Categories *don't* take precedence: Evidence from Welsh

Frances Dowle
Ash Asudeh

LFG'25

1 Overview

- This paper explores complex blocking relationships in L_RFG, 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)
Rwyt	ti	'n	dawel.
r-ʊɪt	tiː	=n	dawel.
POS-be.PRS.2SG	2SG	=PRED	quiet.

'You are quiet.'

(2)
Dwyt	ti	ddim	yn
d-ʊɪt	tiː	ðɪm	ən
NEG-be.PRS.2SG	2SG	NEG	PRED

dawel.
dawel.
quiet

'You aren't quiet.'

(3)
Beth rwyt		ti	'n
bɛθ	**r**-ʊɪt	tiː	=n
what	**POS**-be.PRS.2SG	2SG	=PROG

hoffi?
hɔfi?
like.NF
'What do you like?'

(4)
Beth dwyt		ti	ddim
bɛθ	**d**-ʊɪt	tiː	ðim
what	**NEG**-be.PRS.2SG	2SG	NEG

yn hoffi?
ən hɔfi?
PROG like
'What don't you like?'

- There is also a neutral form, which can occur in either positive or negative clauses, and which lacks any onset consonant. This form is thus functionally and morphologically unmarked.
- Given this (and following Dowle forthcoming), we'll argue that the neutral form of the copula is a default form, arising when the polarity-specific forms are blocked. Capturing the distributions of the relevant forms is then about explaining why the polarity-specific forms are sometimes blocked.
- We are particularly interested in a contrast between declarative and if-clauses. In if-clauses, the sequence 'complementizer + neutral form' form wins. In declarative clauses, the positive form of the copula on its own beats the sequence 'complementizer + neutral form'. We will show that this is because of a difference between how f-structure and c-structure information is treated by the morphology.

- Morphological competition in L_RFG:
  - As a realizational model of morphology, L_RFG assumes that functional and categorical information is introduced by phrase structure rules into a kind of c-structure tree pretty familiar from LFG.
  - The job of the morphology is to then find the forms that best express the information in the tree. It assesses the lexicon 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.
  - **Because VIs may express only a subset of the information in the tree, there are usually multiple possible candidate VIs available**. According to the Elsewhere Principle (Kiparsky, 1993), which lots of theories adopt, where the grammar provides a choice of expressions, the expression that is most informative 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).

Given a c-structure c and two sets of vocabulary items, α and β,

$$\textbf{MostInformative}_c(\alpha, \beta) =$$
$$\alpha = \{x \mid x \text{ is a VI } \wedge \textbf{features}(x) \subseteq \textbf{targets}(c) \wedge \forall y \exists z.[y \in \textbf{categories}(x) \wedge z \in \textbf{labels}(c) \wedge \pi_2(z) = y]\}$$
$$\beta = \{x \mid x \text{ is a VI } \wedge \textbf{features}(x) \subseteq \textbf{targets}(c) \wedge \forall y \exists z.[y \in \textbf{categories}(x) \wedge z \in \textbf{labels}(c) \wedge \pi_2(z) = y]\}$$
$$\begin{cases} \alpha \textbf{ if } |\alpha| < |\beta| \\ \beta \textbf{ if } |\beta| < |\alpha| \\ \bot \textbf{ otherwise} \end{cases}$$

Figure 1: **MostInformative**$_c$ according to Asudeh and Siddiqi (2024)

Given two VIs, α and β,

$$\textbf{MostInformative}_f(\alpha, \beta) = \begin{cases} \alpha \textbf{ if } \exists f.f \in \Phi(\pi_2(\pi_1(\alpha))) \wedge \forall g.g \in \Phi(\pi_2(\pi_1(\beta))) \rightarrow g \sqsubset f \\ \beta \textbf{ if } \exists f.f \in \Phi(\pi_2(\pi_1(\beta))) \wedge \forall g.g \in \Phi(\pi_2(\pi_1(\alpha))) \rightarrow g \sqsubset f \\ \bot \textbf{ otherwise} \end{cases}$$

Figure 2: **MostInformative**$_f$ according to Asudeh and Siddiqi (2024)

- This paper will show that when **MI**$_f$ and **MI**$_c$ select different forms according to their particular criteria, it is **MI**$_f$ that matters rather than **MI**$_c$.

2 Data

2.1 Forms of the copula

- Some paradigms:

	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-ɪχ	ma(ɪ)n
NEG	d-ədu	d-ʊɪt	see below	d-əd-ɪn	d-əd-ɪχ	d-əd-in
NEUT	əd-u	ʊɪt		əd-ɪn	əd-ɪχ	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

– 3SG forms are actually completely unspecified for person and number (Dowle, 2022): these forms occur with noun phrases irrespective of number, because person-number inflection in Welsh is restricted to occurring with (definite) pronouns. '3SG' is a convenient label for us to use in an informal way. This also explains why there is no definite–indefinite split in the 3PL

2.2 Restrictions

- 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 (answers to polar questions).

(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 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 iawn?
r-ʊɪt tiː =n jaun?
__POS__-be.PRS.2SG 2SG =PRED 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.

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

- But Welsh has focus-fronting, and we see that this restriction does not hold when focused material is fronted:

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

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

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

(11)
Os	dydw	i	ddim	yn	gallu	dod,	bydda	i'n	dweud
ɔs	**d**-əd-u	i	ðɪm	=n	gaɬi	dɔd,	bəð-a	i	=n
if	**NEG**-be.PRS.1SG	1SG	NEG	PROG	can.NF	come.NF,	be.FUT-FUT.1SG	1SG	PROG

wrthot ti.
dwəɪd wrth-ɔt tiː.
say.NF at-2SG 2SG
'If I can't come, I will tell you.'

2.2.4 Declarative clauses

- Declarative clauses are a kind of special case. Instead of **being blocked**, positive forms are the ones that **block** the positive declarative main clause (PDMC) complementizer *vɛ*.

(12)
(*Fe)	mae	Sam	yn
(*vɛ)	maɪ	Sam	ən
C_{PDMC}	POS.be.PRS	Sam	PROG

pobi bisgedi.
pɔbi bɪsgɐd-i.
bake.NF biscuit-PL
'Sam is baking biscuits.'

(13)
(*Fe)	rydw	i	'n
(*vɛ)	r-əd-u	i	=n
C_{PDMC}	POS-be.PRS-1SG	1SG	=PROG

pobi bisgedi.
pɔbi bɪsgɐd-i.
bake.NF biscuit-PL
'I'm baking biscuits.'

- *vɛ* is fine with other verbs or other forms of the copula (those 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.'

(15)
Fe	fydda	i	'n
vɛ	vəða	i	=n
C_{PDMC}	be.FUT.1SG	1SG	=PROG

pobi bisgedi heno.
pɔbi bisgɐd-i henɔ.
bake.NF biscuit-PL tonight
'I will be baking biscuits tonight.'

- But it is not possible to say *vɛ əd-i* (i.e. to combine the positive declarative main clause complementizer with the neutral form of the copula):

(16)
* Fe	ydw	i	'n	pobi	bisgedi	heno.
vɛ	əd-w	i	=n	pɔbi	bisgɐd-i	henɔ.
C_{PDMC}	be.PRS-1SG	1SG	=PROG	bake.NF	biscuit-PL	tonight

Intended: 'I am baking biscuits tonight.'

- Note 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. There is no blocking of this complementizer by positive forms of the copula:

(17)

Dw	i	'n	meddwl	taw	$[Mari]_{FOC}$	rwyt	ti	wedi alw.
du	iː	=n	mɛðul	**taʊ**	$[mari]_{FOC}$	r-ʊɪt	tiː	wɛdi alu
be.PRS.1SG	1SG	PROG	think.NF	$\mathbf{C_{SUB.FOC}}$	$[Mari]_{FOC}$	POS-be.FUT	PRED	next

'I think that you called *Mari*.'

- One possibility 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.
- No equivalent data for the negative form. [1]

3 Analysis

3.1 Generalizations

- The negative form is generally allowed - responsives are the only context in which they are blocked.[2]
- The positive forms have a more complex distribution.
  - Sometimes, there seems to be a restriction against positive forms of the copula occurring in clauses of particular types. This is true of responsives and polar interrogatives.
  - Sometimes, there is a problem with positive forms occurring adjacent to a complementizer. 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.
  - The latter kind of restriction also 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*.

[1] There is no negative main clause complementizer so there are no equivalent interactions with the negative form of the copula in main clauses. The subordinate clause negative complementizer is *na(d)*, with final *d* surfacing before vowel-initial forms. Thus, the segmentations *na d-əd-i* and *nad əd-i* are equally possible.

[2] ...that I've mentioned in the main text —they are blocked when subjects are fronted too, which is also true for positive forms.

3.2 Main Vocabulary Items

- Luckily we can capture everything we need to with just a few super simple VIs! \s
- Neutral forms are sequences of a t node, expressing tense, and an Agr node, expressing subject agreement (there isn't an expression of Agr for indefinite forms):

(18) copula stem:
$\langle$ [t], $\quad$ (↑ TENSE) = PRES $\quad\rangle \xrightarrow{\nu}$ *əd*
$\langle\langle$ (↑ SUBJ DEF) $=_c \oplus \rangle\rangle$

(19) 1SG ending:
$\langle$ [Agr], (↑ SUBJ INDEX) = 1SG $\rangle \xrightarrow{\nu}$ *-u*
$\langle\langle$ (↑ TENSE) $=_c$ PRES $\rangle\rangle$

(20) '3SG/DEF' ending:
$\langle$ [Agr], $\langle\langle$ (↑ TENSE) $=_c$ PRES $\rangle\rangle$ $\rangle \xrightarrow{\nu}$ *-i*
$\langle\langle$ (↑ SUBJ DEF) $=_c \oplus \rangle\rangle$

(21) 3PL ending:
$\langle$ [Agr], $\quad$ (↑ SUBJ INDEX) = 3PL $\quad\rangle \xrightarrow{\nu}$ *-in*
$\langle\langle$ (↑ SUBJ PRED FN) $=_c$ pro $\rangle\rangle$
$\langle\langle$ (↑ TENSE) $=_c$ PRES $\rangle\rangle$

(22) INDEF copula stem:
$\langle$ [t], (↑ TENSE) = PRES $\rangle \xrightarrow{\nu}$ *ois*
$\langle\langle \neg$(↑ SUBJ DEF) $\rangle\rangle$

- These can be made into negative or positive forms by the addition of a Pol morpheme:

(23) prefix to make negative forms:
$\langle$ [Pol], $\quad\langle\langle$ (↑ NEG) $=_c \oplus \rangle\rangle \quad\rangle \xrightarrow{\nu}$ *d-*
$\langle\langle$ (↑ TENSE) $=_c$ {PRES|IMPF} $\rangle\rangle$

(24) prefix to make positive forms:
$\langle$ [C, Pol], $\quad\langle\langle \neg$ (↑ NEG) $\rangle\rangle \quad\rangle \xrightarrow{\nu}$ *r-*
$\langle\langle$ (↑ TENSE) $=_c$ {PRES|IMPF} $\rangle\rangle$
$\langle\langle$ (↑ FORCE) $\neq$ {RESP | POL-Q} $\rangle\rangle$

- *maɪ*, the suppletive positive 3SG form of the copula, spans C, Pol, t and Agr, and blocks **r-əd-i* and *r-ois* (more on this later).

(25) $\langle$ [C, Pol, t, Agr], $\quad$ (↑ TENSE) = PRES $\quad\rangle \xrightarrow{\nu}$ *maɪ*
$\langle\langle \neg$ (↑ NEG) $\rangle\rangle$
$\langle\langle$ (↑ FORCE) $\neq$ {RESP | POL-Q} $\rangle\rangle$

- Positive forms are restricted to non-negative f-structures (f-structures in which there is no NEG feature). Negative forms are restricted to contexts in which there is some (other) negative element, which has contributed the privatively-valued (↑ NEG) = ⊕ feature to the f-structure.

3.3 Definiteness marking and suppletion

- Remember this?

	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

- The suppletive form *maɪ* replaces hypothetical **r-əd-i and* hypothetical **r-ois*, despite the fact that we have the VIs needed to make **r-əd-i* and *r-ois*.
- Note that the negative, neutral, and hypothetical non-suppletive positive 3SG forms must *constrain* the value of their subjects, not *define* the value of their subjects.
  - Indefinite NPs in Welsh are literally *unmarked*, in terms of morphological exponence. NPs without any determiner in Welsh are usually indefinite (Borsley et al., 2007, 155). If they are accompanied by a definite article or definite possessor, then they are definite:

(26) a. draenog
draɪnɔg
hedgehog
'a hedgehog'

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

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

  - This suggests that unmarked NPs are underspecified, and are interpreted as indefinite unless a definite feature is added by something in the environment.
  - But you cannot combine a definite negative or neutral form of the copula with an unmarked NP to coerce a definite reading:

(27) * Dydy draenog ddim yn cysgu.
d-əd-i draɪnɔg 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.
- Because *maɪ* blocks *r-əd-i* as well as *r-ois*, we can conclude that constraining equations cannot override $\mathbf{MI}_c$. Otherwise, we'd expect *r-əd-i* to block *maɪ*.

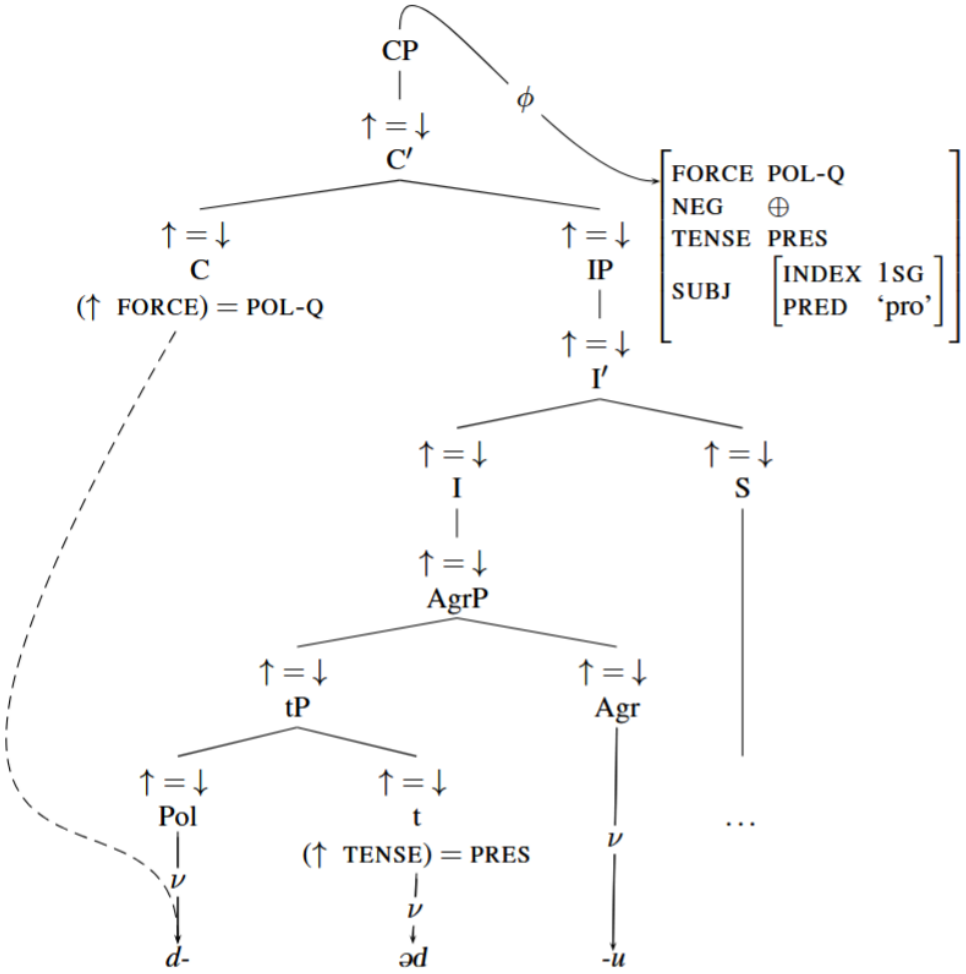
maɪ	**r-əd-i*	**r-ois*
1 VI	3 VIs	2 VIs
C, Pol, t, Agr	C, Pol, t, Agr	C, Pol, t
(↑ TENSE) = PRES	(↑ TENSE) = PRES	(↑ TENSE) = PRES
⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑ NEG) ⟩⟩	⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑ NEG) ⟩⟩ ⟨⟨ (↑ SUBJ DEF) =$_c$ ⊕ ⟩⟩	⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑NEG)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

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

3.4 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.

(28)



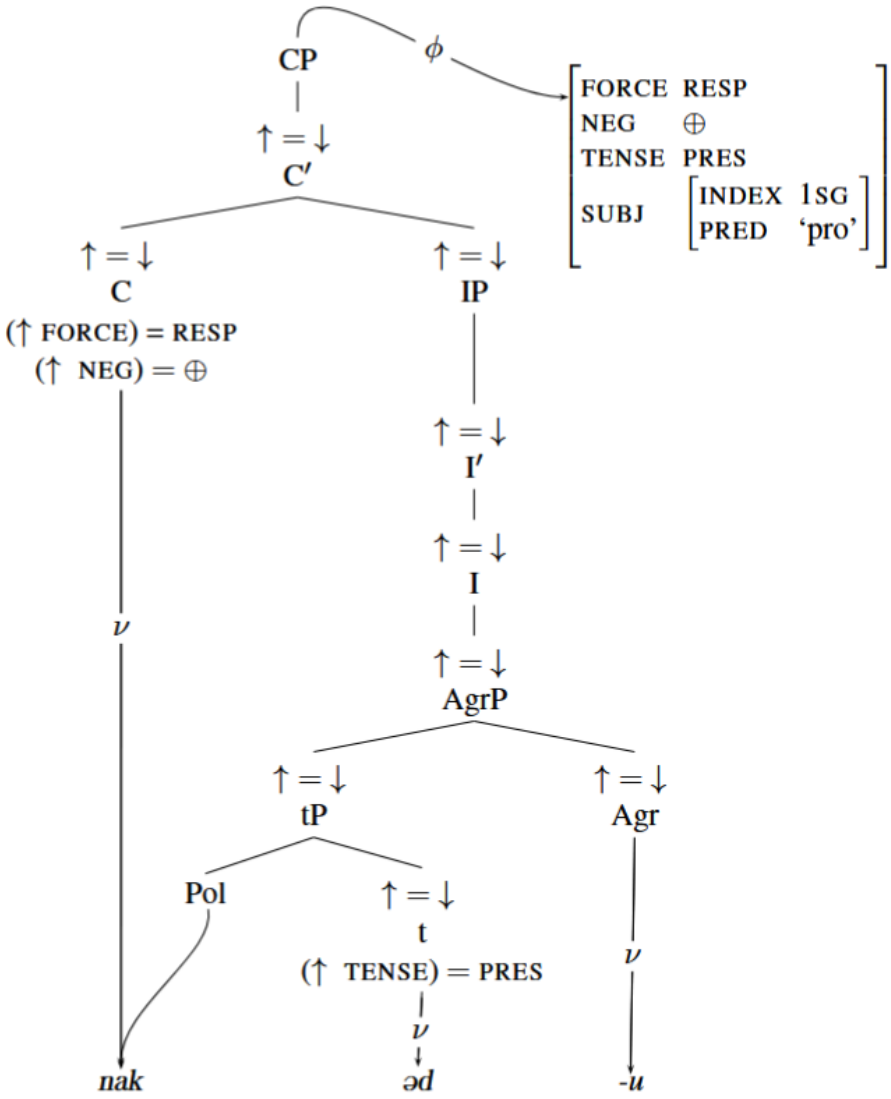
- There is no VI for a C in a polar question in Modern Welsh, so the C head is Pac-man spanned here. The same also happens in main-clause negative declaratives. (In subordinate clauses, the C is spelled out as *na*).
- Negative forms are blocked from responsives not because of a functional restriction but because the negative responsive complementizer spans the Pol head.

(29) $\langle$ [C, Pol], (↑ POL) = – (↑ FORCE) = RESP $\rangle \xrightarrow{\nu}$ *nak*

(30) $\langle$ [Pol], $\langle\langle$ (↑ NEG) $=_c \oplus \rangle\rangle$ $\langle\langle$ (↑ TENSE) $=_c$ {PRES|IMPF} $\rangle\rangle$ $\rangle \xrightarrow{\nu}$ *d-*

- This gives the following structure:

(31)



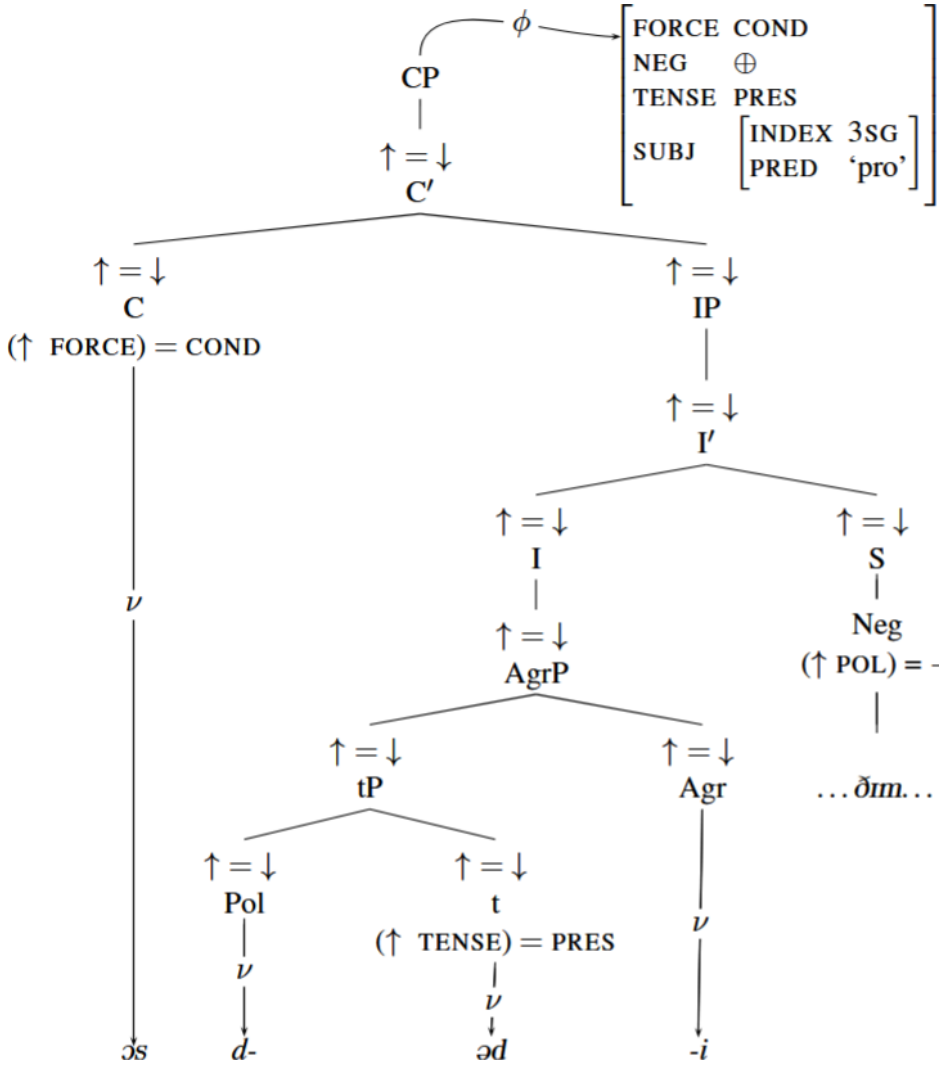
- *nak əd-u* is chosen over *d-əd-u* in this context, because it is both $\mathbf{MI}_f$ and $\mathbf{MI}_c$. The *nak əd-u* sequence realises more nodes (C, Pol, t, Agr vs. Pol, t, Agr), and it expresses more functional information:

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

Table 6: Comparison of functional information in *nak əd-u* vs *d-əd-u*

- In negative if-clauses, *os* is just an ordinary C, (not a C-Pol span like *nak*) and so the negative copula is perfectly fine here.

(32)



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. *taʊ* occurs with fronted material directly precede the fronted material.

(33)
Fe	fydd	y	dyn	yn	nesaf.
vɛ	viːð	ə	diːn	ən	nesav
C_{PDMC}	be.FUT	the	man	PRED	next

'The man will be next.'

(34)
Dw	i	'n	meddwl	taw	[y	dyn]$_{FOC}$	fydd	yn	nesaf.
du	iː	=n	mɛðul	**taʊ**	[ə	diːn]$_{FOC}$	viːð	ən	nesav
be.PRS.1SG	1SG	PROG	think.NF	**$C_{SUB.FOC}$**	[the	man]$_{FOC}$	be.FUT	PRED	next

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

- 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 (no flexibility as to its position).

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

urth-ɔt tiː.
at-2SG 2SG
'If I can come, I will tell you.

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

(35) $$\text{CP} \rightarrow \left(\begin{array}{cc} \text{Foc} & \text{XP} \\ ((\uparrow \text{FORCE}) = \text{COND}) & (\uparrow \text{UDF}) = \downarrow \end{array} \right) \quad \text{C}'$$

(36) $$\text{C}' \rightarrow \text{C} \quad \text{IP}$$

(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) $\rangle\rangle$

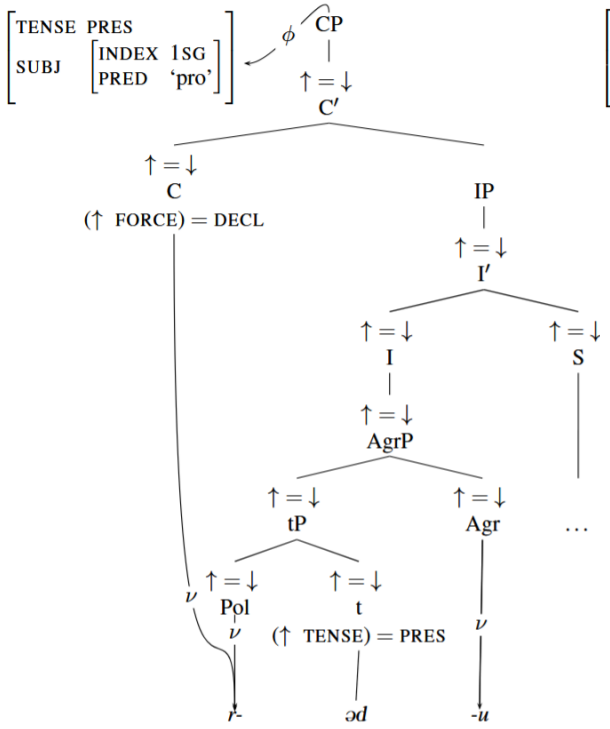
- ɔs has a disjunctive category specification because it can appear in either position:

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

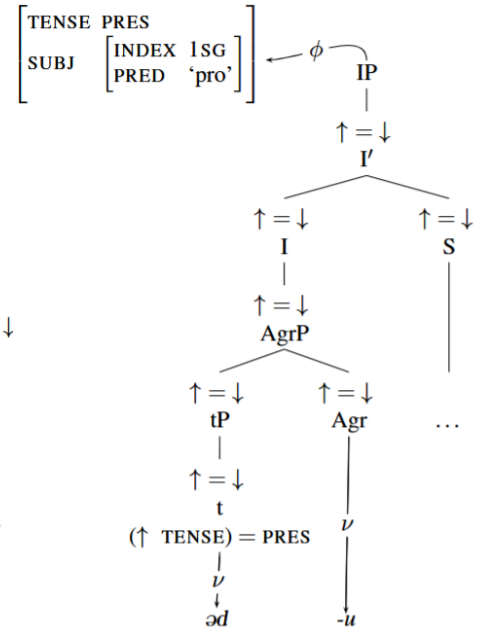
- If we couple this with the assumption that positive forms of the copula span the C node of the tree, we now have an explanation for the blocking of positive forms in if-clauses next to ɔs and in positive declaratives next to to vɛ.

- Compare the kinds of structures spelled out by *r-ɔd-u* with those spelled out by *ɔd-u*:

(40)

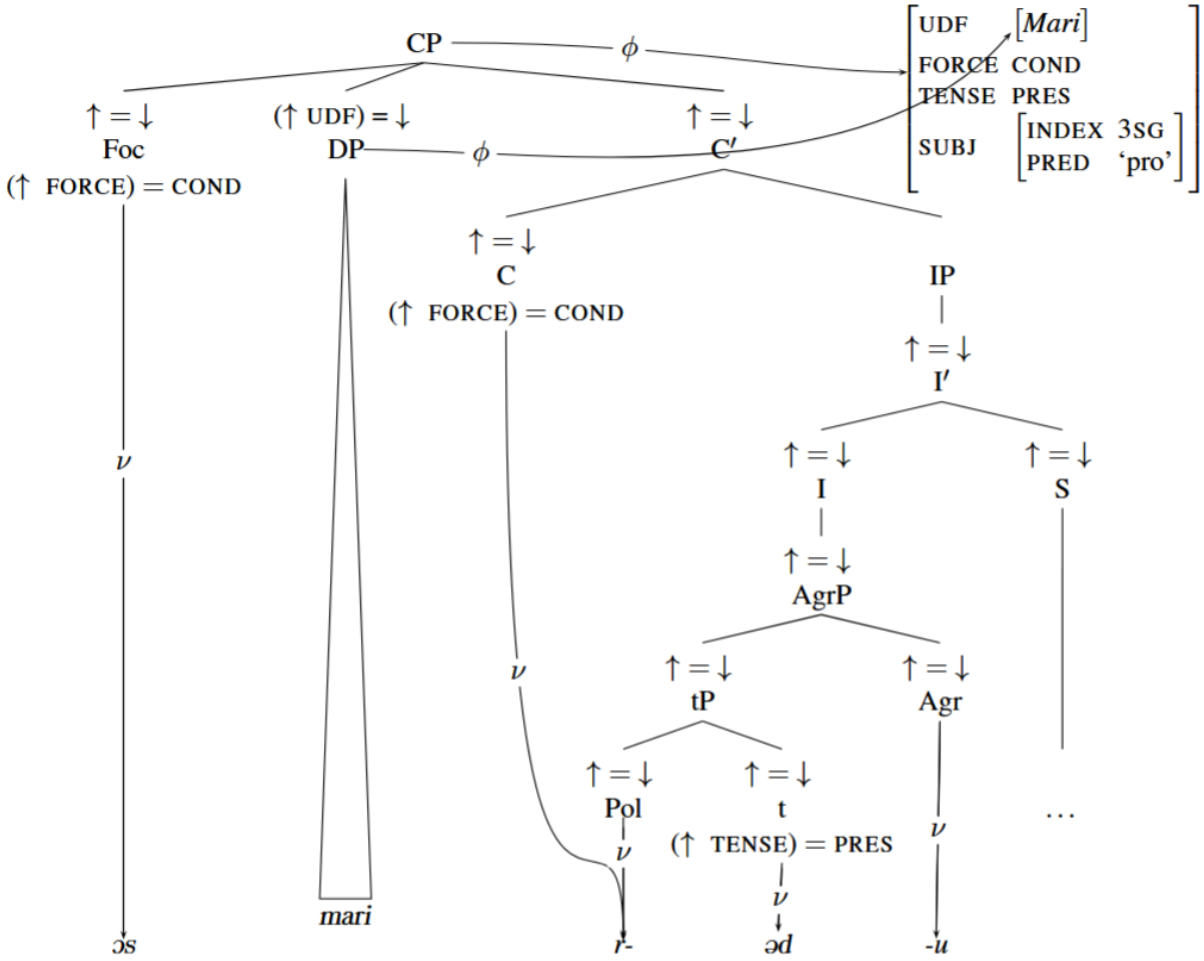


(41)



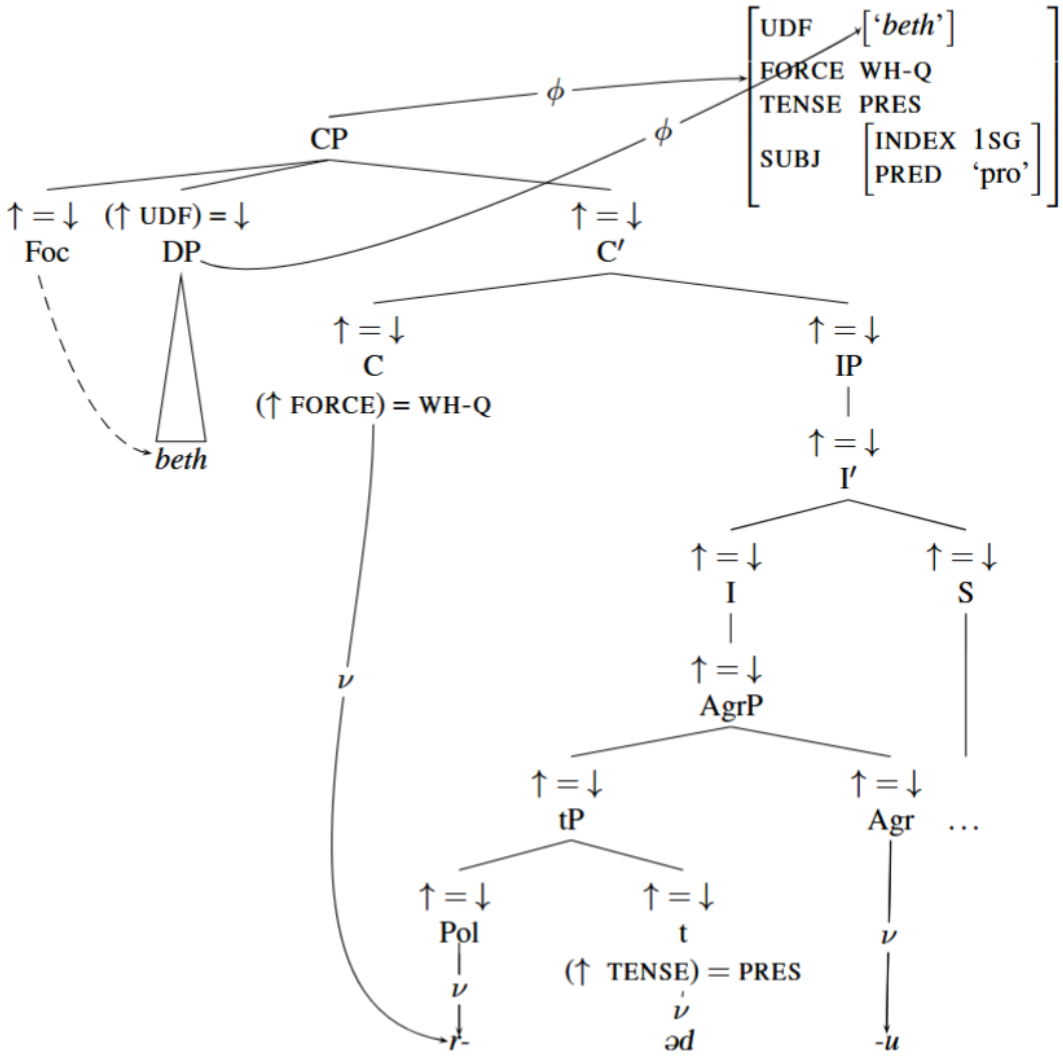
- We now have an explanation as to why the presence of focus-fronting material in 'if'-clauses prevents blocking between *os* and positive forms. *ɔs* is in competition with the positive copula in sentences without fronted material, but not when there is fronted material.
- Here's an if-clause with fronting,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

(42)



- Note that the structure is very similar to a declarative clause with fronting or a constituent question.

(43)

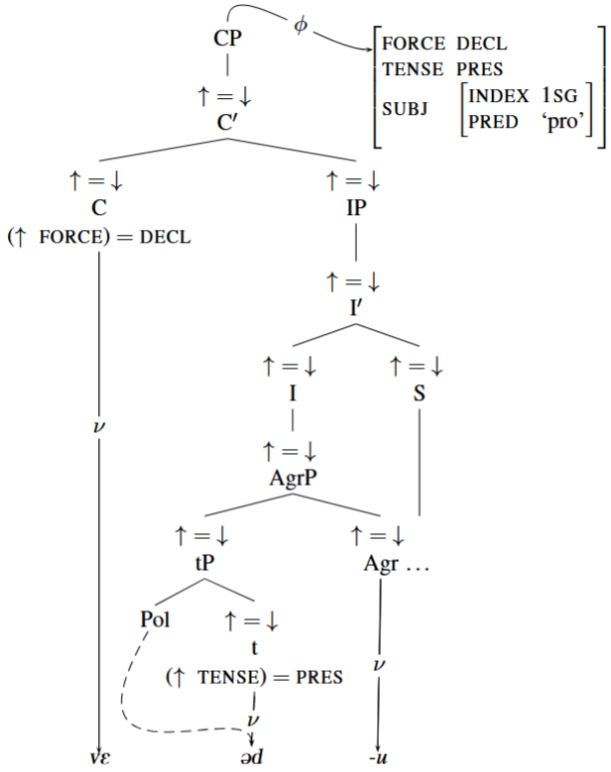


- At this stage, our analysis predicts the ungrammaticality of *os r-əd-u* and *vɛ r-əd-u*. It does not yet explain why the ungrammaticality of these sequences has differnt outcomes in if-clauses as compared to declaratives.

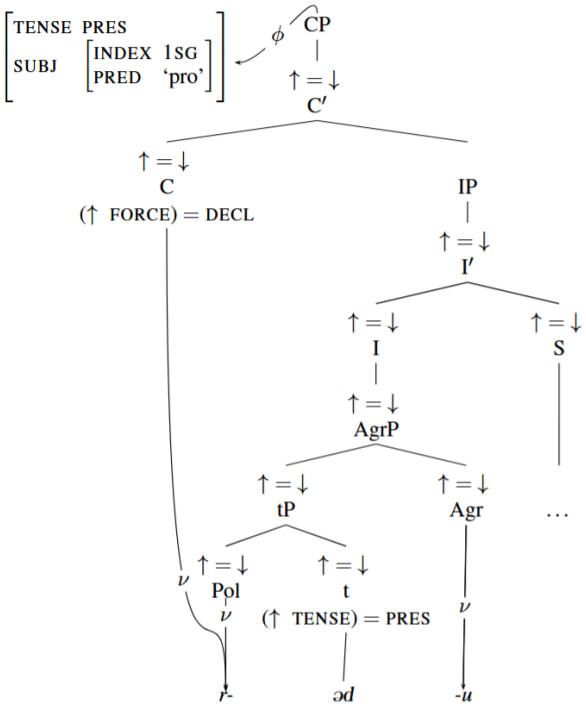
3.6 The difference between if-clauses and declaratives

- Key question: why do positive forms like *r-əd-u* beat *vɛ əd-u* but not *ɔs əd-u*?
- As far as $\mathbf{MI_c}$ is concerned, *r-əd-u* should always beat these sequences as it expresses more c-structure nodes. The sequences involve the unexpressed Pol node being Pac-man spanned. But this only happens in declaratives, not in if-clauses.

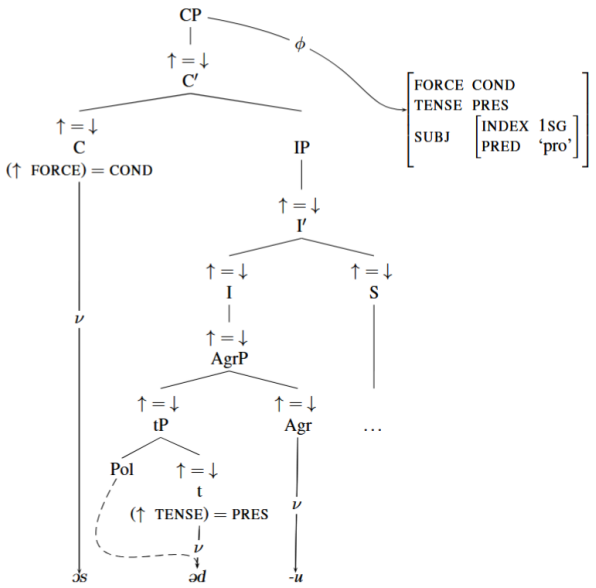
(44) (ungrammatical)



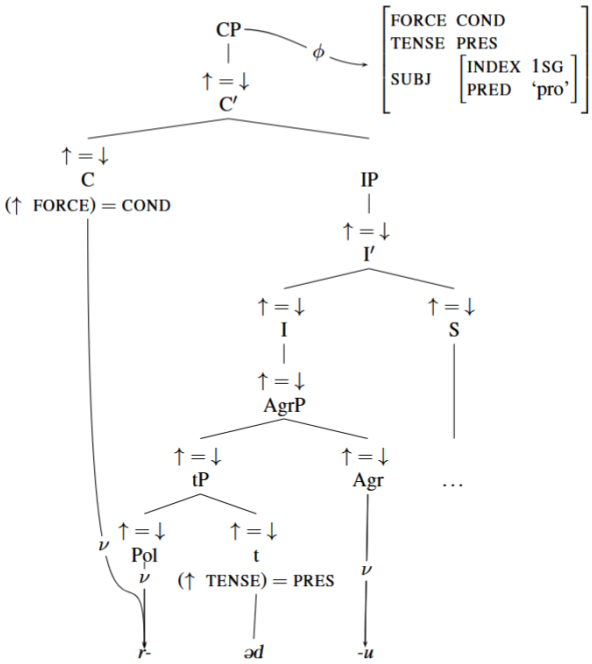
(45) (grammatical)



(46) (grammatical)



(47) (ungrammatical)



- The fact that, in Welsh, the sequence *ɔs əd-u* blocks *maɪ* in conditional clauses, tells us that *something* must be able to $\mathbf{MI}_c$. An obvious candidate is $\mathbf{MI}_f$—in keeping with previous intuitions that $\mathbf{MI}_f$ is more important than phonological specificity (captured by MostSpecific), because "concepts tend to find a way to be expressed" (Asudeh and Siddiqi, 2024).[3]

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

Table 7: Comparison of functional information in *maɪ* vs *ve əd-i* and *ɔs əd-i*

- But 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 (↑ SUBJ INDEX) = 1SG	(↑ TENSE) = PRES (↑ SUBJ INDEX) = 1SG (↑ FORCE) = DECL	(↑ TENSE) = PRES (↑ SUBJ INDEX) = 1SG (↑ FORCE) = COND
⟨⟨ (↑ FORCE) ≠ {RESP | POL-Q} ⟩⟩ ⟨⟨ ¬ (↑ NEG) ⟩⟩	⟨⟨ ¬ (↑ NEG) ⟩⟩ ⟨⟨ ¬ (GF ↑) ⟩⟩ ⟨⟨ ¬ (↑UDF) ⟩⟩	

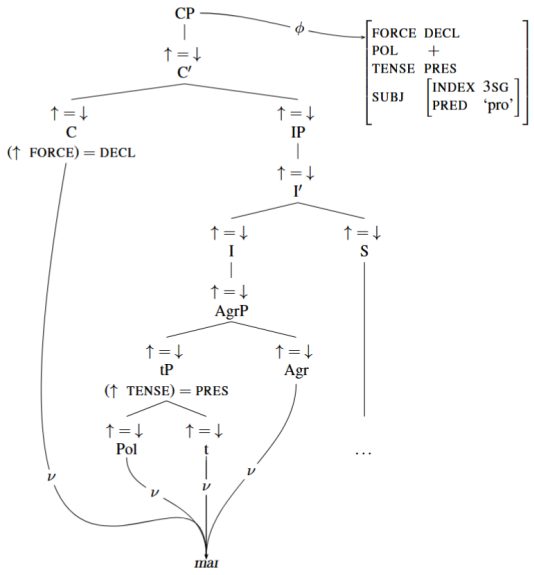
Table 8: Alternative comparison of functional information in *maɪ* vs *ve əd-i* and *ɔs əd-i*

- We have to reject this approach, because it would not be able to predict the selection of *r-əd-u* over **vɛ əd-u* regardless of the ordering of $\mathbf{MI}_c$ or $\mathbf{MI}_f$ (or MI_s?). This approach is also supported by the fact that clauses can be interpreted as declarative by default —*vɛ* is actually optional.
- If $\mathbf{MI}_c$ 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*.
- We've shown that constraining equations can't override $\mathbf{MI}_c$. Now, we've shown that defining equations *have to*.

[3] We might also want to model declaratives as the absence of an overt force feature, in which case, *vɛ* would necessarily contain negative constraints with respect to Force.

4 Summary

- The morphological component in L_RFG has (at least) two mechanisms for comparing forms, $\mathbf{MI}_f$ and $\mathbf{MI}_c$. These are used to select the maximally informative VI possible, to express a sentence's 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 derived by the addition of an *r-* or *d-* prefix to the neutral form.
- The present tense '3SG' copula also shows 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 $\mathbf{MI}_c$.
- 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*. Focus-fronting material allows positive forms to occur in if-clauses, because the focus-fronting material forces *Os* into a higher position.
- In if-clauses without fronting, the sequence *ɔs əd-u* beats *r-əd-u* because it wins on $\mathbf{MI}_f$ grounds - it has more defining equations. Thus, defining equations can override $\mathbf{MI}_c$.
- In positive declaratives, *vɛ əd-u* is blocked by *r-əd-u*, which means that its force specification is a constraining equation —otherwise we could have posited it as a defining equation.
- Constraining equations are not evaluated as part of $\mathbf{MI}_f$, and $\mathbf{MI}_f$ takes precedence over $\mathbf{MI}_c$.



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.
- Just because something is formally possible, this does not mean we have to implement it. Ash thinks it should be a principle of L_RFG that the feature make-up of a grammar should reflect morphological markedness. Thus, where possible, negative constraints condition the absence of morphological marking, not the occurrence of an equally-morphologically-marked (overt) form.
- Note that 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 (↑ 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. The ability to use negative constraints might still be useful for getting *ə*d-*i* to block *ois*.
- What about polarity?
  - In the Welsh data, we have a situation in which positive concord and negative concord are equally capable of being part of the overt expression of a form: thinkg about *nak* (negative responsive marker) and *vɛ* (positive declarative main clause marker). We also see this in the all important r- and d- forms.
  - This could suggest that we want to have a binary feature for polarity, based on Ash's principle, even though it is not strictly *necessary* to capture the data (we didn't use a binary valued feature here!). Ash's principle guides us to adopting a binary feature system for polarity based on morphological markedness... or does it?
  - Taking each context separately, we actually never have a minimal pair in which positive and negative marking are equally morphologically marked. Negation is always more marked:
    - Context 1: no copula of right tense in the sentence sentence:
      - positive polarity: unmarked
      - negative polarity: marked by *ðɪm* on its own. [4]
    - Context 2: relevant copula form present, no blocking of r or d based on c-sturcture:
      - positive polarity: marked by *r*- on its own
      - negative polarity: marked by *d*- and *ðɪm*
    - Context 3: *r*- or *d*- is blocked, either because the Pol node is blocked, or, for postiive forms only, because the C node is blocked:
      - positive polarity: unmarked
      - negative polarity: still marked by some overt negator e.g. *nak*

[4] or some other negator

- As an addendum here, we propose to notationally distinguish a positively-valued (i.e. present) privative feature from a positively-valued binary feature.
  - At f-structure, a feature must have a value if it is present: this is a formal requirement of AVMs. This means privative features are values that have only one possible value, not features that are simply present or absent without any value.
  - There is an ambiguity in using (↑ FEAT) = + for both privative and binary features.
  - Thus, we propose to use + only for binary features, meaning that the use of + implies the possibility of the feature being valued as –.
  - For privative features, we use $\oplus$; there is no $\ominus$ equivalent, because privative values are either present with the value $\oplus$ or absent entirely.

References

Asudeh, Ash and Siddiqi, Dan. 2024. The story of *er*. In Miriam Butt Butt, Jamie Y. Findlay and Ida Toivonen (eds.), *Proceedings of the LFG'24 Conference*, pages 69–89, Publikon.

Borsley, Robert D, Tallerman, Maggie and Willis, David W. E. 2007. *The Syntax of Welsh*. Cambridge: Cambridge University Press.

Dowle, Frances. 2022. A new analysis of Welsh agreement. In Miriam Butt Butt, Jamie Y. Findlay and Ida Toivonen (eds.), *Proceedings of the LFG'22 Conference*, pages 124–143, Publikon.

Dowle, Frances. forthcoming. Neutral Forms of *be* as Default Forms: The utility of underspecification and blocking in a Welsh morphosyntactic phenomenon, accepted by the Transactions of the Philogocial Society.

Kiparsky, Paul. 1993. Paninian Linguistics. In Ronald Asher and Seumas Simpson (eds.), *Encyclopedia of Languages and Linguistics*, Oxford: Pergamon.