of LFG adopts the extension of  $X'$  syntax to the minor categories, which are treated as functional heads projecting a maximal projection in their own right: a functional head is one which is (f-structural) co-head with its c-structure complement. In some languages, inflected subclasses of lexical categories (e.g. inflected verbs), as well as the closed class elements (determiners, pronouns, complementisers, particles), may exhibit functional characteristics. In Welsh, for example, all finite verbs forms are I, projecting to IP. Functional and lexical categories are not distributed in a random fashion but stand in some sort of selectional relationship to each other. One implementation of this idea is the proposal that a functional category (and its projection) shares the categorial features of a particular lexical category - thus D may be identical in categorial features with N, and I with V, while a functional category may select as its complement only that or those lexical projections with which it is categorially compatible (hence I selects verbal projections, and D selects nominal projections).

What possible analyses of syntactic (pronominal) cliticisation emerge in this model? Clearly, the pronominal status of the clitic suggests that it is a D. The principles of endocentricity then

---

<sup>1</sup>Thanks to Bob Borsley, Joan Bresnan, Andrew Spencer and participants at LFG97 for helpful comments and suggestions.

require it to project a functional maximal projection. In principle, this FP may itself appear as complement or specifier of a lexical head, or adjoined to a projection, or the F may itself take a sister XP as complement (this latter is essentially the sort of analysis adopted in Head movement accounts of cliticisation such as that of Rouveret (1994)). In section 2.2 we consider further these structural possibilities, but we begin by giving a basic outline of the facts of Welsh cliticisation.

## 2 Welsh clitics and C-structure

The Welsh clitic pronouns we are concerned with are phonologically pro-clitic to their (morphosyntactic) host, and the singular forms exert a mutation effect on the following element.<sup>2</sup>

PNG	clitic	mutation effect
1S	fy	nasal mutation
2S	dy	soft mutation
3SM	ei	soft mutation
3SF	ei	aspirate mutation
1PL	ein	
2PL	eich	
3PL	eu	

The pro-clitic pronouns of (2) occur obligatorily in three environments: to express the pronominal object of a non-finite verb (in tenseless and periphrastic clauses), the pronominal possessor, and the subject of the one non-finite form *bod* (*be*) which appears in I.<sup>3</sup>

These environments are exemplified in (3) - (8), which also show that despite the otherwise rigidly head initial character of Welsh, the clitic pronoun appears in a non-canonical pre-head position (and is optionally doubled by a post-head pronoun).

- (3) Mae Emrys wedi **gweld y plentyn**.  
 is.3s Emrys Asp see the children  
 Emrys has seen the children.
- (4) Gwadodd ei fod wedi dweud hynny, a minnau wedi **ei glywed**.  
 deny-pt.3s 3sm be Asp say that, and 1s Asp 3s hear  
 He denied that he had said that, and I having heard it.
- (5) **car y meddyg**  
 car the doctor  
 the doctor's car
- (6) **fy mhen (i)**  
 1s head (1s)  
 my head
- (7) Dywedodd Gwyn **fod Emrys yn ddiog**.  
 say-pt.3s Gwyn be Emrys pt lazy

<sup>2</sup>A set of (closely related) variant forms are phonologically en-clitic to the preceding word (in traditional terminology, the infixed forms). Assessing the morphosyntactic status of these forms (as syntactic words or as affixes) is complex, and we do not dwell on it here, being concerned essentially with the analysis of the forms in (2).

<sup>3</sup>The phonologically en-clitic forms also occur optionally to express the object of an inflected verb in I, but then only in the presence of a presentential particle.

Gwyn said Emrys is lazy.

- (8) Dywedodd Gwyn **ei fod ef** yn ddiog.  
 say-pt.3s Gwyn 3S be 3sm pt lazy  
 Gwyn said he is lazy.

The data in (7) and (8) illustrates an interesting phenomenon: in certain embedded contexts, a morphologically non-finite form of *bod* (*be*) may appear in I. This is the only case in which a non-finite verb occurs in sentence initial position, and the non-finite form of *bod* alternates with overtly finite forms according to the particular combination of tense/aspect selected.<sup>4</sup> In (8) the pronoun appears as a clitic rather than a pronominal inflection because the verb is a non-finite form rather than a finite form — cliticisation of the subject is possible in just this case.

## 2.1 Prefixed Pronominals are Syntactic Terminals

The term clitic is used loosely to refer to a range of morphosyntactic elements which fail to form a separate phonological word, but which differ significantly in (morphosyntactic) status. Some (pre-theoretical) clitics are canonically positioned *affixes*, combining with their structural (and prosodic) host in the morphology, subject to the rules of the word-formation component and of the lexical phonology. These elements display no mismatch between their structural (morphosyntactic) and phonological or prosodic behaviour. As affixes they are attached in the lexical component, and their lexical status may be indicated by lexical exceptions, idiosyncratic allomorphic variation, haplogy, suppletion, ordering with respect to other affixes, indeed, by any behaviour symptomatic of a lexical origin. A clitic which is in fact a word-internal affix will, given Lexical Integrity, have no syntactic representation at all (see (Miller 1992) for an analysis of French object clitic pronouns along these lines).

Others are (true) *syntactic clitics* or *bound words*. This is a syntactic  $X^0$  element forming a transparent syntactic construction with its (structural) host, but lacks the phonological status of a word. A bound word will show no morphological or lexical phonological interaction with either its structural or its prosodic host, since its phonological and its constructional (structural) attachment is post-lexical. It is expected to participate as other  $X^0$  categories do in syntactic processes. A single bound word may have different syntactic and prosodic hosts, being e.g. (syntactically) pro-clitic and thus initial in its syntactic constituent, and prosodically en-clitic on the preceding word.

It has been argued that some (pre-theoretic) clitics show a mixed behaviour — such *lexical clitics*, *phrasal affixes*, or *edge inflections*) appear to combine the morphophonological interactions of an affix with the syntactic positioning and low selectivity of a syntactic clitic. Lexical integrity may be preserved if these elements are analysed as phrasally positioned lexical affixes: their distribution is guided syntactically by *trigger* and *marker* features, but they are spelt out by the word-internal morphology (see (Zwicky 1987, Lapointe 1992a, Lapointe 1992b) for an analysis of the English possessive marker along these lines). If there are phrasal affixes, their existence complicates the application of diagnostics to clitics. Evidence for morphologically conditioned allomorphy in a case of cliticisation is consistent with both canonical and phrasal affixation. Likewise, since affixes may not be prosodically detached from within the morphological word, prosodic evidence will not distinguish between phrasal and canonical affixation. Unambiguous evidence of a syntactic relationship between clitic and host is evidence for a bound word analysis over affixation, since the

---

<sup>4</sup>Tense neutralisation occurs elsewhere in the grammar of Welsh: the non-finite verbform may stand in for a tensed form in a narrative sequence, where the subject is identical across a sequence of sentences. In this case, the first verb will be tensed and then in every subsequent clause the verb will appear in citation (nonfinite) form.

clitic-host relationship is syntactically transparent. But this sort of evidence is difficult to find. Tests such as low selectivity do not distinguish properly between phrasal affixation and syntactic cliticisation, since obviously phrasal edge phenomena are not sensitive to their host in syntactic terms.

There is very little evidence of the sort of morphological or phonological dependency that could only be captured by an affixal analysis of Welsh pre-head clitics. It is very striking that there are no lexical exceptions to pre-head cliticisation in Welsh. The clitic is not restricted to (a subclass of) V but occurs with V, I and N heads. The fact that the clitic causes initial consonant mutation of the following element is perfectly consistent with syntactic status, since ICM is not a purely lexical process. Significant evidence that the clitic attaches syntactically comes from the fact that it may be separated from the head by other  $X^0$  elements: certain Adjectives and numerals may intervene between pre-head pronominal possessors and N, and single Adjectives may intervene between pre-head pronominal objects and the non-finite V (and all these are stressed as syntactic rather than lexical constructions):

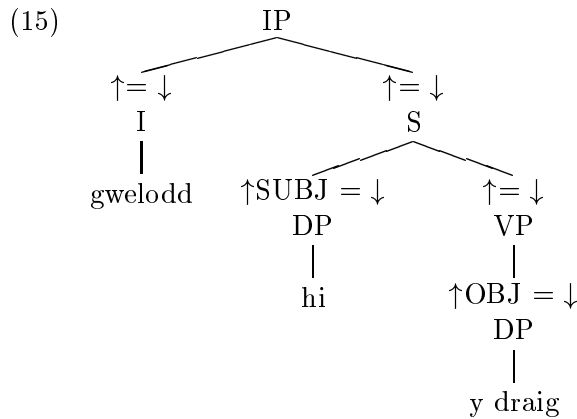
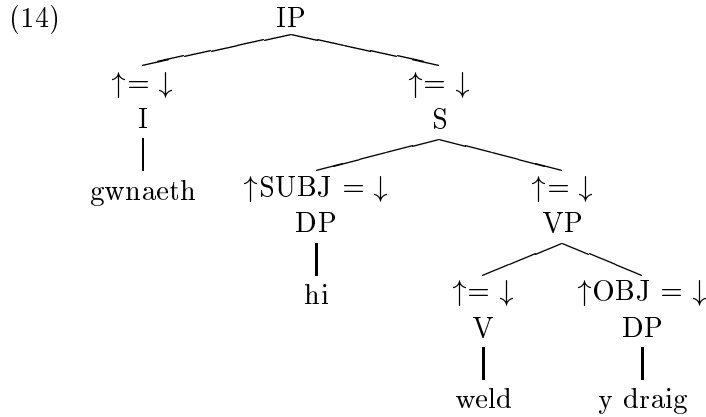
- (9) ei dri chi (ef)  
 3sm three dogs (3sm)  
 his three dogs
- (10) ei lawen gyfeillion (ef)  
 3sm happy friends (3sm)  
 his happy friends
- (11) ei wir weld  
 3sm true see  
 truly seeing him

Evidence such as this argues in favour of the view taken in all generative accounts that these are syntactic clitics. On the other hand, the clitic + host structure shows some morphological characteristics. Clitic-host structures are head-final: Welsh syntax is overwhelmingly head-initial, but head-final structures are common in the morphology. Moreover in coordinate structures, the clitic must be repeated on every coordinated head – that is, it cannot take scope over a coordinate structure. I take this mixed behaviour to indicate that Welsh clitic-host structures are transparent to the syntax, but have some “intermediate” status with respect to the syntax/morphology divide (these matters are discussed more extensively in (Sadler 1997)). The rest of the paper shows how this analysis allows an account of the properties of cliticisation in Welsh within LFG.

## 2.2 Configurational Structure

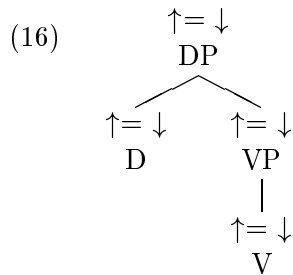
Finite clauses in Welsh are head initial, either VSOX or VSVOX. Adopting the essential insight of the long-standing V-to-I analysis of Welsh clause structure gives the structures (14) for (12) and (15) for (13) ((Kroeger 199x),(Bresnan 1997)):

- (12) Gwnaeth hi weld y draig.  
 do-pt.3s 3sf see the dragon  
 She saw the dragon.
- (13) Gwelodd hi y draig.  
 see-pt.3s 3sf the dragon  
 She saw the dragon.



As a pronominal, the clitic is a D. The structural assumptions of our model and the principles in (1) are consistent with a number of possible structural analyses of the clitic: as a functional head of a projection with the host as co-head, as adjunct to a projection, and as an element occurring in specifier position (recall that the clitic does not occur in canonical complement position).<sup>5</sup>

The **functional projection analysis** is problematic in a number of respects.



It treats the constituency of cliticised and non-cliticised non-finite VPs (and IPs) as radically different: a cliticised VP is a DP, otherwise, it is a VP. A cliticised IP is a DP, otherwise it is an IP; furthermore the assumption that the functional category D can extend a verbal projection in Welsh should be supported by independent evidence. Since D is a head, it bears the annotation  $\uparrow = \downarrow$ . The grammatical function may be introduced on the lexical item itself (rather than the pre-terminal node) by associating the clitic lexically with the disjunctive specification in (17), as well as the PRED and agreement features.<sup>6</sup>

<sup>5</sup>We rule out the possibility that an DP node dominates the clitic alone: as a clitic, we treat it as a D.

<sup>6</sup>Arguably, by failing to associate a grammatical function annotation with the D node this analysis does not reflect the continuum from pronoun through clitic to affix (an affix introduces its own functional annotation while a

- (17) **ei**:  
 $\downarrow = (\uparrow\text{SUBJ}) \vee (\uparrow\text{OBJ}) \vee (\uparrow\text{POSS})$

This function specification can be eliminated by postulating a (non-overt) DP in canonical (object) position, with the clitic providing agreement and features, treating the clitic D as the extended head of a null DP projection. Since this variant requires us to extend the use of null categories in a wholly unmotivated manner, we do not discuss it further.

The **specifier analysis** (18) is inconsistent with the structure-function mapping principles: the clitic is a (non-canonically positioned) argument while the specifier of a lexical category is an adjunct. We do not consider this analysis further.

- (18)
- 
- ```

graph TD
  VP --> D
  VP --> V_prime[V']
  V_prime --> V
  
```

This leaves the **adjunction analysis** in (19).

- (19)
- 
- ```

graph TD
  Z --> D
  Z --> Z
  
```
- where Z ranges over VP, V' and V

Structural evidence to distinguish between the variants of this analysis is difficult to come by: with strictly posthead complementation, the evidence permitting us to choose between X' and X<sup>0</sup> adjunction would concern the positioning of the clitic relative to known X<sup>0</sup> or X' adjuncts, while adjunction to XP would be distinguished from lower adjunction points only by the presence between the clitic and the lexical head of adjuncts known to adjoin to XP. The fact that the clitic cannot have scope over a coordination is suggestive of an X<sup>0</sup> analysis (though this itself does not account for all the restrictions on coordination).

Introducing prenominal clitics in X<sup>0</sup> adjunctions entails extending the same analysis to prenominal modifiers (see (21)), eschewing the rather more standard view of prenominal adjectives as X' or X'' adjuncts. In fact, a number of arguments can be made for the unconventional structure in (20).

- (20)
- 
- ```

graph TD
  NP --> N_prime1[N']
  NP --> DP
  N_prime1 --> N_prime2[N']
  N_prime1 --> AP
  N_prime2 --> N
  AP --> mawr
  N --> D
  N --> N2[N]
  D --> ei
  N2 --> A
  N2 --> N3[N]
  A --> hen
  N3 --> gi
  DP --> hi
  
```

---

pronominal DP (in Welsh) is configurationally associated with the functional annotation).

- (21) ei hen gi mawr hi  
 3sf old dog big 3sf  
 her big old dog

This analysis correctly predicts the impossibility of multiword APs in pre-head position. The head final nature of pronominal modification is less anomalous in a lexical level structure (or small construction), than it would be in a true phrasal structure. Phrasal syntax (ie syntax at the X' level and above) is rigidly head initial, but morphological structures are head final in the language.

We are thus positing a structural type which is syntactically transparent but which has much in common with morphological structures (see (Poser 1992), (Sadler and Arnold 1994) and (Sells forthcoming) for argumentation (from different perspectives) concerning such small constructions). This structure makes sense of the linear precedence facts for pre-head elements, the restrictions on pre-head modification, and the status of the clitic as intermediate between a fully incorporated pronominal affix and an independent pronoun. As we will see in a later section, the postulation of small constructions which are syntactically transparent but in competition with (other) morphological structures, permits us to extend an account of morphological blocking to the pre-emptive character of Welsh cliticisation.<sup>7</sup>

To summarise, I have proposed that Welsh bound word clitics correspond to functional categories that do not project a maximal projection. They occur in lexical level (small) constructions which show a number of characteristics of morphological structures, while remaining syntactically transparent. These proposals leave the principles in (1) intact, but posit in addition a class of (adjoined) c-heads which are not necessarily f-heads (that is, not necessarily annotated  $\uparrow = \downarrow$ ) within syntactically transparent constructions built along morphological lines. In the following section, I consider the structure-function mapping of bound word clitics.

### 2.3 Welsh clitics and f-structure

We suggested above that it was incorrect to treat the (clitic) D as an f-head (annotated  $\uparrow = \downarrow$ ), although this analysis trivially preserves the generalisation that c-heads are f-heads. Such an approach would assimilate Welsh clitics to the dependent marking mode of grammatical function assignment, despite the fact that there is no evidence of a case-encoded system operating synchronically in the language, and strong evidence of a configurational system of grammatical function assignment.

The alternative, that grammatical function assignment in clitic structures is configurational also makes sense of the use of a single set of clitic pronouns in Welsh to express the OBJ, SUBJ and POSS functions: this is consistent with the general lack of (grammatical function assignment by) case marking in the language.<sup>8</sup>

I propose to extend the principles of structure-function association to permit lexical level adjunctions (within small constructions) to correspond to argument functions, while categories adjoined to phrasal categories correspond only to non-argument functions.<sup>9</sup> The three cases of

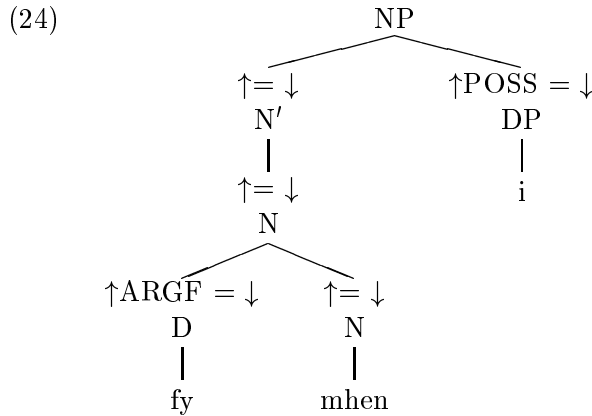
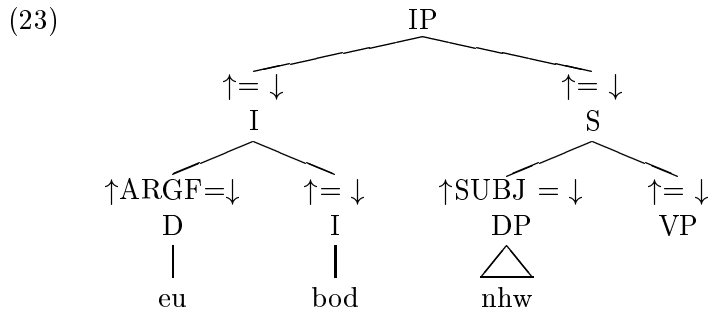
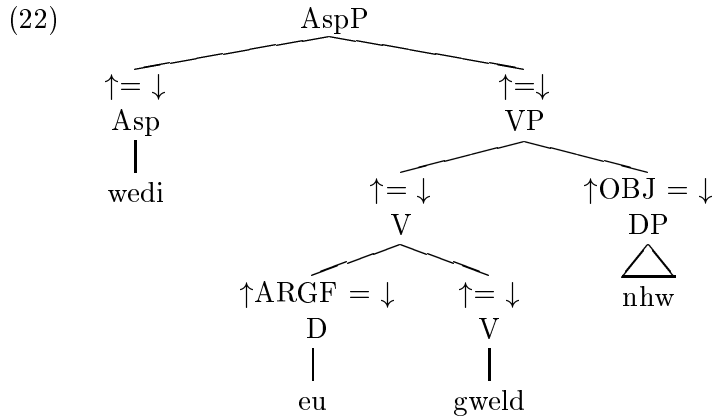
<sup>7</sup>Poser is explicit in proposing that small, syntactically transparent, constructions which are in competition with morphological constructions and in particular are subject to blocking phenomena are *morphological constructions*.

<sup>8</sup>The simplicity of the clitic system is in contrast with the incorporated pronominal affixes: the morphological system distinguishes prepositionally incorporated pronominals from the verbally incorporated pronominals — the latter combine pronominal and tense features. This argues (though somewhat obliquely) against the incorporation analysis of bound word clitics.

<sup>9</sup>Even if we restrict attention to Welsh, it is clear that they do not always provide argument functions, because of the phenomenon of lexical attachment of adjectives and numerals.



obligatory cliticisation thus involve the following structures:



How is the choice of argument function constrained?<sup>10</sup> I propose that the choice of argument function is regulated by a principle limiting cliticisation to those functions assigned within the f-structure which the c-head maps into. This principle imposes the sort of locality condition on cliticisation that we observe.

<sup>10</sup>In his HPSG analysis of Welsh cliticisation, (Borsley 1995) proposes that a cliticisation lexical rule marks as a clitic form the pronominal first complement of non-finite V and N. This approach requires an analysis whereby N subcategorises for complements and no SUBJ/POSS function. Furthermore, Borsley also assumes that finite verbs subcategorise for complements alone and no subject (and thus violate the Subject Principle of LFG) - this allows for the SUBJ cliticisation which occurs with sentence-initial *bod* (which we have argued is morphologically a non-finite form, occurring in I). There appears to be little independent motivation for Borsley's assumptions which also cannot in principle be extended to cover the case of optional cliticisation of VSO objects (which are second complements in his analysis).

- (25) A language may make use of a lexical adjunction structure to express argument functions. The argument functions which can be so expressed are limited to those generated within the  $X^{max}$  projection of the c-head  $X^0$  by the endocentric mapping principles

This principle, combined with the endocentric mapping principles, predicts the three cases of obligatory cliticisation presented in Section 2, namely OBJ clitics within VP, SUBJ clitics within IP and POSS clitics within NP.<sup>11</sup>

In the following section we turn to a number of fundamental questions concerning the distribution of the cliticisation strategy.

### 3 Clitics, Co-expression and Pre-emption

In this section we argue that our view of clitic structures as small constructions provides a straightforward account of their distribution. The facts are as follows:

1. A pronominal argument *almost always* fails to appear solely in canonical argument position, but is (obligatorily) expressed by means of either cliticisation and pronominal incorporation (*pre-emption*).
2. The clitic/incorporated pronoun may be optionally doubled by a pronominal element in canonical position, but cliticisation or pronominal incorporation never doubles a full lexical argument (*mutual exclusion*).
3. These two strategies are essentially in complementary distribution across the c-heads governing these arguments (*distribution of strategies*)

There are a number of complexities in the picture which should be signalled at this point. Firstly, there are two exceptions to the generalisation that pronominals do not appear in canonical position: the pronominal argument of a defective (non-inflecting) preposition occurs only in canonical post-head argument position, and the pronominal object of a finite verb in a VSO structure may remain uncliticised, occurring in canonical OBJ position (pronominal OBJ in VSO structures may be *optionally* en-cliticised, in the presence of a pre-sentential particle: we discuss this phenomenon in Section 3.3). Secondly, there is one exception to the complementarity of the cliticisation and incorporation strategies: a subject of a predicate in I generally occurs as an incorporated pronominal, but appears as a clitic on (uninflected) *bod* (*be*) in I.

#### 3.1 Blocking and Pre-emption

In an important paper on morphological blocking, Andrews (1990) demonstrates that LFG permits a straightforward account of both pre-emption and mutual exclusion as they occur in the distribution of the analytic/synthetic verbforms in Irish. In Irish, as in Welsh, the synthetic verb form incorporates the subject pronoun and it used obligatorily to express pronominal subjects. When available, it pre-empts the analytic form. As in Welsh, the synthetic form cannot co-occur with a

---

<sup>11</sup>It is because our principle (25) makes reference to the endocentric mapping principles (which associate SUBJ with the specifier of I) that we predict SUBJ cliticisation to *bod* in I: the SUBJ function is also associated with DP within the exocentric S construction in Welsh. If (25) placed an alternative requirement, that limited the functions to those in the  $\phi$  image of  $X^{max}$ , we would incorrectly predict the possibility of cliticising the OBJ in a VSVOX structure (thanks to Joan Bresnan for clarifying the wording of (25)).

full lexical subject (mutual exclusion); Irish differs from Welsh however in failing to permit copy pronouns with incorporated pronominals.

Andrews observes that the uniqueness of PRED values in LFG provides a straightforward account of mutual exclusion. If incorporated pronominals introduce a  $(\uparrow\text{PRED}) = \text{'PRO'}$  equation then it follows automatically (from PRED uniqueness) that an incorporated pronominal may not be doubled by a lexical element bearing a PRED value.

This insight may be extended to the case of Welsh pronominal inflections by associating the  $(\uparrow\text{PRED}) = \text{'PRO'}$  equation obligatorily with the inflection and *optionally* with the copy pronoun. These equations will permit the inflection to occur with a copy pronoun, but correctly prevents it occurring (as an agreement marker) with full lexical NPs.

This approach to mutual exclusion also extends straightforwardly to Welsh clitics:<sup>12</sup>

|                                        |                                          |
|----------------------------------------|------------------------------------------|
| ei:D                                   | hi:D                                     |
| $(\uparrow\text{PRED}) = \text{'PRO'}$ | $((\uparrow\text{PRED}) = \text{'PRO'})$ |
| (26) $(\uparrow\text{PERS}) = 3$       | $(\uparrow\text{PERS}) = 3$              |
| $(\uparrow\text{NUM}) = \text{SG}$     | $(\uparrow\text{NUM}) = \text{SG}$       |
| $(\uparrow\text{GEN}) = \text{FEM}$    | $(\uparrow\text{GEN}) = \text{FEM}$      |

The pre-emption of analytic by synthetic forms follows from Andrews' Principle of Morphological Blocking, which states that a lexical element is blocked if the contribution that it makes to the f-structure of the sentence subsumes that made by an alternative choice of lexical element:

- (27) Suppose the structure S has a preterminal node P occupied by a lexical item  $l_1$ , and there is another lexical item  $l_2$  such that the f-structure determined by the lexical entry of  $l_1$  properly subsumes that determined by the lexical entry of  $l_2$ , and that of  $l_2$  subsumes the f-structure associated with P in S (the complete structure, after all unifications have been carried out). Then S is blocked. (Andrews 1990: 519).

What is being compared here is contribution that two alternative lexical entries (the synthetic and analytic verbforms) make to the f-structure associated with the V/VP/S nodes. The principle of morphological blocking essentially requires us to pack into the morphological form as much as possible of the information that we want to express. It prefers the more informative word over the less informative word. In the case at hand, it thus *obliquely* favours a morphological means of expression (of the pronominal subject) over a syntactic means of expression (as a syntactic construction), but it does so by evaluating the contribution of competing morphological constructions (verbforms). It does not itself rule out a copy pronoun (in Andrews' analysis, this option is ruled out in Irish by the simple fact that there are no lexical entries for pronouns lacking a PRED value - PRED uniqueness thus prevents doubling).

As we have seen, pronominal cliticisation in Welsh displays precisely the same characteristics of mutual exclusion and pre-emption as pronominal inflection, strongly suggesting the relevance of Andrews' Morphological Blocking account to these constructions.

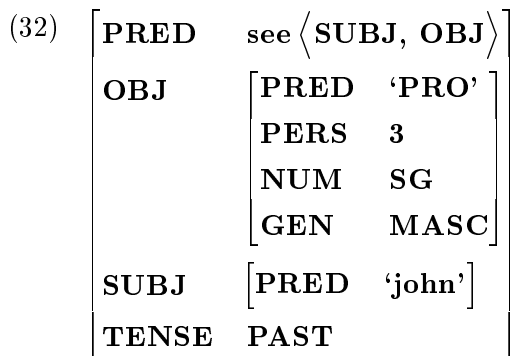
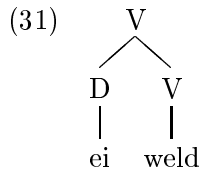
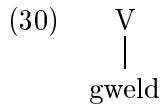
Consider how this would work in the following pair of sentences:

---

<sup>12</sup>In other languages, inflections/clitics alternate in status between incorporated pronominal and agreement marker — then, the optional PRED equation is associated with the inflection or clitic. All things being equal, this would predict (incorrectly for Welsh and Irish) that an agreement inflection could accompany a full lexical NP.

- (28) Mae Emrys wedi ei weld (ef).  
is.3s Emrys asp see 3sm see (3sm)  
Emrys has seen him.
- (29) \*Mae Emrys wedi gweld ef.  
is-3s Emrys asp see 3sm  
Emrys has seen him.

The competition is between (30) and the “phrasal” X<sup>0</sup> (31) and their associated f-structures. The f-structure corresponding to (30) subsumes that corresponding to (31) which in turn subsumes (32). Given the subsumption relation between the associated f-structures, the X<sup>0</sup> wins out, as more informative.



This behaviour constitutes strong evidence in favour of our analysis of clitic structures as small constructions. On the one hand there is a significant amount of evidence that Welsh clitics are not affixes but syntactic clitics forming syntactically transparent clitic-host constructions (Sadler 1997). On the other hand, the complete parallel between pronominal incorporation and pronominal cliticisation with respect to pre-emption strongly suggests that the latter as well as the former are subject to Morphological Blocking. The apparent contradiction between these properties may be reconciled by adopting the small construction analysis. Small constructions are syntactically transparent but subject to at least some of the principles of the morphological component — specifically, they enter into blocking relations with other lexical elements.<sup>13</sup>

In extending Andrews (1990) principle of Morphological Blocking to clitic structures in this way our proposal is very much in the spirit of work by Poser (1992). Poser suggests that those syntactically transparent constructions containing only X<sup>0</sup> material which enter into morphological blocking

<sup>13</sup>Given the syntactic transparency of clitic constructions, one might wonder whether our account (erroneously) predicts that the clitic and the full pronoun should to enter directly into a blocking competition as alternative lexical fillers of the canonical position. This does not arise because the clitic is not itself a fully-fledged syntactic terminal, but, as we have shown, enters into a small construction with its (morphosyntactic) host. Note also that the clitic has specific categorial selectional requirements (it selects an X<sup>0</sup> host).

relations with lexical items should themselves be considered to be *morphological constructions*, and we have essentially adopted this view of clitic-head structures here. Morphological Blocking is concerned with competition *within* the morphology (however the latter is construed). Bresnan (1997) formulates a more general condition governing the competition between the morphological and syntactic means of expression: her Principle of Economy of Expression, which states:

- (33) All syntactic phrase structure nodes are optional and are not used unless required by independent principles (completeness, coherence, semantic expressivity) (page 83)

in which principle syntactic phrase structure nodes refers to non-preterminal c-structure nodes (ie those which do not dominate lexical material). The Principle of Economy of Expression, as formulated, will not block the occurrence of the copy pronoun in clitic structures (or vice versa), since it does not govern the insertion of pre-terminal nodes. However the relation between these principles of Blocking and Economy remains an open and interesting question for further work.

### 3.2 Choice of Strategy

A remaining issue concerns the distribution of the two strategies of pronominal incorporation and cliticisation across the different c-heads: pronominal OBJs of P are inflectional, while those of V are clitics; pronominal SUBJs of (morphologically finite) I are inflectional, while those of (morphologically non-finite) *bod* in I are clitics, and so on. One possibility is that the clitic - head structure is available across the grammar, that is, for all c-heads, but that it is simply pre-empted by the availability of pronominal inflection (for objects of prepositions and subjects of finite verbs).

An indication that this is not the case is provided by the behaviour of the (small) number of inflectionally defective prepositions in Welsh, which lack pronominal incorporating forms. In these cases, the preposition takes a full pronominal object rather than a clitic: that is, the clitic does not stand in for the gaps in the inflectional paradigm. This suggests that there is an additional, morphotactic aspect of the distribution of the clitic strategy.<sup>14</sup> I assume therefore that clitics (as functional heads) select a sister which is disjunctively specified as [ +N ∨ +V ]. This disjunctive specification of the selectional requirements of the clitic permits cliticisation to N, V and I (I assume that the latter is specified as +V, as a verbal functional projection). On this view then, the choice between inflection or clitic for a c-head is fixed morphotactically, while the interpretation of clitics (grammatical function assignment), as we have seen, is a matter for the syntax.

Recall that although the two strategies are essentially in complementary distribution across the c-heads, the category I permits both clitic SUBJs (when non-finite) and inflectional SUBJs (when finite).

- (34) Dywedodd y dyn nad oedd neb yno y pryd hynny.  
say-pt.3s the man not was.3s no-one there the time that

The man said that no-one was there at that time.

- (35) Dywedodd y dyn ei fod yn mynd i'r dref y diwrnod hwnnw.  
say-pt.3s the man 3sm be pt go to-the town the day that

The man said that he was going to town that day.

---

<sup>14</sup>An alternative view would be that the grammar specifies once and for all that the affixal strategy rather than the clitic strategy is used for pronominal arguments to prepositions.

The question that arises is why the clitic strategy occurs *only* with non-finite forms in I. This follows without further stipulation from the fact that the verbal inflectional paradigms in Welsh express both tense and subject-related information in a single affix (3rd person plural past, 1st person singular future, and so on). This (morphological) fact ensures that pronominal incorporation accompanies the selection of TENSE. If I is non-finite, the clitic strategy emerges because the clitic + head structure (*ei fod* in (35)) blocks the alternative form *bod*.

Note finally that there is one tensed form in each paradigm which occurs with full lexical SUBJs. This form is homophonous with the 3s form, but specifies no NUM or PRED values for the SUBJ.

- (36) Darllenodd y dynion y llyfr.  
 read-3 the men the book  
 The men read the book.

The occurrence of this form with third person plural pronouns (as in (37)) is excluded by morphological blocking: the more informative *darllenasant* (3rd person plural), which specifies both NUM and PRED values, blocks *darllenodd*.

- (37) \*Darllenodd nhw y llyfr.  
 read-3 3pl the book  
 They read the book.

To summarize, I have suggested that the clitic strategy is limited to N, V and I heads by a morphotactic constraint. The limitation of cliticisation to non-finite forms in I follows without further stipulation from the fact that tense and agreement features are expressed by means of a single affixal form. In the final section, we turn to a further case of cliticisation which we have not yet discussed, involving the *optional* cliticisation of OBJ in VSO structures.

### 3.3 A Problem Case: Finite Object Cliticisation

Condition (25) permits an OBJ pronoun to occur as a clitic to the V head within VP and the principle of morphological blocking makes cliticisation obligatory in this context. Condition (25) correctly rules out the possibility of “long” cliticisation of the OBJ of a V (in VP) to I: in ISVOX structures, a pronominal OBJ cliticises to V, not to I.

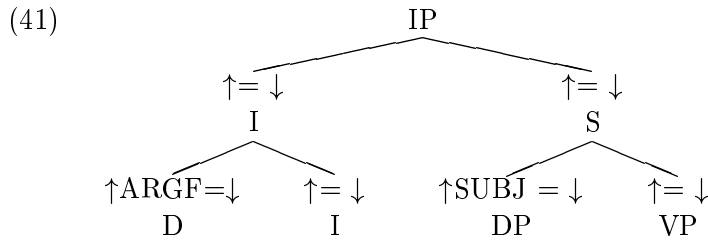
However if the *main verb* is finite (that is, in a VSO structure), the OBJ may cliticise, provided that it is immediately preceded by a (vowel final) presentential particle. This last requirement entails that only the en-clitic forms (and not the forms that have been the focus of discussion in this paper) occur as (finite verb) OBJ clitics:

- (38) Oni'ch gwelodd?  
 COMP-2pl see-pt.3s  
 Did he/she not see you?
- (39) Fe'i gwelais ef.  
 pt-3s saw-1s 3sm  
 I saw him.
- (40) Pwy a'i cred hi?  
 who pt-3sf believe.3s 3sf

Who believes her?

Nothing so far has accounted for either the *availability* or the *optionality* of this (en-)cliticisation process.

In VSO structures, the finite main verb is in I: I is the extended head of the VP projection within which the grammatical function (OBJ) of the clitic is normally assigned. Space precludes any detailed discussion of the en-clitic pronominal paradigm. It seems likely however, that most of these forms should be analysed as syntactically pro-clitic but *phonologically* en-clitic elements - that is, that a rule of the post-lexical phonology attaches them prosodically to the preceding element.<sup>15</sup> If this is correct the structure is as in (41):



The notion extended head is defined as follows in (Bresnan 1997):

- x is an extended head of y if
  - (i) x is a categorial head of y *or*
  - (ii) y lacks a categorial head and x and y are mapped into the same functional structure, x is categorially similar to y and every node in the tree which dominates  $x^{max}$  also dominates y

Reformulating (25) as follows accounts for the availability of finite object cliticisation:

- (42) A language may make use of a lexical adjunction structure to express argument functions. The argument functions which can be so expressed are limited to those generated within the  $X^{max}$  projection of the (extended) c-head  $X^0$  under the (universal) endocentric mapping principles

The remaining question, then, is why cliticisation is optional in this case. One possibility is that these forms are exceptional in failing to trigger morphological blocking: (Andrews 1990) introduces a flag OPTIONAL to account for a similar failure of certain synthetic forms to block analytic forms in Irish. He notes that the use of some synthetic forms is register specific and thus that an OPTIONAL form may be present in the grammar of one register but not another: it is significant to note in this context that finite object cliticisation is quite marginal in contemporary Welsh, being essentially limited to the written literary language.

An alternative line of investigation also presents itself, which we cannot pursue here. In finite object cliticisation, the host category is an extended head which is not also the categorial head. This is not the case where cliticisation is obligatory (of the subject pronoun to *bod* in I, of the object to a non-finite verb, and of the possessor to a nominal head). It is possible that the extended head relationship is related to the failure of morphological blocking: in some mysterious way, the fact

---

<sup>15</sup>A number of these enclitic forms, however, show evidence of a lexical relationship with the phonological host, suggesting that they have been reanalysed as suffixal inflectional morphemes.

that the licensing of a function corresponding to the clitic involves the extended head relationship somehow commits the structure as more syntactic and removes it from the domain of blocking.<sup>16</sup>

## Bibliography

- Andrews, A. 1990. Unification and Morphological Blocking. *Natural Language and Linguistic Theory* 8(4):507–558.
- Borsley, R. 1995. On some similarities and differences between Welsh and Syrian Arabic. *Linguistics* 33:99–122.
- Bresnan, J. 1997. Lexical Functional Syntax. Unpublished, Stanford University.
- Kroeger, P. 199x. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: CSLI.
- Lapointe, S. 1992a. EDGE Features in GPSG. In *Papers from the 26th Chicago Linguistic Society*. CLS.
- Lapointe, S. 1992b. Life on the Edge: Argument in favor of an Autolexical Account of Edge Inflections. In *Papers from the 28th Chicago Linguistic Society*. CLS.
- Miller, P. 1992. *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland.
- Poser, W. 1992. Blocking of Phrasal Constructions by Lexical Items. In I. Sag and A. Szabolsci (Eds.), *Lexical Matters*, 111–130. Stanford: CSLI.
- Rouveret, A. 1994. *Le syntaxe du gallois*. Paris, France: Editions CNRS.
- Sadler, L., and D. Arnold. 1994. Prenominal adjectives and the phrasal/lexical distinction. *Journal of Linguistics* 30:187–226.
- Sadler, L. 1997. Welsh Clitic Pronominals. Unpublished, University of Essex.
- Sells, P. forthcoming. Case, Categories and Projection in Korean and Japanese. In Y.-S. K. H. D Ahn, M-Y Kang and S. Lee (Eds.), *Morphosyntax in Generative Grammar*. Seoul: Hankwuk Publishing Co.
- Zwicky, A. 1987. Suppressing the Z's. *Journal of Linguistics* 23/1:133–48.

---

<sup>16</sup>Note that the puzzle posed by the *optionality* of cliticisation in this one case is to some extent independent of our proposal that these en-clitic forms are in I. If it should turn out that they are syntactically attached to the pre-sentential particle (complementiser), forming [<sub>C</sub> C D ] structures, then blocking should prefer the particle + clitic combination over the bare particle.