

A new analysis of Welsh agreement

Frances Dowle

University of Oxford

Proceedings of the LFG'22 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2022

CSLI Publications

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

Abstract

INDEX agreement (Wechsler & Zlatić, 2003) in Welsh occurs only with weak (or ‘dependent’) pronouns, and not with nouns. This poses challenges for the standard model of agreement in LFG, in which agreement is modelled as feature unification (Haug forthcoming). This paper presents data which is not addressed in the previous LFG account of Welsh agreement (Sadler 1997), including data regarding focus-fronting constructions. It introduces the independent-dependent pronoun distinction and overcomes an empirical challenge to the previous account: in the previous analysis, pronouns have optional PRED values, which means they are not prevented from occurring in the ‘gap’ of long-distance dependencies. Through the use of lexical sharing schemata (Wescoat 2002, Bresnan 2021), the distribution of the dependent pronouns is appropriately restricted.

1 Introduction

Many INDEX (Wechsler and Zlatić 2003) agreement targets in Welsh are sensitive to the pronominal or nominal status of their controllers, such that agreement does not occur with a nominal.^{†,1} If the controller is pronominal, agreement is obligatory. Standard Modern Spoken Welsh allows pro-drop (Borsley et al. 2007, 307–8), and agreement occurs with a pronominal controller even if the pronoun is null.² If overt, a pronoun with which there is agreement will appear in its weak form (or ‘dependent’ form as it is known in the Celtic tradition). Strong (‘independent’) pronouns and agreement cannot co-occur.³ For example, Welsh prepositions can be index agreement targets, controlled by their objects:

- (1a) **am-danyn**(=nhw)
 about-**3PL**=**3PL_{DEP}**
 ‘about them’

[†]I would like to thank my supervisors, Dr Louise Mycock and Professor David Willis; my language informants Lois Williams, Osian Elis, and Llewelyn Hopwood; and the hybrid audience of the LFG22 Conference for their contributions to this paper. Remaining errors my own.

¹In this paper, the following abbreviations are used in addition to the those found in the Leipzig Glossing Rules: AFF = affirmative; NF = non-finite; PROG = progressive; PRD = predicate marker; PRS = present.

²This supports the modelling of agreement at f-structure – where null pronouns are represented – rather than c-structure.

³The independent-dependent distinction in Welsh is parallel but not identical to the strong-weak distinction between pronouns in other languages. Willis (2007) concludes that independent pronouns are strong, whilst dependent pronouns are weak. However, the Welsh dependent pronouns differ in some ways to the weak pronouns of Cardinaletti and Starke (1999), in that they can be coordinated (25) and modified (27). The principal reasons to consider the dependent pronouns weak are that they are morphophonologically reduced compared to independent pronouns and they are more restricted in their distribution than their independent counterparts (cf. Cardinaletti and Starke 1999, 149). Independent pronouns are used only where the dependent pronouns are not possible, which is typical behaviour for strong-weak pronoun pairs (Cardinaletti and Starke 1999, 153).

- (1b) **am-dana(=i)**
 about-**1SG**=1SG_{DEP}
 ‘about me’

The inflection of the preposition *am* ‘about’ (as *amdana* or *amdanyn*) is obligatory in (1a) and (1b), whether the dependent pronoun is overt or not:

- (1c) * **am(=nhw)**
 about=3PL_{DEP}
 intended: ‘about them’

If the object of the preposition is nominal, agreement becomes ungrammatical, even if the INDEX values match:

- (1d) **am** y=merched
 about the=girls(3PL)
 ‘about the girls’
- (1e) * **am-danyn** y=merched
 about-**3PL** the=girls(3PL)
 intended: ‘about the girls’

Some prepositions lack inflectional forms, and so cannot agree even with pronominal objects. In such cases, pro-drop is not licensed, and pronominal objects of these prepositions appear in their independent form. For example, *gyda* ‘with’, has no inflectional paradigm, and hence never agrees; *i* ‘to’ has only a partial paradigm, and so agrees only with certain person and number combinations. Hence (2a) and (2b) both contain independent pronouns:

- (2a) gyda ni
 with 1PL_{IND}
 ‘with us’ cf. *gyda ti* (2SG), *gyda hi* (3SG.F)
- (2b) i fi
 to 1SG_{IND}
 ‘to me’ cf. *i ti* (2SG), *i ni* (1PL), *i chi* (2PL)

Wherever the combination of agreement and dependent pronoun is possible, the choice of non-agreement and independent pronoun is ungrammatical:

- (2c) * i hi
 to 3SG.F_{IND}
 intended: ‘to her’
- (2d) i-ddi(=hi)
 to-3SG.F=3SG.F_{DEP}
 ‘to her’

In LFG, agreement is usually modelled at f-structure (Haug forthcoming, Börjars et al. 2019) and is symmetrical. When the target and the controller contribute compatible person, number, gender (PNG) or other attribute values to the controller’s f-structure, feature unification occurs. This standard model, which usually identifies targets’ controllers by their grammatical function, is not sufficient to account for the Welsh data, since pronouns and nouns with the same grammatical function and the same set of PNG features behave distinctly with respect to agreement.

Sadler (1997) proposed an LFG solution to the incompatibility of nouns and agreement in Welsh, proposing the obligatory contribution of a PRED = ‘pro’ value by the agreement target to the agreement controller. This renders agreement incompatible with any element already contributing its own PRED value, such as a noun. Sadler (1997) analysed pronouns in Welsh as making an optional PRED contribution. In this paper, I evaluate the suitability of Sadler’s analysis, taking into account a wider array of data, and propose a modification using lexical sharing. The new analysis proposes separate entries for independent and dependent pronouns, and restricts the distribution of the dependent pronouns such that they do not encounter the same challenges as the pronouns in Sadler’s original analysis.

This paper has the following structure: first, a background to Welsh sentence structure is given (§2); then, the key data on index agreement (§3) is presented; the existing analysis of agreement in Welsh (Sadler 1997) is discussed (§4); and an alternative analysis with constraining equations is proposed (§5); the paper ends with some discussion and a conclusion (§6).

2 Sentence structure in Welsh

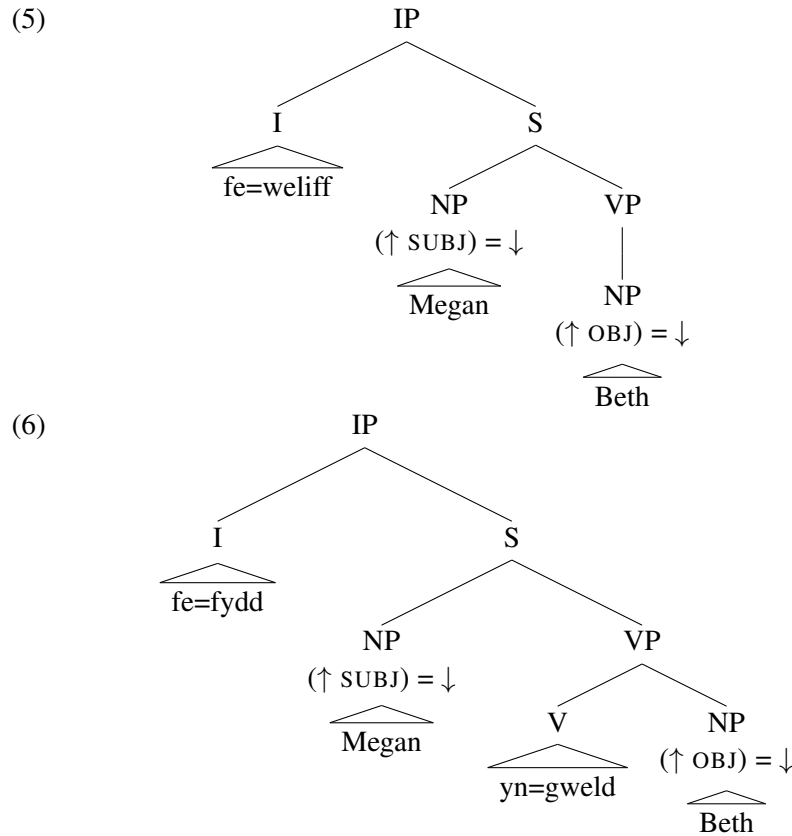
Welsh is highly configurational, and has no morphological case marking (King 2014, 30; Willis 2017, 44). It is usually described as being a VSO language, although in addition to the basic VSO word order, in which the verb is a finite lexical verb, there are many constructions in Welsh which make use of an auxiliary verb and a non-finite lexical verb ($V_{\text{FIN-AUX}}SV_{\text{LEX}}O$). Compare the two possibilities for expressing ‘Megan will see Beth’:⁴

- (3) $Fe^S = \text{wel-iff}$ Megan Beth.
 $AFF = \text{see-FUT}$ Megan Beth
 ‘Megan will see Beth.’

⁴Welsh is a language with initial-consonant mutation, meaning that the phonological form of the initial segments of words and morphemes can change due to lexical and morphosyntactic triggers; this is why there is no /g/ in the stem of *wel-iff* in (1), but there is a /g/ in *gweld* in (2). As standard in Modern Welsh, mutations are reflected in the orthography. Additionally, in the examples presented in this paper, mutation triggers have been marked. This is done with superscript letters (S=Soft, N=Nasal, A=Aspirate) after lexical triggers, or, where syntactic context is the trigger, superscript letters appear at the beginning of the affected word. For more on initial mutation in Welsh in LFG, see Mittendorf and Sadler (2006).

- (4) Fe^S=fydd Megan yn=gweld Beth.
 AFF=be.FUT Megan PROG=see.NF Beth
 ‘Megan will see Beth.’

These sentences are truth-conditionally synonymous, with register being the primary motivation for selecting one over the other (Borsley et al. 2007, 41). As is apparent from this pair of sentences, finite verbs and non-finite verbs have different positions in Welsh. We shall also see that they behave differently with respect to agreement, since finite verbs agree with their subjects and non-finite verbs agree with their objects (via an agreement clitic). Following Bresnan (1997) and Sadler (1997), these different word orders are captured in LFG by an analysis of finite verbs as category I and non-finite verbs as category V; Welsh phrase structure rules (PSRs) then constrain the relative orders of I and V. (3) and (4) are represented in simplified form by the c-structures (5) and (6) respectively.⁵ Node annotations of type $\uparrow=\downarrow$ are omitted.

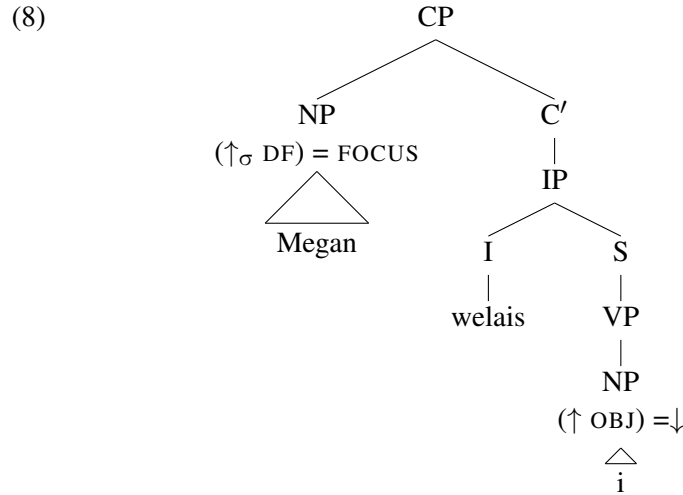


Other word orders are also found in Welsh, motivated by differences at information structure. An XP that is not in an island can be focused in Welsh by occurring in sentence-initial position as the specifier of CP (Borsley et al. 2007,

⁵On the use of NPs rather than DPs in Welsh, see Mittendorf and Sadler (2011).

33). The complementiser head of the CP may (in more formal registers) or may not (in less formal registers) be overt:⁶

- (7) [Megan]_{FOCUS} (a) ^swel-ais=i.
 Megan COMP see-PST.1SG=1SG_{DEP}
 ‘I saw *Megan*’



I refer to sentences that contain an XP which is focused in this way as Focus Fronting Constructions (FFCs). FFCs are not the only means of focusing an element in Welsh. Phrases, words and parts of words can also be focused by phonetic means, that is, using contrastive stress. In sentences containing focus of this type, normal word order is preserved. As we shall see, agreement between verbs and their arguments is suspended when the argument is focused in an FFC.

3 Index agreement

3.1 Index agreement targets in normal word order constructions

There are a number of agreement targets in Welsh that behave in much the same way as the preposition in (1), in that they exhibit agreement with pronominal controllers (overt or not), and not with nominal ones. These include finite verbs, which are controlled by their subjects. Hence in (9a) and (9b) respectively, *fydd* ‘be.FUT’ is inflected for 3PL and 1SG, as indicated by the agreement suffixes *-an* and *-a*.

- (9a) Fe^S=fydd-**an**(=nhw)=n mynd.
 AFF=be.FUT-**3PL**=3PL_{DEP}=PROG go.NF
 ‘They will go.’

⁶For further detail on representing discourse-configurationality in LFG, see Dalrymple and Nikolaeva (2011).

- (9b) Fe^S=fydd-**a**(=i)=n mynd.
 AFF=be.FUT-**1SG**=1SG_{DEP}=PROG go.NF
 ‘I will go.’

As before, agreement with a nominal is ungrammatical:⁷

- (9c) Fe^S=fydd(*-**an**) y= merched yn=mynd
 AFF= be.FUT-**3PL** the=girls(3PL) PROG=go.NF
 ‘The girls will go.’

Agreement with the objects of non-finite verbs (10) and possessors (11) behaves in much the same way, but rather than the non-finite verb or noun agreeing directly, the agreement target is a syntactically independent, phonologically clitic marker (glossed with _{AGR}), which often, but not always, procliticizes to the non-finite verb or noun. This is *eu* and *fy* in examples (10/11a) and (10/11b) respectively. As with the other paradigms demonstrated so far, there is no agreement with nominal objects or possessors, which means that the inclusion of the clitic agreement markers when the object or possessor is nominal (as in the (c) examples) is ungrammatical.

- (10a) Fe^S=fydd Bethan yn=**eu**=gweld(=nhw).
 AFF=be.FUT Bethan PROG=3PL_{AGR}=see.NF=3PL_{DEP}
 ‘Bethan will see them.’
- (10b) Fe^S=fydd Bethan yn=**fy**^N=ngweld(=i).
 AFF=be.FUT Bethan PROG=1SG_{AGR}=see.NF=1SG_{DEP}
 ‘Bethan will see me.’
- (10c) ... yn>(***eu**)gweld y=merched
 ... PROG=3PL_{AGR}=see.NF the=girls(3PL)
 ‘... see the girls’
- (11a) Roedd **eu**=cath(=nhw)...
 was **3PL**_{AGR}=cat=3PL_{DEP}
 ‘Their cat was ...’

⁷ The status of the form of finite verbs such as *fydd* in (9c) is open to some debate. This form is sometimes referred to as the 3SG form, since it is identical to the form found with a 3SG pronoun:

- (i) Fe=fydd hi=n mynd.
 AFF=be.FUT 3SG.F=PROG go.NF
 ‘She will go.’

This means that different generalisations are then made about the finite verb agreement paradigm compared to all the other INDEX agreement paradigms. In particular, that finite verbs agree with singular nouns, but only fail to agree with the number of plural nouns. The other paradigms – where there are distinct 3SG.M and 3SG.F agreement forms, are not amenable to such an analysis. As an anonymous reviewer points out, it may be desirable to posit a zero 3SG morpheme in (i), thus distinguishing the form in (i) from the non-agreement form in (9c); this would mean we have two different analyses of *fydd* depending on whether it occurs with a pronoun or noun.

- (11b) Roedd **fy^N=nghath(=i)**...
 was **1SG_{AGR}=cat=1SG_{DEP}**
 ‘My cat was ...’
- (11c) Roedd **(*eu=)**cath y=merched...
 was **3PL_{AGR}=cat** the=girls(3PL)
 ‘The girls’ cat was ...’

Evidence for the syntactic independence of these agreement markers comes from the fact that any numeral and/or any of the small class of pronominal adjectives can occur between the marker and the noun in NPs (12), and a small number of adverbs can occur between the marker and the verb in VPs (13):

- (12) **fy^N=nwy^S** gath(=i)
 1SG_{AGR}=two.F cat(=1SG_{DEP})
 ‘my two cats’
- (13) Dw=i wedi **ei^S=rannol^S** gau.
 be.PRS.1SG=1SG PERF 3SG.M=partially close
 ‘I have partially closed it.’ (e.g. the door).

To summarise the data presented thus far: targets – finite verbs, prepositions, the agreement clitics of non-finite verbs and possessums – obligatorily agree with their subjects, objects, objects, and possessors (respectively) only if the relevant grammatical function is pronominal, even if the pronominal argument is non-overt. Agreement does not occur with nouns. Where there is agreement with them, pronouns appear in their dependent forms. Pronouns can be used without agreement (e.g. with uninflecting prepositions), in which case they occur in their independent form.

3.2 Pronouns and agreement in Focus Fronting Constructions

Agreement in FFCs behaves differently to agreement in normal word order sentences like those just demonstrated. FFCs are not used to focus objects of prepositions or possessors (these positions seem to be embedded in islands), but the divergent agreement behaviour of FFCs can be observed with focus-fronted subjects or objects of non-finite verbs. Rather than having a combination of agreement and the dependent pronoun, as in (9a) above, when a pronominal subject of a finite verb is focused in this way, the pronoun appears in its independent form and there is no agreement:

- (14) [Nhw] fydd-(***an**) yn=mynd.
 3PL_{IND} be.FUT-3PL PROG=go.NF
 ‘They will go.’ (cf. 9a)

If the object of a non-finite verb is focused in an FFC, it appears as if third person singular marking occurs, irrespective of the PNG features of the focused argument,

and irrespective of the nominal or pronominal status of the object. Once again, independent pronouns are found in these constructions:

- (15) Fi/Nhw/Bethan fydd Megan yn=(ei=)^Salw.
 1SG/3PL_{IND}/Bethan be.FUT Megan PROG=EI=call.NF
 ‘Megan will call *me/them/Bethan*.’

The EI marker is identical in form and mutation effects to the 3SG.M clitic agreement marker, but there are reasons to consider it a separate lexical item (contrary to e.g., Willis 2011). Firstly, the marker is optional, whereas 3SG.M agreement never is in Standard Modern Spoken Welsh. Even if the marker is not included, soft-mutation occurs, which cannot happen with 3SG.M marking. Furthermore, the fact that this marker also occurs when the displaced object is a noun, or a pronoun with mismatching PNG features, motivates an analysis of it as a different lexical item to the 3SG.M agreement marker (although the two are probably diachronically related). Ultimately, the fact that the marker is invariant seems to preclude its analysis as agreement in the strictest sense. It also seems unlikely to be default agreement, since non-agreement is a better candidate for the default in Welsh, on the basis that, when a particular agreement inflection is unavailable for a particular PNG combination, in normal word order constructions it is not the 3SG.M form that occurs but the non-agreement form. Since the marker also occurs when some element from the verb’s complement clause is focused (Willis 2011, 198), it seems to be a marker of long-distance dependencies rather than a marker of agreement with an ex-situ object of a non-finite verb. Detailed analysis of this marker is set aside in this paper, but the reason for the lack of true agreement with focus-fronted pronominal objects requires explanation within the wider explanation of agreement in Welsh.

Note that the restriction against dependent pronouns appearing in FFCs applies even if the FFC is embedded as a subordinate clause:

- (16) Dw=i=n meddwl [taw fi/*i fydd yn=mynd].
 I think [that 1SG_{IND}/1SG_{DEP} be.FUT PROG=go.NF]
 ‘I think that *I* will go.’

4 An existing LFG analysis

4.1 Sadler’s (1997) analysis

Sadler (1997) proposes an analysis of agreement and pronouns in Welsh which is the inverse of the usual analysis of pro-drop languages, such as Spanish. This analysis is based on the data in (1), (2) and (9)–(11), i.e., it does not include the data regarding FFCs.

Under Sadler’s (1997) analysis, Welsh agreement forms (whether inflecting verbs or prepositions, or agreement clitics) obligatorily provide PRED = ‘pro’ values to their controllers, which are identified by their grammatical function; thus,

the preposition in (17a) can contribute a PRED to its object. In the lexical entries of the pronouns (e.g. (17b)) the PRED = ‘pro’ is marked as optional.⁸

- (17a) Lexical entry for an inflected preposition (*am-danyn* ‘about-3PL’)

amdanyn	P ⁰
(↑ PRED)	= ‘about⟨OBJ⟩’
((↑ OBJ) _L PRED)	= ‘pro’
((↑ OBJ) _L PERS)	= 3
((↑ OBJ) _L NUM)	= PL

- (17b) Lexical entry for a pronoun (*nhw* ‘3PL.F’)

nhw :	N ⁰
((↑ PRED)	= ‘pro’)
(↑ PERS)	= 3
(↑ NUM)	= PL

- (17e) Lexical entry for an uninflected preposition (*gyda* ‘with’)

gyda :	P ⁰
(↑ PRED)	= ‘with⟨OBJ⟩’

This analysis licenses pro-drop (completeness can be satisfied by the contributions from the agreement form alone), stand-alone pronouns, and the use of overt pronouns with agreement forms (the pronoun only optionally supplies a PRED = ‘pro’ value). The PRED value of nouns in Welsh is (necessarily) obligatory, and thus, given Sadler’s analysis of agreement forms, nouns cannot be controllers. If a noun were to occur as the object of the preposition *am-danyn* ‘about-3PL’, for example, a violation of uniqueness at f-structure (a ‘clash’ of PRED values) would result, since two PRED values would be contributed to the same grammatical function, one by the agreement form and one by the noun. Contrast with this the typical analysis of a pro-drop language like Italian (cf. Börjars et al. 2019, 69-74), or Spanish (as Sadler 1997 does). In a language like Italian, inflected verb forms optionally provide a PRED = ‘pro’ value for their subjects, while pronouns have obligatory PRED values.

On the basis of the lexical entries that Sadler proposes for Welsh, pronouns are just as compatible with non-agreement as with agreement. This is necessary

⁸The use of Dalrymple & Hristov’s (2010) symbol *L* in this entry indicates that all the index information and the PRED value provided by an agreement head correspond only to the left conjunct, in the event that the agreement controller is one conjunct in a coordinated phrase. Example (25) demonstrates the phenomenon of left conjunct agreement in Welsh. The use of the *L* symbol prevents a clash from occurring between the values supplied by the agreement head and those supplied by the right conjunct. Sadler’s (1997) original analysis does not address the issue of left-conjunct agreement, but the inclusion of this variable reflects subsequent work on single-conjunct agreement in LFG.

to account for examples like (2a, 2b), where no agreement occurs. But, it is not desirable in the case of phrases like (2c), where non-agreement is ungrammatical because of the possibility of using agreement (2d). Given the types of lexical entries in (17), (2c) is predicted to be grammatical, because even though the preposition does not contribute a PRED value for its OBJ, the pronoun itself still can, via the optional equation in (17b).

To address this, Sadler completes her analysis – extending it from simply allowing agreement with pronouns to requiring it where possible – using Morphological Blocking (Andrews 1990). Morphological Blocking requires that the more specific form of the target (with agreement) is chosen over the less specific form (without agreement) wherever possible. Sadler (1997) analyses agreement clitics as non-projecting heads (cf. Toivonen 2003), so she extends the application of Morphological Blocking to explicitly allow it to apply to structures involving non-projecting heads. This accounts for the fact the clitics display the same blocking behaviour as inflectional agreement.

Sadler’s analysis has the desired outcome of making agreement compatible only with pronouns, and not with nouns, and ties this to pro-drop in Welsh. It also explains why agreement is compulsory in normal word order constructions.

4.2 Empirical challenges

There are three aspects of the Welsh data that are not addressed in Sadler’s (1997) analysis. Each of these are discussed here.

4.2.1 The dependent–independent pronoun distinction

In sections 1 and 2, pronouns were labelled as ‘dependent’ or ‘independent’, a distinction which is recognised by Borsley et al. (2007, 26-7), Willis (2017, 44) and Willis (2020, 8). Sadler’s analysis does not address this covariation in form and agreement behaviour that Welsh pronouns exhibit.

Although the two series of Welsh pronouns are often segmentally identical, they *can* be distinguished segmentally, as shown in Table 1. The 1SG independent pronoun is *fi* /vi/, whilst the dependent pronoun is usually *i* /i/ but has a post-vocalic allomorph *fi* /vi/. The 2SG independent pronoun in most of Wales is *ti* /ti/, whilst the dependent pronoun is usually *di* /di/, except that it has an allomorph *ti* /ti/ that occurs after verbs and prepositions ending in /t/. In north-western Welsh (NwW) dialects, there is an entirely distinct 2SG independent pronoun *chdi* /χdi/ (Willis 2017, 45–6). For other person/number combinations, the independent pronouns can still be identified, as all independent pronouns can be reinforced by an initial /ə/. Furthermore, the dependent pronouns are often described as obligatorily clitic (although, unusually, they can bear contrastive stress). Their classification as clitics seems to stem from the fact that they never bear ordinary (i.e., non-contrastive) lexical stress, and they never appear without phonological material to their left.

	Independent		Dependent	
1S	fi	/vi/	i (fi)	/i/, /vi/
2S	ti	/ti/	di (ti)	/di/, /ti/
2S (NwW)	chdi	/χdi/	di (ti)	/di/, /ti/
3SF	hi	/hi/	hi	/hi/
3SM	fe	/ve/	(f)e	/e/, /ve/
1P	ni	/ni/	ni	/ni/
2P	chi	/χi/	chi	/χi/
3P	nhw	/nhu:/	nhw	/nhu:/

Table 1: Welsh independent and dependent pronouns

Sadler proposes just one type of pronoun entry, which licenses pronouns in both agreement (dependent) and non-agreement (independent) contexts. This kind of approach does not capture that the dependent or independent uses of the pronoun also correspond to changes in the form of the pronoun (prosodically and sometimes segmentally as well). This patterning together of form and agreement behaviour could be captured if two distinct sets of pronominal lexical entries were proposed for Welsh: one for the dependent and one for the independent pronouns.

4.2.2 FFCs

Sadler’s analysis does not consider the behaviour of agreement in FFCs. Because the data on FFCs is not considered, there is no explanation as to why Morphological Blocking does not apply in FFCs. On the basis of Sadler (1997), the reverse predictions are made regarding the grammaticality of the following two sentences:

- (18) Fi fydd yn=mynd.
1SG_{IND} be.FUT PROG=go.NF
‘I will go.’
- (19) * I fydd-a=n mynd.
1SG_{DEP} be.FUT-1SG=PROG go.NF
‘I will go.’

4.2.3 Long distance dependencies and gaps

Finally, Sadler’s analysis does not account for why pronouns are blocked from the ‘gap’ in long-distance dependencies (LDDs). When a subject or object in Welsh is involved in a LDD, Welsh uses a gap strategy (as opposed to a resumptive strategy). No overt argument, including a pronoun, can occur in the position which corresponds (configurationally) to the primary grammatical function of the material in DIS (Dalrymple et al. 2019, 653). This would violate coherence, because two PRED values are contributed to the same embedded f-structure. For example, the

English example in (21) is ungrammatical because two PRED values are supplied for the SUBJ/DIS, one by the phrase ‘the girl’ and one by the phrase ‘her’.

(20) The girl [I saw __ at the party] ...

(21) * The girl [I saw her at the party] ...

Because Sadler argues that Welsh pronouns only optionally contribute a PRED value, the same explanation for the ‘gap’ in LDDs cannot be extended to Welsh. This leaves the ungrammaticality of including a pronoun (dependent or independent) in the ‘gap’ of the following sentences unexplained:

(22) * Fi fydd=(f)i yn=mynd.
 1SG_{IND} be.FUT=1SG PROG=go.NF
Intended: ‘I will go.’ (cf. 19)

(23a) * Fi fydd Megan yn=ei=weld=(f)i
 1SG_{IND} be.FUT Megan PROG=EI=see.NF=1SG
Intended: ‘Megan will see me.’ (cf. 13)

(23b) * Fi wel-iff Megan=(f)i
 1SG_{IND} see.-FUT Megan=1SG
Intended: ‘Megan will see me.’

Taking this extra data into account, we can now move towards a new analysis of Welsh agreement which expands on the insights of Sadler’s (1997) analysis.

5 Towards a new analysis

In this section, it is argued that the appropriate distinction between the dependent and independent pronouns is not a c-structure distinction. Instead, an analysis is pursued in which they are distinguished functionally, with independent and dependent pronouns differing in whether they bear a PRED value. The new analysis makes use of lexical sharing Wescoat (2002, 2005) to appropriately constrain the PRED-less dependent pronouns.

5.1 Separate entries for dependent and independent pronouns

In this section, data will be presented which shows that both types of pronouns are NPs. Despite their divergent phonology, prosody, and agreement behaviour, there are good reasons to consider that both classes of pronoun in Welsh should be analysed as having the same representation at c-structure. This is motivated by the fact that both classes of pronoun can coordinate with full lexical NPs:

(24) Mae gwallt brown gyda fi a Megan.
 be.PRS.AFF hair brown with 1SG_{IND} and Megan.
 ‘Megan and I have brown hair.’

- (25) Fe=^Sfydd-wch=chi a Megan yn=mynd.
 AFF=see-FUT.2PL=2PL_{DEP} and Megan PROG=go.NF
 ‘You and Megan will go.’

It is also motivated by the fact that both types of pronoun can be modified by relative clauses, just like a noun:

- (26) ’Dych=chi [CP ^Sfydd yn=mynd] yn=lwcus.
 be.PRS.2PL=2PL_{DEP} be.FUT PROG=go.NF PRD=lucky
 ‘You who will go are lucky.’
- (27) Chi [CP ^Sfydd yn=mynd], dych=chi=n lwcus.
 2PL_{IND} be.FUT PROG=go.NF be.PRS.2PL=2PL=PRD lucky
 ‘You who will go, you are lucky.’
- (28) Mae=r dyn [CP ^Sfydd yn=mynd] yn=lwcus
 be.PRS.3S=the man be.FUT PROG=go.NF] PRD=lucky
 ‘The man who will go is lucky.’

The final motivation for this shared c-structure representation is the configurational assignment of grammatical functions in Welsh (cf. 5, 6). Grammatical functions are assigned to a phrase in Welsh via the c-structure (formalised as annotations on the PSRs for Welsh). Nouns, independent pronouns and dependent pronouns in the same position (e.g. after a prepositional) will be structurally assigned the same grammatical function (e.g. OBJ); this generalisation can be straightforwardly captured if all three types of phrase are NPs. Hence, the two pronouns’ differences lie not in their c-structure representation, but in their interactions with agreement.

To distinguish the dependent and independent pronouns, it is possible to extend Sadler’s mechanism of using competing obligatory PRED values to block cooccurrence with agreement. Since, like nouns, independent pronouns do not occur with agreement, they can be analysed as having obligatory PRED values, just as nouns do. Combined with the types of entries for agreement forms that Sadler proposes (17a), this makes agreement and independent pronouns incompatible. In principle, dependent pronouns could be analysed as either contributing an optional PRED value, or no PRED value at all; both possibilities maintain their compatibility with agreement and a distinction with the independent pronouns. The latter option is preferable since the dependent pronouns never occur without agreement, and so there is no evidence that they should ever have the option to contribute a PRED value:

- (29) Independent pronoun: *fi* ‘1SG_{IND}’

$fi : N^0$ $(\uparrow \text{PRED}) = \text{‘pro’}$ $(\uparrow \text{PERS}) = 1$ $(\uparrow \text{NUM}) = \text{SG}$
--

- (30) Dependent pronoun: i ‘1SG_{DEP}’ (to be revised)

i :	N^0
	($\uparrow$ PERS) = 1
	($\uparrow$ NUM) = SG

5.2 Welsh agreement with lexical sharing

The PRED-less dependent pronoun entries then need to be adjusted so that the dependent pronouns are excluded from FFCs and the ‘gap’ of LDDs.

Excluding the dependent pronouns forces the use of the independent pronouns, which in turn blocks the use of agreement in FFCs. The dependent pronouns can be excluded from the focus position of FFCs using lexical sharing (Wescoat 2002, 2005). In so doing, this analysis links the inability of dependent pronouns to occur in FFCs to their clitic status. Wescoat (2005) argues in favour of using lexical sharing to analyse clitics, irrespective of whether clitics show morphophonological evidence of having formed a lexicalised unit with their host. Clitic-host sequences are represented as single lexical items which correspond to two (or more) nodes at c-structure. Lowe (2015, 2016) raises concerns around the potential for this to result in an exponential increase in the size of the lexicon, given the productivity of such sequences and the fact that such an analysis requires that all host-clitic sequences appear in the lexicon. In response, Bresnan (2021) proposes the use of lexical sharing schemata to represent clitics. The schemata allow for a kind of lexical production of host-clitic sequences, without every combination of host and clitic being individually listed in the lexicon.

Bresnan’s schemata allow the direction of attachment of a clitic to be specified, along with any categorical or functional restrictions that a clitic places on its host. This makes the schemata useful in restricting the distribution of the dependent pronouns. The use of the lexical sharing schemata can be motivated for the Welsh dependent pronouns on the basis of their being clitic, and at the same time be used to restrict their distribution.

Firstly, the use of schemata for the dependent pronouns will allow for the specification that they require a host to their left. This will exclude them from the focus position of unembedded FFCs (19), where they would appear sentence-initially.

Secondly, the schemata should specify that only certain categories of words can act as hosts for the dependent pronouns. Recall from section 2 that dependent pronouns can occur after finite verbs (as their subjects), after prepositions (as their objects), after non-finite verbs (as their objects), and after nouns (as their possessors). In addition, dependent pronouns can occur after post-nominal adjectives, if the adjective modifies the possessum of the pronoun:

- (31) $fy=N$ ghath oren= i
 1SG_{AGR}=cat orange=1SG_{AGR}
 ‘my orange cat’

The lexical schema of a dependent pronoun should therefore allow hosts of the following categories: I^0 (finite verbs), V^0 (non-finite verbs), P^0 , N^0 and Adj^0 :

(32) Lexical Schema for the 1SG pronoun (to be revised)

=i :	$\{I^0 \mid V^0 \mid N^0 \mid Adj^0\}$	N^0
		($\uparrow$ PERS) = 1
		($\uparrow$ NUM) = SG

This excludes the dependent pronoun from embedded FFCs (16), where it would have to follow a complementiser (category C^0).⁹ By specifying as possible hosts only types of host that are attested, the dependent pronouns are excluded from FFCs. This then means that the independent pronouns will be used. Since independent pronouns are incompatible with agreement, the requirement to use independent pronouns in FFCs explains the lack of agreement in these constructions, an issue that Sadler’s original analysis did not address.

At this stage of the analysis, the lexically-sharing, PRED-less dependent pronouns face the same challenges that Sadler’s optionally PRED-less pronouns faced with respect to the ‘gap’ of LDDs (cf. section 4.2.2). That is, nothing currently prevents them from occurring after the non-agreeing tensed verb (I^0) in (22), or after the non-finite verb (V^0) in (23a) or noun (N^0) in (23b). (Only the independent pronouns are blocked – by virtue of their obligatory PRED value.) By adding further detail on the types of possible hosts that dependent pronouns can have, it is possible to exclude the dependent pronouns from these positions. Specifically, the pronouns should be restricted to taking as hosts only I^0 words that agree with them. This is achieved by adding functional restrictions to the schema, such that only correctly agreeing I^0 words will satisfy them (see 34).

The exclusion of the pronoun from the gap in (23a) can be achieved if we assume that when a dependent pronoun seemingly attaches to a V^0 element (i.e. when it is the object of the V), it in fact attaches to the clitic-host unit which consists of the agreement clitic and V. That is, in a sentence like (10b), the pronoun *i* attaches to *fy=ngweld*, not *ngweld*:¹⁰

(33) Representation of *fy=ngweld* ‘1SG_{AGR}=see.NF’

fy=ngweld :	$\hat{D}$	V^0
	($\uparrow$ PRED) = ‘pro’	($\uparrow$ PRED) = ‘see(SUBJ, OBJ)’
	($\uparrow$ PERS) = 1	
	($\uparrow$ NUM) = SG	

⁹This seems preferable to a restriction on dependent pronouns from bearing FOCUS at i-structure, because contrastive stress may be used to focus dependent pronouns.

¹⁰The analysis of agreement clitics as category $\hat{D}$ follows Sadler (1997). Again following Sadler (1997), I assume that these non-projecting $\hat{D}$ nodes bear annotations at c-structure which ensure that their functional contributions are directed to either OBJ or POSS at f-structure – this information is therefore not required in the entries of these elements, and a single entry is used for the agreement clitics in both VPs and NPs.

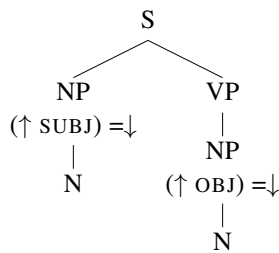
As the lexical instantiation in (33) shows, the agreement clitic-verb host unit itself involves lexical sharing. The lexical schema in (32) can be modified so that rather than taking hosts of type V^0 , the dependent pronoun actually selects hosts of type $\hat{D}V^0$. We want to disallow the pronoun from attaching to units of this type ($\hat{D}V^0$ units) if the $\hat{D}$ element is the EI marker discussed in section 3.2 and found in sentence (23a). Since the whole $\hat{D}V^0$ is now considered as the possible host candidate for the pronoun, it is possible for the dependent pronoun to place functional restrictions on the $\hat{D}$ element even though it is not immediately adjacent:

(34) Lexical schema for the 1SG pronoun

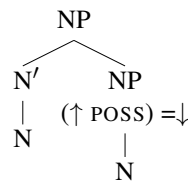
=i :	{	I^0		$\hat{D}$	V^0		N^0		Adj^0	}	N^0
		($\uparrow$ SUBJ PERS) = 1		($\uparrow$ PERS) = 1							($\uparrow$ PERS) = 1
		($\uparrow$ SUBJ NUM) = SG		($\uparrow$ NUM) = SG							($\uparrow$ NUM) = SG

Finally, we wish to prevent the dependent pronoun from appearing in the object gap even when there is no overt V, as in the FFC in (23b). In other words, we wish for the pronoun to only accept a N^0 or Adj^0 as a host if the pronoun appears as the possessor in the same NP. For now, I assume that the pronoun is prevented from attaching to the noun (or any modifying adjectives, should there be any) in a sentence like (23b) because this kind of attachment would involve crossing a phonological phrase boundary; it may prove to be the case that this restriction should be recast in syntactic terms, i.e., that the noun in (23b) is a unsuitable host because it is in a different NP to the pronoun (which is in the object NP). Compare the (partial) structure of (23b) shown in (35), in which lexical sharing of the two N^0 s is not allowed, with the structure of a sentences like (11b) in which N^0 sharing is allowed (36):

(35)



(36)



To summarise this analysis:

1. Independent pronouns will not occur with agreement forms (because of the PRED value clash).
2. Dependent pronouns will not occur in the focus position of FFCs, because they are clitics, and as clitics, they can specify certain requirements about their host; here, they specify that they must have a host to their left and it cannot be a complementiser.

3. As a result of dependent pronouns being blocked here, independent pronouns are used, and since these are incompatible with agreement, there is no agreement in FFCs.
4. Dependent pronouns, although PRED-less, do not appear in the gaps of LDDs, because in these positions they would have unsuitable hosts. They require I^0 and $\hat{D}V^0$ hosts to coindex them, but such coindexing forms are blocked because of the agreement clash which would arise.

This lexical sharing analysis therefore restricts the PRED-less dependent pronouns, not by listing many exclusions as to when they can occur, but by listing only a restricted set of possible hosts for the clitic elements. Accepting only a restricted set of hosts is typical behaviour for clitics.

6 Discussion

One potential problem with the preceding analysis is that it makes use of lexical sharing. Belyaev (2021) describes lexical sharing as “a controversial analytic device in LFG, as it constitutes a violation of the strict version of Lexical Integrity.” He says that, “the broad consensus is that, if it can be used at all, it should only be applied as a last resort for those cases where no other solution is adequate.”

Another problem that the preceding analysis faces is that it misses a cross-linguistic generalisation regarding the gap of LDDs. Specifically, it does not allow us to capture the similarities in the English data in (21) and (22), and the Welsh data in (23) and (24). We would ordinarily expect pronouns to be PRED-bearing elements, and as PRED-bearing elements, it would be no surprise that they are blocked from occurring in gaps, but the lexical sharing approach blocks the pronouns from the gap in a very different way. It could also be argued that it misses a generalisation within Welsh as well, specifically that the pronoun is blocked from the position corresponding to the primary grammatical function of the element in focus.

If dependent pronouns are assumed to be PRED-bearing elements like independent pronouns, then we can no longer use an obligatory PRED value in the lexical entries of agreement forms to prevent their co-occurrence with nouns and independent pronouns. In a revised analysis, the fact that agreement occurs only with pronouns and not nouns would have to be accounted for in a separate part of the lexical entry: a constraining equation requiring the controller of an agreement form to have a PRED = ‘pro’ value:

(37) *am-danyyn* ‘about-3PL’

amdanyn : P^0 $(\uparrow \text{ PRED}) = \text{‘about’} \langle \text{OBJ} \rangle$ $((\uparrow \text{ OBJ})_L \text{ PRED}) = \text{‘pro’}$ $((\uparrow \text{ OBJ})_L \text{ PRED FN}) =_c \text{‘pro’}$ $((\uparrow \text{ OBJ})_L \text{ PERS}) = 3$ $((\uparrow \text{ OBJ})_L \text{ NUM}) = \text{PL}$

Pursuing this line of analysis leads to a number of difficulties, however. In particular, it is unclear how to block agreement and independent pronouns co-occurring. In many cases the use of the independent pronoun seems linked to the unavailability of the dependent pronoun, in which case we wish to capture a requirement for dependent pronouns to be conindexed, and then some kind of order of preference in the selection of the pronouns. The dependent pronoun can be revised as follows:

(38) Dependent pronoun: *i* ‘1SG_{DEP}’ (alternative)

i : N⁰
 (↑ PRED = ‘pro’
 (↑ PERS) =_c 1
 (↑ NUM) =_c SG

But the way in which the dependent pronoun blocks the independent one remains unclear. As a reviewer points out, the correct subsumption relationships between the f-structures corresponding to the different pronoun entries do not hold, and so this cannot be Morphological Blocking. Such an analysis also eventually seems to lead to a degree of stipulation, as it becomes necessary to find an independent explanation for the lack of agreement in FFCs; it would no longer be the case that this follows from the use of the independent pronoun.

There may yet be promise in pursuing an alternative approach, but there is an additional benefit to the lexical sharing approach that should be considered. Recall that in a normal word order construction, the use of a non-agreement form and a dependent pronoun is usually ungrammatical. However, in some Welsh dialects, the use of the modifier *mond* ‘only’ before the pronoun renders such a combination grammatical:¹¹

(39) Aeth *(mond) fi i=r barti.
 go.PST *(only) 1SG_{IND} to=the party
 ‘I went to the party.’

(40) Es (*mond)=i i=r barti.
 go.PST.1S (*only)=1SG_{DEP} to=the party.
 ‘I went to the party.’

This data follows from the lexical sharing analysis, because prenominal adjectives are non-projecting $\widehat{\text{Adj}}$ elements (following Sadler 1997), and the lexical schemata for dependent pronouns specify that only projecting Adj^0 elements can be hosts to dependent pronouns. The dependent pronouns are therefore blocked from sentences like (40) in much the same way as they are blocked from FFCs. It is difficult to see why an approach which restricts the dependent pronouns using constraining equations would lead to the blocking of the dependent pronoun in (40).

¹¹In other dialects, the use of this modifier prompts the use of a FFC, so sentences of the type in (39) do not arise.

7 Conclusion

This paper has shown that, despite hesitancy around the use of lexical sharing in LFG, it may be a very useful analytic tool in the analysis of Welsh dependent pronouns. With one analytical device, the use of which it has been argued can be justified based on their clitic nature, the dependent pronouns can be blocked from a number of positions. Using many of Sadler's (1997) original insights, it is possible to extend the LFG account of agreement in Welsh with the addition of lexical sharing. Without a mechanism like lexical sharing, pronouns without PRED values or with only optional ones are not suitably restricted. The Welsh data therefore shows that there is some promise in the idea put forward by Wescoat (2005) and Bresnan (2021) that lexical sharing can be extended to the analysis of all clitics, even if they show no sign of morphophonological idiosyncracies in their attachment to a host. The analysis also showed that clitics might need to impose restrictions on elements which are not immediately adjacent to them; in this case, the dependent pronouns required that any non-finite verb they attach to be itself preceded by an agreement clitic marker.

References

- Andrews, Avery D. 1990. Unification and morphological blocking. *Natural Language & Linguistic Theory* 8(4), 507–557.
- Belyaev, Oleg. 2021. Paradigm structure influences syntactic behaviour: Ossetic case inflection. In Tracy Holloway King I. Wayan Arka, Ash Asudeh (ed.), *Modular design of grammar: linguistics on the edge*, pages 251–281, Oxford University Press.
- Borsley, Robert D., Tallerman, Maggie and Willis, David. 2007. *The syntax of Welsh*. Cambridge University Press.
- Bresnan, Joan. 1997. Mixed categories as head sharing constructions. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG'97 Conference*, CSLI Publications.
- Bresnan, Joan. 2021. Formal grammar, usage probabilities, and auxiliary contraction. *Language* 97(1), 108–150.
- Börjars, Kersti, Nordlinger, Rachel and Sadler, Louisa. 2019. *Lexical-Functional Grammar: An introduction*. Cambridge University Press.
- Cardinaletti, Anna and Starke, Michael. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk (ed.), *Clitics in the languages of Europe: volume 5/part 1*, De Gruyter Mouton.
- Dalrymple, Mary and Hristov, Bozhil. 2010. Agreement patterns and coordination in Lexical Functional Grammar. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG10 Conference*, pages 186–206, CSLI Publications.
- Dalrymple, Mary, Lowe, John J. and Mycock, Louise. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford University Press.

- Dalrymple, Mary and Nikolaeva, Irina. 2011. *Objects and information structure*. Cambridge Studies in Linguistics, No. 131, Cambridge University Press.
- Haug, Dag. forthcoming. Agreement. In Mary Dalrymple (ed.), *The Handbook of Lexical Functional Grammar*, Language Science Press.
- King, Gareth. 2014. *Basic Welsh: A grammar and workbook*. Routledge.
- Lowe, John J. 2015. Degrees of degrammaticalization: A lexical sharing approach to the English possessive. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG'15 Conference*, pages 208–228, CSLI Publications.
- Lowe, John J. 2016. English possessive 's: clitic and affix. *Natural Language & Linguistic Theory* 34(1), 157–195.
- Mittendorf, Ingo and Sadler, Louisa. 2006. A treatment of Welsh initial mutations. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG'97 Conference*, pages 343–364, CSLI Publications.
- Mittendorf, Ingo and Sadler, Louisa. 2011. Welsh prenominals: At the syntax-morphology interface. In Olivier Bonami and Patricia Cabredo Hofherr (eds.), *Empirical Issues in Syntax and Semantics* 8, pages 383–408.
- Sadler, Louisa. 1997. Clitics and the structure-function mapping. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG'97 Conference*, CSLI Publications.
- Toivonen, Ida. 2003. *Non-projecting words: A case study of Swedish verbal particles*. Dordrecht.
- Wechsler, Stephen and Zlatić, Larisa. 2003. *The many faces of agreement*. CSLI Publications.
- Wescoat, Michael T. 2002. *On lexical sharing*. Ph.D. thesis, Stanford University.
- Wescoat, Michael T. 2005. English nonsyllabic auxiliary contractions: An analysis in LFG with lexical sharing. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG'05 Conference*, pages 468–486, CSLI Publications.
- Willis, David. 2007. Specifier-to-head reanalyses in the complementizer domain: Evidence from Welsh. *Transactions of the Philological Society* 105(3), 432–480.
- Willis, David. 2011. The limits of resumption in *wh*-dependencies. In Alain Rouveret (ed.), *Resumptive pronouns at the interfaces*, pages 189–221, John Benjamins.
- Willis, David. 2017. Investigating geospatial models of the diffusion of morphosyntactic innovations: The Welsh strong second-person singular pronoun *chdi*. *Journal of Linguistic Geography* 5(1), 41–66.
- Willis, David. 2020. Using social-media data to investigate morphosyntactic variation and dialect syntax in a lesser-used language: Two case studies from Welsh. *Glossa* 5(1), 1–33.