Syntactic Clitics, the structure-function mapping and morphological blocking

April 10, 1998

Abstract

It is well known that the notions of phonological, morphological and syntactic word do not precisely coincide to pick out precisely the same elements. Such mismatches are not necessarily problematic to lexicalist theories if it can be shown that the differing definitions of wordhood are relevant to different levels of linguistic structure (thus the phonological word is not relevant to constitutent structure, but is relevant to prosodic structure). This paper focusses on a class of clitic pronouns in Welsh which appear to exhibit both properties diagnostic of syntactic wordhood and of morphological status and considers how they might be accommodated within a model which embraces the principle of lexical integrity and eschews syntactic movement.

Bresnan (1997) proposes a highly constrained model of constituent structure based on two (alternative) universal modes of organisation - endocentricity and lexicocentricity. For endocentric structures, a set of universal principles of structure-function association provide 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.

This paper offers an account of Welsh pronominal cliticisation within this version of Lexical Functional Grammar. We first demonstrate that the pronominals in question are indeed syntactically rather than morphologically positioned and show how the endocentric principles of Bresnan (1997) can naturally extend to these structures. The mapping to f-structure is accomodated by extending the universal principles of structure-function association to permit argument functions to be associated with clitic positions.

We then turn to questions of selection and the choice between different forms of pronominal expression. Where independent pronouns compete with clitic pronouns, the latter are preferred. This parallels the competition between synthetic (inflected) and analytic (syntactic) forms which is familiar from other domains, and which is often considered to fall under some variant of a Priciple of Economy. This similarity to morphological structures might be taken as evidence that clitic-host structures are after all morphological rather than syntactic in Welsh, but we argue that this conclusion would be incorrect. Instead, we show how a natural account of this can be given within LFG by extending the domain of Andrew's Morphological Blocking Principle to syntactic clitic host structures.

1 Introduction

The formal framework of LFG models syntax in terms of *parallel* structures related by co-descriptional constraints which state correspondences between these structures. These distinct syntactic representations model the external (c-structural) and internal (f-structural) properties of linguistic objects, the former by means of Phrase Structure trees of the familiar sort and the latter by means of attribute value structures. The theory embraces the lexical integrity principle, which posits a strict separation between morphological and syntactic structures and principles of composition, but recognises the fact that syntactic and morphological structures can have the same grammtical functions at f-structure.

The observed variability of phrasal syntactic structures is captured by postulating two modes of phrasal organisation - endocentricity and lexocentricity (most obviously exhibited by the non-configurational languages) as the unmarked options made available universally (Bresnan 1997). Endocentricity is modelled by a version of X' syntax in which the categorial signature includes a set of functional categories. As is common, category labels may be thought of as a convenient abbreviatory device for a set of features pertaining to level, categorial feature matrix and type (functional or lexical). This rich view of phrase structure is able to accommodate typologically diverse languages while also permitting structural variation within a single language: for example, languages may exhibit a mixed endocentric/exocentric organisation, being endocentric in one part of the phrasal backbone and exocentric in another, and may mix configurational and non-configurational modes of expression.

Bresnan (1997) proposes a (universal, unmarked) set of principles of structure-function correspondence for endocentric constructions:

(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

So-called 'minor' categories are treated as functional heads projecting to 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, as we will see, all finite verbs forms are I, projecting to IP. [1]

[1]Clearly, functional and lexical projections are not randomly associated in phrasal structure, but stand in some sort of natural, selectional relationship to each other (thus, for example, DP is sometimes thought of as (one of) *the* functional projection(s) of NP). It is not entirely obvious (in LFG or in other frameworks) how this should be captured. One possibility is the proposal that a functional category (and its projection) shares the categorial features of a particular lexical category

One might wonder what space of analyses is available for syntactic (pronominal) cliticisation in this model, which, as noted above, assumed the projection to X″ of all categories. Clearly, the pronominal status of the clitic suggests that it is a D. The principles of endocentricity then require it to project a functional maximal projection. As a functional category, it is sister to, and f-structure co-head with, a lexical 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 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 we consider further these structural possibilities, but we begin by giving a basic outline of the facts of Welsh cliticisation.

2 The Morphosyntactic Status of Welsh clitics

The Welsh clitic (traditionally, the prefixed pronoun) paradigm is shown in (2). In common with the other Celtic languages, Welsh has a complex set of initial consonant mutations, and the table also shows the mutation effect of clitics on the following element.

(2)

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 pronouns in (2) are phonologically proclitic to their (morphosyntactic) host. A set of (closely related) variant forms are phonologically enclitic to the preceding word (in traditional terminology, these are the infixed forms). Assessing the morphosyntactic status of this second set of clitic pronominal 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.

The clitic pronouns in (2) occur obligatorily in three environments: to express the pronominal object of a non-finite verb (in tenseless and periphrastic clauses), the

- thus D may be identical in categorial features with N, and I with V, and functional categories may select as its complement only that or those lexical projections with which it is categorially compatible (hence I select verbal projections, and D selects nominal projections). The theory of the relation between functional and lexical projections is of great interest in the grammar of Welsh, where e.g. the behaviour of the Verbnoun may be thought of as evidencing the crucial role of the functional shell in determining the behaviour of a lexical projection, but we leave this matters to one side.

pronominal possessor within DP and the pronominal subject of the one non-finite form *bod: be* which appears in I. [2]

Although Welsh is otherwise rigidly head initial, the clitic pronoun appears in a non-canonical pre-head position and is optionally doubled by a post-head pronoun. It thus differs in surface position from the related non-pronominal argument, as shown by the italicized portions of the pairs of sentences below.

(3) a. A oedd hi'n medru *darogan y dyfodol*?
pt was.3s she-asp able foretell the future
Was she able to foretell the future?

b. Wyt ti'n meddwl y bydd y brenin am *ein lladd ni*?
are.3s you-asp think Comp be.fut.3s the king asp 1pl kill 1pl
So you think that the king will wish to kill us?

(4) a. Gwadodd ei fod wedi dweud hynny, a minnau wedi *ei glywed*.
deny.pt.3s 3s be asp say that, and 1s.conj asp 3s hear
He denied that he had said that, and I having heard it.

b. Mae Emrys wedi *gweld y plentyn*.
is.3s Emrys asp see the children
Emrys has seen the children.

(5) a. *car y meddyg*
car the doctor
the doctor's car

b. *fy mhen (i)*
1s head (1s)
my head

(6) a. Dywedodd Gwyn *fod Emrys* yn ddiog
say.pt.3S Gwyn be Emrys pt lazy
Gwyn said Emrys is lazy.

b. Dywedodd Gwyn *ei fod ef* yn ddiog
say.pt.3s Gwyn 3s be 3sm pt lazy
Gwyn said he is lazy.

There is strong evidence that the apparently non-finite form of *bod* in example (6) *is* in fact in an I position and bears finite syntactic features despite its morphological form, which is that of a non-finite verb). Only *bod* shows this behaviour, and it is only on this assumption that these sentences are consistent with Welsh word order. Crucially, the non-finite form of *bod* in examples such as these alternates with overtly finite

[2]The infixed (phonologically enclitic) forms also occur in these environments and also optionally to express the object of an inflected verb in I after a presentential particle.

forms according to the particular combination of tense/aspect selected.[3] These cases therefore appear to involve procliticisation of a subject pronoun to a (morphologically non-finite) form in I.

2.1 Prefixed Pronominals are Syntactic Terminals

The term clitic is often used somewhat confusingly in the literature in a pre-theoretical sense to refer to a range of morphosyntactic elements which actually differ significantly in morphosyntactic status under standard assumptions (such as the Lexical Integrity Hypothesis and the principles of morphology free and phonology free syntax). What is common to all these cases is simply the failure to form a separate phonological word. In this section we briefly review the range of analyses which may be appropriate for a given pre-theoretical 'clitic' and then consider the morphosyntactic status of the Welsh prefixed pronominal.

At one end of the spectrum some (pre-theoretical) clitics turn out to be 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. Such (pre-theoretical) clitics, when correctly viewed as affixes turn out to display no mismatch whatever between their structural (morphosyntactic) and phonological or prosodic behaviour. The affixal status of such 'clitics' may be evidenced by the existence of lexical exceptions and idiosyncratic allomorphic variation, haplology, 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. An analysis along these lines, within a lexicalist framework, is proposed for the French (object) clitic pronouns in (Miller 1992a).

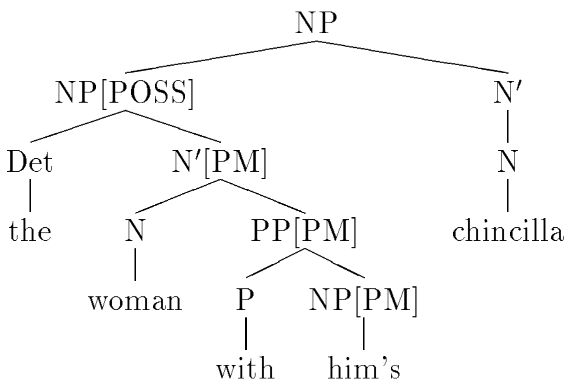
At the other end of the spectrum are clitics which turn out to be (true) *syntactic clitics* or *bound words*. A syntactic clitic is a syntactic X° element which forms a transparent syntactic construction with its (syntactic or structural) host, but which does not have the phonological status of a word. Bound word clitics involve interactions at the boundary between syntax and the phrasal phonology. A syntactic clitic (bound word) will show no morphological or lexical phonological interaction with its (structural or prosodic host), since its phonological and its constructional (structural) attachment is post-lexical. It can be expected to participate as other X° categories do in syntactic processes. As is well established, a single bound word may have different syntactic and prosodic hosts, being for example, (syntactically) proclitic and thus initial in its syntactic constituent, and prosodically enclitic on the preceeding word.

It has been argued that a number of (pre-theoretic) clitics do not fit into this simple picture. Essentially these clitics (variously termed *lexical clitic, phrasal affix, edge*

[3]Tense neutralisation such as this is not otherwise unknown in Welsh. In a narrative sequence it is possible to substitute the non-finite verbform for a tensed form, 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.

inflection) appear to show a mixed behaviour, combining the morphophonological interactions of an affix with the syntactic positioning and low selectivity of a syntactic clitic. A series of articles ((Zwicky 1987, Lapointe 1992a, Lapointe 1992b)) suggest that the English possessive marker is one such element: a typical account is Halpern's (1995) analysis which uses two sorts of feature, a *trigger* and a *marker* feature to introduce and spell out the possessive.

(7)



Analyses along these lines allow for phrasally positioned affixes (edge inflections) while preserving lexical integrity. They thus allow for morphological and (lexical) phonological dependency to co-occur with essentially syntactic positioning. The affix is attached to the host by the word formation rules, as with other affixes, and this relationship is not transparent in the syntax, but the distribution of the marker is syntactically conditioned by the introduction of the trigger feature and the percolation of the marker feature. On the downside, lexical integrity is preserved at a certain cost in terms of otherwise unmotivated book-keeping features. The elements for which an edge inflection analysis has been proposed clearly present a challenge to our current understanding of lexical integrity and any alternative analysis would appear to require us to weaken our assumptions about lexical integrity and the separation of syntax from morphology and phonology.

Note that the availability of phrasal affixation or edge inflection 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/attachment of clitic and host is evidence for a bound word analysis over an affixation analysis, since the 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.

To my knowledge, all generative accounts to date of Welsh pronominals have assumed without discussion that the prefixed pronoun is a syntactic clitic rather than an affix. In the rest of this section we review several pieces of evidence and conclude that the syntactic clitic analysis is correct.[4]

1. Regularity: It is very striking that there are no lexical exceptions to pre-head (prefixal) cliticisation in Welsh.[5] Moreover the clitic is not restricted to co-occur with (a subclass of) verbs but occurs with uninflected verbs and with nouns alike. This widespread and exceptionless distribution is suggestive of a syntactic rather than a morphological source.

2. Initial Consonant Mutation: Welsh has a complex system of initial consonant mutation processes, which bring about a set of related phonological changes. While the phonological effect of initial consonant mutation (ICM) is relatively straightforward, there is no fully adequate analysis of the set of morphosyntactic triggers, and thus ICM remains something of a descriptive challenge (but see REFS (Ball, Tallerman)) for some relevant accounts. It is clear that ICM may be triggered *syntactically*, that is, in a set of syntactic environments. (8) exhibits what is sometimes (misleadingly) referred to as Direct Object mutation in the literature, in which the second XP after a finite verb form undergoes soft mutation. (9) shows that adjectives after feminine nouns undergo soft mutation and thus shows that ICM is also sensitive to morphosyntactic features such as *gender*. (10) is an example of lexically triggered mutation - here the presentential negative particle *ni* causes soft mutation of the finite verb. As shown in (11), it also occurs word-internally, and therefore applies in the morphology.

(8) Gwelais i gi (ci) mawr
see-pt.1s 1s dog big
I saw a big dog

(9) merch dal (tal), gref (cref)
girl tall strong
a tall strong girl

(10) Ni ddangosodd (dangosodd) Siôn y llun i Mair
Comp show-pt.3s Sôn the picture to Mair
Siôn did not show the picture to Mair

(11)

af-	rhaid	afraid	SM	*unnecessary*
arch-	derwydd	archdderwydd	SM	*archdruid*
en-	mawr	enfawr	SM	*huge*
cam	deall	camddeal	SM	*misunderstand*

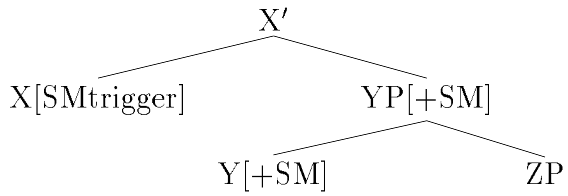
As the table in (2) summarises, certain proclitic pronouns trigger ICM, However, because there is clear evidence of syntactically, lexically and mrophologically triggered ICM, this provides no theoretical leverage at all on the analysis of proclitic. pronominals.

[4]Note that we are not concerned here with the enclitic counterparts of these clitic pronouns, for at least some of which we believe the balance of evidence is in favour of an affixal analysis.

[5]The *preceding* element may require selection of the so-called infixed form rather than the prefixed pronoun, but this is an independent matter.

Zwicky (1984,1986) and Lapointe (1988) argue that initial consonant mutation itself should be analysed essentially as phrasal affixation or lexical cliticisation. Initial consonant mutation evidently does emerge at the edge of a consitutent. Under the view that they adopt, the trigger determining the feature in question must be adjacent to the affected word and c-command it, and the edge phenomenon is spelt out as for a morphophonemic rule.[6]

(12)



This sort of account of syntactic triggering of morphophonemic alternations can be accommodated within a model of the syntax/morphology interface which respects lexical integrity, but does not force us into any particular analysis of pre-head clitics. If they are bound words, those pre-head pronominals which cause ICM are simply lexical triggers. If they are canonical affixes, then these affixes are also triggers of a left-edge phenomenon on its sister. This is not unexpected, if there is other evidence in the language of initial consonant mutation applying within the word formation component. All that this requires is that ICM is also evidenced within the morphology. If prehead clitics are themselves inflections appearing at the left edge of a phrase (and therefore at the left edge of the leftmost word in a phrase), then we have a (syntactic) left edge phenomenon (phrasal or non-canonical affixation) itself exerting a left-edge alternation (ICM) on the element to which it is lexically attached in the word formation component. This is also unproblematic, because, there is independent evidence of the application of ICM in the word formation component.

I conclude, therefore, that the phenomenon and distribution of ICM does not provide crucial evidence for the analysis of pronominal clitics.

3. Aspirate H as an initial mutation: One phenomenon which may potentially involve allomorphic variation and thus imply a morphological relationship between clitic and host is the insertion of aspirate *h*-. A vowel initial word takes an initial *h*- following a number of elements including *ei(3sf), ein(1pl), eu(3pl)* (and their infixal counterparts)::

(13) Edrychodd ar ei *h*oriawr.
looked-3s.pt at 3SF watch
She looked at her watch

(14) Y mae hyn yn ein *h*atgoffa o'r syniad.
pt is this pt 1pl remind of-the idea
This reminds us of the idea

[6]The behaviour under coordination, where the ICM is in some cases spelled out on each conjunct, is in line with similar exceptions under coordination for posessive marking.

(15) ein *h*afon a'n psygota
1pl river and-1pl fish
our river and our fish

Instead of seeing this as lexically conditioned allomorphy I would suggest that traditional grammars are correct in viewing this as a special case of ICM, in this case, syntactically or lexically triggered. If this analysis is correct, there is no independent evidence for the affixal status of pre-head pronouns from this phenomenon.

4. Vowel Reduction: The Welsh proclitic pronouns are open or closed monosyllables with a single short vowel. Thomas (1984) argues that some clitics (including the proclitic pronouns *fy, ein, dy, eich*) are derived by vowel reduction from an underlying high vowel. He relates this to the systematic phonological alternation lowering a high vowel to a schwa (as in *dyn* ('man'): *dynion* ('men')) attested in morphological complex words: lowering always occurs in the prefinal syllable of a word, and never in a (phonologically) stressed monosyllable (normal word stress falls on the penultimate syllable in Welsh). He argues that the same process of lowering applies in some morphologically noncomplex polysyllabic words, and in the phonological word formed by a clitic and its host (first and second person clitics, the clitic determiner *yr* and the aspectual particle *yn*). This evidence is consistent with a morphological analysis of clitics but it is not conclusive since, even if such a lowering rule were to be correct, there would need to be evidence of interaction with rules of the lexical phonology. Since very little work has been done on the phonology of Welsh, I conclude that this phonological argument is inconclusive.

5. Separation: Although the clitic appears pre-head, it is not always adjacent to the head. This appears to pose serious difficulties for an analysis of the clitic as a canonically positioned affix. Pre-head pronominal possessors can be separated from the head N by a relatively small number of premodifiers, including simple numerals and ordinals, quantifiers and certain (senses of) adjectives such as *hen*: 'old'; *ambell*: 'occasional'; *prif*: 'chief'; *holl*: 'all'; *pob*: 'each, every'; *pa*: 'which'; *unig*: 'only'[7].

(16) ei dri chi (ef)
3sm three dogs (3sm)
his three dogs

(17) ei lawen gyfeillion (ef)
3sm happy friends (3sm)
his happy friends

(18) ei thrydedd briodas
3sf third marriage
her third marriage

(19) hen gi maer y dref
old dog mayor the town

[7] Prenominal adjectives cause soft mutation of the following noun

the town mayor's old dog

(20) rhai merched bach
some women small
some small women

If the possessive clitic were an affix, it would have to be categorially rather unselective about its host.

In addition and in similar fashion, a single *adjectival* modifier sometimes intervenes between a nonfinite verb (the verbnoun) and a pronominal proclitic object:

(21) ei wir weld
3sm true see
truly seeing him

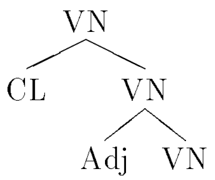
Many such combinations are completely transparent (*cyflym redeg* 'run quickly'; *llwyr ddeall* 'totally understand'; *ôl-redeg* 'run backwards' and alternative with verb + adverbail phrase constructions:

(22) rhedeg yn gyflym
run prt quick
run quickly

(23) cyflym redeg
quick run
run quickly

If the canonical affix hypothesis is to be maintained, in the light of these data, the pre-head modifiers must also be word internal elements, that is, positioned by the (compounding) word formation rules of the language, with argument-incorporating affixation occurring outside this compounding, giving a (word-internal) structure along the lines of (24):

(24)



For prenominal adjectival modification, there is certainly some lexical idiosyncracy in the sense that (readings of) adjectives must be lexically marked as occurring in pre- or post-head position. But this does not in itself require a lexical (word-internal) analysis. Furthermore the set of adjectives that occur premodifying an uninflected verb as in (22) is not co-extensive with the set of prenominal adjectives. This removes any argument for a step of zero morphological derivation in the analysis of these constructions along the lines shown in (25).

(25) $[_V[_N \text{ A N }]]$

If these observations are correct, there is no selectional evidence requiring us to postulate a lexical (word internal) source for [A N] and [A V] constructions, and yet such an analysis would be required if prehead pronominal clitics are canonical affixes.

6. Head Final: Clitic-host structures are head-final, but Welsh is overwhelmingly head initial in the syntax. The adjective/numeral/quantifier - noun constructions discussed above are of course also head final. These head-final structures are somewhat anomalous from a syntactic point of view. In the morphology, however, head-final structures predominate, since derivation and inflection is mainly suffixal, and compounds are also head-final. The head-final nature of these clitic-host structures is therefore highly consistent with a morphological, rather than a syntactic source, but it does not rule out the latter. Note also that if treated as (agreement) inflections, clitics would be the only example of prefixal inflection in the language.[8]

7. Adjective Noun constructions and stress: The adjective noun constructions discussed above exhibit a syntactic rather than a lexical stress pattern: in cases such as (16)- (21), both head and premodifiers are accented and thus stressed as separate syntactic entities.[9] These (syntactically stressed) modifier-head constructions should be distinguished from so-called strict compounds, which are stressed as single words, with accent on the penultimate syllable.

In interpreting the stress data, one must make clear assumptions about the relationship between morphological word (lexical item), phonological word and syntactic terminal. I am assuming that normally, the phonological word corresponds to the morphological word and to the syntactic terminal, with one main stress per phonological word. If an element is lexicalised, however, it may be lexically associated with single word or multi-word stress pattern (and thus may correspond to more than one phonological word). The existence of a "syntactic" stress pattern, then, does not permit us to conclude unambiguously that a construction is syntactically transparent, though it will strongly suggest this conclusion unless there is independent evidence in favour of lexicalisation. On the other hand, if a construction forms a phonological word, this will generally mean that it is syntactically opaque - the exception to this being precisely the cases of syntactic cliticisation, where a weak element forms a phonological word with its neighbour. Improper compounds are either listed (phrasal) constructions with single word stress, or they are restressed as single words by the phrase level phonology.

[8] In Welsh, finite verbs incorporate their pronominal subjects, which appear as suffixes, and (the majority of) prepositions incorporate their pronominal objects (again, as suffixes). If pronominal proclitics were canonical affixes, we would be forced to conclude that the language uses both incorporation by suffixation and incorporation by prefixation (procliticisation).

[9] These modifier head constructions are traditionally referred to as loose compounds, which may indicate some unease about their syntactic status. Orthographically, most such syntactically stressed constructions are treated as two separate words, although some are hyphenated or written as a single word. Most examples of the latter are clearly lexicalised, given their meaning, and thus must be listed together with their stress pattern.

8. Infixed Paradigm: Some (rather oblique) evidence that the pro-clitic pre-head pronominals are syntactic clitics rather than lexical affixes may be adduced from the existence of the (largely parallel) set of enclitic pronominals:

(26)

PNG	infixed form	mutation effect
1S	'm	
2S	'th	soft mutation
3SM	'w, 'i	soft mutation
3SF	'w, 'i	aspirate mutation
3S	-s	
1PL	'n	
2PL	'ch	
3PL	'w, -s, 'u	

The fundamental difference between affixes and syntactic clitics is that the former are morphologically, phonologically and prosodically integrated with their structural host, while the latter may attach prosodically (that is, form a phonological word with), an element outside the domain within which the functional effects apply. For these cases, the proclitic/enclitic alternation is entirely predictable in terms of phonological features visible in the syntax/phonology interface. This sort of alternation is of course impossible for affixes, since they form a phonological word with the stem to which they attach.

The Welsh enclitic pronouns form a phonological word with the preceding element, and in the majority of cases are related to the proclitic set by a simple process of vowel deletion. The very fact that they attach phonologically to a different word that the pro-clitic pronouns is inconsistent with an analysis of the latter as affixes, on the natural assumption that the grammar does not permit affixes to be detached from one head and re-attached to another on purely phonological grounds.

In addition, at least some of the enclitic forms (from the so-called infixed paradigm) exhibit a high degree of lexical selection with the preceding (phonological) host, suggesting that they have or are being reintepreted as inflectional suffixes. Although a detailed study of the enclitic pronouns falls outside the scope of this paper, I believe there is persuasive evidence that some (if not all) of the enclitic forms are actually suffixes. Firstly, one of these forms *'w* ([w]) occurs only with one specific lexical iterm *i* 'to' (a preposition and a complementiser or pre-sentential particle):

(27) o'i gorun i'w sawdl
from-3sm tip to-3sm toe
from his tip to his toe

(28) i'w chartref hi
to-3sf home 3sf
to her home

(29) Roedd ganddo lawer storu ddifyr i'w hadrodd
was-3s with-3sm many stories interesting to-3sf tell

He had many interesting stories to tell her

(30) y baned wedi'i hyfed
the cupful after 3sf drunk
the contents of the cup having been drunk

Secondly, there is at least one other case in which lexical absorption has clearly taken place in the third person. Infixed clitic pronominals corresponding to the object of a finite verb may occur optionally, positioned between the presentential particle and the inflected verb. In the third person, the pronominal object is morphologically incorporated (as *-s*) by the particles *ni, na, oni, pe* whereas the infix clitic form *'i* also attested in other environment (and appears e.g. with the particles *fe, a, y*). Substitution of *-s* for *'i* is clearly a case of lexical selection.

(31) Nis gwelsoch (ef)
neg-3s see-pt.2pl 3sm
You did not see him

(32) Fe'i gwelais ef
Comp-3s saw-1s 3sm
I saw him

Under the affixation analysis of the pro-clitics, the existence of such a case alongside the pro-clitic forms is quite unexpected, because it is difficult to see what diachronic or synchronic process would (selectively) detach a canonical (or phrasal) affix and re-attach it to the preceeding head.

If these remarks are along the right lines, the infixed paradigm is actually a mixture of syntactic clitics and affixes. The attachment of the clitic forms to the preceding or following phonological words on phonological grounds is entirely predictable under a postlexical cliticisation analysis, while the lexicalisation of certain PNG- host combinations is also consistent with this analysis. On the other hand, treating the pro-clitics as affixes requires us to postulate an unintuitive process of de- and re-affixation. I therefore conclude that there is evidence from the infixed paradigm for the analysis of the prefixed pronouns as syntactic clitics (bound words).

9. The Coordination Criterion: It is often suggested that coordination can be used as a criterion to distinguish syntactic clitics from affixes. Miller(1992b) proposes the following criterion:

> If an item must be repeated on each conjunct in a coordinate structure, then it must be an affix and cannot be a PLC [post lexical clitic]
>
> If an item must fail to be repeated on each conjunct in a coordinate structure, then it must be a PLC and cannot be an affix

Since the definite article differs from other determiners in French in ability to take scope over a coordination, Miller argues that the definite article is an affix:[10]

(33) *le père et/ou fils
the father and/or son

(34) le père et/ou le fils
the father and/or son

(35) chaque homme ou femme
every man or woman

(36) Douze filles et garçons
12 girls and boys

In Welsh, the data are less than straightforward, and more work on coordination needs to be done before this (putative) test can be applied with any confidence. The prefixed clitic pronouns is obligatorily repeated when nouns are coordinated:

(37) ei wasanaeth, ei gyflog, ei weision, ei lywodraeth, ei gwmni
3sm services, 3sm wages, 3sm servants, 3sm government, 3sm company
a'i wlad
and-3sm country

(38)

Prenominal numerals may coordinate with a single nominal head, thus both the examples below are possible:

(39) dwy gyllell neu dair
two knife or three
two or three knives

(40) dwy neu dair cyllell
two or three knife
two or three knives

Significantly, such coordinated numerals may intervene between proclitic pronoun and noun:

[10] Even in French the data are not completely unambiguous, however. As he himself notes, there are a number of cases in which a (supposedly) lexically attached affix may take wide scope over a coordination. Miller essentially suggests that these are vestiges of an earlier system, but one might equally well conclude that they undermine the validity of the coordination criterion itself:
(i) mon collègue et ami (coreferent)
my colleague and friend
(ii) les officiers, sous-officiers et soldats
the officers, non-commissioned officers and soldiers

(41) ei ddwy neu dair cyllell
3sm two or three knife
two or three knives

This behaviour therefore appears to give contradictory results with respect to the coordination criteria. If the data in (37 is interpreted as confirming affixal status for the possessive clitic, then to accommodate (41) we would be led to permit word-internal coordination.

The repetition of the possessive pronoun in (37) is parallel to the obligatorily repetition of the definite article with coordinated Ns:

(42) oherwydd y dŵr a'r llaid a'r tyllau
because the water and-the mud and-the holes
because of the water, mud and holes

The definite article is a phonological clitic. But the significance of phonological weakness as the causative factor here is undermined by the fact that other unstressed monosyllabic phonological clitics, the preposition *yn* and the predicative particle *yn*, are able to take wide scope:

(43) yn Llundain a Berlin
in London and Berlin
in London and Berlin

(44) Lladdasom bawb ym mhob dinas, yn ddynion, gwragedd a phlant
kill-1PL.pt everyone in every city, pt men, women and children
We killed everyone in every city, men, women and children

(45) Bydd yr afon yn hesb a sych
will.be.3s the river pt wasted and dry

The river shall be wasted and dried up

In summary, then, the coordination criteria does not appear to produce unambiguous results for the Welsh proclitic. A number of restrictions on coordination in Welsh exist (for example, it is impossible to coordinate finite sentences (IPs) under a single pre-sentential particle (Comp)) and these are not well studied or understood.

Additionally, there is some reason for thinking that the coordination criterion itself is incorrect. Note that it simply rules out the very existence of phrasal affixation: as an affix, an edge inflection or phrasal affix is required (by the criterion) to be repeated on every conjunct. It is abundantly clear, however, that a phrasal affix may take scope over a coordinate NP (as in English *John and Mary's book*).

2.2 Conclusion

The preceding subsection reviewed a number of aspects of the distribution and behaviour of proclitic pronominals. There is ample evidence that these clitic host structures are syntactically transparent, and little or no evidence of the sort of morphological or phonological dependency or selectional idiosyncracy that could only be captured by an affixal analysis. These observations motivate an analysis which takes Welsh proclitic pronouns to be syntactic elements. On the other hand, one puzzle remains, in that head final syntactic structures are at least rather unexpected in the language, which is strongly head-initial. In the nexy section, we turn to the configurational analysis of clitic host structures, but we will return to this aspect in due course and suggest an interpretation.

3 Configurational Structure

Finite clauses in Welsh are verb initial. The structure of the simple clause (containing a finite verb and its arguments and adjuncts) is basically VSOX. The structure of the periphrastic clause (containing a finite verb in initial position and one or more nonfinite verbnouns) may be given schematically as VSVOX.[11] Following a long-standing analysis of Welsh clause structure, we assume that the finite verb is in I, giving a structure along the lines of (48) for (46) and (49) for (47) ((Kroeger 1995),(Bresnan 1997)):[12]

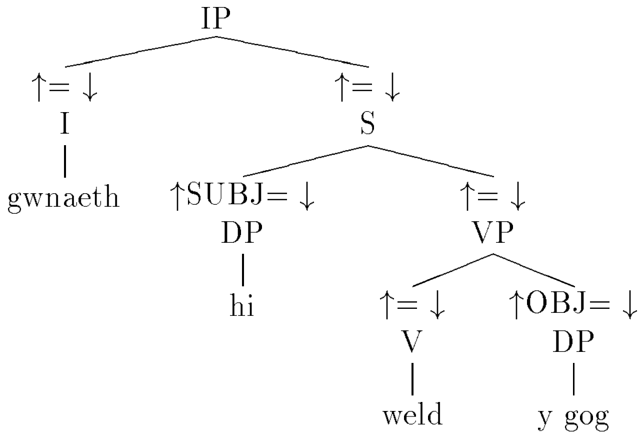
(46) Gwnaeth hi weld y gog
do-3s.past 3sf see the cuckoo
She saw the cuckoo.

(47) Gwelodd hi y gog
see-3s.past 3sf the cuckoo
She saw the cuckoo

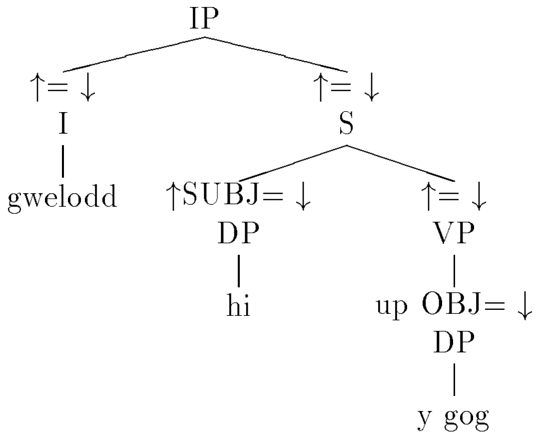
[11]As will become clear, the nonfinite verbnouns may themselves be preceded by as variety of aspectual particles to express certain tense and aspect combinations.

[12]These structures use the (universally available) exocentric but configurational phrase structure constraint S $\rightarrow$ DP XP which introduces a subject function as sister to a predicate. This essentially models VP internal subjects without movement. For further discussion and motivation, see (Bresnan 1997).

(48)



(49)



In (49) the main (finite) verb is in I. The principles of X′ syntax require the VP to project a head, and in some generative models this requirement is satisfied by postulating a syntactic movement of the head from V to I, leaving a trace. In LFG, the relationship between I and the head of VP is expressed non-derivationally (and indeed, movement to pick up affixes or check features is outlawed by the principle of relativized lexical intergrity which bans the reification of pieces of morphology in the syntactic tree). By the definition below (Bresnan 1997), I serves as the extended head of VP, obviating the need to postulate empty c-structure nodes:

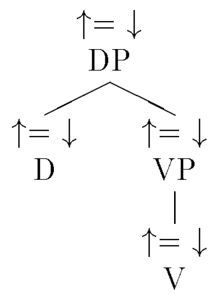
> x is an extended head of y if
> (i) x is a categorial head of y *or*
> (ii) y lacks a categorial head and x and y are mapped into the same functional structure, x is categorially similar to y and every node in the tree which dominates x^{max} also dominates y

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).[13]

[13]We rule out the possibility that an DP node dominates the clitic alone: as a clitic, we treat it as a D.

Consider first the functional projection analysis (illustrated with respect to the cliticisation of the object of the non-finite verb):

(50)



This analysis is strongly reminiscent of the Principles and Parameters analysus of (Rouveret 1994) (although the latter also postulates head movement in the derivation of the non-finite verb (or verbnoun) itself. This is problematic in a number of respects. In terms of consitutent structure and categorial compatibility, we must open up the relationship of consistency between functional and lexical categories to permit D to extend the verbal projection, either by bearing verbal categorial features and/or by selecting a verbal projection (VP and IP) as well as NP. There is evidence from other languages that a D can take a VP or a CP complement sister, of course, but these are cases in which what we have is essentially a nominalisation of a verbal projection (the resulting phrase has nominal semantics, for example). This is quite different from the case at hand, of course, where the projection must be associated with verbal semantics. This analysis also involves a significant divergence between the constituency of cliticised and non-cliticised non-finite VPs (and IPs): a cliticised VP is a DP, otherwise, it is a VP. A cliticised IP is a DP, otherwise it is an IP. Moreover, there is no independent evidence that in Welsh a VP can be sister to a D - definite articles, for example, are completely excluded in this position.

With respect to the mapping to f-structure, the functional head itself, D, will bear the annotation $\uparrow = \downarrow$, as a zero level projection. In fact, this seems slightly unnatural, given that the clitic may be thought of as representing a stage in the progression from full pronoun to affix. An affix bears its own functional annotation while a pronominal DP (in Welsh) is *configurationally* associated with the functional annotation. By failing to associate an annotation with the clitic this analysis fails to reflect the morphosyntactic continuum from pronoun through clitic to affix.

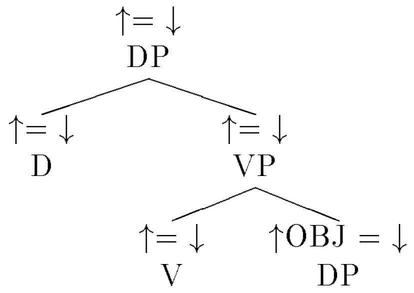
The grammatical function specification may be introduced on the lexical item itself (rather than the pre-terminal node) by associating the clitic lexically with the disjunctive specification in (51):

(51) **ei:**
$\downarrow$= ($\uparrow$SUBJ) OR ($\uparrow$OBJ) or ($\uparrow$POSS)
($\downarrow$PRED) = 'PRO'
($\downarrow$PERS) = 3rd
($\downarrow$NUM) = sing

(↓GEN) = fem

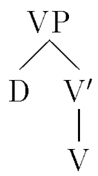
Postulating a (non-overt) DP in canonical position permits us to eliminate the function assignment,, but requires us to treat the clitic D as the extended head of a null DP projection that it c-commands, extending the use of null categories in a wholy unmotivated manner.[14]

(52)



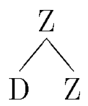
A further possibility is to take the proclitic to be a zero-level specifier, as shown in the specifier analysis, (53). However this inconsistent with the structure-function mapping principles: the clitic is a (non-canonically positioned) argument while the specifier of a lexical category is an adjunct according to these principles. Accordingly, we do not consider this analyses further.

(53)



This leaves the **adjunction analysis** in (54) to consider.

(54)



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 chose between X' and X^0 adjunction would concern the positioning of the clitic relative to known X^0 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 occupy XP. As we have seen above, the only element which can intervene within the VP between the nonfinite verb and the proclitic is a single premodifying adjective, which seems most plausibly to be analysed as a lexical adjunction. If this is right, then there is no evidence from the positioning of (other) adjuncts to determine the level of adjunction in clitic structures.

[14]The notion of extended head will be discussed below.

Adjunction to XP or X′ would in principle permit the clitic to have scope over coordinated VPs, V′s or V^0s. The fact that the clitic cannot have scope over a coordination (as discussed above) is on the other hand suggestive of an X^0 analysis (though this itself does not account for all the restrictions on coordination).

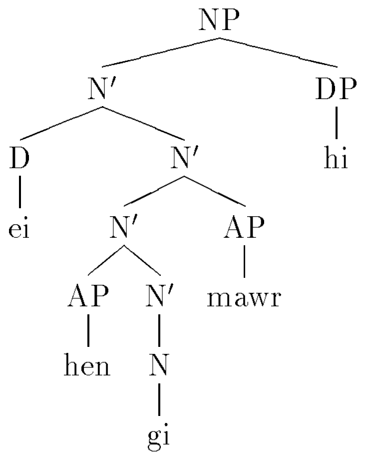
To sum up so far, the functional projection analysis is not well motivated, though it is accommodatable within the model of syntax that we adopt, while the specifier analysis is ruled out by the structure-function mapping principles. The adjunction analysis, and specifically adjunction to V^0 appears to be the most promising analysis. We now turn to consideration of the adjunction analysis of possessive proclitics within nominal structures.

We begin by considering adjectival modification, and concretely an example such as (55):

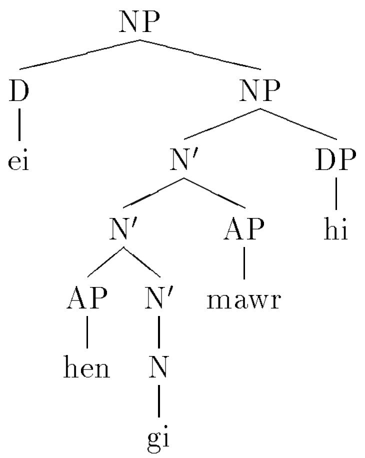
(55) ei hen gi mawr (hi)
3sf old dog big (3sf)
her big old dog

A standard view would be that N-modifying adjectives are AP adjunctions to either N′ or NP: (56) and (57) reflect the N′ adjunction option.

(56)



(57)

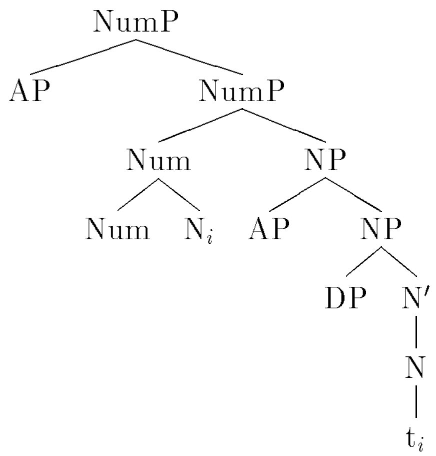


Note that the fact that postnominal adjectives (when present) intervene between the head noun and a (canonically positioned) possessor NP (see (58) is not neccessarily inconsistent with an analysis which adjoins AP to NP, however.

(58) llyfr newydd Dafydd
book new Dafydd
Dafydd's new book

In fact, several variants have been proposed of a head movement analysis which generates postnominal APs as adjunctions to NP with possessor phrases as Spec of NP and involves raising the nominal head to a higher functional projection (NumP) (with prenominal adjectives adjoined to that functional projection), schematically:

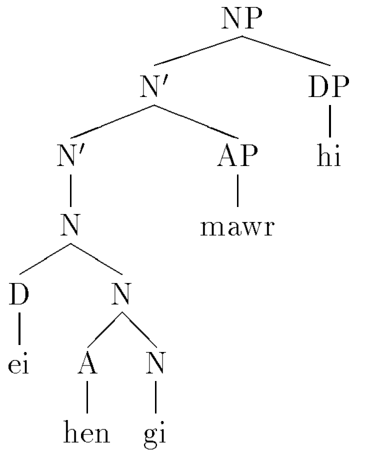
(59)



Analyses along these lines are found, for example, in (Duffield 1996) and (Rouveret 1994), but are not without problems (for some discussion, see (Sadler 1998)). These analyses generally take proclitic pronouns to be generated in a further functional head position D, but the N-to-Num analysis is logically distinct from the analysis of the clitic as the functional head of the phrase, and can be combined with the view of cliticisation as adjunction (specifically, to NumP).

Despite the conventionality of treating prenominal modification as some sort of phrasal adjunction, a number of arguments can be made for the less conventional structure in (60).

(60)



Treating prenominal adjectives as bare A correctly predicts the impossibility of multiword APs in pre-head position. The head final nature of prenominal 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 (Sells forthcoming), (Poser 1992) and (Sadler and Arnold 1994) for argumentation (from different perspectives) concerning 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. [15]

In summary, I have propose that Welsh bound word clitics correspond to functional categories that do not project to a maximal projection. They occur in lexical level (small) constructions which show a number of characteristics of morphological structures, but are syntactically transparent. Such lexical level constructions, participating in both morphological and syntactic systems, are not without precedent in the literature. These proposals leave the principles in (1) largely intact, although we admit the existence 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 turn to this matter and consider the structure-function mapping of bound word clitics.

[15] 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*.

4 Welsh clitics and f-structure

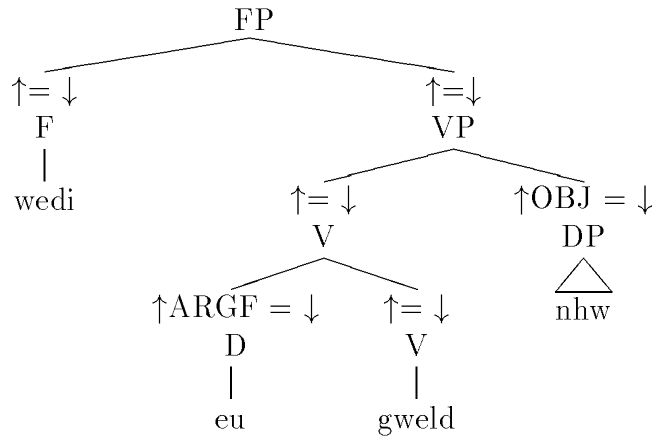
In this section we consider the structure-function mapping for (the syntactic) clitic pronouns. There are a number of reasons why it seems incorrect to treat the D as an f-head (that is, annotated $\uparrow = \downarrow$), although this analysis (somewhat trivially) preserves the generalisation that c-heads are f-heads. A serious problem with the analysis is that it assimilated Welsh clitics to the dependent marking mode of grammatical function assignment, despite the fact that there is no evidence at all of a case-encoded system operating synchronically in the language, while there is strong evidence of a configurational system of grammtical function assignment.

If grammatical function assignment in clitic structures is configurational, it is not surprising that, somewhat unusually from a comparative perspective, a single set of clitic pronouns in Welsh may be associated with variety of different grammatical functions (OBJ, SUBJ and POSS). This is fully consistent with the general lack of (grammatical function assignment by) case marking in the language.[16]

I propose to extend the principles of structure-function association to permit lexical level adjunctions to head categories (within small constructions) to correspond to argument functions. That is, categories adjoined to phrasal categories correspond only to non-argument functions, but lexical level adjunctions may provide argument functions. Clearly, the phenomenon of lexical attachment of adjectives and numerals indicates that lexical adjunction does not always introduce argument functions. The three cases of obligatory cliticisation we are concerned with therefore involve the structures below.

(61) Mae Emrys wedi eu gweld nhw.
is Emrys asp 3pl see 3pl
Emrys has seen them

(62)

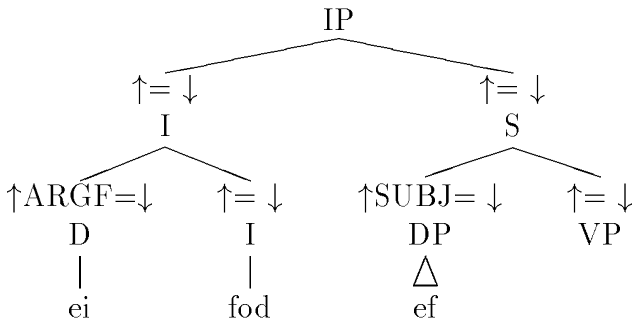


[16]The simplicity of the clitic system (in this regard) is in sharp contrast with the incorporated pronominal affixes: affixally, we must distinguish prepositionally incorporated pronominals from the verbally incorporated pronominals (verb inflections combine pronominal and tense features, while prepositional inflections do not). This also argues, though admittedly somewhat obliquely, against the incorporation analysis of bound word clitics.

In (62 we treat the aspectual particle as a functional category, which we have notated F and FP in the tree. It is an open question whether we should distinguish for Welsh a further functional category, Asp (a view embraced for Socts Gaelic by (Ramchand 1997), for example), or not. It is an intriguing fact about the language that the set of aspectual particles include adjectives as well as prepositions. This raises the question of whether it is appropriate and adequate to simply distinguish between lexical and functional *uses* of lexical elements, rather than reifying this distinction in terms of category labels. In a view of categorial structure which treats atomic category labels as abbreviations for collections of features, one or more of which serves to distinguish a functional from a lexical use of a category, we can potentially dispense with the category label distinction and treat the aspectual particle *wedi* as a functional preposition. We leave this question open.

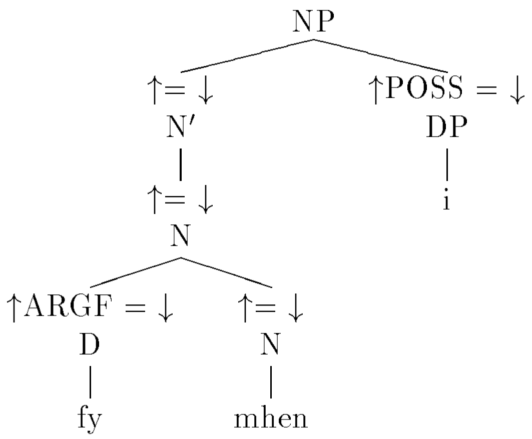
(63) Dywedodd Gwyn *ei fod ef* yn ddiog
say.pt.3s Gwyn 3s be 3sm pt lazy
Gwyn said he is lazy.

(64)



(65) fy mhen i
1s head 1s
my head

(66)



How is the choice of argument function constrained? [17] I propose that the choice of argument function is constrained by a principle limiting cliticisation to those functions

[17]In his HPSG analysis of Welsh cliticisation, (Borsley 1995) proposes that a cliticisation lexical

assigned within the f-structure which the c-head head maps into. This principle imposes the sort of locality condition on cliticisation that we observe.

(67) 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. [18]

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

5 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. We begin by recapping the facts concerning the distribution of pronominal arguments, for clarity:

1. A pronominal argument *almost always* fails to appear solely in canonical argument position, but is (obligatorily) expressed by means of either cliticisation or 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*)

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 about which grammatical functions are projected in Welsh which also cannot in principle be extended to cover the case of optional cliticisation of VSO objects (which are second complements in his analysis).

[18] It is because our principle (67) 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 (67) placed an alternative requirement, that limited the functions to those in the ϕ 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 (67).

Although we have been concerned up to now in the paper solely with pronominal clitics, in this section we dwell at some length on the striking distributional similarities between clitic pronouns and morphologically incorporated pronominals in the language. It will therefore be useful to exemplify the latter briefly.

Most Welsh prepositions inflect for the person and number, and in the 3rd person singular, for the gender of their pronominal objects. The occurrence of full, canonically positioned, object pronouns (ie doubling) is generally optional. Prepositions occur in uninflected form with nonpronominal objects.

(68) Soniais (i) am y dyn.
speak-pt.1s 1s about the man
I spoke about the man.

(69) Soniais (i) amdano (ef).
speak-pt.1s 1s about-3sm (3sm)
I spoke about him.

Finite verbs in Welsh inflect for the person and number of their pronominal subjects. The occurrence of full, canonically positioned, subject pronouns (ie doubling) is generally optional. Finite verbs occur in 3rd person singular tensed forms with with all nonpronominal subjects.

(70) Darllenodd y dynion y llyfr.
read-pt.3s the men the book
The men read the book.

(71) Darllenasant (hwy) y llyfr.
read-pt.3pl (3pl) the book
They read the book.

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* encliticised, in the presence of a pre-sentential particle: we discuss this phenomenon in Section 5.3). 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.

5.1 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. Where it is available it pre-empts the analytic form. As in Welsh, the synthetic form cannot co-occur with a full lexical subject (mutual exclusion) — Irish differs from Welsh however in failing to permit copy pronouns to occur with incorporated pronominals.

Andrews observes that the unique character of PRED values in LFG provides a straightforward account of mutual exclusion. If incorporated pronominals (or strong agreement) are lexically associated with a (↑PRED) = '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 (↑PRED) = 'PRO' equation obligatorily with the affix and *optionally* with the copy pronoun. This permits the inflection to occur with a copy pronoun, but correctly prevents it occurring (as an agreement marker) with full lexical. Under this analysis, what is unusual about Welsh (and different from (most dialects of) Irish) is that full pronominals have optional PRED values. Recall that the same form (of the finite verb occurs with 3rd person singular pronominal subjects and with all lexical NP subjects. Clearly, two different lexical descriptions must be associated with this form in each tensed paradigm. We illustrate with the past tense:

(72) darllenodd:
$(\uparrow \mathrm{PERS}) = 3$
$((\uparrow \mathrm{PRED}) = \text{'PRO'} \wedge (\uparrow \mathrm{NUM}) = \text{sing}\ ) \vee$
$(\ \uparrow \mathrm{NUM}) = \text{plur}$

This approach to mutual exclusion extend straightforwardly to Welsh clitics. In parallel fashion, mutual exclusion between a pronominal clitic and a full lexical NP follows from the fact that the proclitic has a PRED value (recall that, like pronominal inflections, copy pronouns can (and in some cicumstances, must) fail to double pronominal clitics).[19]

[19]It is likely that in some cases full pronouns doubling proclitics are excluded to avoid Binding violations. For example, if the object is coindexed with the subject in a transitive clause, the proclitic is obligatorily doubled by the reflexive full pronoun (*ei hun*, literally 'his self') (i), while a copy pronoun is absolutely excluded in passive structures (ii).

(i) Mae ef yn ei daro ei hun.
is.3s 3sm pt 3sm hit 3sm self

He is hitting himself

(ii) Cafodd ei ladd yn y chwarel.
get-pt.3s 3sm kill in the quarry

He was killed in the quarry

(73) **the proclitic**: ei:
$(\uparrow\text{PRED}) = \text{`PRO'}$
$(\uparrow\text{PERS}) = 3$
$(\uparrow\text{NUM}) = \text{sing}$
$(\uparrow\text{GEN}) = \text{fem}$

(74) **the full pronoun**: hi:
$((\uparrow\text{PRED}) = \text{`PRO'})$
$(\uparrow\text{PERS}) = 3$
$(\uparrow\text{NUM}) = \text{sing}$
$(\uparrow\text{GEN}) = \text{fem}$

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:

(75) 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).

Consider the role of the Morphological Blocking Principle in the case of pronominally inflected finite verbforms. What is being compared here is contribution that two alternative lexical entries (the synthetic and the analytic verbforms) make to the f-structure associated with the V/VP/S nodes. The principle of morphological blocking essentially requires us to pack as much of the information that we eventually want to express into the morphological form. 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

In addition, copy pronouns are absent when proclitics act as resumptive pronouns in long distance dependency constructions. Thus it seems likely that a number of factors as well as the Binding Theory are at work in determining the distribution of the copy pronoun.

suggesting the relevance of Andrews' Morphological Blocking account to these constructions.

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

(76) Mae Emrys wedi ei weld (ef).
is.3s Emrys asp see 3sm see (3sm)
Emrys has seen me.

(77) *Mae Emrys wedi gweld ef.
is.3s Emrys asp see him
Emrys has seen him.

The competition is between (78) and the "phrasal" X° (79) and their associated f-structures. The f-structure corresponding to (78) subsumes that corresponding to (79) which in turn subsumes (80). Given the subsumption relation between the associated f-structures, the X° wins out, as more informative.

(78) [V gweld]

(79) [V [D ei] [V weld]]

(80)
$$\begin{bmatrix} \text{PRED} & \text{see}\left\langle\text{SUBJ, OBJ}\right\rangle \\ \text{OBJ} & \begin{bmatrix} \text{PRED} & \text{`PRO'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SG} \\ \text{GEN} & \text{MASC} \end{bmatrix} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{`john'} \end{bmatrix} \\ \text{TENSE} & \text{PAST} \end{bmatrix}$$

This behaviour consitutes 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.[20]

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^0 material which enter into morphological blocking 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).

The principle of Morphological Blocking in in some ways similar to the more general condition formulated by Bresnan (1997) which governs the competition bewteen the morphological and syntactic means of expression, her Principle of Economy of Expression:

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

This principle applies to non-preterminal (ie phrasal) c-structure nodes and thus prefers lexical over phrasal means of expression. It does not apply directly to the case at hand. Nonetheless, the relation between these principles of Blocking and Economy remains an interesting open question for further work. In particular, it seems fruitful to consider whether they might be viewed as second level generalisations, emerging from a set of ordered constraints in an OT approach to syntax.

5.2 Choice of Strategy

A remaining issue concerns the distribution of the two strategies of pronominal incorporation and cliticisation across the different c-heads. Recall that pronominal objects of prepositions are inflectional, while pronominal objects of (nonfinite) verbs are clitics, and pronominal subjects of (morphologically finite) I are inflectional, while those of (morphologically non-finite) *bod* 'be' in I are clitics. 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 affixal option 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 in-

[20]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^o host).

corporating 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.[21] I assume therefore that clitics (as functional heads) select a sister which is disjunctively specified as [+N $\vee$ +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).

(82) 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.

(83) 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.

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 (83) blocks the alternative form *bod*).

Recall 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.

(84) 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 (85)) is excluded by morphological blocking: the more informative *darllenasant* (3rd person plural), which specifies both NUM and PRED values, blocks *darllenodd.*

(85) *Darllenodd nhw y llyfr.
read-3 3pl the book

[21] 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.

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 (found in formal, literary Welsh) which we have not yet discussed, involving the *optional* encliticisation of the OBJ in VSO structures.

5.3 A Problem Case: Finite Object Cliticisation

Condition (67) permits an 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. This condition correctly rules out the possibility of "long" cliticisation of the OBJ of a nonfinite verb (in VP) to I: in periphrastic structures, a pronominal object 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 or conjunction. 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) object clitics:

(86) Oni'ch gwelodd?
Comp-2pl see-pt.3s
Did he/she not see you?

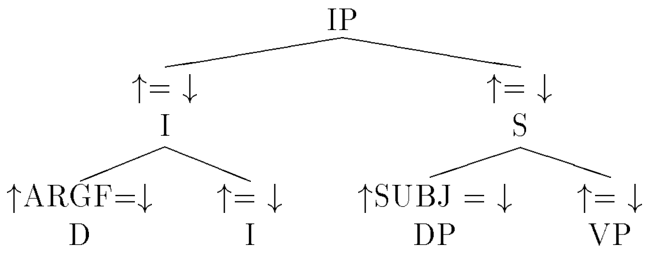
(87) Fe'i gwelais ef.
pt-3s saw-1s 3sm
I saw him.

(88) 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: in the framework we adopt, I is the extended head of the VP projection within which the grammatical function (OBJ) of the clitic is normally assigned. As we signalled above, there are a number of issues which need further investigation concerning the enclitic pronominal paradigm, and thus any analysis must be explored in very tentative fashion. It seems likely however, that at least some of these forms should be analysed as syntactically proclitic but *phonologically* enclitic elements - that is, that a rule of the post-lexical phonology attaches them prosodically to the preceding element. [22] If this is correct the structure is as in (89):

(89)



It is tempting to make use of the notion of extended head in developing an analysis of object cliticisation to I, since it is only possible when there is no overt verb within VP. Reformulating (67) as follows accounts for the availability of finite object cliticisation:

(90) 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): in these cases, the extended head *is* also the categorial head. It is possible that the extended head relationship is related to the failure of morphological blocking: the fact 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.[23]

[22] On the other hand, a number of these enclitic forms show evidence of a lexical relationship with the phonological host, suggesting that they have been reanalysed as suffixal inflectional morphemes.

[23] Note that the puzzle posed by the *optionality* of cliticisation in this one case is to some extent independent of our proposal that these enclitic forms are in I. If it should turn out that they are syntactically attached to the pre-sentential particle (complementiser), forming $[_C$ C D $]$ structures, then blocking should prefer the particle + clitic combination over the bare particle, and thus optionality is also unaccounted for under this structural analysis.

6 Conclusion

This paper has been concerned essentially with the morphosyntactic status of proclitic pronouns in Welsh. Clitics pose a descriptive challenge because they often exhibit an intriguing mix of syntactic and lexical properties. We have argued that Welsh proclitic pronouns are not affixes, but form syntactically transparent constructions with their host. On the other hand, we have argued that these constructions are X^0 level (small constructions) and are treated by part of the grammar as morphological constructions. In particular, they enter into morphological blocking relations. The analysis has been couched within the version of LFG presented in (Bresnan 1997) which offers a rich theory of phrase structure variation while maintaining the principle of lexical integrity.

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.

Duffield, N. 1996. On structural invariance and lexical diversity in VSO languages: arguments from Irish noun phrases. In R. Borsley and I. Roberts (Eds.), *The Syntax of the Celtic Languages*, 314–340. Cambridge University Press.

Halpern, A. 1995. *On the Placement and Morphology of Clitics*. Stanford,CA: CSLI.

Kroeger, P. 1995. *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. 1992a. *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland.

Miller, P. 1992b. Postlexical cliticization vs. affixation: coordination criteria. In G. C. C. Canakis and J. Denton (Eds.), *Papers from the 30th Chicago Linguistic Society*, 382–96. CLS.

Poser, W. 1992. Blocking of Phrasal Constructions by Lexical Items. In I. Sag and A. Szabolsci (Eds.), *Lexical Matters*, 111–130. Stanford: CSLI.

Ramchand, G. 1997. *Aspect and Predication.* Oxford: Clarendon Press.

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.

Sadler, L. 1998. On the Head Movement Analysis of Celtic Noun Phrases. 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.

Thomas. 1984. A Lowering Rule for Vowels in a Dialect of North Wales. In M.J.Ball and G.E.Jones (Eds.), *Welsh Phonology.* Cardiff: University of Wales Press.

Zwicky, A. 1987. Suppressing the Z's. *Journal of Linguistics* 23/1:133–48.