Categories don't take precedence: Evidence from Welsh

Frances Dowle
University of Oxford

Ash Asudeh
University of Rochester

Proceedings of the LFG'25 Conference

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

2025

PubliKon

lfg-proceedings.org

Abstract

The morphological component of an L_RFG grammar is responsible for selecting the word forms (Vocabulary Items, or VIs) which express a given sentence. VIs may realize only a subset of the information in a sentence, but **MostInformative**$_f$ (MI_f) and **MostInformative**$_c$ (MI_c), which evaluate the f- and c-structure information a VI realizes (respectively), ensure that sentences are expressed using the fewest VIs that realize the most information possible. Welsh has positive, negative and neutral forms of the copula in the present and imperfect paradigms which compete to realize copula-containing structures. The Welsh data regarding which form is selected in different contexts is used to show two important conclusions: (1) that constraining equations have a different status to defining equations in VIs, the former not contributing to the evaluation of MI_f, and (2) that MI_f outranks MI_c when the two metrics return different VIs according to their evaluation.

1 Overview

This paper explores complex blocking relationships in Lexical-Realizational Functional Grammar (L_RFG; among others, Melchin, Asudeh & Siddiqi 2020; Asudeh & Siddiqi 2022; Asudeh, Bögel & Siddiqi 2023; Asudeh & Siddiqi 2023; Asudeh, Melchin & Siddiqi 2024b; Asudeh & Siddiqi 2024, 2026). These relationships are explored with reference to the distribution of different forms of the copula in Welsh.[†]

In the present and imperfect, the Welsh copula has polarity-sensitive forms.[1]

(1) Rwyt ti 'n dawel.
r-ʊɪt tiː =n dawel.
POS-be.PRS.2SG 2SG =PRED quiet.
'You are quiet.'

(2) Dwyt ti ddim yn dawel.
d-ʊɪt tiː ðɪm ən dawel.
NEG-be.PRS.2SG 2SG NEG PRED quiet
'You aren't quiet.'

(3) Beth rwyt ti 'n hoffi?
beθ **r**-ʊɪt tiː =n hɔfi?
what **POS**-be.PRS.2SG 2SG =PROG like.NF
'What do you like?'

(4) Beth dwyt ti ddim yn hoffi?
beθ **d**-ʊɪt tiː ðim ən hɔfi?
what **NEG**-be.PRS.2SG 2SG NEG PROG like
'What don't you like?'

The positive and negative forms of the copula sometimes fail to occur where they might otherwise be expected to do so based on the polarity of the clause. We will later demon-

[†]We would like to thank the audience at LFG25 and our reviewers for their invaluable questions and comments. We would also like to thank the L_RFG Lab for detailed feedback on particular points of analysis, especially Danil Alekseev, Oleg Belyaev, Bronwyn Bjorkman, and Dan Siddiqi.

[1]No other verb or tense-inflection of the copula displays this behaviour, though see Borsley & Jones (2005: 50) for two verbs with negative forms.

strate this, showing that the forms are sensitive to clause-type and other contextual factors in addition to polarity.

In this paper, we will recap arguments that the neutral form's distribution can be successfully captured if we treat it as an unmarked/default form, relative to the positively and negatively specified forms, as argued in an LFG analysis of the data in Dowle (2025). Then, we'll show what implications this data has for the nature of blocking relationships in L_RFG. There are two points of particular interest. Firstly, 3SG neutral forms are sensitive to their subject's definiteness but not the clause's polarity, while the reverse is true for 3SG positive forms. There is an apparent tension in that one form seems more specific about its subject, and the other seems more specific about its polarity. Secondly, the sequence of complementizer and positive form is ungrammatical. This has different consequences depending on clause type. In declarative clauses, positive forms block the complementizer. In if-clauses, the complementizer blocks the positive form (occurring instead with the neutral form). This paper explores what these data entail for the nature of blocking in L_RFG.

As a realizational model of morphology,[2] L_RFG assumes that functional and categorical information is introduced by phrase structure rules into the c-structure tree of LFG. The job of the morphology is to then find the forms that best express the information in the tree. It assesses the lexicon (a.k.a. Vocabulary in L_RFG) to find possible exponents among the Vocabulary Items (VIs). It can only select forms that either match or are subsets of the information in the tree. Often there are multiple forms that meet this criterion but only one form can be chosen. This is done according to the Elsewhere Principle (Kiparsky 1993), i.e. the most informative form wins. This notion has been formalised in L_RFG by a family of constraints, including **MostInformative**$_f$, which compares f-structure information, and **MostInformative**$_c$, which compares c-structure information (MI_f and MI_c for short; see Section 3.2). This paper will demonstrate that constraining equations are not evaluated as part of MI_f, making them importantly distinct from defining equations in VIs, and also that when MI_f and MI_c select different forms according to their own criteria, it is MI_f that matters rather than MI_c. This has implications for the formalisations of MI_f and MI_c within the theory.

2 Data

2.1 Forms of the copula

Table 1 contains some paradigms, illustrating the relevant copular forms (Jones & Gilby 2004). These are the citation forms, but in spoken Welsh more reduced forms are typically found (King 2016: 180-1, §226-7).

[2]For some approaches to realizational morphology interfaced with LFG, see Butt & Sadler (2004), Bögel et al. (2008), Dalrymple (2015) and Thomas (2021).

	1SG	2SG	3SG	1PL	2PL	3PL
POS	r-ɔið-un	r-ɔið-et	r-ɔið	r-ɔið-en	r-ɔið-eχ	r-ɔið-en
NEG	d-ɔið-un	d-ɔið-et	d-ɔið	d-ɔið-en	d-ɔið-eχ	d-ɔið-en
NEUT	ɔið-un	ɔið-et	ɔið	ɔið-en	ɔið-eχ	ɔið-en

Table 1: Modern Welsh tripartite system of imperfect forms of the copula

	1SG	2SG	3SG	1PL	2PL	3PL
POS	r-ədu	r-ʊɪt		r-əd-ɪn	r-əd-rχ	ma(ɪ)n
NEG	d-ədu	d-ʊɪt	see below	d-əd-ɪn	d-əd-rχ	d-əd-in
NEUT	əd-u	ʊɪt		əd-ɪn	əd-rχ	əd-in

Table 2: Modern Welsh tripartite system of present-tense forms of the copula

	Definite subject	*Indefinite subject*
POS	ma(ɪ)	
NEG	d-əd-i	d-ɔis
NEUT	əd-i	ɔis

Table 3: Third-person singular present-tense forms of the copula

Note that there is generally a very regular pattern: neutral forms are the most basic; positive forms consist of a prefixed *-r* plus the neutral form; negative forms consist of a prefixed *-d* and the neutral form. The third-person stem (in singular and plural) is suppletive. The same suppletive stem is used in the third-person singular both when the subject is definite and when it is indefinite.

The 3SG forms are, unsurprisingly, underspecified for person and number (Dowle 2022; Borsley 2009). When a subject pronoun is fronted, it is this form that occurs, regardless of the person and number specifications of the fronted pronoun. It is also these forms which occur with lexical noun phrases, whether singular or plural, because person-number agreement in Welsh is restricted to occurring with personal pronouns. The restriction on person–number forms only occurring with (definite) pronouns also explains why there is no definite–indefinite split in the 3PL part of the table. In this paper we will still use the term '3SG' for the relevant forms, as a convenient label, since we are primarily interested in considering the ways in which neutral forms are underspecified compared to positive and negative forms.

2.2 Distribution of Forms

Sometimes the positive and negative forms fail to occur where we might otherwise expect them to, and the polarity-neutral form of the copula occurs instead.

2.2.1 Responsives

Positive and negative forms are both blocked in responsives. Responsives are the most basic possible answers to polar questions in Welsh; they typically consist of a form of

the copula and no other material (See Jones 1999 for in-depth discussion).

(5) Wyt ti 'n iawn?
ʊɪt tiː =n jaun?
be.PRS.2SG 2SG =PRED okay?
'Are you okay?'

a. (*R-)Ydw.
(***r**-)əd-u.
(***POS**-)be.PRS-1SG
'Yes.'

b. Nac (*d)ydw.
nak (***d**-)əd-u.
NEG (***NEG**-)be.PRS.1SG
'No.'

2.2.2 Polar questions

As the above example also shows, positive forms fail to occur in polar questions, whilst negative forms are fine here:

(6) * Rwyt ti 'n
r-ʊɪt tiː =n
POS-be.PRS.2SG 2SG =PRED
iawn?
jaun?
okay
Intended: 'Are you okay?'

(7) Wyt ti 'n iawn?
ʊɪt tiː =n jaun?
be.PRS.2SG 2SG =PRED okay
'Are you okay?'

(8) Dwyt ti ddim yn gallu dod i 'r parti?
d-ʊɪt tiː ðɪm ən gaɬi dɔːd i =r parti?
NEG-be.PRS.2SG 2SG NEG PROG can.NF come.NF to =the party
'Can't you come to the party?'

2.2.3 If-clauses

Positive forms appear to be blocked in if-clauses.[3] 'If-clause' refers to the antecedent's clause, rather than the consequent's clause, in antecedent–consequent pairs like the following:

(9) Os (*r)ydw i 'n gallu dod, bydda
ɔs (***r**-)əd-u iː =n gaɬi dɔːd, bəð-a
if (***POS-**)be.PRS-1SG 1SG =PROG can.NF come.NF, be.FUT-FUT.1SG
i 'n dweud wrthot ti.
iː =n dwəɪd urth-ɔt tiː.
1SG =PROG say.NF at-2SG 2SG
'If I can come, I will tell you.'

However, positive forms of the copula are in fact only blocked from occurring adjacent to the complementizer, *os* 'if'. Welsh has focus-fronting constructions, in which an XP

[3] Although see for some nuance on this point, regarding recent dialect changes.

appears pre-verbally and receives mild contrastive focus. The restriction on positive forms in if-clauses does not hold when the if-clause contains a focus-fronted phrase, which intervenes between the complementizer *os* 'if' and the copula:

(10) Os [Mari]$_{FOC}$ rwyt ti wedi alw, bydd popeth yn
ɔs [mari]$_{FOC}$ **r**-ʊɪt tiː wɛdi alu, bið pɔpeθ ən
if [Mari]$_{FOC}$ **POS-**be.IMPF.2SG 2SG PERF call.NF be.FUT everything PRED
iawn.
jaun.
okay
'If it's *Mari* you've called, everything will be okay.'

Some types of fronted phrases (e.g. subjects and predicates) require neutral forms to follow them in all clause types, including if-clauses.

Negative forms are fine in if-clauses even in the absence of fronting:

(11) Os dydw i ddim yn gallu dod, bydda
ɔs **d**-əd-u iː ðɪm ən gaɬi dɔd, bəð-a
if **NEG**-be.PRS.1SG 1SG NEG PROG can.NF come.NF be.FUT-FUT.1SG
i 'n dweud wrthot ti.
iː =n dwəɪd wrth-ɔt tiː.
1SG =PROG say.NF at-2SG 2SG
'If I can't come, I will tell you.'

2.2.4 Declarative clauses

Ordinary declarative clauses are a special case. Instead of *being* blocked, positive forms are the ones that *do* the blocking here. Specifically, these forms block the positive declarative main clause (PDMC) complementizer, which in southern dialects takes the form *vɛ*.[4]

(12) (*Fe) mae Sam yn pobi bisgedi.
(*vɛ) maɪ Sam ən pɔbi bɪsged-i.
C$_{PDMC}$ POS.be.PRS Sam PROG bake.NF biscuit-PL
'Sam is baking biscuits.'

(13) (*Fe) rydw i 'n pobi bisgedi.
(*vɛ) r-əd-u i =n pɔbi bɪsged-i.
C$_{PDMC}$ POS-be.PRS-1SG 1SG =PROG bake.NF biscuit-PL
'I'm baking biscuits.'

Vɛ does not appear to convey any functionally-incompatible information. It is fine with other verbs, and even with other copula forms that don't show polarity sensitivity:

(14) (Fe) hoffwn i fynd.
(vɛ) hɔf-un i vɪnd.
C$_{PDMC}$ like-COND.1SG 1SG go.NF
'I would like to go.'

[4]See Borsley et al. (2007: 35) for some further discussion on the forms and distribution of positive declarative main clause complementizers in Modern Welsh.

(15) (Fe) es i.
(vɛ) ɛs iː.
C_{PDMC} go.PST.1SG 1SG
'I went.'

(16) (Fe) fydda i 'n pobi bisgedi heno.
(vɛ) vəðа iː =n pɔbi bisged-i henɔ.
C_{PDMC} be.FUT.1SG 1SG =PROG bake.NF biscuit-PL tonight
'I will be baking biscuits tonight.'

There is a contrast here with if-clauses. In if-clauses, the complementizer (*ɔs*) and the positive copular form are incompatible, but the result is that the complementizer *ɔs* 'if' is used with a neutral form; hence *ɔs əd-w* rather than *ɔs r-əd-w* in (9). In declarative clauses, the complementizer and positive copular form are again incompatible, but the result is that the positive form occurs on its own; hence *rydw* in (13) instead of *vɛ r-əd-w*. The solution to complementizer–positive-form incompatiblity in if-clauses (to use the neutral instead of positive form) is not available in a declarative clause:

(17) * Fe ydw i 'n pobi bisgedi heno.
vɛ əd-w iː =n pɔbi bisged-i henɔ.
C_{PDMC} be.PRS-1SG 1SG =PROG bake.NF biscuit-PL tonight
Intended: 'I am baking biscuits tonight.'

Recall that the presence of a focus-fronted phrase resolved the incompatibility between *os* 'if' and the positive copular form. There is no positive *main clause* complementizer that occurs alongside focused phrases, but there is a subordinate clause one (which *only* occurs when there is a focused phrase).[5] In a declarative subordinate clause with focus-fronting, we can see that there is no incompatibility between the complementizer and the positive form of the copula:

(18) Dw i 'n meddwl taw [Mari]$_{FOC}$ rwyt ti
du iː =n mɛðul **taʊ** [mari]$_{FOC}$ r-ʊɪt tiː
be.PRS.1SG 1SG =PROG think.NF **$C_{SUB.FOC}$** [Mari]$_{FOC}$ POS-be.PRS 2SG
wedi alw.
wɛdi alu.
PERF call.NF
'I think that you called *Mari*.'

One possibility to account for this data is to propose a functional restriction: *vɛ* cannot occur with imperfect or present verb forms. Only the copula morphologically expresses these tenses (such that they are distinct from the future or conditional), and so this achieves the desired distribution of *vɛ*. We will later reject this idea, instead preferring a unified account of the restriction against complementizer-adjacency for positive forms in both if-clause and declarative clauses.[6]

[5] See Borsley et al. (2007: 128) for more detail.

[6] It is not clear whether the negative form of the copula behaves the same way in declarative clauses. There is (in spoken Welsh) no negative main clause complementizer, and the subordinate clause negative complementizer takes the form *na(d)*, with final *d* surfacing before vowel-initial forms. Thus, the segmen-

3 Analysis

3.1 Generalizations

The above data demonstrates that although both positive and negative forms exhibit restrictions beyond the simple polarity restrictions indicated by their labels, the negative form is much less restricted than the positive one. Responsives are the only context in which negative forms are blocked. We attribute this to the special complementizer, *nak*, which occurs in these contexts.[7] The positive forms have a more complex set of restrictions. Sometimes, there is a restriction against positive forms of the copula occurring in clauses of particular types, a kind of global restriction. This is true of responsives and polar interrogatives. Sometimes, positive forms are prevented from occurring adjacent to a complementizer, a local restriction. This is true in if-clauses and declaratives. Unlike a clause-type restriction, this kind of restriction is characterised by (1) existence of an overt complementizer for the clause and (2) cancellation of the effect if there is an intervening focus-fronted phrase. We showed that the local kind of restriction appears to have different effects in different contexts. In if-clauses, the blocking of **ɔs r-əd-u* leads to occurrence of *ɔs əd-u*. In declaratives, the blocking of **vɛ r-əd-u* leads to the occurrence of *r-əd-u* on its own, not **vɛ əd-u*. There is also the matter of the same suppletive form *maɪ* blocking the (hypothetical) regular positive-definite and positive-indefinite forms. These are the facts to be accounted for in this section.

Some of the data appears to support a semantic or pragmatic approach to the distribution of the forms. After all, positive forms are blocked in two non-assertive contexts: if-clauses and polar interrogatives. However, blocking positive forms of 'be' from non-assertive contexts does not account for the data. Firstly, it does not explain why the positive forms are blocked from responsives. Secondly, it does not explain why the presence of fronted material in an if-clause has an ameliorating effect on the restriction against positive forms in if-clauses. Indeed, close inspection of the data shows that positive forms are not blocked from if-clauses: they are blocked from being adjacent to the complementizer found in if-clauses (as well as the one found in declarative clauses). Therefore, we treat this as a syntactic phenomenon.

3.2 Blocking Principles

L_RFG assumes a set of blocking principles to choose the actual exponent from the set of possible exponents. The final formalization of these principles will appear in Asudeh & Siddiqi (2026), but for a recent formalization, see Asudeh & Siddiqi (2023). Here we present just the intuitions behind the two of the five principles that are relevant to this paper.

tations *na d-əd-i*, in which the complementizer and negative form co-occur, and *nad əd-i*, in which the complementizer and neutral form co-occur, are equally plausible. The same is true for the now-literary main complementizer *ni(d)*. Indeed, it is the ambiguity of segmentation that led to the development of negative forms in the first place. We assume the *na d-əd-i* analysis, because it is not necessary to assume that negative forms are incompatible with complementizers more generally, and so assuming they are incompatible in these cases adds unnecessary complexity.

[7] Negative forms and positive forms are also blocked when subjects are fronted, although we set aside this data in this paper; see Dowle (2025) for discussion of how this could also be addressed in a way which does not disrupt the analysis presented here.

(19) **MostInformative$_c$**: Of the candidate vocabulary items for exponence, which match the categorial information of the node(s) to be exponed, choose the vocabulary item that contains the most categories; if there is no such candidate do nothing.

For example, choose a VI whose category list is [A,B] over a VI whose category list is [A].

(20) **MostInformative$_f$**: Of the candidate vocabulary items for exponence, whose f-descriptions define an f-structure that subsumes the f-structure of the node(s) to be exponed, choose the vocabulary item that defines the f-structure at the top of the subsumption lattice, i.e. the f-structure that is properly subsumed by all other candidate f-structures.

For example, choose a VI whose defining equations are $\{(\uparrow \text{ATTR}_1) = \text{VAL}_1, (\uparrow \text{ATTR}_2) = \text{VAL}_2\}$ over $\{(\uparrow \text{ATTR}_1) = \text{VAL}_1\}$; if there is no such candidate do nothing.

In short, MI_c privileges vocabulary items that contain more c-structural category information, whereas MI_f privileges vocabulary items that contain more f-structural information. Thus, MI_c is about categories and prefers VIs that express more categories from the c-structure, while MI_f is about function and prefers VIs that express more function, in terms of functional structure. A key contribution of this paper is to show that constraining equations are not relevant to the calculation of MI_f, since they do not *define* but rather restrict possible f-structures. The difference between defining and constraining equations in relation to MI_f has not previously been considered in L_RFG, nor has it been discussed in relation to blocking in mainstream LFG (Andrews 1990). A further contribution is some evidence to suggest that MI_f may override MI_c if they provide conflicting selections.

3.3 Main Vocabulary Items

Here we present the main VIs relevant to the data under consideration. The copula as a whole sits under the I node of the c-structure tree in L_RFG, just as in mainstream LFG. In L_RFG, each morpheme, not just each word, has its distribution represented in the c-structure and so the I node is further decomposed into the different segmentable morpheme positions, with associated functional information. VIs expone these terminal nodes. We consider neutral forms of the copula to be sequences of a T morpheme, expressing tense, and an Agr morpheme, expressing subject agreement (although the indefinite form lacks any Agr exponent). Positive and negative forms additionally consist of a Pol morpheme. Note that this is the position associated with the expression of positive or negative concord, not the syntactic category associated with the main marking of negation (which could be Neg, Adv or some other category). Pol is probably a minor category not found in many languages. Together, these morpheme categories sit under I and form the copula.[8]

[8]Templates like @TENSE! and @AGREEMENT! abbreviate the possible values for TENSE and agreement-related f-structure features (such as SUBJ INDEX). For example, one of the options that @TENSE! provides is (↑ TENSE) = PRES, which is appropriate for a present-tense clause. We don't define them in this paper, because to do so would be to provide a full account of tense and agreement in Welsh.

(21) $\text{I} \rightarrow \left\{ \begin{array}{c} \text{Pol} \\ \uparrow{=}\downarrow \end{array} \mid \begin{array}{c} \sqrt{} \\ @\text{ROOT}(\_) \end{array} \right\} \begin{array}{c} \text{T} \\ \uparrow{=}\downarrow \\ @\text{TENSE}! \end{array} \begin{array}{c} \text{Agr} \\ \uparrow{=}\downarrow \\ @\text{AGREEMENT}! \end{array}$

The rule additionally captures that Pol does not occur with root-containing verbs in I.[9]

Here are the two VIs for the *d-* and *r-* prefixes.

(22) prefix to make negative forms:
$\langle$ [Pol], $\langle\langle\ (\uparrow \text{NEG}) =_c \oplus\ \rangle\rangle$ $\rangle \xrightarrow{\text{v}}$ *d-*
$\langle\langle\ (\uparrow \text{TENSE}) =_c \{\text{PRES} \mid \text{IMPF}\}\ \rangle\rangle$

(23) prefix to make positive forms:
$\langle$ [C, Pol], $\langle\langle\ \neg\ (\uparrow \text{NEG})\ \rangle\rangle$ $\rangle \xrightarrow{\text{v}}$ *r-*
$\langle\langle\ (\uparrow \text{TENSE}) =_c \{\text{PRES} \mid \text{IMPF}\}\ \rangle\rangle$
$\langle\langle\ (\uparrow \text{FORCE}) \neq \{\text{RESP} \mid \text{POL-Q}\}\ \rangle\rangle$

Both VIs contain only constraining functional equations. The notation '$\langle\langle$ ' and ' $\rangle\rangle$' is used to overtly indicate constraining equations.[10] Both of these prefixes are constrained by negation, but do not themselves define the polarity of the clause. Recall that sentential negation is expressed later in the sentence by e.g., *ðim* (cf. 11); if overt negation is found, *d-* is licensed, but in its absence, clauses are positive, the NEG feature is absent from the structure, and *r-* can occur. We treat negation as a privatively-valued feature, and we use $\oplus$ as the only possible feature value for a privative feature, to notationally distinguish it from a positive value (+) for a binary feature. At f-structure, a feature must have a value if it is present, as a formal requirement of AVMs; this is why even a privative feature must have a value specified. In addition to negation-sensitivity, both *r-* and *d-* are restricted to present or imperfect clauses; as only the copula expresses these tenses in a way which is morphologically distinct from the future or conditional, this also restricts the polarity markers to occurring with the copula.

Unlike *d-*, *r-* shows sensitivity to force, being prevented from occurring in a structure that is responsive or polar-interrogative. The f-structure feature FORCE is a re-naming of the familiar STMT-TYPE feature of LFG.[11] FORCE has the following possible values in Welsh: {RESP | POL-Q | WH-Q | COND | DECL}. These are morphosyntactic features, reflecting Welsh morphosyntactic distinctions, which go beyond the distribution of forms of 'be'. For example, Colloquial Welsh disfavours pro-drop in many contexts including ordinary declarative clauses, but still requires it in responsives.[12] Finally, the *r-* exponent spans both the C and Pol nodes; this captures the adjacency restriction between

Without lexical items in the c-structure, L_RFG in its current formulation uses phrase-structure rules to introduce the necessary f-structure features for a sentence; the VIs then expone these category–feature pairs.

[9] Although this may be too restrictive for those speakers who allow the prefix with certain content verbs (Borsley & Jones 2005: 50). These minor cases can be dealt with via a third disjunction.

[10] This notation is first introduced in L_RFG by Asudeh et al. (2024a), in response to feedback from practitioners of Minimalism/Distributed Morphology that the constraining equations were not always readily identifiable to those outside the theory; LFGers can simply disregard the extra notation.

[11] Again, this naming is intended to make the role of the feature more readily interpretable to people less familiar with LFG.

[12] Readers may wonder why a single feature INT isn't used for both polar and interrogative questions. This allows for a more efficient expression of the effect of fronting phrases on the form of 'be' that occurs. Certain fronted phrases (subjects, predicates) require the neutral form to follow them, and this is true

positive forms and overt complementizers, as we shall see in more detail in subsequent sections.

The following VIs capture the present-tense stem (T-node) of the copula, according to whether the subject is definite or indefinite (see Section 3.4):

(24) copula stem:
⟨ [T], (↑ TENSE) = PRES ⟩ $\overset{\text{ν}}{\rightarrow}$ *əd*
⟨⟨ (↑ SUBJ DEF) =$_c$ ⊕ ⟩⟩

(25) INDEF copula stem:
⟨ [T], (↑ TENSE) = PRES ⟩ $\overset{\text{ν}}{\rightarrow}$ *ois*
⟨⟨ ¬(↑ SUBJ DEF) ⟩⟩

The following are some examples of the kind of VIs used to expone the Agr node. Note that none of these can occur with the indefinite copula stem, due to conflicting definiteness requirements. The templates of the type @1SG.SUBJECT abbreviate (↑ SUBJ PRED) = 'pro', (↑ SUBJ DEF) = ⊕, (↑ SUBJ INDEX PERS) = 1, and (↑ SUBJ INDEX NUM) = SG, with different person and number values as applicable.[13]

(26) 1SG ending:
⟨ [Agr], @1SG.SUBJECT ⟩ $\overset{\text{ν}}{\rightarrow}$ *-u*
⟨⟨ (↑ TENSE) =$_c$ PRES ⟩⟩

(27) 3PL ending:
⟨ [Agr], @3PL.SUBJECT ⟩ $\overset{\text{ν}}{\rightarrow}$ *-in*
⟨⟨ (↑ TENSE) =$_c$ PRES ⟩⟩

(28) 3SG/DEF ending:
⟨ [Agr], (@3SG.SUBJECT) ⟩ $\overset{\text{ν}}{\rightarrow}$ *-i*
⟨⟨ (↑ TENSE) =$_c$ PRES ⟩⟩
⟨⟨ (↑ SUBJ DEF) =$_c$ ⊕ ⟩⟩

The suppletive present-tense, 3SG positive form, *maɪ*, being a portmanteau form, spans C, Pol, T and Agr. This portmanteau form blocks the analytic 3SG positive forms that could otherwise be derived by the above VIs: **r-əd-i* and **r-ois* (see Section 3.4).

(29) ⟨ [C, Pol, T, Agr], (↑ TENSE) = PRES ⟩ $\overset{\text{ν}}{\rightarrow}$ *maɪ*
⟨⟨ ¬ (↑ NEG) ⟩⟩
⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩

in both constituent questions and declaratives with fronting. If constituent questions are treated as the same clause type as polar questions, in selecting the neutral forms by default, then those fronted phrases which are *not* followed by the neutral form must be listed, alongside those that require the neutral form in declaratives being listed. Previous LFG work has also assumed a syntactic distinction between constituent and polar questions, using the QUESTION-TYPE feature with possible values {CONSTITUENT | POLAR} (Butt et al. 2017); we simplify by representing WH-Q and POL-Q as distinct values of a single feature type, FORCE, rather than using both STMT-TYPE and QUESTION-TYPE, though this is not essential to the analysis.

[13] We assume that verb-subject agreement involves INDEX features, as per Wechsler & Zlatić (2003). We follow Sadler (1997) in treating PRED = 'pro' contributions in Welsh verb-agreement forms as obligatory, contra most pro-drop languages. For further discussion see Dowle (2022) and Dowle (2026).

3.4 Definiteness marking and suppletion

The VIs provided reflect the fact that, unlike the suppletive positive form, the negative and neutral third-person singular forms of the copula are sensitive to the definiteness value of their subjects. Table 2 is repeated here as Table 4.

	Definite subject	*Indefinite subject*
POS	ma(ɪ)	
NEG	d-əd-i	d-ɔis
NEUT	əd-i	ɔis

Table 4: Third-person singular present-tense forms of the copula

It is important to establish that they *constrain* rather than define this value. NPs are indefinite by default; a definite article or definite possessor marks them as definite:

(30) a. draenog
draɪmɔg
hedgehog
'a hedgehog'

b. y draenog
ə draɪmɔg
the hedgehog
'the hedgehog'

c. draenog y dyn
draɪmɔg ə diːn
hedgehog the man
'the man's hedgehog'

In contrast, the definiteness value of a NP cannot be set by a copular form. A definite negative or neutral form with an unmarked NP does not result in a definite reading:

(31) * Dydy draenog ddim yn cysgu.
d-əd-i draɪmɔg ðɪm ən kəsg-i
NEG-be-PRES hedgehog NEG PROG sleep.NF
Intended: 'The hedgehog isn't sleeping.'

Thus, a definite article or possessor in some way defines the definiteness of the whole NP, but the copula verb form does not. This is captured by the use of a constraining equation for definiteness in the VIs for the definite and indefinite forms.

The suppletive form *maɪ* blocks hypothetical forms **r-əd-i and *r-ois*. This leads us to conclude that despite the extra constraining equations in **r-əd-i* and **r-ois*, the fact that *maɪ* is a portmanteau is sufficient for it to be selected instead. Thus, constraining equations cannot override MI_C.

maɪ	**r-əd-i*	**r-ois*
1 VI	3 VIs	2 VIs
C, Pol, T, Agr	C, Pol, T, Agr	C, Pol, T
(↑ TENSE) = PRES	(↑ TENSE) = PRES	(↑ TENSE) = PRES
⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑ NEG) ⟩⟩	⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑ NEG) ⟩⟩ ⟨⟨ (↑ SUBJ DEF) =$_c$ ⊕ ⟩⟩	⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑NEG) ⟩⟩ ⟨⟨ ¬(↑ SUBJ DEF) ⟩⟩

Table 5: Comparison of *maɪ* vs *r-əd-i* and *r-ois*

3.5 Complementizer positions in Welsh, and if-clauses with fronting

As we have seen, we have several complementizer-like elements in Welsh: *vɛ*, *taʊ* and *ɔs* (among others); *vɛ* occurs without fronted material and directly precedes the verb, while *taʊ* occurs with fronted material and directly precedes said fronted material.

(32)
Fe	fydd	y	dyn	yn	nesaf.
vɛ	viːð	ə	diːn	ən	nesav
C$_{\text{PDMC}}$	be.FUT	the	man	PRED	next

'The man will be next.'

(33)
Dw	i	'n	meddwl	taw	[y	dyn]$_{\text{FOC}}$	fydd	yn
du	iː	=n	mɛðul	**taʊ**	[ə	diːn]$_{\text{FOC}}$	viːð	ən
be.PRS.1SG	1SG	=PROG	think.NF	**C$_{\text{SUB.FOC}}$**	[the	man]$_{\text{FOC}}$	be.FUT	PRED

nesaf.
nesav
next
'I think that *the man* will be next.'

The conditional complementizer *ɔs* is unusual because it can occur with fronted material or without it. When it occurs with fronted material, it occurs before it; when it occurs without it, it is next to the verb (there is no flexibility as to its position).

(9′)
Os ydw	i	'n	gallu	dod,	bydda		i	'n
ɔs əd-u	iː	=n	gaɬi	dɔːd,	bəð-a		iː	=n
if be.PRS-1SG	1SG	=PROG	can.NF	come.NF	be.FUT-FUT.1SG		1SG	=PROG

dweud wrthot ti.
dwəɪd urth-ɔt tiː.
say.NF at-2SG 2SG
'If I can come, I will tell you.'

(10)
Os [Mari]$_{\text{FOC}}$	rwyt	ti	wedi	alw,	bydd	popeth	yn
ɔs [mari]$_{\text{FOC}}$	r-ʊɪt	tiː	wɛdi	alu,	bið	pɔpeθ	ən
if [Mari]$_{\text{FOC}}$	POS-be.IMPF.2SG	2SG	PERF	call.NF	be.FUT	everything	PRED

iawn.
jaun.
okay
'If it's *Mari* you've called, everything will be okay.'

We assume that there are two complementizer-like positions in Welsh, one high one, which we call Foc, and one lower one, which we call C. Most things, like *taʊ* (Foc) and *vɛ* (C) belong to only one or the other category and are distributed accordingly:

(34) FocP → Foc (($\uparrow$ FORCE) = COND) XP ($\uparrow$ UDF) = $\downarrow$

(35) CP → FocP $\uparrow$=$\downarrow$ C′ $\uparrow$=$\downarrow$

(36) C′ → C @FORCE! IP $\uparrow$=$\downarrow$

The @FORCE! lists all of these as possible values for (↑ FORCE), discussed above: {RESP | POL-Q | WH-Q | COND | DECL}.

(37) $\langle$ [C], $\langle\langle$ (↑ FORCE) $=_c$ DECL $\rangle\rangle$ $\rangle \xrightarrow{\nu} v\varepsilon$
$\langle\langle \neg$ (GF ↑) $\rangle\rangle$
$\langle\langle \neg$(↑ NEG) $\rangle\rangle$
$\langle\langle \neg$(↑ UDF) $\rangle\rangle$

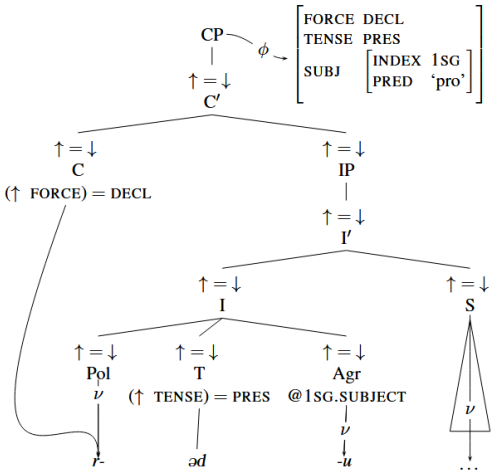
(38) $\langle$ [Foc], $\langle\langle$ (↑ FORCE) $=_c$ DECL $\rangle\rangle$ $\rangle \xrightarrow{\nu} ta\upsilon$
$\langle\langle$ (GF ↑) $\rangle\rangle$
$\langle\langle \neg$(↑ NEG) $\rangle\rangle$

The conditional *ɔs* has a disjunctive category specification because it can appear in either position (the feature (↑ FORCE) = COND can appear in both C and Foc accordingly):

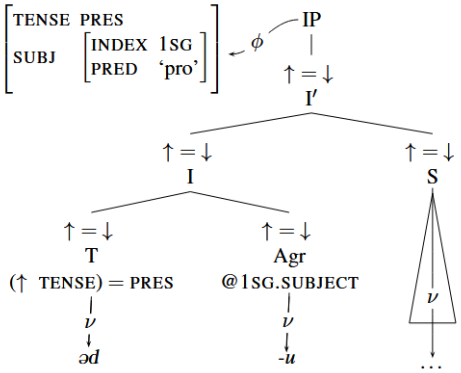
(39) $\langle$ [{ C | Foc }] , (↑ FORCE) = COND $\rangle \xrightarrow{\nu} ɔs$
$\langle\langle \neg$(↑ UDF) $\rangle\rangle$ $\langle\langle$ (↑ UDF) $\rangle\rangle$

Coupled with the assumption that positive forms of the copula span the C node, we have an explanation for the blocking of positive forms next to the complementizers *ɔs* and *vɛ*. Compare the kinds of structures spelled out by *r-ɘd-u* with those spelled out by *ɘd-u*:

(40)

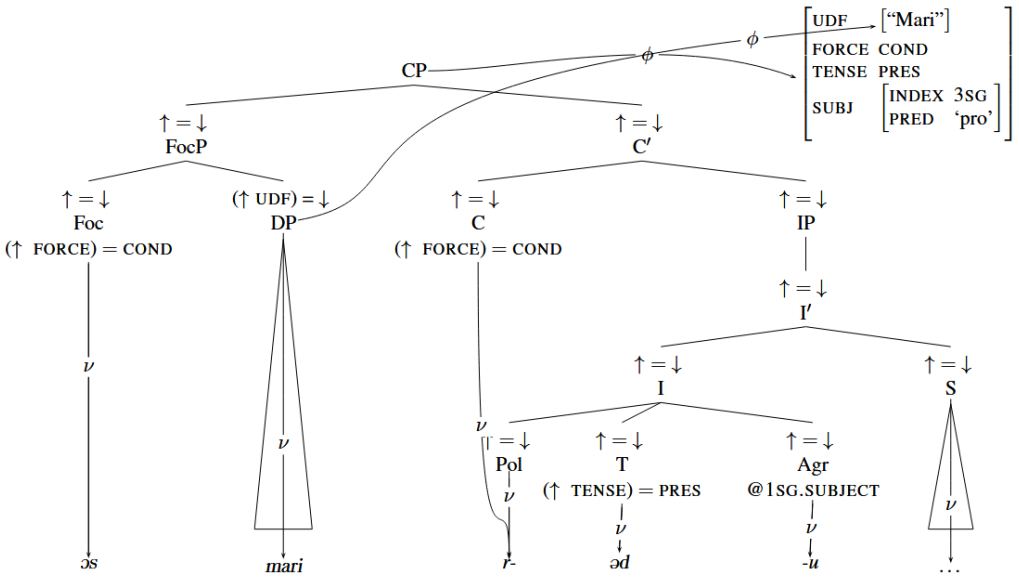


(41)



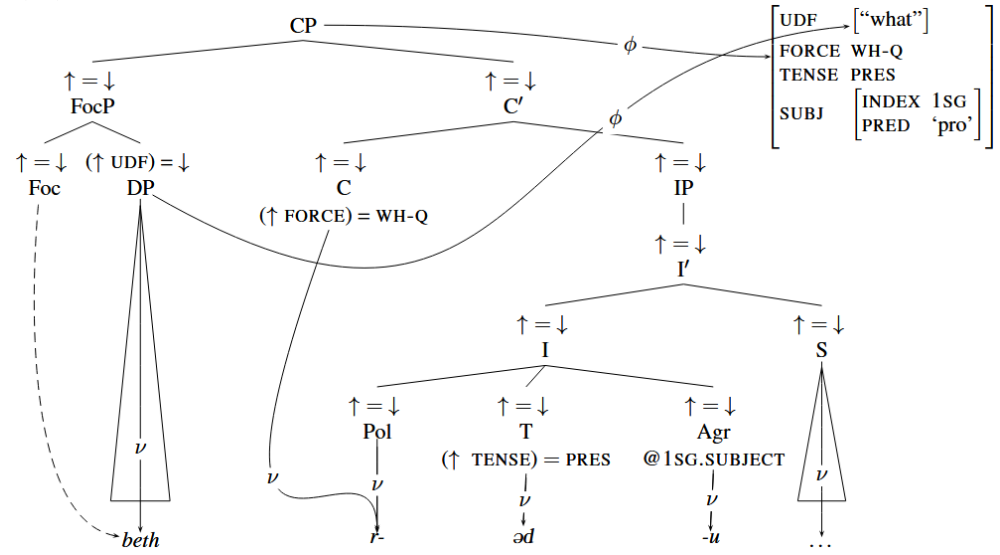
We now have an explanation as to why the presence of focus-fronting material in 'if'-clauses prevents blocking between *ɔs* and positive forms: *ɔs* is in competition with the positive copula in sentences without fronted material, but not when there is fronted material. Here's an if-clause with fronting, in which *r-ɘd-u* isn't blocked:

(42)



The structure is very similar in a declarative clause with fronting or a constituent question.

(43)



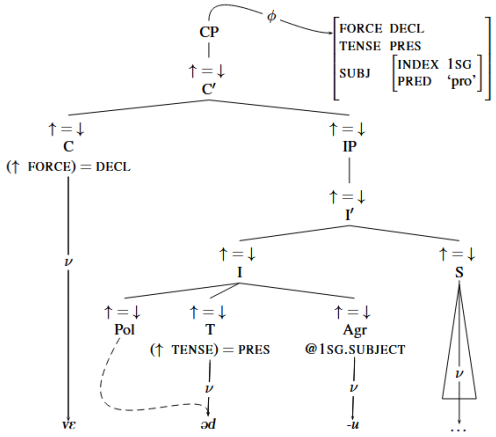
At this stage, our analysis predicts the ungrammaticality of *os r-əd-u* and *vɛ r-əd-u*. It does not yet explain why the ungrammaticality of these sequences has different outcomes in if-clauses as compared to declaratives.

3.6 The difference between if-clauses and declaratives

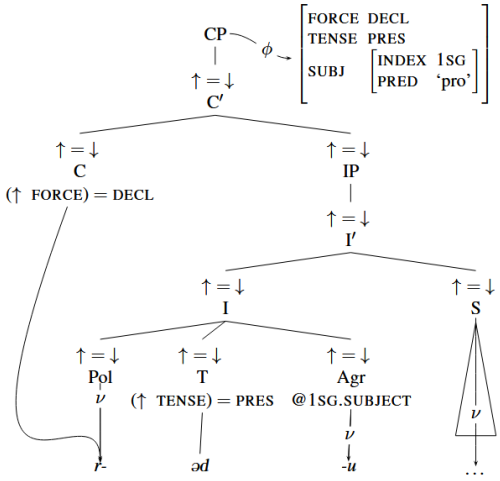
One of the key questions to be resolved is why positive forms like *r-əd-u* block *vɛ əd-u* (in declarative clauses) but do not block *ɔs əd-u* (in if-clauses). As far as MI_c is

concerned, *r-əd-u* should always beat these sequences as it expresses more c-structure nodes. The complementizer-containing sequences involve the unexpressed Pol node being Pac-Man spanned, thus making them less informative with respect to MI_c. Pac-Man spanning is an operation that ensures no c-structure nodes are left unexpressed by requiring that a node with no actual exponent in the Vocabulary gets "gobbled up" by the exponent of a neighbouring node, thus avoiding null exponence.[14] Despite the MI_c evaluation being the same in both cases, we have reverse grammaticality judgements in the two contexts:

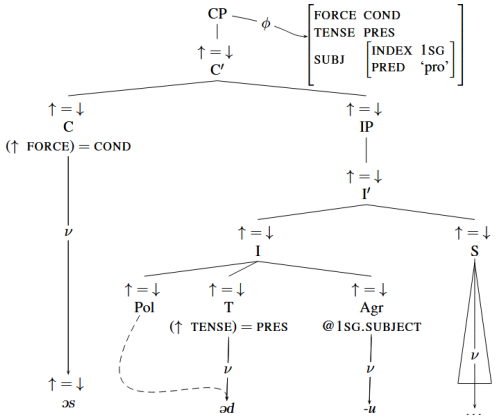
(44) (ungrammatical)



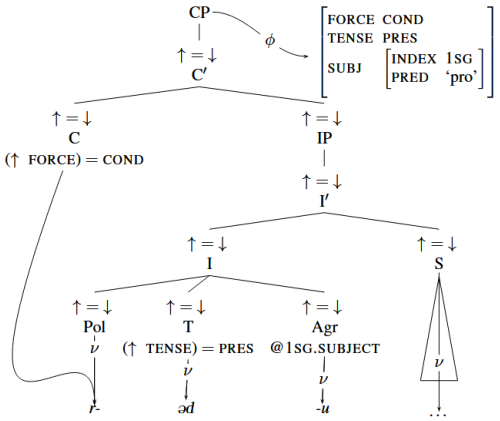
(45) (grammatical)



(46) (grammatical)



(47) (ungrammatical)



The fact that the sequence *ɔs əd-u* blocks *r-əd-u* in if-clauses tells us that *something* must be able to override MI_c. An obvious candidate is MI_f—in keeping with previous intuitions that MI_f is more important than phonological specificity (called **MostSpecific**) because "concepts tend to find a way to be expressed" (Asudeh & Siddiqi 2024).

[14] For a brief discussion of Pac-Man spanning, see Asudeh, Bögel & Siddiqi (2023: 24, 39). Asudeh & Siddiqi (2026) contains more extended discussion, and formalization, of Pac-Man spanning.

r-əd-u	**ve əd-u*	*os əd-u*
C, Pol, T, Agr	C, T, Agr	C, T, Agr
(↑ TENSE) = PRES @1SG.SUBJECT	(↑ TENSE) = PRES @1SG.SUBJECT	(↑ TENSE) = PRES @1SG.SUBJECT (↑ FORCE) = COND
⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑ NEG) ⟩⟩	⟨⟨ (↑ FORCE) =$_c$ DECL ⟩⟩ ⟨⟨ ¬ (↑ NEG) ⟩⟩ ⟨⟨ ¬ (GF ↑) ⟩⟩ ⟨⟨ ¬ (↑UDF) ⟩⟩	

Table 6: Comparison of functional information in *r-əd-u* vs *ve əd-u* and *ɔs əd-u*

The data makes clear that MI_c can be overridden. If MI_c were to always take precedence, there would be no way to account for the fact that the C, T, Agr sequence *os əd-i* ever wins over *r-əd-u*. The if-clause data thus leads us to the conclusion that MI_f takes precedence over MI_c. Unlike *r-əd-u*, *ɔs əd-u* overtly marks the conditional force of the clause, and this is why *ɔs əd-u* is not blocked by the span. Although constraining equations can't be violated, they cannot override MI_c, whilst defining equations must.

Because of the way feature unification works in LFG, meaning that two defining equations specifying compatible values for the same feature can resolve into an f-structure with a single value for the feature in question, we could also have chosen to represent the VIs in the following way, with *vɛ* containing a defining equation:

r-əd-u	**ve əd-u*	*os əd-u*
C, Pol, T, Agr	C, (Pol), t, Agr	C, T, Agr
(↑ TENSE) = PRES @1SG.SUBJECT	(↑ TENSE) = PRES @1SG.SUBJECT (↑ FORCE) = DECL	(↑ TENSE) = PRES @1SG.SUBJECT (↑ FORCE) = COND
⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑ NEG) ⟩⟩	⟨⟨ ¬ (↑ NEG) ⟩⟩ ⟨⟨ ¬ (GF ↑) ⟩⟩ ⟨⟨ ¬ (↑UDF) ⟩⟩	

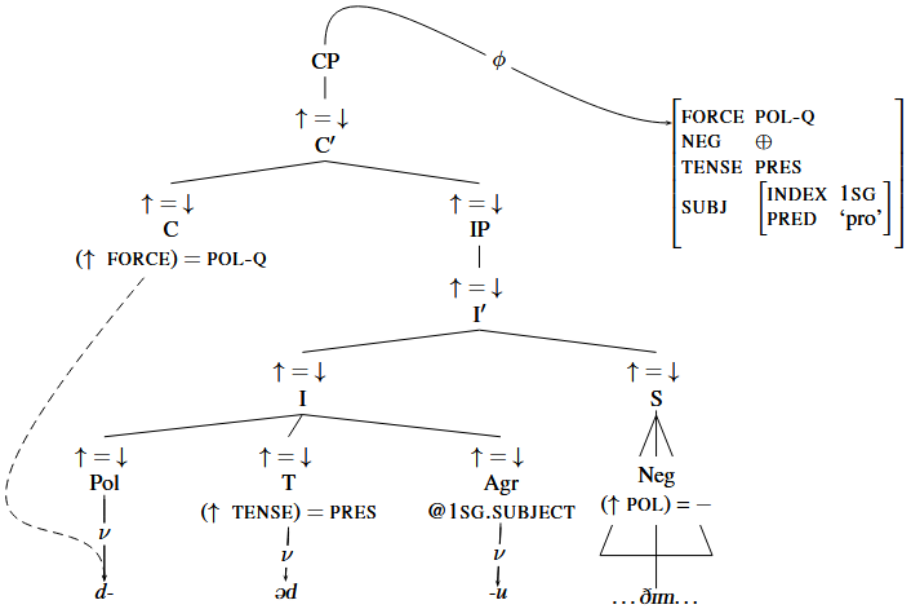
Table 7: Alternative comparison of func. information in *r-əd-u* vs *ve əd-u* and *ɔs əd-u*

Clauses can be interpreted as declarative by default; *vɛ* is optional in declarative clauses; see (14), (15), and (16). This does not mean that *vɛ* cannot contribute a defining equation for this feature, but it does mean that it cannot be the only thing to do so, and that a constraining equation possibility exists instead. As the defining equation approach is the only approach which makes it impossible to predict the selection of *r-əd-u* over **vɛ əd-u*, our observations of the data lead us to adopt the equally-available constraining approach in Table 6 instead. Blocking relationships thus provide us with a principled means for choosing between defining equations and constraining equations for VIs. We illustrated in Section 3.4 that constraining equations do not outcompete MI_c; here we use that insight to arrive at the analysis which can capture the observed blocking relationship.

3.7 Negative forms

Recall that negative forms are blocked in responsives, but fine in polar questions, declaratives, and if-clauses. Negative Pol doesn't contain any clause-type restriction on its distribution. It's fine in negative polar questions.

(48)



There is no VI for ⟨ C, (↑ FORCE) = POL-Q ⟩ in Modern Welsh, so C is Pac-Man spanned, as indicated by the dashed line. This is also true in main-clause negative declaratives.[15]

We propose that negative forms are blocked from responsives not because of a functional restriction but because of the negative responsive complementizer. *ŋak əd-u* is chosen over *d-əd-u* in this context, because it is both MI_f and MI_c: it realises more nodes (C, Pol, T, Agr vs. Pol, T, Agr), and it expresses more functional information:

(49) ⟨ [C, Pol], (↑ POL) = − (↑ FORCE) = RESP ⟩ $\xrightarrow{\nu}$ *nak*

(50) ⟨ [Pol], ⟨⟨ (↑ NEG) =$_c$ ⊕ ⟩⟩ ⟨⟨ (↑ TENSE) =$_c$ {PRES|IMPF} ⟩⟩ ⟩ $\xrightarrow{\nu}$ *d-*

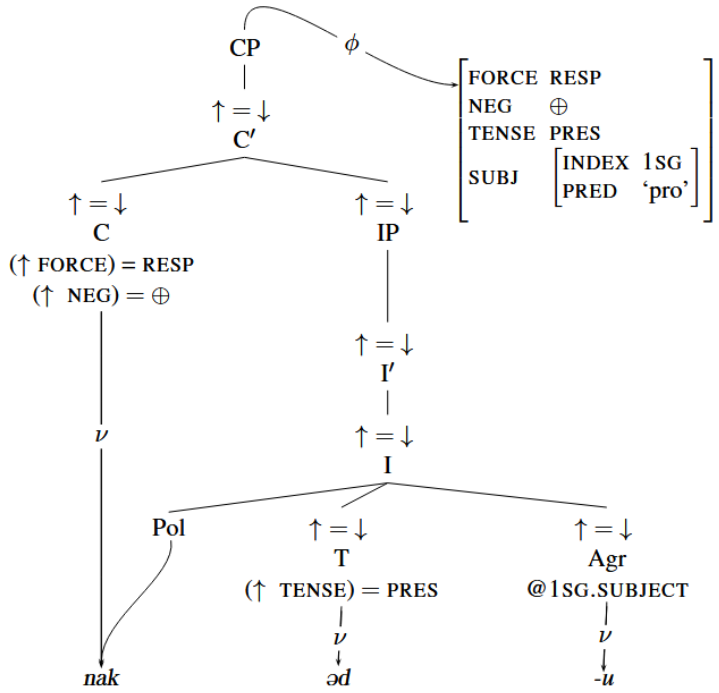
nak əd-u	*d-əd-u*
(↑ TENSE) = PRES	(↑ TENSE) = PRES
@1SG.SUBJECT	@1SG.SUBJECT
(↑ FORCE) = RESP	
(↑ NEG) = ⊕	
	⟨⟨ (↑ NEG) =$_c$ ⊕ ⟩⟩

Table 8: Comparison of functional information in *nak əd-u* vs *d-əd-u*

[15] In subordinate clauses, the C is expressed as *na*.

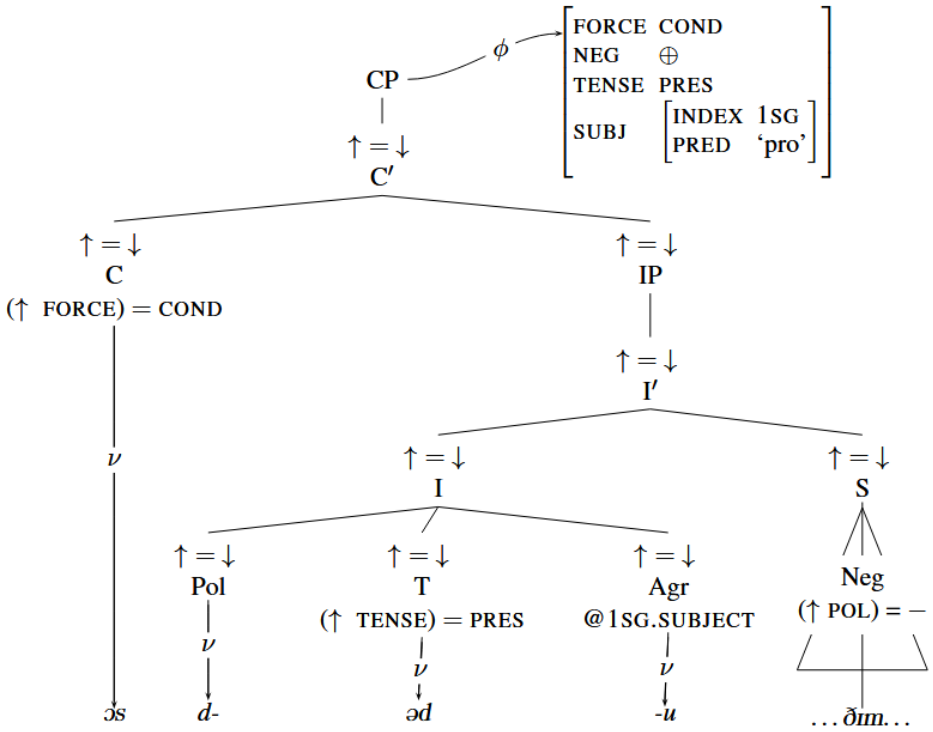
This gives the following structure:

(51)



In negative if-clauses, since ɔs is just an ordinary C, (not a C-Pol span like *nak*) the negative copula is correctly predicted to occur.

(52)



4 Conclusion

The morphological component in L_RFG has two mechanisms for comparing forms based on syntactic information, MI_f and MI_c. These are used to select the maximally informative VI possible, to express a sentence's c-structure tree. Welsh has positive, negative and neutral forms of the copula, but only in the present and imperfect tense. The positive and negative forms are usually derived by the addition of an *r-* or *d-* prefix to the neutral form, although there are some suppletive third-person forms in the present-tense paradigm. The negative form of the copula is allowed in many more contexts than the positive–responsives are the only context in which they are blocked. They get blocked in responsives because *nak*, the negative responsive marker, spans the Pol head that *d-* would normally express.

The positive forms are completely blocked from responsives and polar interrogatives via a functional restriction. Positive forms of the copula span the C node, which blocks them from occurring with *vɛ* and next to *ɔs* (C elements). Focus-fronting material allows positive forms to occur in if-clauses, because the focus-fronting material forces *ɔs* into a higher position (Foc). This is also why a universal functional restriction is not appropriate for the blocking of the positive forms next to *ɔs* 'if', because positive forms are not blocked from if-clauses entirely; it is only an adjacency restriction.

The negative and neutral present tense 3SG copular forms show a subject definiteness sensitivity, which must be captured by a constraining equation. The fact that the same suppletive form *maɪ* blocks both the definite-agreement form and the indefinite-agreement form shows that constraining equations don't override MI_c.

In if-clauses without fronting, the sequence *ɔs ədu* beats *rədu* because it wins on MI_f grounds — it has more defining equations. Thus, defining equations do override MI_c. In positive declaratives, *vɛ ədu* is blocked by *rədu*. This can be explained if its force specification is treated as a constraining equation, making it unable to override *rədu* on MI_f grounds. *rədu* then wins on MI_c grounds. Treating the force specification as a constraining equation is justified because *ve* is optional, suggesting that declarative force can be contributed by independent means (e.g. as a default value) and does not necessarily require the feature to be defined (and then exponed) in C. Thus, given the insights regarding the relative ordering of MI_c, constraining equations and defining equations established on the basis of the suppletive third-person stem, the interactions of the positive declarative complementizer and positive forms can be captured.

In sum, MI_f and MI_c both operate independently of constraining equations. This allows MI_c to override constraining equations that might otherwise make a form count as 'more informative'. Constraining equations are not evaluated as part of MI_f, either. MI_f takes precedence over MI_c, whereas constraining equations do not. This distinction between how defining and constraining equations are evaluated by the MI principles reflects their differing effects on f-structure in LFG: constraining equations only check whether f-structure information has been independently added by some defining equations and must thus be evaluated only once the f-structure has been defined. Here we do not assume ordering between defining versus constraining equations in VIs, but nevertheless similarly accord constraining equations a secondary role of expressing *morphological conditioning* (see Asudeh et al. 2024a for discussion in the context of L_RFG) rather than expressing morphological information.

A Appendix: Discussion on binary and privative features

In L_RFG, the same features present in the f-structure are used in the morphology, to organise the distribution of VIs. We might therefore expect the features in the f-structure to match the requirements of the morphological component i.e. a privative feature is sufficient to capture a morphological situation in which there are two possible allomorphs: one conditioned by the presence of the privative feature, and one conditioned by its absence. Even if the absence of the feature is overtly morphologically marked in some context, this is easy to capture in L_RFG because L_RFG inherits from LFG the possibility of stating negative constraints.

This creates a situation in which syntactic/featural markedness and morphological markedness do not necessarily align: the absence of a feature at f-structure can be overtly marked morphologically. We propose that it should be a principle of L_RFG that the feature make-up of a grammar should reflect morphological marking. Thus, negative constraints should condition the absence of morphological marking, not the occurrence of an equally-morphologically-marked (overt) form. We partially used this line of argumentation to argue that indefinite NPs in Welsh are characterised by the absence of a DEF feature, not the presence of a $(\uparrow \text{DEF}) = -$ feature value. We said that unmarked NPs are underspecified, and are interpreted as indefinite unless a definite feature is added by something in the environment.

However, we also adopted a privative approach to polarity, which may seem contradictory, given the presence of both the *r-* and *d-* prefixes, and both the *vɛ* and *nak* complementizers. The existence of such forms shows that positive and negative polarity concord are equally capable of being part of the overt expression of a form; under the above principle, this means that a binary feature should be used, even though it is not strictly *necessary*, formally, as demonstrated.

Nevertheless, a more fine-grained approach suggests that a privative feature is still justified: if each type of context is considered separately (Table 9), we find that we never have a minimal pair in which positive and negative marking are equally morphologically marked. Negation is always more marked. There is no positive equivalent of the sentential negator *ðim*, for example. Thus, the privative approach is justified, especially given that actual polarity marking (as opposed to polarity concord) only overtly marks the negative, and not the positive.

Context	positive marking	negative marking
sentence without PRES/IMPF copula	unmarked	*ðim*
sentence with PRES/IMPF copula	*r-* on its own	*d-* and *ðɪm*
responsives	unmarked	*nak*

Table 9: Positive utterances are always less marked than negative counterparts

References

Andrews, Avery D. 1990. Unification and morphological blocking. *Natural Language & Linguistic Theory* 8(4). 507–557. https://doi.org/10.1007/bf00133692.

Asudeh, Ash, Bronwyn Bjorkman, Frances Dowle, Neil Myler, Daniel Siddiqi & Lisa Sullivan. 2024a. Fusional morphology, metasyncretism, and secondary exponence: A morphemic, realizational approach to Latin declension. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'24 Conference*, 45–68. Konstanz: PubliKon.

Asudeh, Ash, Tina Bögel & Daniel Siddiqi. 2023. Modelling exponents. In Miriam Butt Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'23 conference*, 22–44. PubliKon. https://lfg-proceedings.org/lfg/index.php/main/article/view/38.

Asudeh, Ash, Paul B. Melchin & Daniel Siddiqi. 2024b. Constraints all the way down: DM in a representational model of grammar. In Robert Autry, Gabriela de la Cruz, Luis A. Irizarry Figueroa, Kristina Mihajlovic, Tianyi Ni, Ryan Smith & Heidi Harley (eds.), *Proceedings of the 39th West Coast Conference on Formal Linguistics*, 468–476. Cascadilla Press. https://www.lingref.com/cpp/wccfl/39/paper3661.pdf.

Asudeh, Ash & Dan Siddiqi. 2024. The story of *er*. In Miriam Butt Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'24 conference*, 69–89. PubliKon. https://lfg-proceedings.org/lfg/index.php/main/article/view/47.

Asudeh, Ash & Daniel Siddiqi. 2022. Realizational morphosemantics in L_RFG. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'22 conference*, 20–40. CSLI Publications. https://lfg-proceedings.org/lfg/index.php/main/article/view/20.

Asudeh, Ash & Daniel Siddiqi. 2023. Morphology in LFG. In Mary Dalrymple (ed.), *Handbook of Lexical Functional Grammar*, 855–901. Berlin: Language Science Press. https://doi.org/10.5281/zenodo.10185976.

Asudeh, Ash & Daniel Siddiqi. 2026. *Constraint-based morphosyntax*. Cambridge University Press. Forthcoming; with Oleg Belyaev, Bronwyn M. Bjorkman, Tina Bögel, Frances Dowle, Michael Everdell, Paul B. Melchin, Neil Myler, Will Oxford, and Lisa Sullivan.

Bögel, Tina, Miriam Butt & Sebastian Sulger. 2008. Urdu ezafe and the morphology-syntax interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'08 conference*, 129–149. CSLI Publications. https://web.stanford.edu/group/cslipublications/cslipublications/LFG/13/papers/lfg08boegeletal.pdf.

Borsley, Robert D. 2009. On the superficiality of Welsh agreement. *Natural language and linguistic theory* 27(2). 225–265. https://doi.org/10.1007/s11049-009-9067-3.

Borsley, Robert D. & Bob Morris Jones. 2005. *Welsh negation and grammatical theory*. Cardiff: University of Wales Press.

Borsley, Robert D, Maggie Tallerman & David W. E. Willis. 2007. *The syntax of Welsh*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511759932.

Butt, Miriam, Tina Bögel & Farhat Jabeen. 2017. Polar *kya* and the prosody-syntax-pragmatics interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'17 conference*, 125–145. Stanford: CSLI Publications. https://web.stanford.edu/group/cslipublications/cslipublications/LFG/LFG-2017/lfg2017-bbj.pdf.

Butt, Miriam & Louisa Sadler. 2004. Verbal morphology and agreement in Urdu. In U. Junghans & L. Szucsich (eds.), *Syntactic structures and morphological information*, 57–100. Berlin: Mouton de Gruyter.

Dalrymple, Mary. 2015. Morphology in the LFG architecture. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'15 conference*, 64–83. Stanford: CSLI Publications.

Dowle, Frances. 2022. A new analysis of Welsh agreement. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'22 conference*, 124–143. CSLI Publications. https://lfg-proceedings.org/lfg/index.php/main/article/view/14.

Dowle, Frances. 2025. Neutral forms of *be* as default forms: The utility of underspecification and blocking in a Welsh morphosyntactic phenomenon. *Transactions of the Philogocial Society* https://doi.org/10.1111/1467-968x.70002.

Dowle, Frances. 2026. *Agreement in Welsh*. D.Phil. thesis, University of Oxford. To appear.

Jones, Ann & Meic Gilby. 2004. *BBC learn Welsh grammar guide: The ideal aid to speaking and writing, up-to-date and easy to use*. Talybont: Y Lolfa.

Jones, Bob Morris. 1999. *The Welsh Answering System*. Berlin, New York: Mouton de Gruyter.

King, Gareth. 2016. *Modern Welsh: A comprehensive grammar* (3rd edn.). London & New York: Routledge. https://doi.org/10.4324/9780203987063.

Kiparsky, Paul. 1993. Pāṇinian linguistics. In Ronald Asher & Seumas Simpson (eds.), *Encyclopedia of languages and linguistics*, vol. 6, 2918–2923. Oxford, New York: Pergamon Press. https://web.stanford.edu/~kiparsky/Papers/encycl.pdf.

Melchin, Paul B., Ash Asudeh & Daniel Siddiqi. 2020. Ojibwe agreement in Lexical-Realizational Functional Grammar. In Miriam Butt & Ida Toivonen (eds.), *Proceedings of the LFG'20 conference*, 268–288. Stanford: CSLI Publications. https://csli-publications.stanford.edu/LFG/LFG-2020/lfg2020-mas.pdf.

Sadler, Louisa. 1997. Clitics and the structure-function mapping. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'97 conference*. Stanford: CSLI Publications.

Thomas, Amanda. 2021. *Divergence and mismatches in Portuguese morphology and syntax*. D.Phil. thesis, University of Oxford.

Wechsler, Stephen & Larisa Zlatić. 2003. *The many faces of agreement*. Stanford, CA: CSLI Publications.