Universiteit
Leiden
The Netherlands

Why Jesus and Job spoke bad Welsh : the origin and distribution of V2 orders in Middle Welsh

Meelen, M.

Citation
Meelen, M. (2016, June 21). *Why Jesus and Job spoke bad Welsh : the origin and distribution of V2 orders in Middle Welsh*. *LOT dissertation series*. LOT, Utrecht. Retrieved from https://hdl.handle.net/1887/40632

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/40632

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/40632 holds various files of this Leiden University dissertation.

Author: Meelen, M.
Title: Why Jesus and Job spoke bad Welsh : the origin and distribution of V2 orders in Middle Welsh
Issue Date: 2016-06-21

CHAPTER 7

Diachronic syntactic change

> *"As there are sophisticated methods for its reconstruction, the common ancestor language of Welsh, Cornish, and Breton is so accessible that with a bit of practice we would be able to strike up a conversation with a second-century British Celt in his native language and explain to him how his language had changed - quite dramatically as a matter of fact - by the end of the sixth century."*
>
> (Schrijver, 2014:30-31)

7.1 Introduction

Schrijver's above-sketched optimistic scenario is based on the success of the Comparative Method reconstructing the sounds and words of older stages of languages we no longer have direct access to. Following in the footsteps of Sir William Jones (1746-1794), a philologist and judge of Welsh descent (see Silk (2014)), this method led to a number of late nineteenth-century breakthroughs by Neogrammarians like Hermann Osthoff and Karl Brugmann (see their famous manifesto in the preface to volume 1 of their *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*, Osthoff and Brugmann (1878)). The field of historical linguistics was so influential that it inspired Darwin with the early discoveries leading up to his theory of evolution (see Alter (1999)). Schrijver therefore argues that "every educated human being should be aware of the method", however, "hardly anyone actually is" (Schrijver, 2014:6).

Schrijver then goes on to explain this comparative method, first by providing a step-by-step example, then by reconstructing much of the phonology, morphology and lexicon of Proto-British, the language of his second-century British Celt. Syntax,

however, is not discussed '[f]or practical reasons' (Schrijver, 2014:1). These reasons actually go beyond the 'time frame' and 'nature of the available source material' (or the lack of material altogether). They even go beyond linguistic expertise or preference, as we can read in a study entirely dedicated to diachronic syntax: 'one can no more reconstruct the syntax of a proto-language than one can reconstruct last week's weather[1]' (Lightfoot, 2002:135). It seems then, that we can only converse with our British Celt in a language that, by necessary assumption, has the same syntax as that of his descendants, who did write things down in their vernacular, be it Breton, Welsh or Cornish. However, although these three were closely related, it matters which one we use as an exemplar for our Proto-British grammar.

To illustrate this let us examine one example again that was frequently used in Middle Welsh, the Abnormal Sentence. The agreement patterns observed in Middle Welsh are hardly ever found in Middle Breton: both pronouns and plural nouns yield default third-person singular inflection on the verb in Breton. The question is therefore when speaking to our British Celt, should we Agree or not Agree? If we want to make sure we are not making any syntactic 'errors', we might be better off greeting him in Late Latin, a language conveniently found in various sources around that period. Depending on our Celt's social status and, arguably, his place of birth, there is a good chance he was perfectly capable of speaking or at least understanding this language of the Roman invaders.

Once we are done with formalities and chit-chat about the reconstructed weather, we would like to get down to business and tell this British Celt all about the drastic changes his language will undergo in the next few centuries. Not just the sounds, but also the order of "the magical letters S, V and O" (C. Watkins, 1976:305) will be changed. We could answer the puzzled look on his face reassuring him that those magical new word orders will not last much more than a thousand years. This might be an adequate answer to his first question ('When?'), but can we give any satisfactory explanation as to *how* and *why* it changed so dramatically?

This chapter aims to shed more light on the 'how' question regarding some major syntactic changes in the Middle Welsh period. In section 7.2 I first discuss the main mechanisms of diachronic syntax and which specific challenges it presents from an empirical, theoretical, and - depending on our definition of syntax - ontological perspective. A reconstruction of Proto-British syntax goes beyond the scope of the present study, but I will emphasise the importance of comparative studies illustrated with some examples from Middle Breton. If we want to gain a better understanding of the syntactic history of the Welsh language, a solid methodology for both historical syntax and syntactic reconstruction is indispensable. The challenging examples presented in the previous chapter are addressed again in the context of mechanisms of grammaticalisation and reanalysis in section 7.3. Finally, in section 7.4 I take a closer look at the role of information structure in the study of diachronic syntax.

[1] This frequently-cited metaphor has its origin in early work by Jerzy Kuryłowicz on the laws of analogy (Kuryłowicz, 1949), who observed that historical linguists cannot predict when it would rain, even if the presence of gutters predict that it would.

7.2 Approaches to diachronic syntax

If the object of study in syntactic research is limited to the competence of the individual speaker-hearer, their I-language (see Chapter 1 and Chomsky (1965) and Chomsky (1986)), we need to ask ourselves whether 'diachronic syntax' exists at all. Children are extremely successful in acquiring grammar, but as soon as they grow up speaking a language with, for example, SVO word order, they are unlikely to change to VSO at any given stage during their lifetime. Change over time in the internalised grammar of individuals is generally very restricted[2] (see, amongst others, Clahsen (1991)).

But syntactic innovations have been found in historical documents in various periods of time. Modern French, Romanian and Italian word order, for example, differ from Latin, just as Modern Greek differs from the language spoken by Plato and Socrates. So how do we account for that? Arguably the easiest way out of this apparent paradox is to challenge the premise: maybe I-language is not the right object of study in diachronic syntax? After all, we can only ever study I-language through spoken or written sources of E-language and even then it remains difficult to be sure that we are in fact dealing with the 'real' I-language.

When comparing Middle Welsh texts to Modern Welsh texts, we can indeed observe that the basic word order has changed. If we want to account for either the Middle Welsh or Modern Welsh sentence structures we observe to reach descriptive adequacy, we need to go beyond mere observations. Adequate generalisations can, however, only be made abstracting away from the observed examples in a systematic way. Therefore even to answer the question of how a sentence/construction/word order pattern is derived synchronically, we need a certain level of abstraction and thus a syntactic framework that gives us tools and methods of analysis. Explanatory adequacy then goes even further in addressing the 'why' question: why do we find pattern X (and not pattern Y) or - in our diachronic scenario - why does pattern Y replace pattern X?

In the first part of this section I briefly discuss some approaches to diachronic syntax that have been used to explain various phenomena in historical Welsh syntax. Examples include (Cognitive) Construction Grammar and the loss of V2 in Early Modern Welsh, contact-induced change by language shift in Early Brythonic and generative acquisition-based models of change. Some of these overlap and/or share specific mechanisms proposed to account for syntactic change. Before moving on to the diachronic analysis of the Welsh data presented in the previous chapter in a generative framework, I discuss these mechanisms and the most important challenges in the study of historical syntax.

[2] 'Unlikely' in the previous sentence refers to that fact that such rigorous word order changes are not observed by researchers studying language change. This lack of evidence does not exclude the possibility of such changes occurring in individual grammars. Although Crisma and Longobardi (2009) assert that 'within an I-language, there seems to be no such a thing as change' (Crisma & Longobardi, 2009:4), there are certain subtle changes, mainly in frequencies rather than in the emergence of new structures (see also Walkden (2014:35n20)).

7.2.1 Diachronic Construction Grammar

Within Construction Grammar (CxG) the main focus on diachronic syntax is centred around how constructions (form-meaning configurations larger than morphemes and words) change over time. The interaction of frequency and constructionalisation has played an important role, as well as how constructions develop to become more lexicalised or more schematised (Barðdal, Smirnova, Sommerer, & Gildea, 2015:20). Explanatory adequacy within Cognitive Construction Grammar is achieved through the concept of *motivation*: each construction must be motivated by principles of grammaticalisation, discourse demands, iconic or general principles or appeal to constraints on acquisition (Goldberg, 2006:17). Goldberg (1995) formulates the Principle of Maximized Motivation as follows: "If construction A is related to construction B syntactically, then the system of construction A is motivated to the degree that it is related to construction B semantically ... Such motivation is maximized." (Goldberg, 1995:67).

Currie (2013) employs this principle of motivation in his study on the loss of V2 in Early Modern Welsh. According to Currie, adverb-initial word orders 'motivated' verb-initial word orders, because of the perceived parallelism between sentences with clause-initial adverbs and those without: "[t]he basis for this motivational relationship is the formal similarity between the respective pairs of constructions and the fact that clause-initial adverbial phrase could be analysed as a clause connector, separate from the verbal phrase, so that the following construction - XP + verb or verb - could be perceived as clause-initial" (Currie, 2013:67).

The concept of 'motivation' is criticised in other corners of the field of Construction Grammar. Within Unification Construction Grammar (a non-usage-based version of CxG focussing on unification-based formalism, see Kay and Fillmore (1999)) the concept of 'motivation' is discarded, because it fails to make any testable predictions. According to Goldberg (2006), however, this is a misinterpretation of the concept of 'motivation'. She argues that "[w]hile motivation is distinct from prediction insofar as a motivated construction *could have been otherwise*, it typically could not have had the opposite values of the properties claimed to provide motivation" (Goldberg, 2006:219).

At first glance, however, it seems unclear what this means in the case of the loss of V2 in Welsh, because the two available structures (adverb-initial and verb-initial) are claimed to motivate *each other* (Currie, 2013:67). According to Currie, the lack of verb-initial orders in Middle Welsh is due to the lack of Adverb + Verb orders in that same period. The 'prediction' in this sense must therefore be that because of this correlation of mutually motivating word order patters, verb-initial orders would not develop if Adverb + Verb orders had not increased in frequency. To the extent this makes any predictions concerning the change of word order from V2 to verb-initial in Early Modern Welsh, we are still left with what Roberts (2007) calls a 'Chicken-and-Egg' problem of syntactic reanalysis: which is the cause and which the effect of change?

According to Willis (1998), the loss of preverbal particles *a* and *y* was a crucial factor in the loss of V2. It not only led to an environment in which sentence-initial

subject-pronouns could be reanalysed as clitics (see also Willis (2007a)), but it furthermore led to an increase in Adverb + verb orders (since, as we have seen in Chapters 4 and 5, Adverb + *y* + Verb was by far the most frequently-found word order pattern towards the end of the Middle Welsh period). Currie, however, states that "we cannot say the decline in the use of *y* necessarily caused the increase in use of AIV [absolute verb-initial - MM] order" (Currie, 2013:67). Other factors, such as synchronic variation in word order patterns in Early Modern Welsh and the importance of the Welsh Bible translations were just as much part of the 'motivation' for the change from V2 to VSO (Currie, 2013:71).

This variation, according to Currie, does not correlate with any socio-linguistic factors (e.g. class, dialect, register or genre): "the main parameter of variation appears to be stylistic choice by individual writers" (Currie, 2013:69). This then explains the 'gradual' pattern of the loss of V2 and should thus serve as an argument against Willis's parametric approach since the change took centuries to fully complete (see Willis (1998), but also the discussion on 'discrete' and 'gradual' change from a generative point of view below). The theoretical 'mechanism' behind this pattern of individual variation is borrowed from the Cognitive Sociolinguistic framework. Within this framework, Coupland defines the concept of *styling* where speakers "can frame the linguistic resources available to them in creative ways, making new meanings from old meanings" (Coupland, 2007:84) (as cited by Currie (2013:69-70)). Some Early Modern Welsh authors chose to use more verb-initial sentences in prose, because these verb-initial orders already frequently occurred in poetry and in the first Welsh Bible translations (and they wanted to imitate this elevated poetic style); others did not.

It should be noted, however, that Currie's (2013) conclusions regarding the high frequency of verb-initial orders in various excerpts of the Bible translation are slightly misleading, because he is conflating different types of Biblical genres. Crucially, this high(er) number of verb-initial orders is found in the *Book of Isaiah* (41.0% verb-initial order according to his Table 1) and the *Psalms* (24.8%), neither of which contain the narrative prose found in, for example, the *Book of Esther* (with only 9.4% verb-initial orders) or the *Gospel of Mark* (6.5%). According to both the Christian as well as the Judaic tradition, the *Psalms* belong to the Poetic texts of the Bible along with, for example, *Job* and *Proverbs* (see, amongst others, Vriezen and Van der Woude (2000:96)).[3]

In other diachronic studies within Construction Grammar, usage-based motivation is often specified from a structural, referential, semantic, discourse-pragmatic and/or contextual point of view. This then, in combination with the relative frequencies of various constructions, aims to give a comprehensive explanation of the particular syntactic change under investigation (see Fried (2009) on the development of the subjective epistemic particle *jestli* '[in-my-opinion-]maybe' in conversational Czech and the rise of the dative substitution in Icelandic by Barðdal

[3] See furthermore Watson (1973:2) and Green (2005:60) for the poetic nature of the language of the *Book of Isaiah*. Since verb-initial orders were already (more) frequently found in poetry, this distribution is not at all surprising.

(2011)). In Currie's study of the loss of V2 in Welsh, however, many questions remain. For example, to what extent did the authors' choice to imitate poetic style reflect their daily speech, if at all? Why did they choose to imitate Biblical poetry, rather than Biblical prose (which was still subject-initial V2)? Furthermore, if using verb-initial orders was indeed a stylistic literary choice of some authors, how and why did VSO become the prevalent word order in Modern Welsh?

Overall, it is not only intuitively attractive, but arguably also necessary to look for 'motivations' of syntactic change beyond the structural domain. To the extent it is possible working with limited historical data, evidence from semantic, information-structural and sociolinguistic variation should definitely be taken into account. These factors are built into usage-based Construction Grammar. In theory then this seems a reasonable approach to problems in diachronic syntax. In practice, however, looking at Currie's (2013) account of the loss of V2 in Welsh, many questions remain unanswered and it is not clear why - if at all - this approach achieves more 'explanatory adequacy' than, for example, the arguments originally put forward by Willis (1998) in a generative framework (and Willis (2007a) or, in 'flexible syntax' by Bury (2002).[4]

7.2.2 Sociolinguistic variation and language contact

One of the important factors in diachronic syntax also touched upon in the previous section is 'variation'. The source of variation can lie in sociolinguistic factors, but also in (combination with) situations of language contact. There are several approaches to language change that focus on characterising the exact nature of variation. After all, "[i]t is speakers and not languages that innovate" (Milroy, 1992:169). In what is arguably the most influential study of sociolinguistic variation (Weinreich et al., 1968), language is a form of 'orderly heterogeneity' (see also Nevalainen and Raumolin-Brunberg (2003:12)). Rissanen (2008:56) groups the most important extralinguistic factors that affect the choices of variants in the following way:

1. Sociolinguistic ⇒speaker's/writer's social status, education and the relationship between discourse participants
2. Textual ⇒genre, topic or purpose of text, discourse situation and medium
3. Regional ⇒language contact

He notes that many of these extralinguistic factors overlap. Research into variation and change thus necessarily needs to take a combination of these factors into account as well as "internal processes of change" (like, for instance, grammaticalisation or analogical levelling discussed below) (Rissanen, 2008:57). A balanced corpus with extensive metadata on the origin and philological background of the texts is indispensable in this type of approach.

[4]For more on Construction Grammar and explanatory adequacy, see the series of papers discussing this problem by Adele Goldberg and David Adger in Goldberg (2006) and Adger (2013a) *et seq.*.

Language contact in a historical context is a particularly difficult field of study. Labov presents the 'Principle of Contingency' according to which specific instances of change require specific (rather than universal) explanations (Labov (2001:503) and also Walkden (2014:46) for discussion). Contact can lead to change, but - surprisingly - also to continuity in grammars. Bilingualism and the ability of children to acquire more than one language perfectly if they learn both from a young age, can play a role in this. This is shown by studies of a corpus of Welsh-English bilingual speech in which only one possible instance of convergence (i.e. contact-induced transfer) was found (modifier-head order within noun phrases) (P. Davies & Deuchar, 2010). Although there is a large amount of bilingualism in North Wales (and there has been for a long time), P. Davies and Deuchar (2010) conclude that Welsh grammar - in particular the noun phrase under investigation - exhibits continuity rather than change.

Whenever there *is* contact-induced change, it appears to come in different forms. Thomason and Kaufman (1988:50) present a 'scale of interference' according to which the extent and type of contact determine the type of change from lexical borrowing with minimal contact to structural changes with intensive long-term contact. Winford (2005) characterises this distinction as recipient or source language agentivity. In the case of recipient language agentivity transfer of linguistic material typically includes the borrowing of open class vocabulary items and it is likely to lead to complexification of the recipient language. Cases of source language agentivity, on the other hand, are called 'imposition'. Here the transfer mainly consists of phonological and syntactic features.

In the following section, I describe two cases of language contact and syntactic change in the history of Welsh. First I discuss the proposal of language shift (resulting in 'imposition') in British Celtic put forward by Schrijver (2002; 2007; 2014). Then I briefly discuss proposed cases of Latin influence on Welsh grammar in a later stage (due to literary translations and/or adaptations from Latin originals).

Language shift in early Britain

Schrijver (2002) (and also Schrijver (2007) and Schrijver (2014)) sketches a scenario of language contact, in particular language shift in the history of the Brythonic languages to account for various morpho-syntactic phenomena found in the British Celtic languages (but, crucially, not in Irish). According to Bede's description of Britain (written in the first half of the eighth century), there were five languages present at the time: English, British, Irish, Pictish and Latin. In the centuries after the collapse of the Roman empire, there is evidence (in the form of inscriptions) for three of these in Wales: British, Irish and Latin (see, amongst others, Sims-Williams (2003), Falileyev (2003) and Russell (2012)). The extent to which each of these three was spoken and in daily use is a matter of ongoing debate (cf. Adams (2007), T. M. Charles-Edwards (2013) and Schrijver (2014)), but it is clear that what distinguishes Brythonic languages from Irish is the loss of final syllables and the case system. After 'the departure of Rome', both Latin and Brythonic were spoken and there was probably a high degree of bilingualism

(Russell, 2012:216-218).

The scenario outlined by Schrijver (2002) involves a split between speakers of Celtic in the lowlands and the highlands. Highland British Celtic is argued to be the predecessor of Welsh, Breton and Cornish in the west, whereas Lowland British Celtic (with a more Irish-like phonological system) and Late British Latin influenced the sound system of the Anglo-Saxon invaders in the southeast. During the Roman period, Latin was a superstrate language and as such it donated many lexical items to Brythonic. After the collapse of the empire, however, the situation was reversed rendering Brythonic a superstrate language as opposed to speakers of Latin who then became of lower status. Based on the language contact theory of Thomason and Kaufman (1988), Schrijver (2002) proposes that the observed Latinised morpho-syntactic features in Brythonic are the result of language shift. Speakers of the then substrate-language Late Latin moved to the 'Highland Zone' and rapidly shifted to speaking Brythonic, keeping a Latin accent (and Latin-like morpho-syntax), but avoiding Latin vocabulary (Schrijver, 2014:32).

According to Russell (2012:220-221), there are various geographical and sociolinguistic problems with this scenario. Here I focus on the proposed morpho-syntactic influence from Latin transferred by language shift. The mentioned features include the loss of neuter gender and the case system and the development of the pluperfect in Brythonic languages. The first two are equally problematic, according to Russell (2012). Loss of neuter gender, first of all, also happened in Irish, so this is not necessarily a feature of the grammar of Brythonic languages only (it might have been on its way out in Celtic in general) (Russell, 2012:222). As for the loss of the case system, the nominative and the genitive arguably survived the longest in Brythonic. In Old French, however, the nominative and accusative are both still attested. If British Latin "shared north-western Romance features with the Latin of northern Gaul" (Russell, 2012:222) as Schrijver (2002) suggests, this is a problem. The reconstructed paradigms of Late Spoken British Latin in Schrijver (2014:46-47), however, show that for all five declensions, the genitive survived alongside the collapsed/combined nominative-accusative (or, in the fifth declension type *homō*, the nominative-vocative **omō* was distinguished from the accusative-genitive **om(ɪ)nɪ* and dative **om(ɪ)ni*). If Schrijver's (2014) reconstructions of the Late British Latin nominal paradigms are correct, the loss of the case system in British Celtic indeed followed a parallel development with Late British Latin. This pattern was unlike that found in Old Irish, in which five distinctive cases survived (Thurneysen, 2003 [1946]).

This distinction between Irish and Brythonic languages also exists in the development of the pluperfect in the latter, but not in the former branch of Celtic. MacCana (1976) first proposed that this new paradigm observed in Welsh, Breton and Cornish was influenced by Latin. Russell (2012:223) argues, however, that it is unlikely that the periphrastic origin of the form *amauerat* 'had loved' was still discernible in British Latin, since its pronunciation had developed to /a'ma:rat/. If the periphrastic form *amauerat* still existed on a high literary level, it probably had little impact on spoken British Celtic. Even if it had existed, it could hardly serve as

a model for Brythonic *carassei* 'had loved', because this cannot be decomposed as a form of the preterite + the imperfect of the verb 'to be' (Russell, 2012:223).

Overall, the presented scenario involving language shift with speakers retaining elements of their native Late British Latin grammar is certainly possible. The evidence of syntactic similarities put forward by Schrijver (2002) is in the present state, however, still inconclusive. Out of the three suggested syntactic innovations in Brythonic, Russell (2012) argues only the one about the pluperfect is potentially convincing.

Later Latin influence on Welsh

According to D. S. Evans (2003 [1964]) and D. S. Evans (1971), influence from Latin can also be found in later stages of Brythonic languages, in particular in Welsh translations of Latin texts in the (Early) Middle Welsh period. This type of contact is not language shift by speakers of the substrate Late British Latin, but rather textual influence on a literary level. Examples of these literary Latinised features are third-person plural agreement with plural nouns (going against the 'Complementarity Principle' discussed in the previous chapter) and the use of the definite article + demonstrative as relative pronouns (e.g. *yr hwnn, yr hynn* 'that, which'). With respect to the plural subject-verb agreement, Schumacher (2011) points out that this is the only possible pattern in Old Welsh prose, regardless of whether the subject preceded (as in (1a)) or followed the verb (as in (1b)):

(1) a. *enuein di sibellae int hinn*
names of Sibyllae be.PRES.3P these
'These are the names of the Sibylls' (*MC*)

b. *imguodant ir degion*
beseech.PAST.3P the nobles
'the nobles besought one another' (*Chad* LL xliii)

Strachan (1909:61) already mentioned that agreement in Old and Middle Welsh shows 'certain peculiarities'. Just like in Middle Welsh prose, he argues "[i]n the earlier poetry the plural is quite common, and in corresponding constructions in Old Irish the plural is regular. In Welsh there has been an encroachment of the singular upon the plural, as there has been in later Irish." (Strachan, 1909:62). Koch (1991) notes that default third-person singular agreement must have been well established in Old Welsh, giving examples from, among others, the same Old Welsh marginalia in the Lichfield Gospels Schumacher mentioned above (*Chad* LL xliiff):

(2) *imaliti duch cimarguithejt*
lead.3S you story-tellers
'as the story-tellers would lead you' (*Chad* 3)

For neuter plural subjects, default third-person singular agreement is not unexpected from an Indo-European point of view. Examples of this are found in Hittite,

Greek and Old Avestan that are argued to go back to old collective nouns (see Beekes (1995:173) and Fortson (2010:132)). For masculine and feminine plurals, plural agreement was found in most Indo-European languages, including Celtic. The question is thus exactly when and how the Complementarity Principle came into being in the history of British Celtic. Koch notes that the third-person plural verbal ending *-nt* (in the old conjunct paradigm of the verb[5]) could have been lost by regular sound change (i.e. apocope in Proto-British) in which case the singular and plural ending of the verb coalesced completely.[6] The formal similarity of the singular and plural conjunct forms could be the base for analogical levelling in the rest of the verbal system. This then would explain the lack of agreement with plural nouns following verbs in Middle Welsh and the lack of agreement altogether in Breton and Cornish (although there too, are exceptions). It still does not explain the agreement with preverbal plural nominal subjects in the Middle Welsh Abnormal Sentence. In section 6.4 below I put these cases in a diachronic and cross-linguistic perspective.

Interim Summary variation and contact

Variation no doubt plays a significant role in language change. Language contact and in particular situations in which speakers of one (substrate) language shift to another (superstrate) language can result in more variation and change in the morpho-syntactic domain as well. There is, however, very little data, both linguistic and socio-historical, from the time of intensive contact between speakers of Brythonic and British Latin in the crucial period after the collapse of the Roman empire. The extent of variation and change caused by language contact is therefore difficult to ascertain. We need a comprehensive description of the syntax of Late British Latin and a sound methodology for reconstructing the syntax of Proto-British. Neither of these are provided by the above-mentioned approaches.

The second example of language contact in a later period (from translating Latin) is of a very different kind. Especially if we have the original text in Latin, grammatical similarities between the two languages are easier to expose. Since contemporary native tales are also available in that period, it would be possible to distinguish phenomena that are typically inherited from Proto-British (or even Celtic) from those borrowed from Latin, like the relative pronoun *yr hwnn*. If those constructions are fully incorporated in the language, we still need tools to adequately describe their formal function within Middle Welsh grammar and how (and why) they changed (again) in Early Modern Welsh.

[5] Insular Celtic had two separate paradigms of verbal endings that can still be found in Old Irish. Traces of the old absolute forms can also be found in Old Welsh, so presumably this system was still found in Brythonic.

[6] According to the 'standard doctrine' (VKG §152 and L&P §88), Proto-British word-final *-nt* survived apocope, but it is not altogether clear why this would be the case, since all final consonants except **-r* disappear in Proto-British. Koch's suggestion is, however, impossible to verify - for now, at least - since, according to Peter Schrijver (p.c.) the third-person plural conjunct verb form is the only (reliably) attested example of word-final **-nt*.

In the following section I present both tools to handle syntactic change as well as a proposed methodology for the reconstruction of syntax based on a generative acquisition-based approach.

7.2.3 Syntactic change in generative grammar

The most foundational study on diachronic syntax within generative grammar is Lightfoot's *Principles of Diachronic Syntax* (Lightfoot, 1979). It transfers mechanisms and insights of the first decades of generative grammar to diachronic syntax and, most importantly, identifies the 'source' or 'starting point' for any syntactic change as language acquisition (see also Paul (1920 [1880]) and Harris and Campbell (1995) for a non-generative approach with the same starting point). Subsequent work in the field (in particular Lightfoot (1991), Lightfoot (1999), Roberts and Roussou (2003), Van Gelderen (2004), Roberts (2007) and Van Gelderen (2011)) is built on the same assumption connecting syntactic change to learnability and acquisition.

In section 7.2 I questioned the usefulness of I-language in the study of diachronic syntax, because the research question in the field typically concerns observations in E-language. Within the Minimalist Program (MP), the syntactic component itself is considered to be invariant, therefore 'syntactic change' as such cannot exist (see the introduction of Biberauer and Walkden (2015) and Walkden (2014:31n14) for discussion and M. Hale (1998) who made the original point). A pure I-language approach to diachronic syntax might not exist (Walkden, 2014:31), but the progress and various breakthroughs in the field (see in particular the annual conferences on *Diachronic Generative Syntax* (DiGS) and the volumes resulting from the conferences, e.g. Biberauer and Walkden (2015)) show that it is worthwhile to keep a notion of I-language and thus a generative approach to syntactic change. This allows us to share the tools and mechanisms of the Minimalist Program analysing how language works and it furthermore gives access to related research in language acquisition.

Not all generative syntactic tools and insights can be straightforwardly applied to diachronic syntactic problems, however. In this section, I discuss some of these challenges and the solutions that have been proposed within the field of diachronic generative syntax. Continuing from the previous section, I start with the notion of variation as a possible source for language change. I then move on to various types of syntactic change such as Reanalysis and Grammaticalisation and how they can be accounted for in an acquisition-based model. Finally, I explore the dynamics of change and the possibilities and limitations of syntactic reconstruction.

Variation in generative grammar

What is syntactic change or language change in general? An instance of 'change' can be defined as a case in which the grammar of a language ('Grammar 2' or G2) that is derived from another language (Grammar 1) differs from this G1. We are thus dealing with variation between two grammars (or two languages or dialects)

over time (a historical relation or 'H-relation', as Crisma and Longobardi (2009:5) call it). One great advantage of the early generative Principles & Parameters approach was the reconciliation of the universal principles solving the Poverty-of-Stimulus problem with the parameters attempting to account for cross-linguistic variation. It specified the relation between the language experience (Primary Linguistic Data or PLD, the input for the language learner) and the innate language faculty of Universal Grammar (UG). To illustrate this: a very crude example of a universal principle could be 'combine the verb with a direct object'. An example of a parameter for a particular language could then be 'let the direct object precede the verb' resulting in languages with linear OV order. An example of syntactic change could be the resetting of that parameter, e.g. OV order changed into VO order (the so-called 'Head Parameter', cf. Travis (1984), Koopman (1984) and Pintzuk (1991), Pintzuk (2002) and Lightfoot (1991) for the diachronic example). Kroch (1989) described this as a situation of grammar competition: a language with parameter-setting 'OV order' (Grammar 1) competes with a later stage of that language in which the parameter switched to 'VO order', resulting in Grammar 2.

Upon closer investigation of the data of these and other proposed parameters, this view of a binary setting that must be switched in a catastrophic fashion turned out to be too simplistic (see also the section on *The dynamics of syntactic change* below). Examples found in the history of English OV and VO word orders suggest for example that this change consisted of various different stages. OV order with quantified and negative objects was lost at a later stage, for example, and, most importantly, the major catastrophic switch from OV to VO seems to have taken centuries to complete (see Pintzuk (2002) for evidence from Old English and Van der Wurff and Foster (1997) for surface OV up until the sixteenth-century). Questions arose on whether certain syntactic changes (always) clustered together and, if so, how and why those changes in particular and not others? Were there non-parametric changes as well and, if so, how can they be characterised and formalised within the system?

Various empirical problems with the traditional parametric approach have been put forward by Newmeyer (2005). In addition, there are specific problems of implementation. It is for example first of all controversial what triggers a certain parameter setting (cf. Dresher (1999) and Lightfoot and Westergaard (2007)): what counts as a cue? The parametric approach furthermore suffers from the Linking Problem (cf. Pinker (1984), but also Beekhuizen et al. (2014) on why this particular problem is so far not solved by any linguistic theory and therefore not just a challenge for parametric theory as described in Chapter 1 of this thesis). Finally, from the point of view of acquisition, parameters need to be learned in the right sequence and there seems to be a growing number of parameters that have to be acquired (cf. Gibson and Wexler (1994), J. D. Fodor (1998) and Evers and Van Kampen (2008)).

According to Newmeyer (2004), the parametric approach of syntactic variation has a further major disadvantage: it lacks what Longobardi (2003) termed 'evolutionary adequacy' (see also Gianollo, Guardiano, and Longobardi (2008)).

This is a new level of empirical adequacy added to the well-known three advocated by Chomsky (1964). Beyond the observational, descriptive and explanatory level, a theory of linguistics should also aim to reach 'evolutionary adequacy', i.e. why did we evolve to have precisely the type of language faculty we have today and why do we have the attested variety of languages (and not others)? Newmeyer (2004) proposes a rule-based grammar instead: variation or "language-particular differences can be captured by difference in language-particular rules" (Newmeyer, 2004:183). A major disadvantage of any rule-based system, however, as Holmberg and Roberts (2005) point out, is that it is unrestrictive in the sense that in principle 'anything goes'. This is typically not what we find in human languages, however (see also Biberauer, Holmberg, and Roberts (2014)).

More recent studies on (parametric) variation within the Minimalist Program have therefore moved the source of variation from 'switchboard-style' parameters in UG to functional features in the lexicon. This was first suggested by Borer (1984) and picked up by Chomsky in early Minimalist work:

(3) **The Borer-Chomsky Conjecture (BCC)** (M. Baker, 2008:353)
All parameters of variation are attributable to the features of particular items (e.g. the functional heads) in the lexicon.

From the point of view of first-language acquisition, this is a real advantage, because it puts the burden of learning (back) to acquiring vocabulary with idiosyncratic properties.[7] According to Walkden (2014:22-23) it furthermore makes more (and clearer) predictions about possible languages than, for example, the rule-based alternative put forward by Newmeyer (2004). According to Roberts and Holmberg (2005), parameters represent points of underspecification and as such are not really primitives of UG. The grammatical system becomes operative once these underspecifications are filled. According to Chomsky (2005) there are three factors in language design. Biberauer, Holmberg, Roberts, and Sheehan (2014) argue that parameters arise as a result of the interaction of the three factors:

1. Factor 1 ⇒ innate endowment (UG): basic operations Merge and Agree (plus a formal feature template [iF]/[uF], and a very small subset of [F]s not derivable from the input)
2. Factor 2 ⇒ primary input experience (PLD) giving evidence for movement, doubling, systematic silence and multifunctionality
3. Factor 3 ⇒ non-language-specific innate capacities: general computational conservatism of the learning device, e.g. Feature Economy (FE) and Input Generalisation (IG)

The third factor is perceived as generally applicable learning biases (Biberauer,

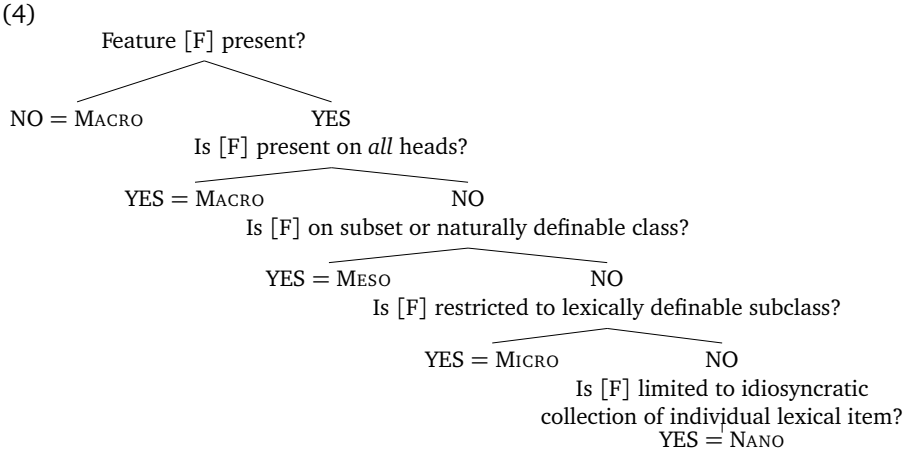
[7] As Walkden (2014) points out, such a lexical approach to variation is somewhat similar to the 'Constructicon' in Construction Grammar discussed in the previous chapter (see also Barðdal and Eythórsson (2012)). The tools and mechanisms in the Minimalist Program, however, differ considerably from those available in Construction Grammar, like the concept of 'motivation' discussed above.

Holmberg, Roberts, & Sheehan, 2014). It consists of a 'minimax' search or optimisation algorithm and is thus wholly in line with Minimalist assumptions making maximal use of minimal means. Feature Economy (FE) is generalised from Roberts and Roussou (2003) as the bias to postulate as few features as possible (i.e. possible to account for the input). Input Generalisation (IG) stipulates that learners maximise the use of the available features (see also Roberts (2007)). This kind of 'emergent parameter' approach is used as a counterargument against Newmeyer's comment on the lack of 'evolutionary adequacy' in a parametric approach to language variation. The underspecification of formal features can appear in three forms (cf. Biberauer, Holmberg, Roberts, and Sheehan (2014:108)):

1. association of formal features with (functional) heads
2. values of formal features, triggering Agree
3. purely diacritic features triggering movement

Clustered syntactic changes can now be thought of in terms of 'cascading parameters' (Biberauer & Roberts, 2008) and networks of parametric changes (Roberts, 2007), or, in line with the third-factor learning biases and the latest output of the project on *Rethinking Comparative Syntax* ('ReCoS') at the University of Cambridge: parameter hierarchies (see Biberauer et al. (2014) and much other work available via the ReCoS project website). The hierarchy consists of different levels, ranging from macroparameters (all (functional) heads share the value v_i of feature [F]), to mesoparameters (all functional heads of a given naturally definable class, e.g. [+V], share v_i), to microparameters (a small subclass of functional heads shows v_i, e.g. pronouns or modal auxiliaries), and finally, nanoparameters (one or more individual lexical items is/are specified for v_i) (Biberauer et al. 2014) and (Ledgeway, 2016):

(4)



Examples of hierarchies and parameters on different levels are given by Roberts

(2012) and Biberauer et al. (2014) (see also the contribution to the volume *Parameter Theory and Linguistic Change* (Galves, Cyrino, Lopes, Sandalo, & Avelar, 2012)). From a diachronic perspective, macroparameters are expected to be highly stable (e.g. rigid head-finality). Null-subjects in the history of Romance are much less stable and would count as a mesoparameter. Microparameters are even more likely to change over time (e.g. English modals). Nanoparameters, finally, could literally come and go with one lexical item (e.g. the relics of the Conditional Inversion in English).

This framework thus gives us very concrete tools to describe and explain variation, either synchronic or diachronic. It leaves the nature of the formal features unspecified, however. This issue is related to a final question concerning variation: does 'free' variation or 'true optionality' in one single grammar exist?

According to Biberauer and Richards (2006), there are indeed cases in which 'the grammar does not mind'. In their study on the EPP feature, they show the *option* of pied-piping of the whole phrase bearing the interpretable φ-feature with examples from auxiliaries in Afrikaans. Since we often have very little or no information about the sociolinguistic situation in earlier stages of languages, it is difficult to make any such claims in diachronic syntactic studies. If two grammars are 'in competition' (as advocated by Kroch (1989) and Pintzuk (1991) among others), we cannot be certain whether this variation is a genuine case of 'true optionality' or whether the variants were distinct on some (sociolinguistic) level, with evidence for this having been obscured over time (Roberts, 2007:331). Walkden (2014) finds some further issues with Biberauer & Richards's necessary rejection of derivational determinism asking why there would be no difference between, in their pied-piping example, moving a small or a big category and what determines which of the two options will be taken. In the end, speakers/writers *do* make a decision, but if an algorithm is non-deterministic it is unimplementable. He therefore concludes that "[f]or a given selection of lexical items, there is only one possible derivational outcome" (Walkden, 2014:23). This means there can be no 'true optionality' or 'free variation' within a single grammar. He furthermore adds that speakers have access to multiple varieties of their language and that there is a 'user's manual' regulating the choice between them (cf. Culy (1996:114)). This variation can be subtly conditioned, not semantically (in the strict truth-conditional sense), but functionally or contextually. Walkden argues that these sociolinguistic factors or 'social knowledge' should be treated as part of the lexicon. As such they can enter the derivation like any other type of formal feature (Walkden, 2014:28-31).

Certain types of formal features are (relatively) uncontroversial, such as referential features (φ-features), negative (polarity) features or features related to questions, such as wh-features. The exact nature of the EPP feature is still an issue of debate, but the fact that there must be some sort of movement-triggering feature (as an 'Edge feature' or simply in the form of a diacritic caret ˆ) is not. As discussed in the previous chapter, a wide variety of information-structural features has been proposed, such as TOPIC, FOCUS or ANAPHOR. Whether there should be more or fewer of those and whether that might be language-specific is still a matter of

ongoing cross-linguistic research. As I have argued in the previous chapter, for Middle Welsh, we need Topic and Focus features at the very least (with possibly, an added distinction between different subtypes of topics, such as Aboutness, Familiar and Contrastive). A further set of 'social' features might exist, as Walkden (2014) suggests to resolve the issue of 'free variation', but the exact nature and effect of those in Middle Welsh is difficult to ascertain on the basis of the corpus under investigation.

In conclusion, within the Minimalist Program, variation can still be thought of as parametric variation with the locus of parameters in the formal (functional) features of the lexicon. Clusters or cascading changes in the grammar of a language can be captured by a hierarchical structure of parameters. With these tools in mind, I first discuss two major mechanisms of syntactic change before moving on to the complex issues concerning actuation and diffusion of syntactic innovations.

Types of syntactic change

In principle, any element of the grammar that can exhibit variation (within a language or between different languages/dialects) could be subject to change. Diachronic changes have been studied in the core domains of argument structure (thematic roles and grammatical functions, e.g. English psych verbs (Allen, 1995)) or passives (Dreschler, 2015) and complementation (e.g. in Latin *ut*-clauses (Vincent, 1988)). The earlier diachronic syntactic descriptions furthermore focussed on major changes in word order. The change from OV to VO in English already discussed above, could for example be seen as a change in head-directionality. But a simple parametric switch from head-final to head-first does not adequately account for the complex data in the history of English. However, parametric change in the much more fine-grained sense of change of functional features in a parameter hierarchy within a Minimalist framework could be the right approach to all these types of change.

Syntactic innovation can also change the underlying structure of a certain pattern without necessarily modifying the surface manifestation. This is called syntactic reanalysis (see, among others, Harris and Campbell (1995)). The preconditions for diachronic reanalysis are structural ambiguity and a preference for simplicity. The hearer assigns a specific parse to the input that is different from the structure assigned by the speaker (Walkden, 2014:39). An often-cited example is the reanalysis of *for...to* in Middle English creating a complementiser marking Case on subjects in nonfinite clauses as presented by Fischer (1992:330-334) and Fischer, Van Kemenade, Koopman, and Van der Wurff (2000:214-200):

(5) a. PREDICATE [$_{PP}$ [$_{P}$ *for* NP] [$_{TP}$ *to* VP] ⇒
PREDICATE [$_{CP}$ [$_{C}$ *for*] [$_{TP}$ NP *to* VP]]
b. It is bad [$_{PP}$ for you] [$_{TP}$ to smoke] ⇒
It is bad [$_{CP}$ for [$_{TP}$ you to [$_{VP}$ smoke]]]

Willis (2016) cites this example under 'spontaneous syntactic innovation' and notes that this standard account is sharply criticised by, among others, Garrett (2012:55-66). If reanalysis becomes a possibility at any time (but is never required) it fails to explain why it actually happens (see also the discussion on triggers of change and acquisition in the next section). Just as the above-mentioned types of syntactic change, diachronic reanalysis of this kind might be reduced to a parametric change. In this particular case the category of the lexical item *for* changed from preposition to complementiser. Since the preposition *for* still exists in other constructions, it looks like a second lexical item *for* was created in the lexicon with a different featural and categorial makeup so that it can function as a complementiser selecting a TP (instead of a preposition selecting an NP or DP).

Another very well-studied area of syntactic change is grammaticalisation. Grammaticalisation is a specific type of reanalysis in which 'less grammatical items', for example simple open class lexical (content) items, become 'more grammatical'. In other words grammaticalisation is "the dynamic, unidirectional historical process whereby lexical items in the course of time acquire a new status as grammatical, morphosyntactic forms, and in the process come to code relations that either were not coded before or were coded differently" (Traugott & König, 1991). The term was first coined by Meillet (1958 [1912]) and presented in comprehensive discussion in, among others, Heine and Kuteva (2002). Apart from being defined as a historical process, the term is also used to describe a research framework (Hopper & Traugott, 2003). According to Campbell and Janda (2000), there are different processes involved in grammaticalisation, such as phonological reduction (e.g. English 'let us' > 'let's'), loss of 'syntactic freedom' (e.g. French *pas* 'step' > *pas* as a negative marker), pragmatic inferencing (e.g. English 'since' from temporal sequence to inferred causation) and semantic bleaching (e.g. German *Mann* 'man' > *man* 'one, some human being').

Campbell (2000:141) argues that grammaticalisation is in itself not a mechanism of change. It relies primarily on the above-mentioned mechanism of reanalysis and also on the extension of the construction in question. As such it could also be viewed as a parametric change or a change in the featural makeup of a lexical item. As Roberts and Roussou (2003) describe it in a formal (generative) account grammaticalisation is a categorial reanalysis driven by change in properties of functional heads. When a new exponent of a functional head F is created, it may also involve creating new parametric properties (triggering Agree or internal Merge) associated with that head. A good example of this type of reanalysis in the history of Welsh is the specifier-to-head reanalysis of personal pronouns becoming complementisers (Willis, 2007a). Another example that I will describe in greater detail in the second part of the chapter is the grammaticalisation of the so-called *sef*-construction in Middle Welsh.

The logical problem of language change

As noted above, within the generative model language change is defined as two distinct grammars in a historical relation. This leaves two logical possibilities for

the locus of change. First, this could be first language acquisition whereby syntactic change is driven by 'abductive' reanalysis (i.e. the result of a transmission failure) (cf. Andersen (1973), Lightfoot (1999) and Van Kemenade (2007) among many others). A second option could be a change in the internalised grammar of an adult speaker. Such change, however, is not considered to constitute a case of 'real' diachronic change "until a future generation of speakers have adopted the mixed system as their own." (Faarlund, 1990:10). Although some cases of this type of change have been argued to be fully completed during one generation, I focus on first-language acquisition here.

If children are generally successfully acquiring the syntactic system of the language of their parents for generations, how can they suddenly be unsuccessful in doing so? This 'logical problem of language change' has received much attention in diachronic syntactic literature (cf. Clark and Roberts (1993:12), Kroch (2000:699-700), Lightfoot (2006:15) and Willis (2016)). Even non-generative approaches must address this question if they want to speak of causation in syntactic change.

Figure 7.1 shows the traditional Z-model of abductive change presented by Andersen (1973:767). The main idea behind this model is that the child may make an error of abduction and mistake a similar case of a structural analysis for the actual case uttered by the speaker of Grammar 1. This can happen because there is no direct link between Grammar 1 and Grammar 2: contact between the two I-languages is mediated by the E-language output. The mismatches that can arise in such situations are in fact the reanalyses we find in syntactic innovation.

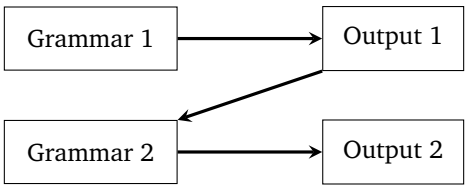


Figure 7.1: Andersen's (1973) Z-model

The difficulty mostly lies in the assumption that first-language acquisition is fully deterministic: children always succeed in acquiring the language perfectly. This Z-model is highly idealised, however, since the primary corpus ('Output 1') is never generated by the grammar of a single individual. It consists necessarily of indeterminate evidence: a finite set of sentences uttered by different individuals, each of whom have a grammar that is not 100% the same as 'Grammar 1' in the model (see Niyogi and Berwick (2009) and Walkden (2012) for further discussion).[8] In an attempt to solve this paradoxical issue in acquisition, Roberts and Roussou (2003) propose a form of 'weak determinism' saying that "the goal of language acquisition is to fix parameter values on the basis of experience; all parameter values must be fixed, but there is no requirement for convergence with the adult grammar" (Roberts & Roussou, 2003:13). Some models of L1 acquisition depart from the

[8] 'Abduction' in itself might furthermore not be the right description of the phenomenon in diachronic syntax (see Lass (1997), Deutscher (2002) and Walkden (2011)).

deterministic assumption altogether. The models proposed by Gibson and Wexler (1994) and C. D. Yang (2002) each contain probabilistic components. In such a scenario, the child can posit more than one grammar (i.e. parameter setting) on the basis of the input she receives. The PLD, in other words, is ambiguous and/or leaves certain options unspecified. A probabilistic model (based on frequency of syntactic structures that count as cues or triggers for a certain grammar, for example) helps the child to determine which grammar to choose.

The type of syntactic innovation or reanalysis based on the ambiguous structure of a sentence in the input could be categorised as 'spontaneous innovation'. Such a purely endogenous solution to the actuation problem lacks explanatory power. Why, for example, does the reanalysis take place at a given time and place, as Weinreich et al. (1968) already pointed out. Willis (2016:3) notes, however, that "if misparsing by children (leading to reanalysis) is distributed randomly in the population (perhaps with some social contexts, such as population mixing, favouring it), it would be pointless to expect more" (of an explanation). Our task then is first of all to accurately describe the conditions and pathways of the reanalysis in a plausible way (i.e. not violating any principles of grammar that are well-established from research into synchronic variation and L1 acquisition). Furthermore, we need to investigate whether the reanalysis is indeed randomly distributed and, to the extent this is possible in our historical context, what the possible social contexts are favouring one pattern rather than the other.

Typological approaches form another kind of endogenous solution proposed already in the earliest stage of historical linguistic research (from the Universals listed by Greenberg (1963) to Indo-Europeanists like Lehmann (1973) and Vennemann (1974)). The core argument consists of applying synchronic restrictions on the ways in which languages combine features to diachronic syntactic changes. These approaches "make system-based predictions about possible and impossible changes" (Willis, 2016:3). In this manner possible pathways for changes are predicted, but the changes themselves do not *have to* occur. If they do, these pathways can still not predict when this will happen (cf. Hawkins (1990:99) and Willis (2016:§3)). Within a generative approach Biberauer, Sheehan, and Newton (2010) argue that the 'Final-over-Final constraint' (FOFC) restricts possible diachronies. For diachronic syntax, FOFC predicts that a change from head-final to head-initial word order must follow a particular order to avoid head-final over head-final structures. A possible explanation for this could lie within the cognitive domain as a processing preference. These types of cognitive preferences may in fact lie behind more (or all?) observed typological universals. For this highly deterministic approach it is first of all important to confirm the cognitive claims with data from thorough psycholinguistic experiments. It is furthermore of crucial importance to have a comprehensive description and adequate analysis of all cross-linguistic data of the phenomenon under investigation.

Another solution to the actuation problem is based on language usage. An increased frequency of use can, for example, explain cases of grammaticalisation. If a particular sequence is often used, but rarely varied, children could acquire it as

a single unit. Within Construction Grammar, this is called 'constructionalization'; in other frameworks, it is simply referred to as 'lexicalisation' (Willis, 2016:6). This moves the 'why'-question to language usage, i.e. why was this particular pattern used more frequently? The answer to this might be irretrievable as long as we do not have access to an accurate description of the sociolinguistic history of the period under investigation.[9] Again we can nonetheless aim to identify the factors that might have aided the increased frequency of a particular pattern.

A further important question arises here (which is also relevant in the context of change in general): how frequent does the pattern have to be to be 'grammaticalised', 'lexicalised' or 'reanalysed'? In other words: what is the so-called 'tipping point' for Grammar 1 to change to Grammar 2 and can this be described in terms of (relative) frequency alone? In order to answer this question, we first need to be clear on the exact cue or trigger for the change. In the case of lexicalisation or grammaticalisation, this is often very straightforward: the frequency of the pattern in the target context (vs. other contexts) could be retrieved from a historical corpus. Assuming we are dealing with a well-balanced corpus that accurately reflects different stages of the language,[10] we can define the frequency required for the change with relative ease. In studies of the acquisition of a particular type of word order, e.g. V2 word order, however, the situation is more complex. Sentences with initial subjects, for example, cannot count as 'triggers' for the child to postulate a V2 grammar (even though subject-initial sentences are V2 in many Germanic languages). The evidence would not be sufficient, however, because an English-like SVO grammar is also possible on the basis of that input. To convince the child to opt for a positive 'V2 setting' of the parameter in question (see section 6.4 below), she needs a significant input of non-subject-initial word orders (followed directly by a finite verb). Lightfoot (1999:154) estimated that roughly speaking, an average of 30% of the sentences should have this type of $XP_{Non-Subject}$-V_{Fin} order to convince the child that her language has a V2 grammar. With syntactically annotated corpora, this estimated number could be compared to a sample of real data. In a corpus of Modern Dutch, C. D. Yang (2000:114) found that 23% of the sentences had $XP_{Non-Subject}$-V_{Fin} order. Since Dutch children successfully acquire V2, he concluded that Lightfoot's estimation of 30% might be too high. On the basis of the Dutch corpus study it seems that 23% should be sufficient. Westergaard (2009:67) conducted a similar study of Norwegian corpora. She finds only 13.6% in her child-directed corpus. These numbers found in spoken corpora might differ in historical written corpora, because it is not always clear to what extent the written data reflect the spoken language at the time. This type of research in first-language acquisition is nonetheless extremely useful in attempting to accurately describe situations of historical change.

The actuation problem in historical syntax can also be 'solved' by considering

[9] See also Lass (1980:101-103) and Walkden (2012:897-898) on Popper's methodological version of the principle of causality (Popper, 1968:67) and why it might not be appropriate to ask ourselves this particular kind of even further-removed or deeper 'why'-questions in the study of historical syntax.

[10] This is a somewhat idealised situation, because there are various practical limitations building a well-balanced historical corpus, as discussed at length in Chapter 2.

changes in other parts of the language. Apocope or the loss of final syllables discussed in section 7.2.2 can, for example, lead to the loss of a case system, because the morphological endings no longer function as distinctive features. 'Phonological erosion' (whatever causes it) is mentioned by Willis (1998) as the crucial trigger for the loss of V2 in Early Modern Welsh. When the preverbal particles *a* and *y* disappeared, the acquisition of the V2 system became obscured at first and then completely impossible. As shown in the previous chapters, Middle Welsh allowed V3, V4 and V5 orders with adjuncts in the preverbal domain alongside the standard V2 'abnormal' and 'mixed' sentences. With the loss of the preverbal particle *y*, adjunct-initial sentences could easily be reanalysed as Adjunct + VSO orders (see also section 7.2.1 above). Pronominal subjects in initial position were reanalysed as main-clause complementisers after the loss of the preverbal particle *a* (see section 7.3.2). Object-initial orders were very infrequent already towards the end of the Middle Welsh period. According to Willis (2016:9) (building on Willis (1998) and Willis (2007a)), this phonological source of change led to other parallel changes as well, such as the reanalysis of the expletive pronouns as affirmative particles. One phonological change can also lead to another, e.g. a change in the stress pattern can lead to the reduction of vowel quality or even syncope or apocope. What ultimately triggers the initial change in this case is difficult to ascertain. Again psycholinguistic experiments on language production could prove revealing, although the question remains why certain changes were not 'triggered' in the same way centuries earlier, for example.

This leads us to language-external approaches to the logical problem of language change. In principle, external sources in the form of language contact do not necessarily lead to language change. Children are perfectly capable of acquiring more than one language if they get the right input in the earliest stages of their lives. They grow up to be bilingual, fluent in two (or even more) languages or dialects and they can distinguish and use the two grammars without any problems (see also the study on Welsh-English bilingual code-switching and the conclusion of grammatical continuity rather than change by P. Davies and Deuchar (2010) cited above). Syntactic change, however, also occurs in contact situations. According to Meisel, Elsig, and Rinke (2013), a change in the core grammar can in fact *only* occur when non-native speakers form a large part of the speech community (see Meisel et al. (2013:171-182) and Willis (2016)). As discussed above in the section about language shift, syntactic changes are often considered to require a specific type of contact. One possible situation would be the shift of speakers of the substrate to the superstrate language, keeping grammatical features of their substrate so that they become embedded in the superstrate language. Since instances of syntactic change have also been reported in situations without language contact, a complete rejection of any kind of endogenous approach to syntactic change seems to be unfeasible (Willis, 2016:10). This finally brings us to the notion of 'inertia', as formulated by Longobardi (2001) and Keenan (2002):

(6) "Syntactic change should not arise, unless it can be shown to be *caused*" (Longobardi, 2001:278)

(7) "Things stay as they are unless acted upon by an outside force or Decay." (Keenan, 2002:327)

In the context of the acquisition-based model, the Inertial Theory stipulates that a grammar can only change if the conditions in the process of acquisition have changed. According to Willis (2016:11), this then solves the timing part of the actuation problem because reanalysis of a particular structure only occurs at the time something else changes (e.g. phonological erosion or loss of a lexical item). To a certain extent, this 'solution' is no more than a shift of locus of the problem: why would phonology or morphology not be equally inert in this theory? Walkden's (2012) thought experiment about a child failing to acquire V-to-C movement in her grammar (G2), because she never hears wh-questions in the language of her parents (G1) is very insightful in this context. The grammar of her parent(s) (G1) did not change in any way, it just happens to be the case that direct questions were never asked when the child was around, so V-to-C-movement was not part of the PLD. Although this situation might be extremely unlikely, the main argument holds: the 'cause' of change (Longobardi, 2001) consists of the non-occurrence of a particular pattern. The ultimate reasons for this non-occurrence could be a wide variety of extralinguistic events and even chance and human intentionality (i.e. the 'planning' of utterances) needs to be taken into account as well. Walkden (2012:896) thus concludes that the notion of causality in the Inertial Theory is so broad it is rendered entirely vacuous, because it cannot make any useful empirical predictions.

To conclude this section, research on processes of first-language acquisition can help historical linguists characterise the changes more accurately. The acquisition-based approach advocated within the Minimalist Program by the ReCoS project includes typological, cognitive and acquisitional biases (e.g. Input Generalisation and Feature Economy) that not only help predict pathways of changes, but might also shed light on 'what has not happened' and why this is the case. Computational models of acquisition and the competing-grammar approach advocated by C. D. Yang (2002) can give us further insights in predicting changes based on frequencies of patterns containing cues or triggers for a certain innovation. From an empirical point of view, historical linguists should not only describe the syntactic innovation itself, but also the necessary change in conditions (in the acquisition process) that 'triggered' the innovation (how did it happen and why did it happen in this particular way and not vice versa). It is furthermore necessary to try to identify both endogenous and exogenous factors "which might have aided a variant grammar in persisting or becoming more prevalent" (Walkden, 2012:899). This last notion is related to the diffusion of 'reactuation' of syntactic innovations, which is the topic of the next section.

Dynamics of change

Parametric change is traditionally described as having two main characteristics. It is:

1. catastrophic ⇒ when it changes suddenly and irrevocably at a given moment
2. internal to the inquirer ⇒ this means that in principle, it is entirely independent of the child's cultural, social or historical background

Since many syntactic changes observed in diachronic data seem to be gradual, rather than abrupt, the 'catastrophic' nature of parametric change has received much criticism. Language can be transferred in two different ways: transmission in first-language acquisition and diffusion from adult to adult. In this section, I discuss the dynamics of change and possible ways to solve the gradual-abrupt paradox of parametric change.

After the introduction of a novel form (the syntactic innovation or 'actuation' of a change), the form can spread through a speech community. From historical corpora we often observe a period of variation until one system (G2) takes over from the other (G1). This period can be described as individuals having two grammars in competition, formal optionality, diglossia and/or diffusion of the syntactic innovation. This rate of replacement from one grammatical option to another often shows the same 'slow-quick-slow' pattern (as observed by, among others, Osgood and Sebeok (1954:155), Weinreich et al. (1968:113-14)). Kroch (1989) analysed syntactic changes such as the replacement of *have* by *have got* in British English from 1700 to 1935 and the loss of the verb-second constraint in Middle French from 1400 to 1700. He concluded that the 'slow-quick-slow' rate of change can be modelled by a logistic function showing an s-shaped curve when the frequency of new vs. old forms is plotted against time as shown by the equation in Figure 7.2:

$$p = \frac{e^{k+st}}{1 + e^{k+st}}$$

Figure 7.2: S-curve logistic function by Kroch (1989:204) with: p = the frequency of the innovation, t = time, s = the slope of the function, k = the y-intercept (the frequency of the innovation at $t = 0$) and e = Euler's number (approx. 2.71828)

This 'Constant Rate Hypothesis' (CRH) shows the grammars in competition change gradually through a population or within individuals who have access to one of these grammars more readily than others over time. This is in effect a situation of syntactic diglossia (Kroch, 2000:722): speech communities (and individuals) synchronically instantiate several grammatical systems. According to Willis (1998), the same actuation process may be triggered in multiple speakers, in which case apparent diffusion through the speech community may actually be an instance of 'multiple reactuation' (Willis, 1998:47-48). The increase in frequency of the syntactic innovation may furthermore be due to sociolinguistic factors: an abrupt parametric change can therefore appear to be gradual. Lexical diffusion and microparametric changes (Kayne, 2000:3-9) may help keep up the 'mirage of gradualness' as a cushioning effect: "a series of discrete changes to the formal features

of a set of functional categories taking place over a long period and giving the impression of a single, large, gradual change" (Roberts, 2007:300). More studies in first-language acquisition in the context of language contact situations, dialect research, and code-switching will play an important role in refining and explaining the 'S-curved' model. One possible way forward is to include geographical factors into historical syntactic models of change. If geospatial information about the distribution of syntactic innovations is available (e.g. texts from different areas over a certain period of time), this can be integrated into a logistic regression model. Willis (2014) shows how this type of geographically weighted regression model in dialect research can be applied to the innovation and diffusion in the pronominal system of northern varieties of Welsh over the last 150 years. He combines the Constant Rate Hypothesis with geospatial data trying to show the diffusion of syntactic innovations in the speech communities of North Wales. This model can be tested and further refined by studies of 'recent' syntactic innovations in dialect areas for which this type of information is available.

Apart from the speed of change and its geographical diffusion, the direction of syntactic innovations has been the topic of various studies in diachronic linguistics. Especially in studies concerning grammaticalisation, these processes often follow well-defined pathways. However, cases of 'degrammaticalisation' have been reported as well, in which case the directionality of change seems to be reversed (e.g. Willis (2007b), Norde (2009) and Rosenkvist (2010)). The diachronic syntactic 'principles' proposed by Van Gelderen (2009) are similarly laying out certain pathways for change:

(8) **Head Preference Principle** Van Gelderen (2009:136)
Be a head, rather than a phrase.

(9) **Late Merge Principle** Van Gelderen (2009:136)
Merge as late as possible.

These 'principles' are not uncontroversial (cf. Motut (2010)) and, according to Walkden (2014:42) if we adopt the I-language perspective on historical syntax, an independent principle governing the direction of change cannot exist. Willis (2011a:421-424) also notes that if there is any form of universal directionality, it can be reduced to 'local directionality' meaning that the interaction of the acquisition algorithm with the PLD leads to predictable reanalyses. Van Gelderen's Principles, to the extent they are universal, might thus be the result of preferences in the acquisition process. Such acquisitional biases were already discussed in the parametric hierarchy approach above. In this context, a 'pathway of change' is equivalent to the child being pressured to postulate the simplest possible system, for example. Input Generalisation as a principle in acquisition states that a generalisation - if possible - is extended over the widest possible domain (until met with counter-evidence). Functional features may also become less transparent, leading to a complete loss and thus simplification of the system. Changes can occur moving up or down the hierarchy: they might be constrained for cognitive reasons, but change is not unidirectional per se.

To conclude, the exact 'dynamics of change' seem to be specific rather than universal. Since the sociolinguistic context can play an important role in the spread of changes, this should be taken into account (to the extent this is possible in a historic context) in describing the process of transfer in speech communities. Syntactic innovations are not inherently unidirectional, although biases in first-language acquisition in the form of 'local directionality' can lead to predictable reanalyses. The S-curve of the Constant Rate Hypothesis, combined with - where available - geospatial data can further help a well-informed analysis of syntactic changes. However, if we do not have access to ample data (because from the time of 'our British Celt' vernacular texts have not survived or were never written down in the first place), our task of analysing the origin of a particular syntactic pattern is severely complicated. In the next section, I therefore discuss the possibilities and limitations of syntactic reconstruction.

Syntactic reconstruction

In this section I finally turn to the successful 'Comparative Method' in phonological reconstruction mentioned in the introduction. In 1900, Berthold Delbrück, one of the greatest early researchers in the field of historical linguistics expressed his doubts about the possibility of reconstructing syntax in the same way this is done for the lexicon, phonology and morphology (Delbrück, 1900 [1982]:v-vi). Further attempts were nonetheless done by Lehmann (1972), Hopper (1975) and Kiparsky (1995). Various problems arise in the reconstruction of syntax, however, as pointed out by, among others, Lightfoot (1999). I first briefly sketch the fundamentals of the method of comparative reconstruction and then discuss the problems it might cause in the field of historical syntax.

The first step of the comparative method consists of finding a set of corresponding words in (potentially) related languages. In (10) and (11) below, I show a somewhat simplified example from the Indo-European language family for the English adjective 'new'. An important part here is both the formal as well as the semantic similarity to form 'cognates' (form-meaning pairs). In this case, the adjectives in the different Indo-European languages all mean 'new' and can thus be considered proper double (form *and* meaning) cognates. The set below thus qualifies as a proper correspondence set (see Beekes (1995:196) and Schrijver (1995:283ff) for the forms in IE and Celtic respectively):

(10) Sanskrit: *návya-*, *náva-*
Gothic: *niujis*
Hittite: *newa-*
Greek: *néos*
Latin: *novus*
OCS: *novъ*
Tocharian B: *ñuwe*
Old Irish: *núae*
Welsh: *newydd*
Breton: *nevez*
Middle Cornish: *noweth*
Gaulish: *Novio-(magus/dunum)*

The next step involves the proper alignment of the examples, starting from the stem of the adjective, as shown by the sample of languages in (11):

	Sanskrit	n	á	v	-
	Latin	n	o	v	-
(11)	Greek	n	é	-	
	Welsh	n	e	w	-
	PIE	*n	?	?	

From the aligned correspondences, we can then reconstruct the sounds by comparing the forms. For the first letter, this is easy: since initial *n-* appears in all languages, we postulate initial **n-* for the form in the proto-language, in this case Proto-Indo-European (PIE). The vowel and second consonant are less straightforward, because the different languages exhibit different phonemes, or, in the case of the Greek consonant, nothing at all (in that position of the word). To reconstruct these PIE phonemes, we have to find regular sound correspondences in the respective languages, i.e. does Sanskrit short *á* always correspond to Greek *é* and Latin *o*? Does that depend on the phonological context and/or can we find regular sound changes? Sanskrit short *a*, for example, regularly responds to either an *e* or *o* in Greek. However, by regular sound law (Brugmann's Law, cf. Beekes (1995:138)), PIE short *o* in open syllable became a long *ā* in Sanskrit. Since we find a short *a* in an open syllable in the adjective 'new' in Sanskrit *náva-*, it is unlikely this goes back to PIE *o*. The *o* in Latin, however, usually means we have to reconstruct an *o* in PIE as well. However, again by regular sound change PIE **e* became *o* in Latin before *u̯*, ł and *mo* (Beekes, 1995:66). This combined evidence from the regularity and 'exceptionlessness' (*Ausnahmlosigkeit*) of sound changes upon which the Comparative Method heavily relies, forces us to conclude PIE **e* can be the only right vowel to reconstruct. Another final point is the question of orthography and to what extent it is representative of the actual sound. The *v* and *w*, for instance, could both represent the glide or semi-vowel *u̯*. In Greek, furthermore, this **u̯* regularly disappears intervocalically (cf. Beekes (1995:135)). To conclude, the reconstructed form of the stem of the adjective that means 'new' in many different Indo-European languages is PIE **neu̯-*.

The reconstructed 'product' of the Comparative Method by definition does not represent a real language: it is timeless and non-dialectal (cf. Walkden (2014:37)). Successful reconstruction does not need a causal explanation per se: the result is valuable nonetheless, since it shows how the phonological (and morphological) systems of languages and their vocabulary has changed. If we want to apply the same method to syntax, however, we run into problems at the very first step: the 'correspondence problem'. As Calvert Watkins already pointed out in the 1970s, "the first law of comparative grammar is that you've got to know what to compare" (C. Watkins, 1976:312). Walkden (2009) (and subsequent work, Walkden (2014:52), amongst others) conclude that the double cognacy condition (the corresponding form-meaning pair in, for example, the vocabulary item 'new' above) cannot be easily met, because sentences are never the same. Certain idioms or

stock phrases might be compared in several Indo-European languages, but it is impossible to compare whole sentences because sentences are not transmitted as such across generations. Further criticism at attempts to reconstruct syntax were formulated as the 'directionality problem'. As discussed in the previous section, syntactic change is not inherently unidirectional. A change from OV to VO word order could in principle also be reversed. This problem, however, is not necessarily restricted to the syntactic domain. There might be phonetic tendencies, for example, to voice consonants in between vowels, but a change of *o* to *a* could in principle also be reversed. The same could be said about the 'reanalysis problem' stating that grammar must be created again by each new learner: phonological systems need to be learned in the same way (cf. Lightfoot (1979)).

Within the Minimalist Program, a possible solution to the correspondence problem again lies in the Borer-Chomsky Conjecture (BCC). Recall from the beginning of this section that the BCC sees the functional features in the Lexicon as the source of all variation. According to Walkden (2014:55-60), if these functional features take a phonological form as functional items we might reconstruct those in the context of appearance in attested sentences of the daughter languages. An example of this is the reconstruction of the free relative in Brythonic, the predecessor of Welsh, Breton and Cornish in Willis (2011a).

It is difficult, if not impossible to formalise syntactic reconstruction in a framework based on phrase-structure rules (Principles & Parameter theory, Newmeyer's rule-based system or Lexical Functional Grammar) or constraints (e.g. HPSG). An item-based approach like this would in principle work for both derivational as well as representational models (cf. the 'Constructicon' in Construction Grammar and discussion in Walkden (2014)). From the perspective of the Minimalist Program, syntactic primitives are considered to be stored in the lexicon. These functional features form the basis of syntactic variation and can be reconstructed if they take a phonological form in the daughter languages of the proto-form we want to reconstruct.

Interim Summary

In this section I presented several problems in the study of diachronic syntax and how they can be tackled by tools and mechanisms within the framework of Generative Grammar, in particular the most recent version of hierarchical parametric theory in the Minimalist Program. If we assume the existence of an innate capacity for acquiring grammars, we can use insights from synchronic research into formal syntax as well as mechanisms from language acquisitions. This is a considerable advantage in the study of diachronic syntax, because the available data is often limited. Understanding how the grammar of a language is *acquired* helps us understand how grammar can *change*. The question of *why* certain syntactic innovations appeared at a given time and spread through the speech community (problems of actuation and transfer, via transmission and/or diffusion) is more difficult to answer. Evidence for detailed sociolinguistic situations in earlier days is often just as scarce as the extant manuscript sources of the language under investigations.

Adopting a generative acquisition-based approach to diachronic syntax can help us define the exact conditions and/or context in which innovation can and cannot occur and how they can trigger further changes. We could furthermore identify endogenous and exogenous factors playing a role in making variant grammars more prevalent. Finally, the concept of 'multiple reactuation' in speech communities as well as models like the Constant Rate Hypothesis (possibly combined with dialectal (geospatial) data) provide us with a much better understanding not only of specific innovations, but also of the processes involved in syntactic change in general. In the remainder of this chapter, I use the tools and mechanisms of the Minimalist Program discussed in this section to analyse two syntactic innovations in the history of Middle Welsh: the grammaticalisation of the *sef*-construction and the rise of the Abnormal Sentence.

7.3 Diachronic syntax in Middle Welsh

In this section I focus on two syntactic innovations in Middle Welsh: the grammaticalisation process of the identificatory copular clause or '*sef*-construction' and the rise of the Abnormal Sentence. The synchronic syntactic analyses of these constructions were already presented in the previous chapter. Here I present a diachronic analysis in a generative (Minimalist) framework.

7.3.1 Grammaticalisation of the *sef*-construction

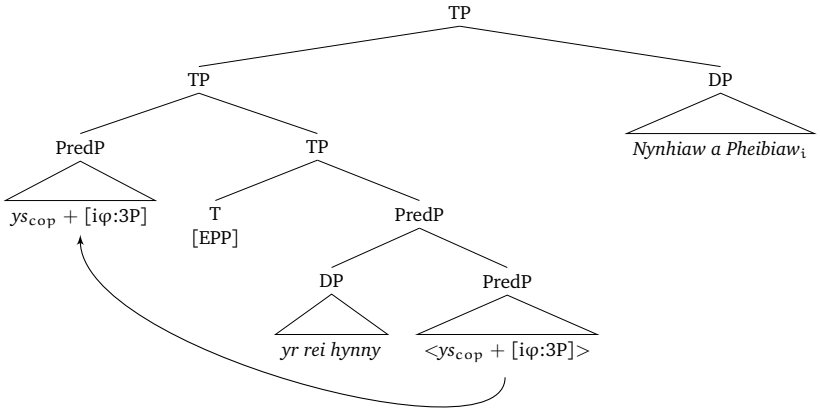
In Chapter 6 I showed various kinds of copular constructions in Middle Welsh. They exhibit different word order patterns and predicate forms (with or without the overt predicate marker *yn*), depending on the information-structural status of the subject or predicate. Predicates that identified the subject could be focussed in Old Welsh by means of a cleft construction, shown in (12a). In Early Middle Welsh, this construction is also attested once with a plural predicate, as in (12b).

(12) a. *issem i anu Genius*
be.PRES.3S.it 3MS name Genius
'that's his name, Genius' (Old Welsh gl. *Genius* in MC - T. A. Watkins (1997:579))

b. *Ys hwy yr rei hynny, Nynhyaw a Pheibyaw*
be.PRES.3S they the ones DEM.P Nynniaw and Peibiaw
'Nynniaw and Peibiaw are those ones' (Lit. 'It's them, those ones, ...)(Middle Welsh CO 598)

I argued that the derivation of (12b) is very similar to the one outlined for Inverted Copular Clauses in Scots Gaelic by Adger and Ramchand (2003). In these sentences, the copula is the head of the Predicate Phrase. It moves to SpecTP to satisfy T's [EPP]-feature pied-piping the complement, in this case the anticipatory predicate third-person plural pronoun *hwy*. The real predicate, co-indexed with the anticipatory predicate, is first-merged adjoined to TP:

(13)



This construction forms the starting point of the reanalyses that occurred in the Middle Welsh period. In texts from this period, we find many variants of this construction. In the following, I argue that these variants show five different stages of the process of grammaticalisation and reanalysis. The examples below represent these five subsequent stages:

Stage 1 - Cleft + focussed predicate (*ys + ef/hwy*)

(14) a. *iss em i anu Genius*
be.PRES.3S it 3MS name Genius
'that's his name, Genius' (Old Welsh gl. *Genius* in MC - T. A. Watkins (1997:579))
b. *Ys hwy yr rei hynny, Nynhyaw a Pheibyaw*
be.PRES.3S they the ones DEM.P Nynniaw and Peibiaw
'Nynniaw and Peibiaw are those ones' (Middle Welsh CO 598)

Stage 2 - Copula + Anticipatory Predicate merge (*ys ef* > *sef*)

(15) a. *Sef gwreic a uynnawd gwreic ieuank*
sef woman PRT want.PAST.3S woman young
'That was the woman he wanted, a young woman.' (YBH 6)
b. *Sef __ a doeth dy nyeint*
sef PRT come.PAST.3S 2S nephews
'That's who came, your nephews.' (WM 89.35)

Stage 3 - Expletive focus marker *sef*

(16) *Sef a wneuthum inheu (...) mynet*
sef PRT do.PAST.1S I (...) go.INF
'This is what *I* did, I went (...)' (WM 492.3 - Watkins 1997:586)

Stage 4 - Loss predicate focus

(17) *Sef a wnaeth pawb yna moli Duw*
sef PRT do.PAST.3S everyone then praise.INF God
'Everyone then did this, they praised God.' (Dewi 4.17)

Stage 5 - Expletive *sef* reanalysed as adverb

(18) *Sef y clywei arueu am ben hwnnw*
sef PRT hear.PAST.3S arms on head that.one
'He could feel armour on that one's head.' (WM 54.28 - Watkins 1997:587)

Schematically, the process with the reanalyses is presented in Figure 7.7:

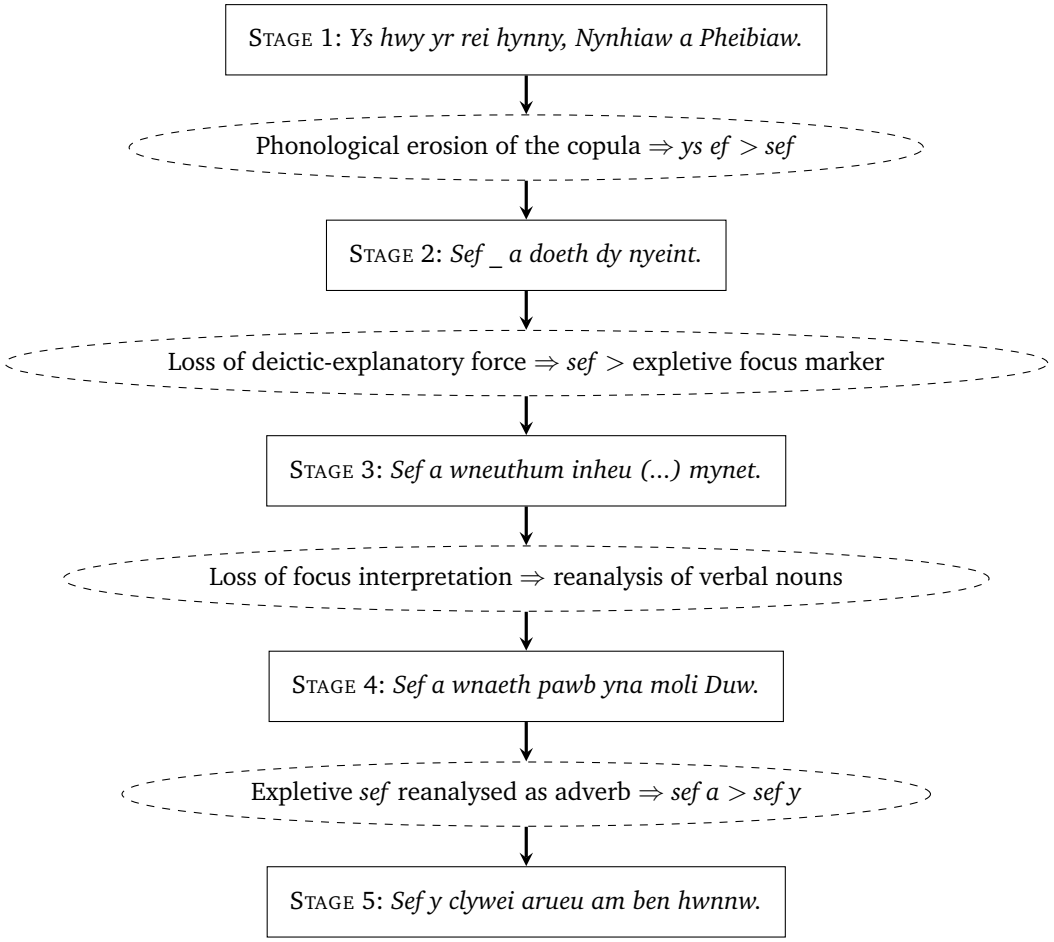


Figure 7.3: Stages of reanalysis of *sef*

From Stage 1 to Stage 2: phonological erosion

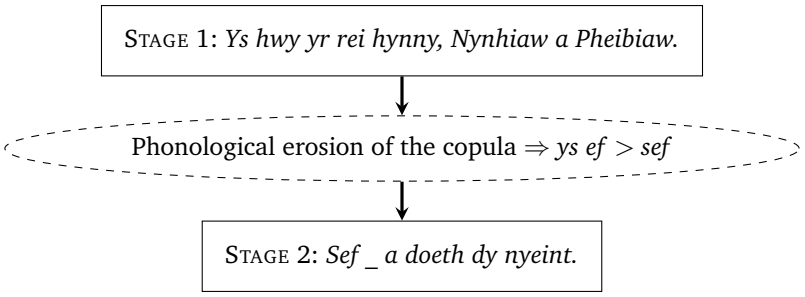


Figure 7.4: Stages 1-2 of reanalysis of *sef*

The derivation of the original cleft sentence with the focussed predicate adjoined to TP was presented in (13) above. T. A. Watkins (1997:579) describes this construction as follows: "In Old Welsh the identificatory copular sentence can be realized as follows: Copula + Anticipatory Predicate + Subject + Postponed Nominal Predicate.". This original *sef*-construction has the following characteristics:
- confined to simple/main clauses of positive declarative sentence types.
- sentences must have nominal (i.e. noun or noun phrase) subject and predicate.
- always identificatory predicates therefore Subject and Predicate must be determinate (definite NPs are inherently so; indefinite NPs may be determinate or indeterminate)
- there is agreement between anticipatory and postponed predicates
- there is agreement between subject and referent
- the only attested tense is present indicative (due to paucity of Old Welsh material, because it is there in Old Irish)
- only attested in 3rd person (since both subject and predicate were obligatorily nominal)
- The subject refers back to a previous (usually immediately preceding) sentence or sentence constituent

Although the full form of the copula is still found in some Early Middle Welsh texts in this construction, there are also signs of phonological reduction. In some cases in Old Welsh already, the copula and anticipatory predicate are written as one word, indicating the start of the merger, as shown in (19a).[11] In Medieval manuscripts, like the *Red Book* of Hergest, the initial vowel of the copula has disappeared, but the double *ss* is still found:

(19) a. *issem* *i* *anu* *Genius*
be.PRES.3S.it 3MS name Genius
'that's his name, Genius' (Old Welsh gl. *Genius* in Martianus Capella - T. A. Watkins (1997:579))

[11] Middle Welsh *ef* 'he, it' was often written as *em* in Old Welsh.

b. *SSef a oruc yr amherawdyr glasowenu.*
sef PRT do.PAST.3S the emperor smile.INF
'The emperor smiled.' (BR 6.25-26)

In most medieval texts, however, the form *sef* is found. This form became structurally ambiguous. It always appeared in the same sequential order *ys + ef* and it was always associated with identificatory predicate focus. The now petrified combination of the copula + anticipatory predicate could thus be reanalysed as one lexical item: the copular focus marker *sef*. This focus marker is then first-merged in the C-domain, satisfying the uninterpretable Focus feature on the C-head.

(20) $[_{\text{PredP}}$ *ys ef*$] > [_{\text{FocP}}$ *sef*$]$

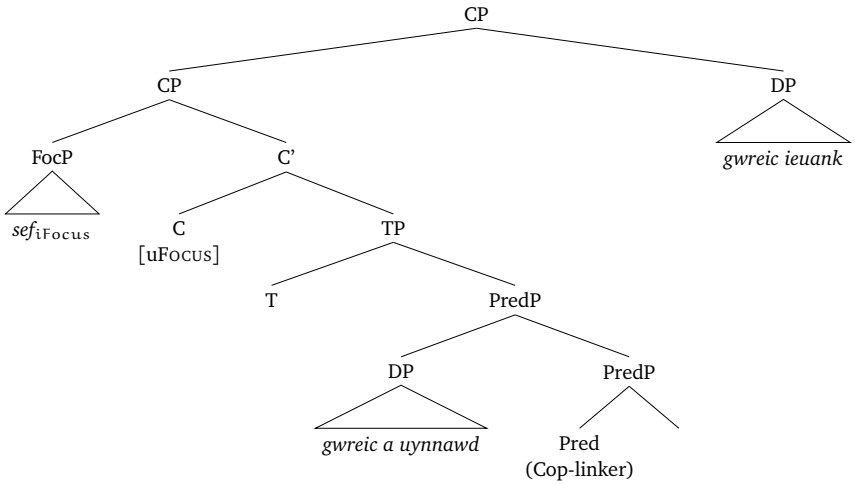
If the subject was not a demonstrative, a relative clause modifying the subject was often used to establish the link with the preceding context. This contextual link was, according to T. A. Watkins (1997) one of the requirements for the *sef*-construction. As shown in example (21), the subject of the clause could be complex, consisting of a DP with a relative clause. The head of the relative could function as the subject, shown in (21a), object (21b) or as an adjunct of the relative verb (21c):

(21) a. *Sef seithwyr a dienghis Pryderi Manawydan (...)*
sef seven.men PRT escape.PAST.3S Pryderi Manawydan (...)
'These were the seven men who escaped, Pryderi, Manawydan (...).' (WM 56.34)
b. *Sef gwreic a uynnawd gwreic ieuank*
sef woman PRT want.PAST.3S woman young
'That was the woman he wanted, a young woman.' (YBH 6)
c. *Sef lle y doethont ygyt y bresseleu*
sef place PRT come.PAST.3P together in Preseleu
'That was the place where they got together, in Preseleu.' (WM 27.28)

In a sentence like (21b), the complex subject DP *gwreic a uynnawd* is in the specifier position of the Predicate Phrase. The head of the PredP is now the phonologically empty copula. This is not a strange stipulation in the context of Middle Welsh, because verbless or 'nominal' copular clauses existed as well (see Chapter 4). The copular focus marker *sef* is then merged in SpecCP and the focussed predicate is adjoined in the same way as before.[12]

[12]Note that adjunction to CP is not necessary to end up with the correct word order *Sef - Subject - Focussed Predicate*. The focussed predicate could also be first-merged (i.e. externally merged) as the complement of the Pred-head and then remain there or be extraposed to end up in the C-domain. I show the derivation with the predicate adjoined to CP here, because adjunction is allowed for the further reanalysis sketched below.

(22)

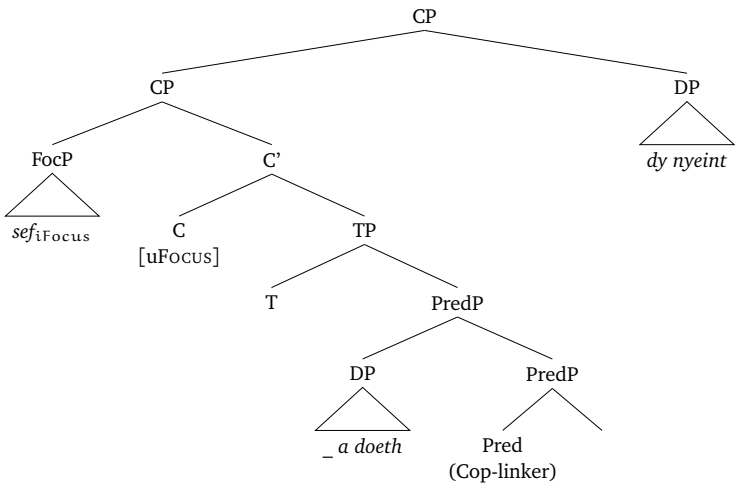


There are furthermore examples of headless relatives.

(23) *Sef __ a doeth dy nyeint*
sef PRT come.PAST.3S 2S nephews
'That's who came, your nephews.' (WM 89.35)

These constructions can be analysed in the exact same way as the above constructions, but they are structurally ambiguous.

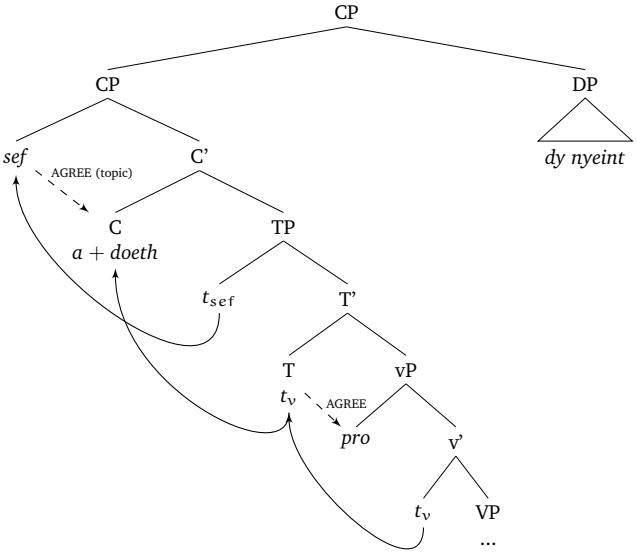
(24)



The ambiguity arises because of the missing head noun in the relative clause that functions as the subject of the copular clause. These subjects were originally in the specifier of the Predicate Phrase. The relative clause *a doeth* 'who came', could

at this stage be reanalysed as the matrix verb. Recall that the most frequently occurring word order pattern in Welsh was the verb-second 'Abnormal Sentence' with the exact same surface structure as relative clauses. The formal focus marker *sef* can now be reanalysed as an expletive merged in SpecTP which subsequently moved up to SpecCP to satisfy C's uninterpretable focus feature. As an expletive, it is considered to be an argument topic and it will thus trigger 'topic agreement', i.e. the complementiser will be realised as *a*, the form it usually takes following core arguments in Abnormal Sentences (instead of *y* following adjuncts). The focussed predicate is then still in the same position adjoining the CP.

(25)



From Stage 2 to Stage 3: loss of deictic-explanatory force

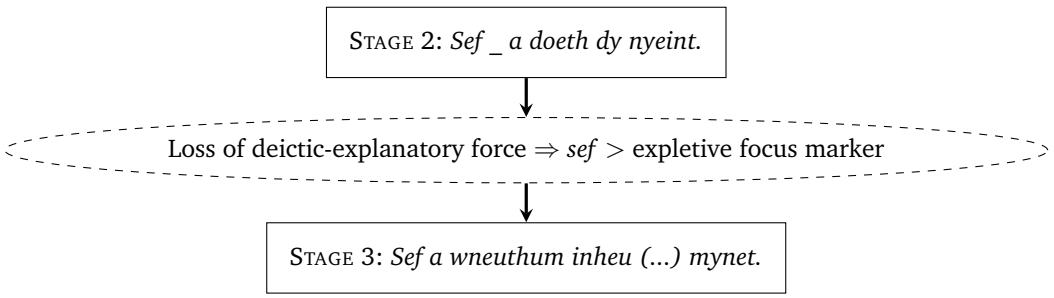


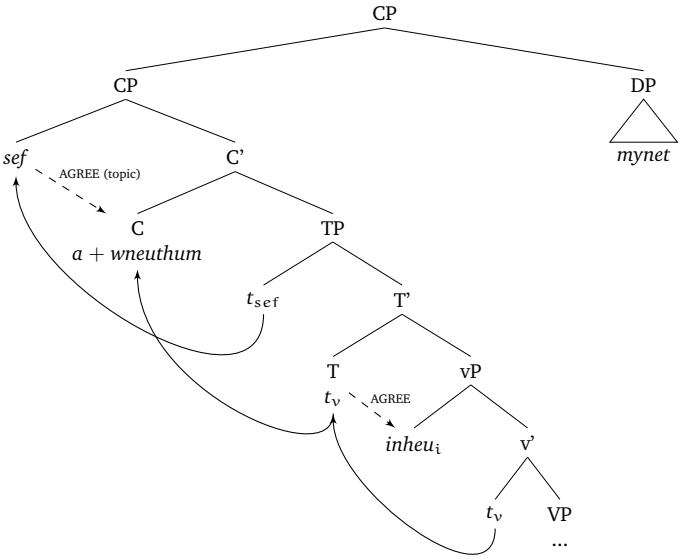
Figure 7.5: Stages 2-3 of reanalysis of *sef*

The next stage of the grammaticalisation process is characterised by the loss of the deictic-explanatory force of *sef*. The new *sef*-construction is no longer necessarily related to the preceding context. The construction could now be used in continuous narrative contexts as well, shown by the example in (26). The construction can in this stage still be parsed in the same way as the examples with headless relative subjects and extraposed predicates above, shown in (27):

Preceding context: "Until it was with difficulty that **I** fled"

(26) *Sef a wneuthum inheu (...) mynet*
sef PRT do.PAST.1S I (...) go.INF
'This is what *I* did, I went (...)' (WM 492.3 - Watkins 1997:586)

(27)



There are two formulaic constructions with unexpressed head-nouns that were very popular and used extensively:

(28) a. *Sef a gausant yn eu kynghor duunaw ar eu llad*
sef PRT get.PAST.3P in 3P council agree.INF on 3P kill.INF
'This is what they decided in their council, they agreed to kill them' (WM 68.8)
b. *Sef a wnaeth y gwraged kyscu*
sef PRT do.PAST.3S the women sleep.INF
'This is what the women did, they slept.' (WM 28.15)

In these sentences, the predicate is a verbal noun: *duunaw* 'agree' or *kyscu* 'sleep'. In non-copular sentences in Middle Welsh, the verbs *cael* 'get' and *gwneuthur* 'do'

could be used as auxiliary verbs. This then paved the way for a further possible ambiguous structure leading to the next stages of the reanalysis.

From Stage 3 to Stage 4: loss of deictic-explanatory force

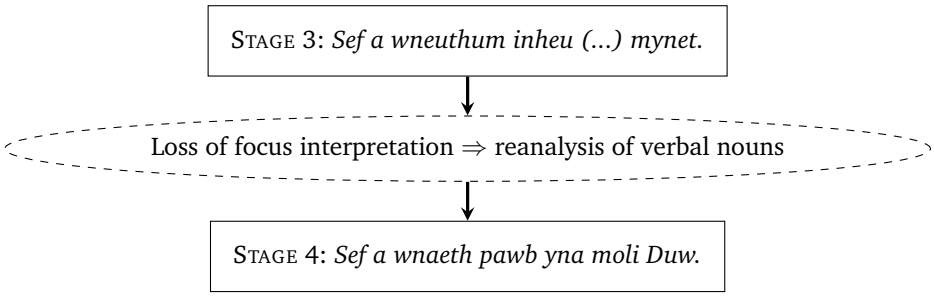
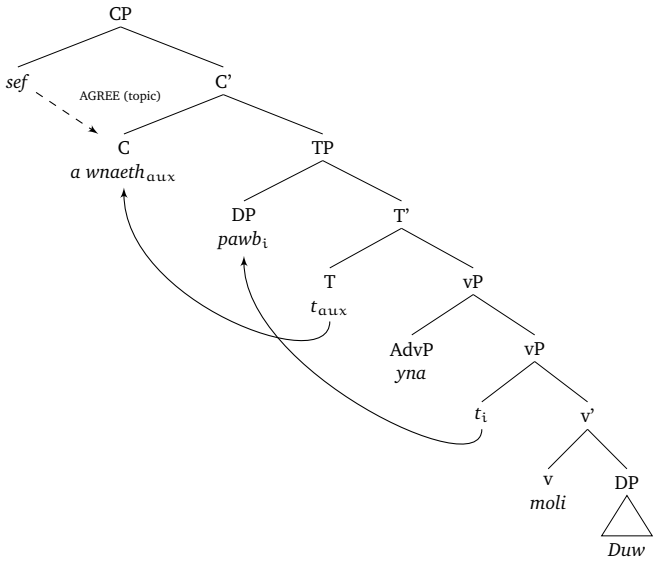


Figure 7.6: Stages 3-4 of reanalysis of *sef*

In the next stage, this structural ambiguity leads to reanalysis of the verbal noun as the matrix verb. The adjoined or extraposed predicate position is lost and along with that the focussed interpretation. The subject moves to SpecTP and agrees with the verb while *sef* is first-merged in SpecCP now. In example (29), the verbal noun *moli* is reinterpreted in this way as the matrix verb and *gwneuthur* 'to do' is the auxiliary (or light verb), resulting in the derivation in (30).

(29) *Sef a wnaeth pawb yna moli Duw*
sef PRT do.PAST.3S everyone then praise.INF God
'Everyone then did this, they praised God.' (Dewi 4.17)

(30)



From Stage 4 to Stage 5: focus marker reanalysed as adverb

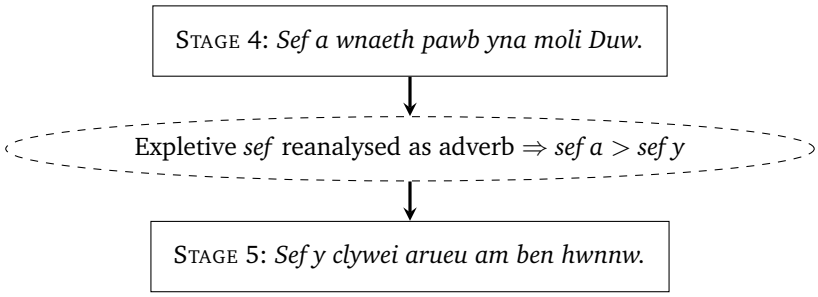


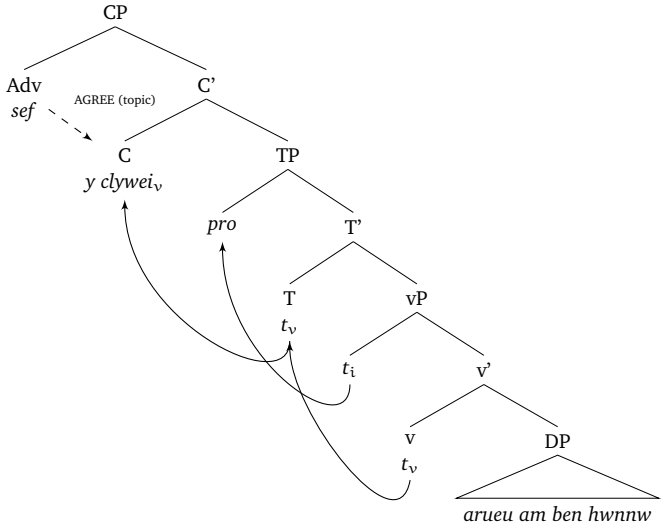
Figure 7.7: Stages 4-5 of reanalysis of *sef*

Eventually the argumental interpretation of expletive *sef* was lost. It was reanalysed as an adverbial element base-generated in SpecCP. Subjects could then move to SpecTP just as they did in any other adjunct-initial Abnormal Sentence (see next section). Adverbs, like all other adjuncts, trigger the pre-verbal particle *y* in the C-head, instead of the particle *a* following argumental DPs as shown in the examples in (31) and the derivation in (32):

(31) a. *Sef y clywei arueu am ben hwnnw*
sef PRT hear.PAST.3S arms on head that.one
'He could feel armour on that one's head.' (WM 54.28)

b. *Sef y kynhelleis inheu y gyuoeth*
sef PRT withhold.PAST.3S I his dominions
'I withheld his dominions.' (WM 394.42 - Watkins 1997:587)

(32)

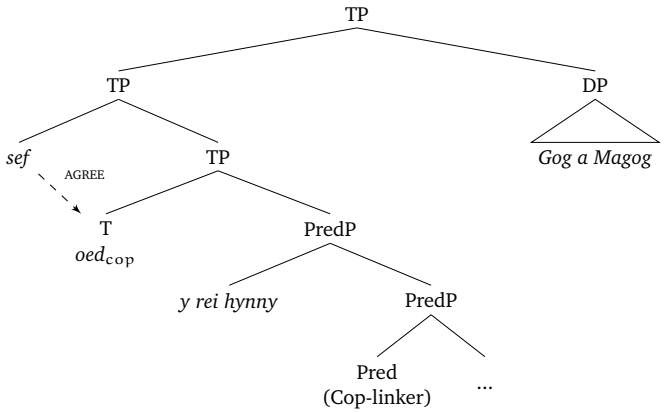


Other *sef*-constructions

The phonological reduction of the copula allowing all subsequent reanalyses described above, also triggered reanalyses of a different kind, creating a further range *sef*-constructions. The cascading pattern of reanalyses described above was specifically possible because of the large number of sentences with subjects consisting of a headless relative (as shown in the section on Stage 2 to Stage 3 above). If the relative clause consisted of a copular clause itself as shown in (33), it could give rise to a further type of reanalysis. Here too the verb from the relative clause could be reinterpreted as the matrix verb as shown by the derivation in (34).

(33) *Sef* ***oed*** *y rei hynny Gog a Magog (...)*
sef be.PAST.3S the ones DEM.P Gog and Magog
'That's what those were, Gog and Magog (...).' (DB 29.11.12)

(34)



Merge of the verb in the T-head could be internal or external, since the verb *bod* 'to be' also functioned as an auxiliary in Middle Welsh. Further movement to the C-head is string-vacuous in this sentence, which is why I only show the TP. The preverbal particle *a* is usually analysed as a complementiser attracting the verb to the C-head, but this *a* could be dropped before *oed*, the imperfect form of the verb *bod* in Middle Welsh. Therefore, with the evidence we have at present we cannot prove it moves up to C or remains in T.

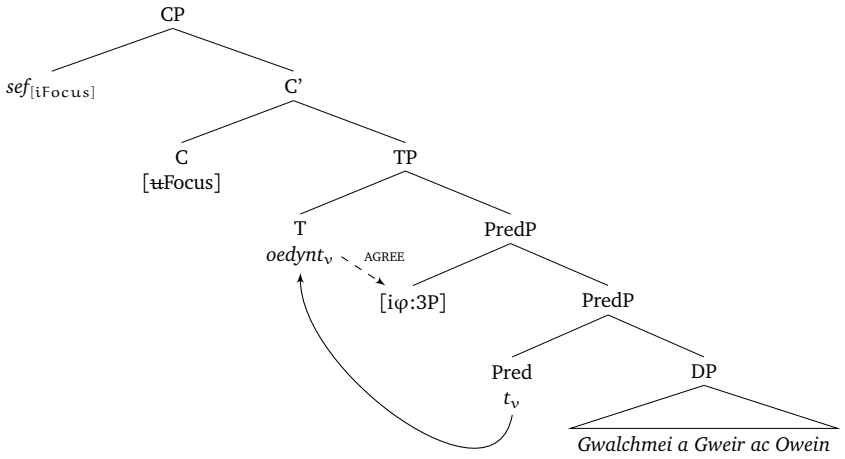
An example with the present-tense verb form *yw* 'is' would now also be a possibility. Note that this could not have been the original form because in the present tense, the verb *bod* 'to be' has a special relative morphology *yssyd* 'that/which is'. Once the verb, either the imperfect form of *bod* or any other verb, was reanalysed as the matrix verb, the medial form of the verb 'to be' could be merged in the T-head as well. This new *sef*-construction with *sef yw/oed...* is called the 'parenthetic-explanatory clause' in traditional Welsh grammars (cf. T. A. Watkins (1997:580-581)):

(35) *Sef **yw*** *honno gwreic doget urenhin*
sef be.PRES.3S DEM.FS wife Doged king
'That's who she is, king Doged's wife.' (WM 453.17 - Watkins 1997:580)

With the advent of medial copular forms like *yw* 'is' above, a further reanalysis could take place: the rise of (dropped) pronominal subjects, as shown in (36). In this construction, *sef* is not interpreted as the expletive. It is externally merged as a focus marker in the specifier of the CP. The verb agrees with the (empty) pronominal subject, as shown in (37). Just as in the above-described stages of reanalysis, here too, the predicate now no longer needs to be in an adjoined position; it can be interpreted in the complement-position of the predicate phrase.

(36) *Sef oed**ynt*** *Gwalchmei (...) a Gweir (...) ac Owein*
sef be.PAST.**3P** Gwalchmei (...) and Gweir (...) and Owein
'That's who they were, Gwalchmei (...) and Gweir (...) and Owein.' (WM 118.19 - Watkins 1997:581)

(37)



Conclusion *sef*-constructions

In this section I presented a detailed analysis of every stage of the process of grammaticalisation of the *sef*-construction in Middle Welsh. For each of the different stages, I presented the characteristics of the ambiguous structures that led to a cascade of new reanalyses. The original trigger was argued to be the phonological erosion of the copula (as already noted by T. A. Watkins (1997)). The predicate + complement *ys ef* first merged into one lexical item that could be externally merged as the expletive in SpecTP or as a focus marker in SpecCP. The relative verb in the complex subject could then be reinterpreted as the matrix verb. From an information-structural point of view, all conditions and characteristics of the original

identificatory focussed predicate were lost. There was no longer a requirement to link the construction to the preceding context and the identificatory interpretation of the predicate as well as its focus marking were lost (semantic bleaching). The *sef*-construction then came to be used in continuous narratives and the variant with the auxiliaries *gwneuthur* 'to do' and *cael* 'to get' became stock phrases. Finally, *sef* lost its argumental status as an expletive and was recategorised as an adverb.

Many of these different forms of the *sef*-construction appear in the same period, sometimes even in the same texts. There has undoubtedly been a period of overlap. We can establish the relative chronology of the different stages in the grammaticalisation process, but since we lack the necessary philological data, it is very difficult to establish a more accurate date for each of the above-sketched stages of reanalysis. There is, however, some supporting evidence for the relative chronology from the *Red Book* of Hergest. The scribe of this manuscript (written around the year 1400) is generally considered to have 'modernised' the text he copied into the *Red Book*. The original was lost, but other older copies of these texts exist, for example, in the *White Book* of Rhydderch, which formed the basis of the present annotated corpus. In comparing certain parallel passages from the *White Book* and the *Red Book*, we see that the 'modernised' *Red Book* more often employs what I described above as the fifth stage of the grammaticalisation process: the adverbial form of *sef* followed by the particle *y* (rather than *a*).

(38) a. *Sef* ***a*** *gausant yn eu kynghor rodi y moch e Wydyon*
sef Csp got in their council give.INF the pig to Gwydyon
'This is what they got in their council: give the pig to Gwydyon' (*White Book*)

b. *Sef* ***y*** *kawssant yn eu kynghor rodi y moch y Wydyon*
sef PRT got in their council give.INF the pig to Gwydyon
'Then giving the pig to Gwydyon was what they got in their council.' (*Red Book*)

To conclude, the identificatory copular clauses with focussed predicates changed dramatically over the Middle Welsh period. Various different forms of this *sef*-construction were found alongside each other, but a careful analysis reveal a clear pattern of a step-by-step reanalysis, with each change triggering the next stage of the process. This relative chronology of the complex grammaticalisation process is to a certain extent confirmed by philological evidence in the form of earlier and later manuscript forms of the same texts.

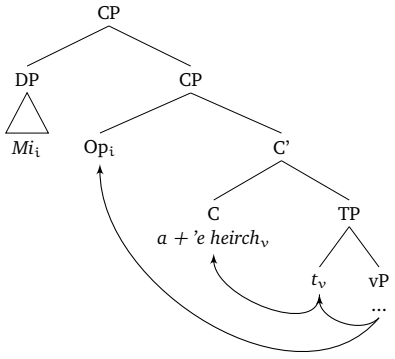
7.3.2 Reanalysis & Extension in the rise and fall of V2

In the previous chapter I discussed the two main types of V2-structures found in Middle Welsh: the so-called Abnormal Sentence and the Mixed Sentence. The traditional distinction between the two is based on Information Structure and agreement patterns: Abnormal Sentences *do* exhibit subject-verb agreement and Mixed Sentences never show subject-verb agreement. Formally, the two can only be

kept apart if the subject preceding the verb is a non-third-person singular pronoun or a plural DP. From an information-structural point of view, the difference is traditionally argued to be Topic (in Abnormal Sentences) vs. Focus (in Mixed Sentences). I have shown in Chapter 6, however, that the IS status of the preverbal constituent cannot be simply divided between these two categories: there are examples of Focus *with* subject-verb agreement and vice versa, examples with preverbal Topics *without* the expected agreement pattern. There are furthermore examples of both agreement patterns in coordinated sentences. A final complication in the data is the 'Complementarity Principle' that holds in all Brythonic languages stating that agreement is only ever found with pronominal elements, never with full DPs. From the point of view of the Complementarity Principle then, agreement with full plural noun phrase subjects in the Abnormal Sentence is unexpected, just as the lack of agreement with pronominal subjects in Mixed Sentences. These Middle Welsh V2-structures are not found in other Celtic languages like Gaulish, Celtiberian or Irish (in any stage of the language). They equally do not occur in Modern Welsh. Modern and Middle Breton as well as Middle Cornish do exhibit the non-agreeing V2 structures equivalent to the Middle Welsh Mixed Sentence. The Abnormal Sentence with subject-verb agreement, however, seems to be a Middle Welsh innovation that was lost again in the Early Modern Welsh period. In the previous chapter, I proposed structures for these 'unexpected' patterns in Abnormal and Mixed Sentences. In the Mixed Sentence, SpecCP is occupied by the relative operator yielding default third-person singular inflection on the verb as shown again in (40):

(39) *Mi a 'e heirch.*
I PRT 3FS seek.3S
'(it is) I who seek her' (Mixed Sentence - WM 479.24)

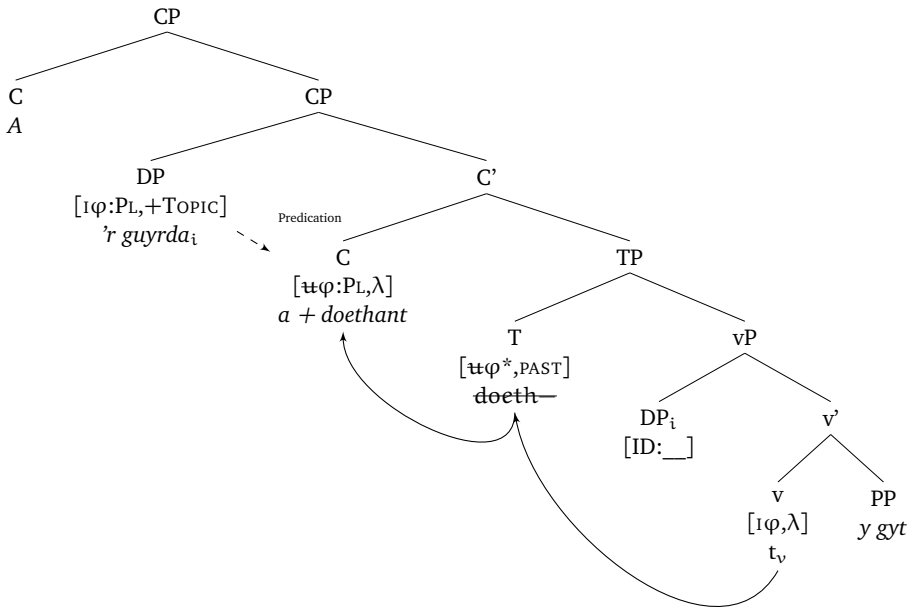
(40)



Plural DPs in agreeing Abnormal Sentences are base-generated in SpecCP. The C-head carries a λ-feature that ensures a predication relation with the DP-topic in its specifier through which agreement can take place. The DP-topic is coindexed with a minimal pronoun subject (a DP without φ-features: [ID:_]).

(41) *A 'r guyrda a doethant y gyt*
and the nobles PRT come.PAST.3P together
'And the nobles came together' (Abnormal Sentence - PKM 90.27)

(42)



The main question from a diachronic syntactic point of view is: where does the Abnormal Sentence with subject-verb agreement come from? Although some Welsh grammarians (e.g. MacCana (1973) and Fife (1991)) have argued that this was merely a literary phenomenon in Middle Welsh, Willis (1998) convincingly argues these V2-structures must have been part of spoken Middle Welsh as well. His arguments are based on language-internal complexity of the V2-rule in various parts of the grammar that would have been hard, if not impossible, to learn as a stylistic feature. Breton and Cornish furthermore also exhibit V2-structures (without subject-verb agreement), so V2 grammar is likely to be inherited from their Common Brythonic ancestor.

A further diachronic question then remains: where do these V2-structures (with and without agreement) come from in general? Richards (1938) and D. S. Evans (1968) already hypothesised that the origin of these Brythonic structure lies in the cleft sentences with contrastive focus. The cleft was followed by a relative clause, introduced by the relative particles *a* or *y*, the exact same particle found in the Mixed and Abnormal V2 orders. Through a process of semantic bleaching, the function of contrastive focus was extended to topics and this then became the basic word order pattern in Middle Welsh (in which the preferred Insular Celtic verb-initial order also found in Irish was lost).

In this section I explore this hypothesis further by examining each of the required syntactic reanalyses and extensions in detail to trace the origin of the Abnormal Sentence. Within the framework of the Minimalist Program, I provide the triggers and linguistic context of every stage in the process that created the right environment for the syntactic reanalyses and extensions we find. In order to describe the first steps that can only be found in reconstructed stages of the languages, it is important to take the sparse Old Welsh data available to us, as well as cross-linguistic evidence from Middle Breton and Middle Cornish into account. Although the focus lies on the rise of V2 structures in Middle Welsh, in the final part of this section I also shed some light on the subsequent loss of V2 with evidence from the 1588 Bible translation.

Overview of syntactic reanalyses & extensions

Figure 7.8 shows an overview of each of the different stages in the process with a description of the possible word order patterns found at that stage specified in the same box. In the following dashed ellipse I describe the trigger(s) that led to a specific change. Any changes in the form of loss/gain of word order patterns in the next stage are presented in the next box. Some of these new patterns may in turn lead to further reanalyses and extensions, until they finally lead to the fifth stage representing Early Modern Welsh when evidence for the acquisition of V2 dropped and the Abnormal Sentence was lost. The Mixed Sentence with contrastive focus on the initial constituent is the only V2-pattern left in positive declaratives in Modern Welsh.

The first stage represents a language that can be reconstructed as the predecessor of Brythonic: 'pre-Common Brythonic'. Following Newton (2006) and Lash (2011), I assume that Insular Celtic had previously lost the articulated CP that was still found in Proto-Indo-European (based on evidence from syntactic reconstruction of Greek, Vedic, Hittite and Latin).

I discuss the labels or languages matching the following stages up to Early Modern Welsh in the context of cross-linguistic evidence from Middle Breton and Cornish. Some word order patterns occur in several stages until they are completely lost or reanalysed. The patterns with optional merger of adjuncts and hanging topics in the C-domain resulting in V2, V3 and V4 orders in Insular Celtic and pre-Common Brythonic, for example, remained until they were replaced by the V2-structures with preverbal particles *a* and *y* in the C-head.

In the same way, patterns with sentence-initial *y(d)* were present from the grammaticalisation of the particle before Stage 2 until Modern Welsh, although during the Middle Welsh period the context in which this sentence-initial *y(d)* was found narrowed down to periphrastic constructions with the auxiliary form of the verb *bod* 'to be'. In the following sections, I discuss each of the stages and the triggers for reanalysis and extension in detail.

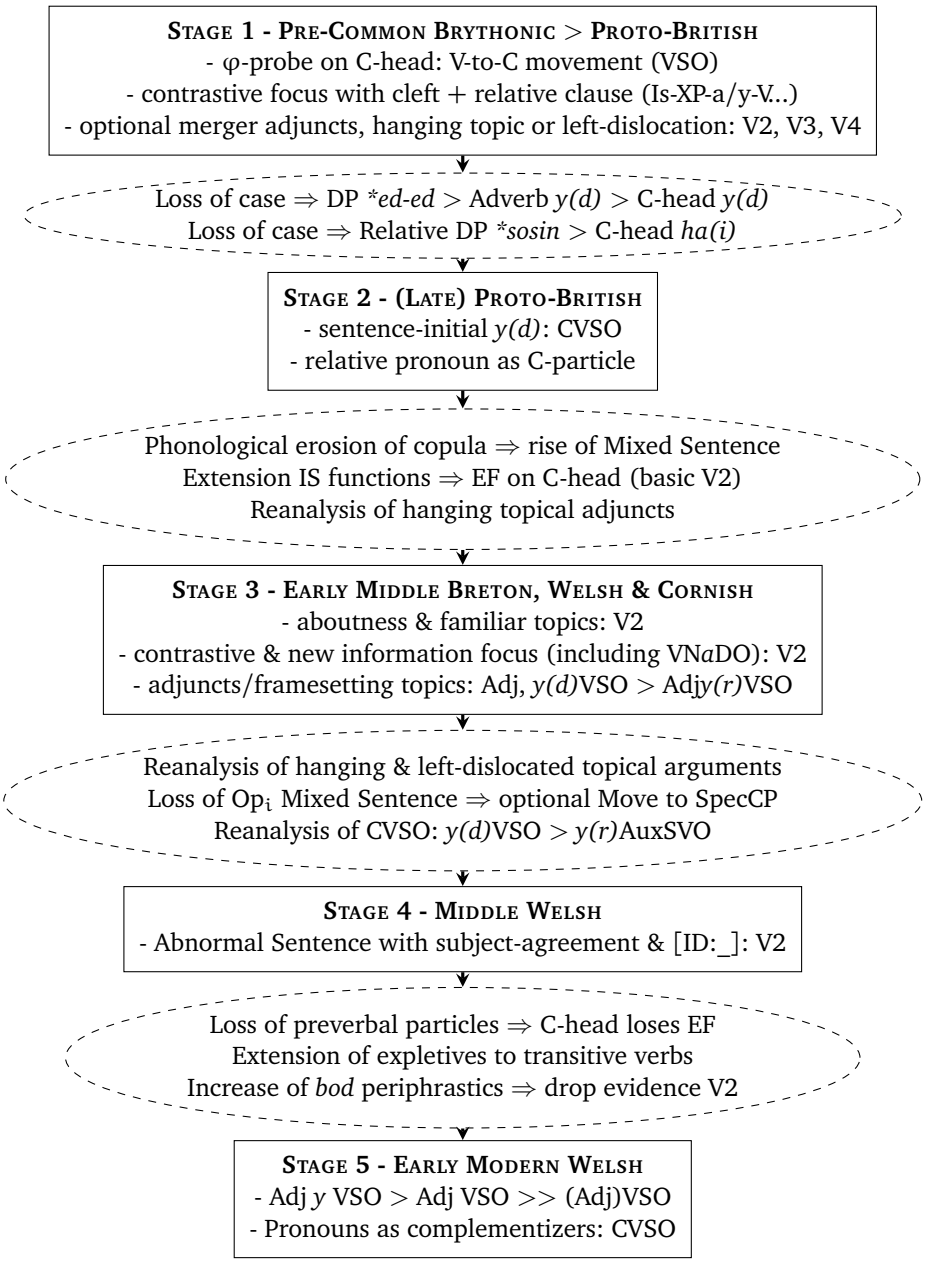


Figure 7.8: Rise & fall V2 from Pre-Common Brythonic to Early Modern Welsh

From Stage 1 to Stage 2: Loss of case

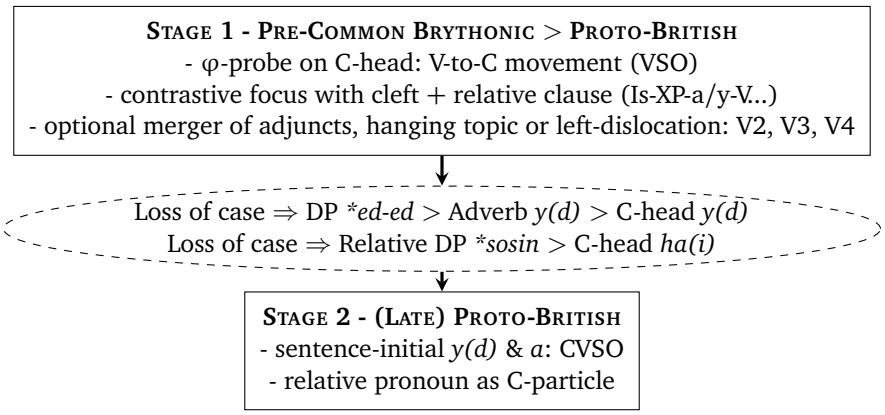


Figure 7.9: Syntactic changes from Stage 1 to Stage 2

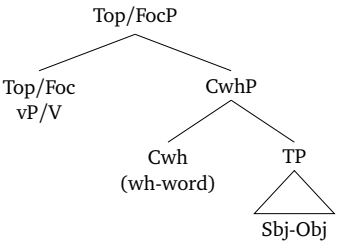
For the first two stages in the process of reanalyses and extensions sketched above we have no written evidence. Stage 1 represents the situation of the language described as Insular Celtic or shortly thereafter, what might be described as 'Pre-Common Brythonic'. This is the form of Celtic spoken in the British Isles before the split of the Irish and the British branches of the Celtic language family. Stage 2 represents the next phase of Common Brythonic, the predecessor of Welsh, Breton and Cornish. Both of these stages can only be described to a certain extent, by means of reconstruction. As discussed in section 7.2.3 above, syntactic reconstruction presents more difficulties than the reconstruction of phonology or morphology. The correspondence problem in particular limits the parts of the grammar that can be reconstructed to those functional items and features that are phonologically overt in the daughter languages. Beyond that, we can still compare syntactic structures and make reasonable assumptions based on plausible patterns of grammaticalisation, reanalysis and local directionality.

Newton (2006) compares the C-domain of Old Irish with that of other Indo-European languages like Greek, Vedic, Sanskrit and Hittite. She concludes that Vedic and Hittite only allow two constituents in the left periphery of the clause: in the Topic/Focus head and the C[+wh] head. Greek and Latin on the other hand allowed multiple topics in the C-domain. Proto-Indo-European as well as Proto-Celtic thus seemed to have a C-domain consisting of at least two functional heads: Top/Foc and C (or 'C[+wh]' as Newton calls it). In the stage of the language she calls 'Pre-Old Irish', this (mildly) articulated CP was lost via "clause truncation". This truncation was established by the reanalysis of relative operator XPs in specCwhP as heads of specCwhP and subsequently as affixes on obligatorily fronted verbs, preverbs or negative elements. The triggering diacritic on Top/Foc was reanalysed as an obligatory movement feature resulting in a 'filled-C condition'. The clause-

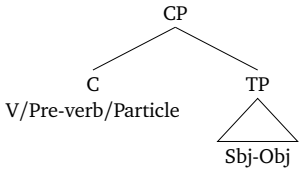
marking suffix **es* linked the verb to the C position. This acquisitional cue then resulted in a reanalysis as V-to-C and pre-verb-to-C movement and conjunct and negative particles occupying the C-head. She then links this new configuration to the development of the Absolute and Conjunct verbal paradigms. Sample derivations for the described sentence structure in PIE and Pre-Old-Irish are given in (43) and (44) below.

Proto-Indo-European ⇒ Pre-Old-Irish

(43)



(44)



The question is: how does Insular Celtic fit in this picture? Does Insular Celtic have an articulated CP like PIE or was this structure already reanalysed in the way Newton has reconstructed for Pre-Old-Irish? The reconstruction of an articulated CP in PIE is based on the possibility of the occurrence of multiple topics or foci alongside other elements in the C-domain (e.g. wh-phrases in CwhP). If we find examples of this in the Brythonic languages, this would be a strong argument to reconstruct an articulate CP in Insular Celtic. The CP truncation could then be postulated as a Pre-Old-Irish innovation only.

As discussed in Chapter 6, in Middle Welsh it was impossible to have both a Topic as well as a Focus constituent preceding the verb. This constraint provides evidence for the strict V2-nature of Middle Welsh word order. The extant data in Old Welsh is extremely limited. Most examples of declarative main clauses exhibit verb-initial order in Old Welsh.

(45) a. *Prinit hinnoid iiii aues*
buy.ABS.3S that four birds
'That buys four birds.' (Old Welsh - Ox 1 B v.234)

b. *Rodesit Elcu guetig equs.*
give.PAST.3S Elcu after horse
'Elcu then gave a horse.' (Old Welsh - Chad2)

There are, however, examples of multiple constituents preceding the verb in Old Welsh yielding V3 or V4 orders, as shown in (46). In these examples, the initial constituents are in fact hanging or left-dislocated topics or adjuncts. As such, they do not provide evidence for an articulated CP. These types of V3 orders are found in Middle Breton and Middle Cornish as well, as shown in (47) and (48) respectively.

(46) a. *Mi telu nit gurmaur*
1S retinue NEG.be.3S very.large
'My retinue, (it) is not very large' (Old Welsh - Juv 3)
b. *Ir pimphet eterin diguormechis lucas hegit hunnoid ...*
the fifth bird add.PAST.3S Lucas go.ABS.3S that.one ...
'the fifth bird that Lucas added, that one goes...'(Old Welsh - Ox 1 B v.234)

(47) a. *breman a crenn me a gouchemen dit*
now PRED express I PRT ask.3S to.2S
'now expressly I ask of you' (Middle Breton - N240)
b. *monet a pret me a preder*
go.INF PRED early I PRT plan.3S
'to go early I plan' (Middle Breton - N64)

(48) a. *Oma ty a ra pedry*
here you PRT do.3S rot
'Here you shall rot.' (Middle Cornish - BMer 3577)
b. *In crist ihesu ny a greys*
in Christ Jesus we PRT believe.3S
'In Christ Jesus we believe.' (Middle Cornish - BMer 1210)
c. *Duk kernov hag oll y dus indan ou threys me as glus.*
Duke Cornwall and all 3MS men under 1S feet I PRT.3P crush
'The Duke of Cornwall and all his men under my feet I shall crush them.'
(Middle Cornish - BMer 2397)

The basic word order in Old Irish was VSO, but similar V2 constructions can be found, as shown in (49):

(49) a. *Cech mab uilc robai ind Éire dochoid chuca.*
every son evil be-rel.PAST.3S in Ireland come.PAST.3S to.3P
'Every son of evil who was in Ireland, he came to them.'(Dindshenchas of Emain Macha - MacCana 1973:96)
b. *Mortlithi márlóchet di doínib dingbatar*
great.plagues great.lightnings from people keep.PRET.PASS.PL.CONJ
'Great plagues and great lightnings are kept from the people.' (AM §12)

It appears then, that in both Brythonic and Irish a specific set of V2, V3 or V4 orders were allowed alongside the basic verb-initial order. There are no overt functional items we can reconstruct for Proto-Insular Celtic, so a perfect correspondence in the form of a double-cognacy condition is impossible to find. We can only compare the extant evidence in the daughter languages and tentatively assume that these V2, V3... orders with adjuncts and hanging and left-dislocated topics were part of the otherwise verb-initial parent language we reconstruct as Proto-Insular Celtic as well. Further comparative evidence could in theory come from Continental Celtic languages like Gaulish in which V2 and V3 orders exist as well.

(50) a. *Ratin briuatiom Frontu Tarbeisonios ieuru*
fort.ACC bridge-dwellers.GEN Fronto Tarbeisu.GEN dedicate.3S
'Frontu, son of Tarbeisu, dedicated the fort of the bridge-dwellers'.(Gaulish OSV - RIG L3)

b. *Buscilla sosio legasit in Alixie Magalu.*
Buscilla this place.3S in Alisia Magalos.DAT
'Buscilla placed this in Alisia to/for Magalos' (Gaulish SOV)

c. *Moni gnatha; gabi budduton imon!*
come.IPV.2S girl take.IPV.2S penis/kiss(?) this
'Come girl; take this penis/kiss(?)!' (Gaulish V1 - St. Révérien[13])

d. *nata vimpi cvrmi da*
girl pretty beer give.IPV.2S
'Pretty girl, bring [me] beer!' (Gaulish OV - spindle-whorl inscriptions)

As the examples in (50) show, however, Gaulish does not only allow hanging and dislocated topics preceding the verb, but also direct objects. These constituents are thus not outside the matrix CP as can be argued for the V2 and V3 structures found in Insular Celtic languages. Instead, these examples show the lack of V-to-C movement (V1 is almost exclusively found in imperatives like (50c)) and cannot tell us much about Insular Celtic. The verb-initial nature seems to be an innovation in the Insular Celtic languages only. For reconstruction of the syntax of Proto-Insular Celtic, we thus have to rely on evidence found in the Irish and Brythonic languages only. In terms of evidence for an articulate CP, we can only reconstruct a phi-probe on C resulting in V-to-C movement and basic verb-initial word order. Since extra-clausal elements such as hanging topics and adverbial phrases can be found in all daughter languages yielding V2, V3 and V4 orders, we can furthermore assume that this was allowed in the Insular Celtic stage of the language as well. It is important to note that allowing these non-verb-initial orders does not exclude the possibility of an articulate CP in Insular Celtic either.

A further reason for Newton (2006) to reconstruct a phi-probe on C yielding verb-initial order in Pre-Old-Irish is the development of the 'double system', i.e. the Absolute-Conjunct paradigms in the verbal system. According to this highly complex system, Old Irish verbs could exhibit different forms according to their position in the sentence. Verbs in absolute sentence-initial position are found with 'absolute' verbal morphology. In Old Welsh, we can still find some examples of absolute verbal endings in the third-person singular. These endings were lost and in Middle Welsh there is no evidence for the Absolute-Conjunct distinction. If we continue to compare Irish and British grammars, we could conclude that this system found in both daughter languages was likely to exist (or to have developed) in their predecessor Insular Celtic as well. However, it is not impossible that the double

[13]There appears to be some discussion on the exact nature and purpose of these sentences with imperative verbs found on spindle-whorls. C. Watkins (1999:542) translates *budduton* as 'penis', but according to Stifter (2011:174n20), the etymology connecting Gaulish *budduton* to Early Irish *bot* 'tail; penis' < *g^{u}*ozdo-* is wrong. Instead, a connection to Middle Irish *bus* 'lip' < **butsu-* "is formally more satisfying". The inscription may thus be of a much more innocent nature, translating 'take that kiss'.

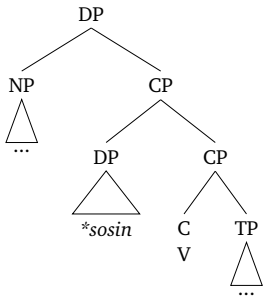
system was a parallel development in both Pre-Old-Irish and Proto-British after the languages split.

To conclude, Old Irish and the earliest attestations in the Brythonic languages show some similar evidence for the development of V-to-C movement in the form of the double inflectional system and sentence-initial particles (declarative and relative). In addition to that, neither Irish nor Brythonic shows explicit evidence for an articulated CP for example in the form of cooccurring topic and focus constituents in preverbal position. However, V2, V3 and V4 orders were possible in both languages with adjuncts and hanging topics. Finally, in Old Irish as well as in Old Welsh, the cleft sentences was the preferred strategy to mark contrastive focus of one single constituent only.

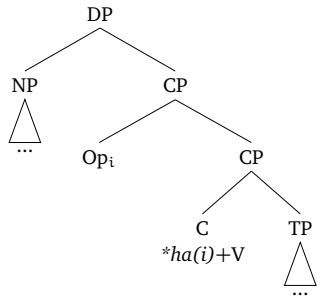
In the course of the Common Brythonic period, some phonological changes triggered further syntactic innovations. Apocope in the 6th century led to a loss of case encoded in morphological endings. As Newton (2006) points out, the loss of morphological endings renders XPs caseless. This in turn can lead to reanalysis of the now indistinct form as a head. This affected the relative pronoun **sosin*, for example. Relative clauses could be formed by the particle **-io* (> Old Welsh absolutive ending *-it*, according to Schrijver (1997)) or with the relative pronoun **sosin* (> Old Welsh *ha(i)*). Since relative pronouns that are inflected (and thus phrases rather than heads) are reconstructed as occupying the specifier of the (lower) CP in PIE (see Newton (2006)), I assume the same position for **sosin* in Brythonic. After case endings were lost the indeclinable relative pronoun was reanalysed as the C-head, as shown in (51) and (52):

Stage 1 relatives ⇒ Stage 2 relatives

(51)



(52)



The loss of case in the neuter pronoun **ed-ed > y(d)* 'it' could have given rise to ambiguity with the adverb *ed* 'thus' (cf. Schrijver (1997:161-176)). Both could be used in relative constructions, for example, in clefts with contrastive focus. Constructions like that are found in Old South-West British (OSWB), shown in (53) and Old Irish as well, as shown in (54).

(53) a. *Is amal it duducer memor.*
COP.3S like thus adduce.REL.3S memory
'It is thus that one adduces memory.' (OSWB - DGVB: Ang477A)

(54) a. *Iss ed dochoid i tir Eogain.*
COP.3S thus go.PAST.REL.3S to land Eogan.GEN
'It is thus that he went into Eogan's land.' (Old Irish - Trip. 150.19)
b. *Is ed fuddera.*
COP.3S this cause.REL.3S
'It is this that causes it.' (Old Irish - Wb 33c12)

The adverbial **ed* > **yd* 'thus' was always found in sentence-initial position. Its reanalysis as a particle in the C-head gave rise to the CVSO orders found in early Breton (*ez*), Cornish (*y(th)* and *as*) and Welsh (*y(d)*) sources.

(55) a. *Yd af i yn agel.*
PRT go.1S I PRED angel
'I shall go as an angel.' (WM 118.27)
b. *Y rodet y march y 'r mab.*
PRT give.IMPERS.PAST the horse to the boy
'The horse was given to the boy.' (PKM 24.4-5)
c. *Y dodym y erchi Olwen.*
PRT come.PAST.1P to seek.INF Olwen
'We have come to ask for Olwen.' (CO 477)

(56) a. *Ez oamp oll, allas, e lastez*
PRT be.1P all alas in suffering
'We are all, alas, in suffering.' (Middle Breton - Nl 328)
b. *Y leferys offeren.*
PRT say.PAST.3S mass
'He said the mass.' (Middle Cornish - BM 4419)
c. *As wrussough cam tremene.*
PRT cause.2P wrong death
'You caused a wrong death.' (Middle Cornish - R40)

There is some further evidence for CVSO orders in this stage in the form of the C-head *a* < *ha(i)* that appears in sentence-initial position in some remnants in Early Welsh poetry, as shown in (57). Schrijver (1997:166) notes, however, that the *a*-particle is merely there to support the cliticised pronoun, which could not occur in sentence-initial position on its own. This could still mean that the particle is the same as the relative marker *a* occupying the C-head in which case we find CVSO order here as well.

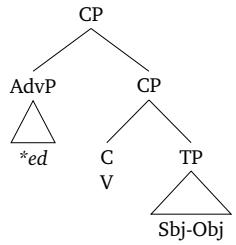
(57) a. *A 's kynnull gwenyn.*
PRT 3MS gather.3S bees
'Bees gather it.' (T 40.8-9)

b. *A 'th uolaf uaur ri*
PRT 2S praise.1S big king
'I praise you, great king.' (HGC II.4)
c. *A 'n gunel iechid*
PRT 1P make.SBJ.3S salvation
'May he make salvation for us.' (HGC III.11)

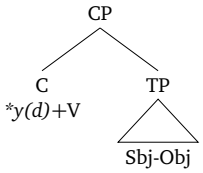
In Early Middle Welsh, the particle *y(d)* could be used in front of any type of verb, although it most commonly appears with *bod* 'to be', *dyuot* 'to come' or *mynet* 'to go'. In Middle Breton, it is only found with those verbs (Hémon, 1975:277). According to Schrijver (1997:171), Old Irish has a very similar adverb *ed* used 'with vb. of motion' (DIL E 58 *ed* VI).

Stage 1 Adv VSO ⇒ Stage 2 CVSO

(58)



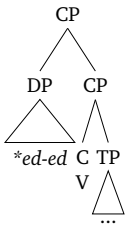
(59)



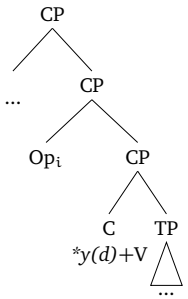
The DP antecedent **ed-ed* also linearly directly preceded the verb in the C-head in all cases. With the loss of the case endings, this phrase could also be reanalysed as a relative marker in the C-head, just like **sosin* in the direct relative clauses. The similarity in form with the adverb-turned-particle *yd* could have given rise to the association of the new relative marker with indirect (rather than direct) relatives, i.e. with non-argumental antecedents. In Middle Breton and Middle Welsh, this relative marker *ez/y(d)* is exclusively found following phrases that are not the subject or object.

Stage 1 (Indirect) relatives ⇒ Stage 2 (Indirect) relatives

(60)



(61)



The main reanalyses between Stage 1 and Stage 2 were triggered by the loss of case morphology due to apocope. Similar to developments in Pre-Old Irish as reconstructed by Newton (2006), phrases occupying the specifier of the CP were reanalysed as particles in the C-head. Relative clauses in the direct predecessors of Welsh and Breton could then be formed in several ways. The relative suffix **-io* became the third-person singular absolute ending (the only absolute ending found in British). Direct relatives were formed by the relative marker **ha(i)* in the C-head. Analogous to this development, the new relative marker **yd* appeared in the C-head following non-argumental antecedents. According to Schrijver (1997), the adverbial phrase **ed* 'thus' was reanalysed as a declarative sentence-initial particle as well. The verb is still moving up to the C-head as well to satisfy the phi-probe. The (relative) particles are like complementisers and have to be merged in the C-head, but they do not carry φ-features and thus cannot satisfy the phi-probe. From a minimalist perspective this means that a similar spec-to-head reanalysis took place here, yielding CVSO orders found in Early Breton and Welsh sources.

From Stage 2 to Stage 3: Loss of copula, rise of V2

A number of changes took place from the first reconstructed stage of the language, (Late) Proto-British, to the earliest attestation in the Brythonic languages. There is evidence for Old Breton and Cornish, but only in the form of lexical glosses (translations) that do not tell us much - if anything at all - about the syntax of these languages. As discussed in Chapter 1, there is more material available in Old Welsh, but even this is very limited. Stage 3 thus also describes the situation as we find in the earliest Medieval stages of the Brythonic languages.

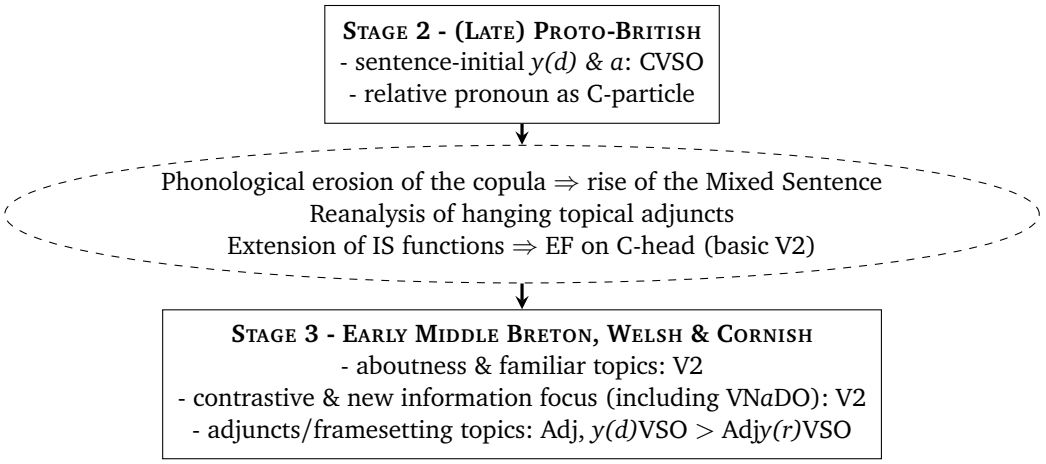


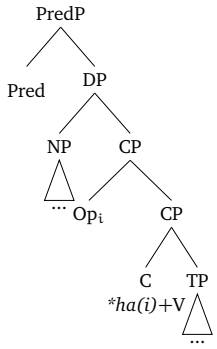
Figure 7.10: Rise & fall V2 from Stage 2 to Stage 3

The phonological erosion of the sentence-initial copula gave rise to the so-called Mixed Sentence, a V2 structure with a relative marker in the C-head carrying an

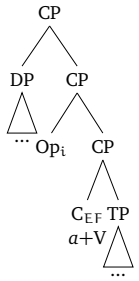
Edge Feature. After the loss of the copula, the sentence was no longer interpreted as a relative clause and the C-head acquired an Edge Feature to ensure its specifier to be occupied at all times.

Cleft Sentence ⇒ Mixed Sentence

(62)



(63)



Initially, the constituent preceding the verb could only be contrastively focussed, but now this information-structural restriction is lost with the emergence of EF on the C-head. Apart from contrastively focussed constituents, contrastive topics can now occupy this first place in the sentence. This was then extended even further to include aboutness and familiar topics until the SpecCP position was a generic position for constituents bearing any kind of IS feature. Non-contrastive focus like new information focus is now also associated with this position. Verbal nouns (with their internal arguments) also belonged to this category now. These sentences with initial verbal nouns (VNs) followed by the inflected form of the verb 'to do' are also frequently found in Middle Breton and Middle Cornish and can were thus likely to exist in Late Proto-British as well, as shown in examples (64), (65) and (66) below.

(64) a. *A dechreu a wnnawn o gyfreith gwlat*
and start.INF PRT do.IPV.1P from law country
'And let us start from the Law of the Country.' (Laws 30)

b. *Agori y drws a oruc ef.*
open.INF the door PRT do.PAST.3S he.
'He opened the door.' (PKM 22.22)

(65) a. *Leuskel a ra hon lestr eun tenn kanol*
fire.INF PRT do.3S 1P boat a gunshot
'Our boat fires a gunshot.' (Middle Breton - MBJJ p.33)

b. *Gervel e zaou vevel a reas ar ronfl.*
call.INF 3MS two servant PRT do.3S the ogre
'The ogre called his two servants.' (Middle Breton - MAV p.34)

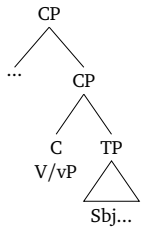
(66) a. *Ty a wra y les.*
you PRT do.3S 3MS width
'You make its width.' (Middle Cornish - O.958)

b. *Oma ty a ra pedry*
here you PRT do.3S rot
'Here you shall rot.' (Middle Cornish - BMer 3577)

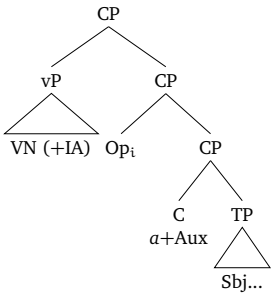
The verb phrases already found in sentence-initial position in earlier stages of the language were now reanalysed as the first constituent in SpecCP with an auxiliary verb in the C-head according to this new V2 requirement.

Celtic verb-initial ⇒ MW VNaDO

(67)



(68)



A further change that took place in Late Proto-British was triggered by the CVSO orders with sentence-initial particle *y(d)* (Breton *ez*, Cornish *y(th)*) in the C-head. Adjuncts in the form of adverbial or prepositional phrases that were originally directly merged outside the matrix CP as hanging topics could now be followed by such a matrix CVSO clause. With the new EF on the C-head, these clause-initial adjuncts could be reanalysed occupying the specifier position of the CP: Adj, *y(d)*VSO > Adj*y(r)*VSO.[14] Schematically, the reanalysis looked like (69) resulting in examples with sentence-initial adjuncts functioning as frame-setting topics followed by the particle *y(r)*, as shown in (70).

(69) $[_{CP}$ PP/AdvP $[_{CP}$ *y(d)* + V $[_{TP}$ ...]]] > $[_{CP}$ PP/AdvP *y(d)* + V $[_{TP}$ ...]]

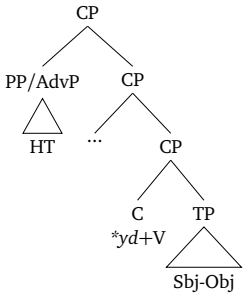
(70) a. *A thrannoeth y talwyt y ueirych idaw.*
and next.day PRT pay.IMPERS.PAST 3MS horses to.3MS
'And on the next day his horses were paid to him.' (PKM 34.23)

b. *Yn Aber Cuawc yt ganant gogeu.*
in Aber Cuawg PRT sing.3P cuckoos
In Aber Cuawg the cuckoos sing.' (CLlH 23.5)

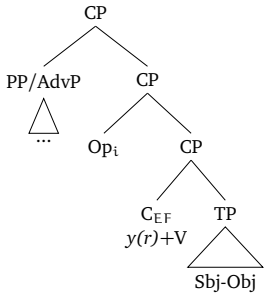
[14]According to Schrijver (1997), *y(d)* changed to *y(r)* in Middle Welsh.

Hanging topic ⇒ Adjunct V2

(71)



(72)



To conclude, the phonological erosion of the copula resulted in a rise of the V2 orders in the so-called Mixed Sentence. The C-head was occupied by the former relative markers *a* or *y(r)* depending on the function of the XP in SpecCP. From an information-structural point of view, there was an extension of the sentence-initial position from contrastive focus to contrastive topic, new information focus and finally also aboutness and familiar topics. The SpecCP position was obligatorily filled by an XP with any of these IS functions because the C-head gained an Edge Feature to attract the verb yielding the preferred verb-second orders in Middle Welsh. When these structures were no longer associated with their relative origin, the Operator that had moved from an adjunct position lower down in the clause to SpecCP was lost and replaced by the PP/AdvP adjuncts.

From Stage 3 to Stage 4: rise of the Abnormal Sentence

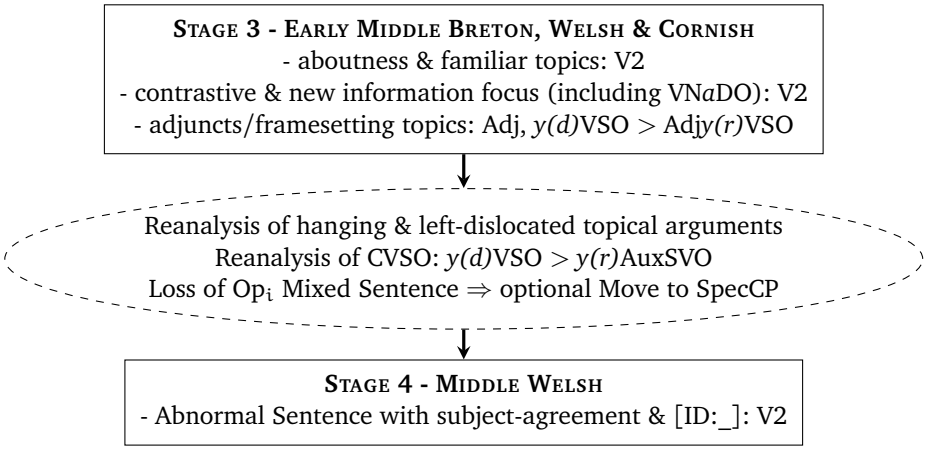


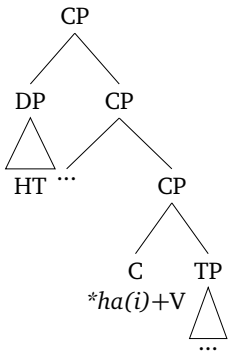
Figure 7.11: Rise & fall V2 from Brythonic to Early Modern Welsh

The changes that took place next resulted in the situation we find in most Middle Welsh literature. The most striking innovation was rise of subject-verb agreement in the so-called Abnormal Sentence. The reanalysis of adjunct phrases formally located outside the matrix CP as hanging topics paved the way for a further reanalysis of arguments as well.

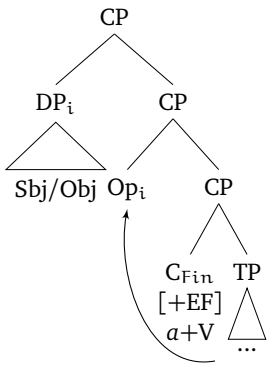
First of all, argumental hanging topics (HTs) that were originally generated outside the matrix CP could be reanalysed as subjects or objects of the matrix. These argumental DPs then occupied the specifier of $C_{Fin}P$, just as their adjunct counterparts.

Hanging topic ⇒ Subject/Object V2

(73)



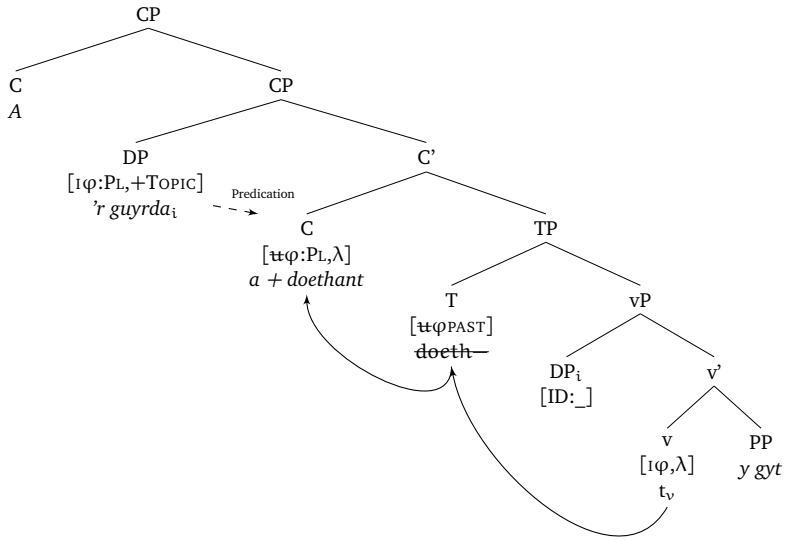
(74)



The original relative markers *a* and *y(r)* that now occupied the C-head had been reinterpreted as positive declarative markers. There was no longer a need to postulate a relative Operator in SpecCP and therefore this was eventually lost as well. Instead, a minimal pronoun [ID:_] entered the derivation as the External or Internal Argument of the verb. With the loss of the relative operator, the base-generated XP in SpecCP could enter a predication relation with the C-head. In addition to the phi-probe and the Edge Feature, the C-head now also bears a λ feature linking the verb in the C-head to the subject DP in its specifier through which the agreement morphology on the verb could be realised. The derivation of these kinds of 'topicalised' Abnormal Sentences is shown again in (76):

(75) *A 'r guyrda a doethant y gyt.*
and the nobles PRT come.PAST.3P together
'And the nobles came together' (Abnormal Sentence - PKM 90.27)

(76)



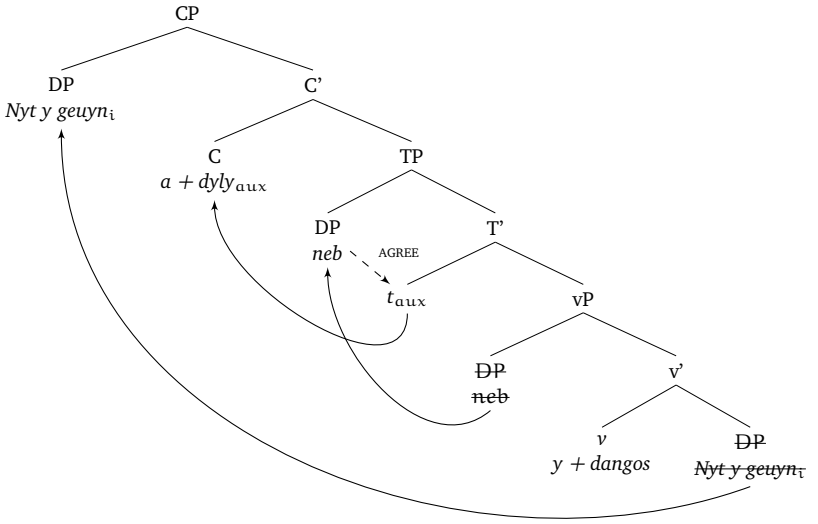
A further development after the loss of the relative operator was the possibility of Moving constituents to SpecCP, rather than externally Merging them in the C-domain with a coindexed minimal pronoun as shown in the Abnormal Sentences above (see Chapter 6 on the minimal pronoun and λ predication in these constructions). Plural DP subjects like *y gwyrda* 'the noblemen' in (75) could not be derived in this way, for the plural agreement goes against the Complementarity Principle that was already well-established in the language by this time. However, pronominal subjects (the most commonly found type of sentence-initial subject) could be analysed either way: both a movement and a base-generated strategy with a minimal pronoun would yield the expected subject-verb agreement as long as no 'trace' of movement is spelled out in the form of an echo pronoun.

From an information-structural point of view, aboutness topics like the full DP subject *y gwyrda* seem to be externally merged at all times, whereas familiar topics like the pronominal subjects could also be internally merged. Constituents representing New Information like verbal nouns or direct objects were gradually lost in the course of the Middle Welsh period. Contrastively focussed constituents are initially externally merged in the typical Mixed Sentence pattern, but in a later stage - after the loss of the relative operator - these could be reanalysed as internally merged constituents as well. This explains the agreement with contrastively focussed pronominal subjects in Late Middle Welsh. Aboutness topics thus seem to be the only constituents towards the end of the Middle Welsh period that were derived via base-generation in SpecCP and coindexed with a minimal pronoun in argument position. These types of topics could remain more associated with their hanging topic origin than familiar topics. Cross-linguistically, there is furthermore

evidence from Italian that indicates a similar base-generation strategy for aboutness topics (cf. Frascarelli (2007)). Constituents with another IS status, like contrastive focus or familiar topics, on the other hand, were fully integrated in the clause and could thus be reanalysed as being derived via a movement strategy instead. An example of such a movement strategy with contrastively focussed constituents from Chapter 6 is given in (78).

(77) *Nyt y geuyn a dyly neb y dangos y*
NEG 3MS.GEN back PRT should.PRES.3S anyone 3SM.GEN show.INF to.3MS.GEN
elynnyon.
enemies
'It is not his back that anyone should show to his enemies' (i.e. 'No one should show his back to his enemies.') (YCM 140.26-7)

(78)



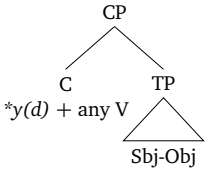
A final syntactic change in this stage was the specification of verbs that were allowed in CVSO contexts. As discussed above, in Breton, CVSO was only possible with verbs of motion and the verb 'to be'. In Middle Welsh there are still some examples with a wider range of verbs like 'to give' or 'to say' etc. In later Middle Welsh, however, the only verb that is allowed to follow the sentence-initial particle *y(d)* is *bod* 'to be'. The particle existed in various forms in front of this verb that was mostly used as an auxiliary, as shown in (79):

(79) a. *Ac y mae matholwch yn rodi brenhinaeth I. y wern*
and PRT be.3S Matholwch PROGR give.INF kingdom I. to Gwern
'And Matholwch is giving the kingdom of I. to Gwern.' (PKM 41.9-10)
b. *Ac y maent yn kyrchu y tir ...*
and PRT be.3P PROGR make.for the land
'And they made for the land...' (PKM 82.16)

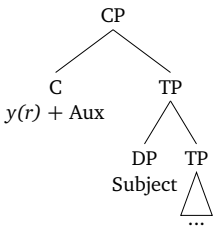
The original CVSO word order pattern thus turned to *y(r)*AuxSVO.

CVSO ⇒ *y(r)*AuxSVO

(80)



(81)



It should be noted that in the above configurations, the C-head does not bear an Edge Feature. If it did, it would trigger the merger of a constituent in SpecCP, which is not what we find in sentences with *y(r)*AuxSVO order. The construction was also found without the particle *y(r)*, but neither of these word order patterns occur frequently in native Middle Welsh tales. The majority of these periphrastic (*y*)AuxSVO orders in the corpus under investigation are found in the 1588 Bible translation in which VSO orders are starting to appear as well. Since there are so few examples in native Middle Welsh literature, it could be argued that those are remnants of the older stage of the language in which C did not yet bear an EF. The increase in frequency in late Middle Welsh could be related to the loss of EF on the C-head again. Alternatively, we have to explain why these Aux-initial orders were possible when C bears a feature that requires its specifier to be filled yielding V2 orders.

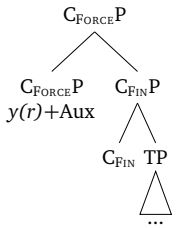
One possibility pursuing this argument could be that the particle *y* and the auxiliary are in fact not in the (same) C-head position, but somewhere higher up in a more articulated left periphery. Recall from the previous section that we had no evidence *for* an articulate CP in Common Celtic or Middle Welsh, but - apart from reasons of economy - we also have no conclusive evidence *against* it. If the (former) relative particle *a* is merged in a lower C-head, say C_{Fin} for example, obligatory merger of XP yielding the observed V2 structures would be in SpecC_{Fin}P. The particle *y(d)* found in absolute sentence-initial position could instead be merged in an even higher position in the left periphery, for instance, the head of C_{Force}. A split-CP analysis like this is in fact proposed by, among others, Roberts (2005) (Tallerman (1998) also proposes multiple layers in the CP, but does not label them as 'Fin', 'Force' or 'Topic/Focus' specifically). The verb *bod* 'to be' in particular then also occupies the highest position in the left periphery. Further evidence for this comes from sentences with negation and subordinate clauses in Modern Welsh (cf. Tallerman (1998), Roberts (2004) and Roberts (2005)).

If this is the case, there are two possible scenarios that account for this particle in C_{Force}: the afore-mentioned two forms that both yielded *y(d)* (the neuter pronoun **ed-ed* 'it, this' and the adverb **ed* 'thus') could actually have resulted in two particles each occupying a different C-head. One of those was reanalysed

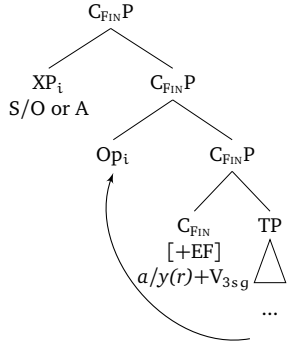
as the head of C_{Force} bearing a phi-probe to attract the auxiliary. The other one was reanalysed as the head C_{Fin} bearing a phi-probe and an Edge Feature yielding the observed V2 structures (just like the other particle in C_{Fin}, the former relative marker *a*). It is important to note in this context that there was another particle *yd* in Middle Welsh (Middle Cornish *ys-*, Breton *ed-*) that was found before present and imperfect forms of the copula that started with a vowel (MW *ydiw*, MC *vsy*, MB *edy* 'is'). According to Schrijver (1997), this particle must be of yet another Celtic source (that nonetheless had the exact same **VdV* structure yielding *yd*). The origin of this particle (and thus its original syntactic function that is of interest to us here) remains obscure.[15]

The exact etymologies of these particles are important if we want to gain a better understanding of Early Welsh syntax, but a comprehensive investigation goes beyond the scope of the present study. Without further evidence from Old Welsh and OSWB sources, their origin might remain 'obscure'. From a syntactic point of view, however, the following two structures were likely to occur alongside each other in Early Middle Welsh: a periphrastic construction with the auxiliary *bod* 'to be' in C_{Force} and a V2 structure with extended IS functions for the sentence-initial constituent (Subject/Object arguments with *a* or Adjuncts with *y(r)*) in SpecC_{Fin}:

(82) *y(r)*AuxSVO



(83) V2



To conclude, Middle Welsh saw the rise of the Abnormal Sentence with subject-verb agreement through the loss of the relative operator and the reanalysis of hanging topics and matrix subjects. The loss of the operator furthermore resulted in a formal split between aboutness topics and constituents with other IS markings. Aboutness topics, for example plural DP subjects, were still base-generated in the C-domain and coindexed with a minimal pronoun in the arguments position of the main clause. Constituents with contrastive focus or familiar topics, on the other hand, were reanalysed as being derived via a movement strategy. Finally, the CVSO order that was possible with all kinds of verbs in early stages of Middle

[15] Schrijver (1997:164) does, however, refer to Pedersen (1913:174, 233) and Morris Jones (1913:288) for what he calls "unconvincing explanations".

Welsh became restricted to constructions with the auxiliary *bod* 'to be'. In these periphrastic sentences, the particle *y(r)* was merged in the head of C_{FORCE} attracting the auxiliary yielding *y(r)*AuxSVO orders as the only remaining alternative to V2 in Middle Welsh positive declarative main clauses. The other declarative particles *a* and *y(r)* occupied the lower C_{Fin}-head bearing an Edge Feature triggering the merger of any XP in its specifier if C_{Force} was not projected (i.e. if the particle *y(r)* associated with C_{Force} was not part of the Numeration.

From Stage 4 to Stage 5: loss of V2

This final stage is without doubt characterised by the loss of V2 word order. The changes involved in this process are described in great detail by Willis (1998) and Willis (2007a). The main triggers for the reanalyses were the loss of the preverbal particles *a* and *y(r)* in the C-head. In combination with the increase in use of periphrastic constructions this led to a significant drop of evidence for the acquisition of V2 word orders.

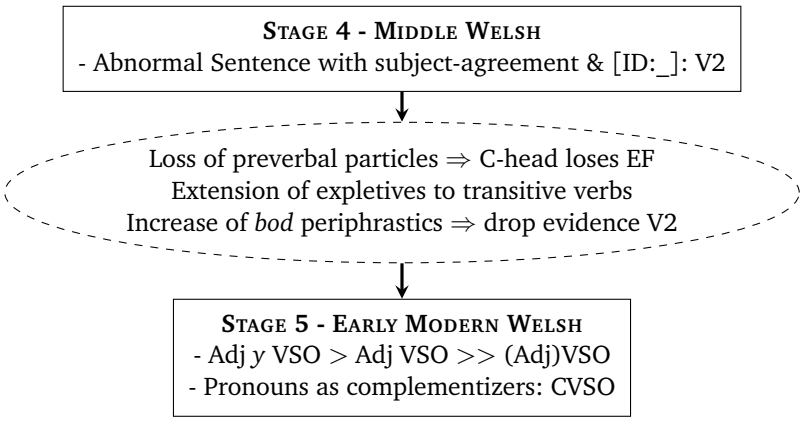


Figure 7.12: Rise & fall V2 from Stage 4 to Stage 5

The loss of the preverbal particles initially resulted in SVO and AdjVSO orders. Recall from Chapter 5 that object-initial sentences were almost completely lost in the late Middle Welsh period, so OVS orders did not arise as the results of the loss of the particles. A further possibility that was more frequently found in the course of the Middle Welsh period put the expletive *ef* in sentence-initial position, even with transitive verbs (see Willis (1998)). Of the SVO sentences, most sentence-initial subjects were pronouns. With the loss of the particle, these pronominal DPs in the specifier of C_{Fin}P were reanalysed as complementisers in the C-head yielding CVSO (again, though now with the former pronouns *mi* and *fe* as C-heads). This type of Spec-to-Head reanalysis was already found in earlier stages of Middle Welsh (the origin of the relative markers *a* and *y(r)*) and is supported by cross-linguistic evidence as well (cf. Willis (2007a)).

Adjunct-initial orders were in turn reanalysed as VSO orders with optional Adverbial or Prepositional Phrases in sentence-initial position. The Edge Feature on the C-head was lost, because children did not receive enough evidence to postulate this feature triggering V2 orders. What constitutes 'enough' in the previous sentence? This brings us back to a point discussed in the introduction: what is the minimum frequency of a cue or trigger needed for a child to postulate a certain grammar? In a language in which the C-head bears a phi-probe, every sentence with a non-subject XP preceding the verb could count as evidence for an EF on the C-head and thus a V2 grammar.

In the Middle Welsh period, there were three different sentence types that could count as this kind of evidence: sentences with initial objects, verbal nouns or adjuncts (adverbial and prepositional phrases). Of those, adjunct-initial orders were most frequently found in almost all Middle Welsh texts in the corpus: as Figure 7.13 shows, these types of non-subject-initial V2 sentences cumulatives make up around 30% of all positive declarative main clauses or even more. In the 1588 Bible translation, however, this is no longer the case: adjunct-initial orders now make up less than 20% and VN-initial and object-initial orders have (virtually) disappeared).

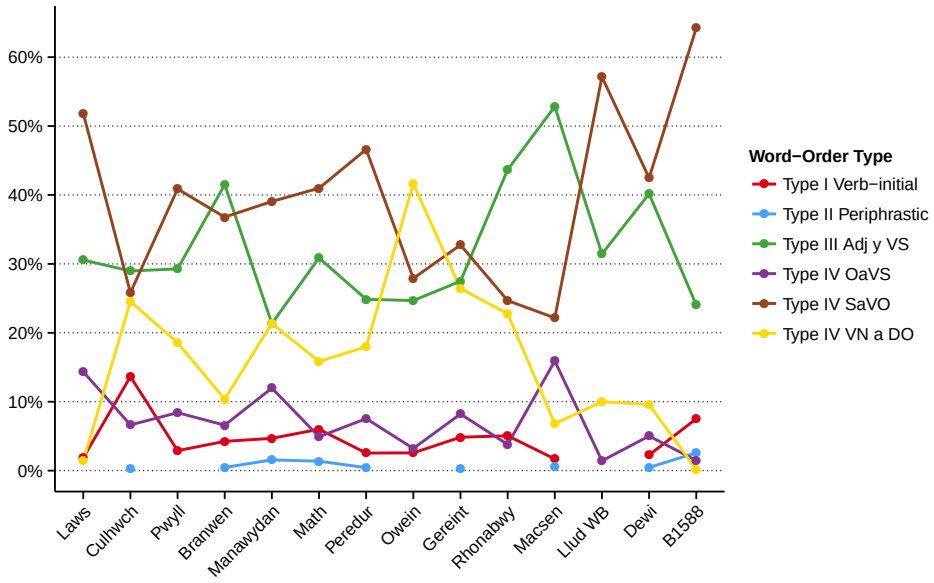


Figure 7.13: Main word order types per text from Early to Late Middle Welsh

Sentences with initial verbal nouns are also frequently found (on average around 20%), but this percentage drops towards the end of the Middle Welsh period. Object-initial orders were never very frequent, remaining around 10%, but again this number drops in the later Middle Welsh texts (BR, LL, Dewi and the Bible translation). If we add up these numbers per text, the 1588 Bible translation already

shows a considerable drop in 'evidence for V2'. Over 60% of the sentences in the Bible are subject-initial, leaving less than 25% adjunct-initial orders. This 25% comes close to the numbers sufficient for V2 acquisition suggested by Lightfoot (1999:154) (30%) and C. D. Yang (2000:114) (23%). With over 60% subject-initial sentences, Late Middle Welsh at first glance looks like it is heading towards SVO like French and English after the loss of V2. Modern Welsh is verb-initial, however, so how did children in the Early Modern Welsh period opt for the loss of the Edge Feature only (keeping the phi-probe and thus verb-initial order)?

Let us first look at the further possible word order types in positive declarative clauses. Object-initial and verb-initial orders in the periphrastic construction with verbal nouns followed by the auxiliary *gwneuthur* 'to do' have almost disappeared completely by the time of the 1588 Bible translation. Absolute verb-initial orders and periphrastic orders with the auxiliary *bod* 'to be' (following the particle *y(r)* in the C_{Force} head) are on the rise, although together they constitute only just over 10% of all positive declarative main clauses. Throughout the Middle Welsh period, however, verb-initial orders were furthermore found in subordinate clauses as well as almost all negative declarative main clauses and yes/no questions. As soon as the pronominal subjects were reanalysed as complementisers, it was no longer necessary to postulate an Edge Feature filling the specifier of C, but the phi-probe on the C-head remained yielding VSO basic word order in Modern Welsh.

Conclusion: the rise & fall of Middle Welsh V2

In this section I have described various processes of reanalysis and extension that led to syntactic innovation from the earliest (reconstructed) stages of the Brythonic languages to Early Modern Welsh. The most striking fact in the history of the Welsh language is that for a period of almost 1000 years (roughly from 600-1600), the grammar seemed to have been defined by a verb-second rule, placing constituents with a specific information-structural status in initial position. Although syntactic change in a generative framework can still be analysed in a parametric context, "V2 grammar" cannot be described as a simple parameter switch. First of all, these verb-second phenomena encompass a wide range of syntactic and information-structural options in the structure of the sentence. This results in varieties within different stages of the language in the case of historical Welsh, but it is also observed in cross-linguistic studies of V2 languages. Not all languages exhibiting a V2 rule have the exact same syntactic structure. This means that a change from 'V2' to 'non-V2' can in fact be the result of a number of smaller reanalyses and extensions in various linguistic domains (see also similar suggestions of changes via small steps by e.g. Haeberli and Ihsane (2015)). In the previous sections, I have given a detailed account of how each of these small syntactic innovations could trigger further extensions and reanalyses, leading to an apparent gradual change in the history of Welsh from verb-initial word order to a preferred V2 order and back again.

I identified possible triggers that led to syntactic innovations, both in the form of reanalyses (e.g. rebracketing or spec-to-head reanalysis) and extensions

(e.g. of the information-structural status of the sentence-initial constituent). I furthermore defined the necessary context in which these changes could take place the way they did, establishing plausible cases for "local directionality" even in reconstructed stages of the language. A prime example of this is the state of the language before the rise of V2. I have argued that a phi-probe on a C-head triggering V-to-C movement, in combination with the existence of clefts indicating contrastive focus as well as optional V2, V3 and V4 orders are a necessary precondition for the development of V2 in the Brythonic languages. In this context, XPs occupying specifier positions in the C-domain (such as the relative pronoun **sosin*) could be reanalysed as functional heads (for instance, triggered by the loss of case morphology due to apocope that turned them into indeclinable relative markers). A change like this is thus wholly in line with Minimalist views on variation stipulated by the Borer-Chomsky Conjecture ("All parameters of variation are attributable to the features of particular items (e.g. the functional heads) in the lexicon." (Baker, 2008:353)). In a hierarchical parametric framework, this would be a 'nanoparametric' change, because it involves the change in the featural make-up of specific lexical items.

The changes that led to generalised V2 in Middle Welsh include further reanalyses in the form of rebracketing of hanging topics to constituents that are part of the matrix occupying the specifier of CP. The extension of the IS function of sentence-initial constituents is a featural change. Alongside an uninterpretable feature probing contrastively focussed constituents, the C-head came to bear a probe for contrastive topics, aboutness topics, familiar topics etc. Along the lines of Minimalist principles of Feature Economy in acquisition, this wide variety of IS-probing features was merged into one generalised Edge Feature (EF) probing any constituent with a specific IS status (i.e. any topicalised or focussed constituent could now be merged in SpecCP).

The combination of a phi-probe and an EF on the C-head yields the so-called 'V2 constraint' that was generalised in this way in Early Middle Welsh (and Breton and Cornish). With an abundance of non-subject-initial V2 orders (in the form of object-initial, verbal-noun-initial and adjunct-initial orders), Middle Welsh children could acquire the V2 rule without any problems. I furthermore argued that the Edge Feature was specifically postulated to be on the lower C-head: $C_{\texttt{Fin}}$, because of the alternative auxiliary-initial periphrastic constructions with the particle *y(r)* in $C_{\texttt{Force}}$. This means that Middle Welsh can be analysed as having reached Stage 3 on the cross-linguistic scale of the Rise of V2 postulated by Wolfe (2015) (see Figure 7.14 below). Modern V2-languages like German or Dutch are characterised by the Edge Feature on the highest C-head: $C_{\texttt{Force}}$. Middle Welsh, however, never reached that stage in the development of V2 in the grammar. The Edge Feature could not be analysed (and thus postulated by children) on the highest C-head, since the periphrastic constructions with the auxiliary *bod* 'to be' were never preceded by other constituents. If Wolfe's 'Stage 3' V2 is a less stable environment for the V2 constraint than his final 'Stage 4' (still existing today in Modern German and Dutch) this could have been a contributing factor in the subsequent loss of V2 in the Early

Middle Welsh period. Much more research in the specific characteristics of each of these stages in a variety of languages is needed, however, before we can draw any such conclusions.

Stage 1	**Stage 2**	**Stage 3**	**Stage 4**
No +EF on C	No +EF on C	C_{Fin} bears an EF	C_{Force} bears an EF
Optional Merger of an XP which is +Top, +Foc, +Neg etc.	Optional Merger of an XP which is +Top, +Foc, +Neg etc.	XP Merger Obligatory	XP Merger Obligatory
CPol, CFoc, CTop probe finite V	C_{Fin} bears an active Phi-Probe	C_{Fin} bears an active Phi-Probe	C_{Force} bears an active Phi-Probe

Figure 7.14: Stages in the Rise of V2 cross-linguistically by Wolfe (2015:44)

Further syntactic innovations in Middle Welsh included the development of the Abnormal Sentence with (unexpected) subject-verb agreement. From the point of view of absolute chronology of the various stages, this is the first structural sign that Middle Welsh is different from its Old South-West British neighbours Breton and Cornish that never developed subject-verb agreement with preposed subjects. A precondition for this further development in Middle Welsh is the existence of sentences with hanging and left-dislocated topics that yielded V2 (and possible V3 and V4 orders in previous stages of the language.[16] These were reanalysed to be in SpecCP position as well satisfying C's Edge Feature, but they were externally merged initially and coindexed with a minimal pronoun lower down in the clause. Apparent subject-verb agreement is the result of the spell-out of the phi-features of the verb in the C-head bearing a λ-feature that allows it to enter into a predication relation with the topical DP in its specifier. It was argued that the situation of Middle Welsh was such that all preconditions were in place allowing this change to happen, including a trigger for the reanalysis of hanging topical arguments analogically to the reanalysis of hanging topical adjuncts. Such an 'analogical trigger' could be viewed as a form of Input Generalisation in which children generalise the structure/interpretation of one construction in all domains or on all levels. This interacted with the extension of IS functions of the sentence-initial constituents at the same time. If we want to answer the question why the situation was such in Middle Welsh and not in Middle Breton or Middle Cornish, a similar thorough investigation of Breton and Cornish word order and information structure is necessary. I leave this - to the extent it is possible with the limited amount of prose data in those languages - for future research at this point.

[16] The Late Latin sources studied by Wolfe (2015) are in Stage 2 of his chronology and these developed into Early Old French, Spanish, Sicilian and Occitan that are argued to be Stage 3 languages with an EF on C_{Fin}. According to this chronology then, Late Latin went through the same process I sketched for Late Proto-British. This might in fact shed some light on the ongoing discussion about language contact after the fall of the Roman empire in Britain (see Schrijver (2002) and Russell (2012) and the discussion on morpho-syntactic similarities in section 7.2.2 above). A detailed investigation of Late British Latin sources is necessary, however, before we can reach any conclusions here.

Since the 'V2 rule' involves at least two different features (both a phi-probe and an EF on the C-head), it is difficult to put this in one single parametric hierarchy. A further complication stems from the fact that the phi-probe is not only associated with attracting the verb, but also with the so-called 'pro-drop' languages. On this highest parametrical level in the hierarchy presented by Biberauer et al. (2014:112), for example, uφ-features are absent from all probes yielding 'radical pro-drop' languages like Chinese or Japanese. Alternatively (and somewhat contrary to their other hierarchies), uφ-features can be present on *all* probes, yielding pronominal arguments, and only then specified to some probes (pro-drop), etc. The phi-probe in the discussion of V2 and verb-initial languages, however, is mainly an indication of verb-movement and can thus be indicative of word order in relation to its subject and direct object. In order to comply with the second condition for V2 (an Edge Feature triggering the merger of an XP to the specifier of its head), we need to complicate the simple hierarchy with further options. A tentative and simplified (i.e. not taking optional/obligatory pied-piping into account, for example) version of this is presented in Figure 7.15:

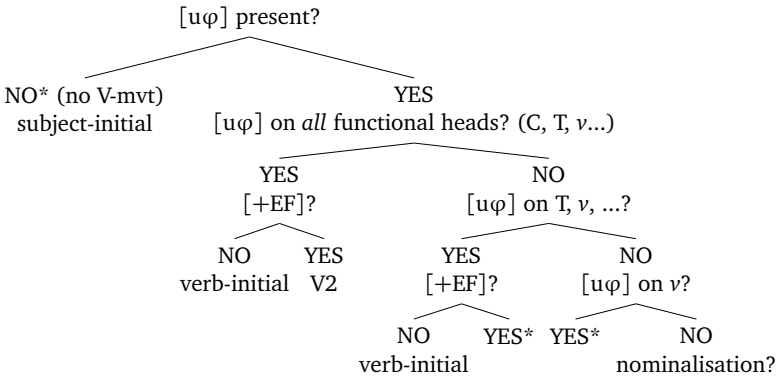


Figure 7.15: Hierarchy for verb-movement via [uφ], including [+EF] yielding V2

This hierarchy shows that a combination of questions need to be answered in order to arrive at V2 word order. In other words, a combination of parametrical settings of the featural hierarchy is necessary to arrive at a grammar with a V2 constraint. Incidentally, if we look at the above tentative figure, we see that the same combination is necessary for verb-initial orders. The subject-initial orders could in theory be further divided into SVO and SOV languages. The asterisk * in the figure here thus actually indicate a link to another parameter hierarchy, namely 'head-finality' (determining, amongst others, OV vs VO orders). Finally, if the verb or V-head is not even probed by little *v*, we could possibly think of languages that involve nominalisation or verb-incorporation. Crucially for our story about the rise and fall of V2 constraints in the grammar, we could arguably insert an extra layer indicating different functional heads in the C-domain. If the Edge Feature is present on the lower C-head (C_{Fin}), for example, the range of languages differs from those

in which EF resides on C_{Force}.

From a diachronic point of view a language can lose or gain two features in this context. The loss of [+EF] on a Macro- or Meso-level is predicted to lead to verb-initial word order (all else being equal), as we see in Welsh. If the phi-probe on the C-head is lost, on the other hand, subject-initial word orders become a possibility as well. The latter arguably happened in the histories of Romance and Present-day English. Much more further research is necessary, however, to test the viability of the above-sketched hierarchy taking more cross-linguistic historical evidence into account. It is furthermore important to investigate the reflexes of possible interaction with other parametrical hierarchies, e.g. the other phi-probe hierarchy for Null Subjects.

For the history of Welsh, the changes in the featural makeup of the C-head facilitated 'the Rise and Fall of V2'. The changes were triggered by the relative markers that came to occupy the C-head and subsequently turned into positive declarative sentence markers. A very similar kind of spec-to-head reanalysis in the CP some 1000 years later then resulted in the loss of the Edge Feature and thus the loss of V2 in Early Modern Welsh.

7.4 Information structure in diachronic syntax

In this chapter I have mainly focussed on structural changes in the history of the Welsh language. One final question that remains concerns the role of information structure in this process of syntactic innovations. In the previous chapter I briefly discussed the 'place' of information structure in the grammar and how it can be encoded in the syntax (rather than other linguistic domains, such as prosody for which we have no historical data). I concluded, following recent Minimalist assumptions that in syntax, information-structural characteristics can be featurally encoded and incorporated as such mainly in the left periphery of the sentence. Do these information-structural features have any influence on changes in the grammar over time? Can they trigger syntactic innovations themselves and/or do they play another role in diachronic syntax?

If we look at the first case study in this chapter on the grammaticalisation of the *sef*-construction, information-structural features definitely played a role in various reanalyses that took place. The original construction only existed in the first place to focus the predicate of identificatory copular clauses. In the course of the process, first the 'identificatory' requirement was lost, leaving a new lexical item *sef* as a specific focus marker. This focussed interpretation was subsequently lost as well and *sef* was eventually reanalysed as the connector of reformulative appositions (like Latin 'id est' still commonly used in abbreviated form in English 'i.e.'). The loss or gain of an information-structural feature like [+FOCUS] could be argued to be the trigger for further syntactic innovations. The question remains, however, what ultimately triggered this loss/gain in the first place.

The same goes for the extension of IS functions of the sentence-initial constituents in verb-second clauses in Middle Welsh. Plausible pathways of extension from, for example, contrastive focus to contrastive topics can easily be identified. The generalisation of probing 'any-IS-marked' constituent rather than probing Focus or Topic specifically is part of a Feature Economy strategy in acquisition. Along such lines, the coexistence of so many different IS features (i.e. Contrastive Topic, Focus, Aboutness Topic, Familiar Topic, etc.) was a necessary prerequisite to postulate a generalised Edge Feature on the C-head. But evidence from languages in which multiple constituents with various IS functions occupy the left periphery shows that this coexistence is not necessarily a trigger for subsequent syntactic innovations. If cross-linguistic diachronic evidence shows that this is a development in more languages, we can get a firmer grip on the role of IS features in this context. For the rise of V2, for example, similar patterns in the history of Romance were discovered by Sam Wolfe. This is a promising start, but much more work is needed before we can reach any final conclusions.

On the basis of much recent literature on diachronic syntax and the case studies presented here in the history of Welsh, we can conclude that information structure definitely plays a role in synchronic variation and thus possible word order patterns. The extent to which it triggers, facilitates or even merely affects changes in the grammar over time is, however, less clear (cf. Taylor and Pintzuk (2015)). As such, this is not a surprising conclusion if we go back to the discussion of possible endogenous and exogenous triggers for language change (see section 7.2.3 and Willis (2016)). In working with historical data (and the extent to which this is available at all) it may not be possible to define the 'ultimate cause' for language change. But in historical syntax, we can describe the exact synchronic state of the grammar in all its detail to explain how and why specific innovations were possible in the first place, how and why they developed in the way they did and why the result is exactly the way we find and not otherwise (cf. Biberauer and Roberts (2015)). Therefore a good understanding of the place of information structure in the syntax of a language as well as a sound methodology to investigate IS functions is important for both synchronic and diachronic research.

7.5 Conclusion

In this chapter I finally turned to diachronic syntax. First of all I discussed various approaches to the study of diachronic syntax, including socio-linguistic variationist, construction grammar and generative approaches. I discussed studies of Welsh historical syntax in these approaches and concluded that they could not give comprehensive accounts or answer all questions in terms of how and why certain changes took place. I argued that adopting a generative acquisitional framework has various benefits in the study of diachronic syntax. First of all it allows us to use insights from various synchronic studies on variation in syntax. The tools and mechanisms tested within the Minimalist Program can furthermore help us define the exact conditions and/or context in which innovation can and cannot occur and

how they can trigger further changes. I have used this to show how innovations were triggered, how children were able to postulate new features or reanalyse the output they are confronted with and why they changed in a certain direction.

I then presented two case studies of syntactic change in the history of Welsh. I described the variation stages and processes of reanalysis and extension and I furthermore examined the role of information-structural features in each of those. The first of these case studies is concerned with a very specific type of focus strategy: identificatory predicate focus. I showed how this construction arose from the cleft construction still found in Old Welsh and how it, after the erosion of the copula, changed in Middle Welsh. First a focus marker *sef* emerged that could be employed in a number of different constructions that were created after reanalysis of the original cleft structure. Then the focussed interpretation was lost and *sef* was reinterpreted as an expletive and, finally, as a linker in reformulative appositional structures ("i.e.").

In the second case study I showed the various stages and innovations involved in the rise of V2 word order in Middle Welsh. A major difficulty in this discussion is the lack of data for the first stages of the language in which the construction originates. This required careful comparison with other Celtic languages such as Gaulish and Irish for the initial stage and other Brythonic languages like Breton and Cornish for the second stage. Since syntactic reconstruction suffers from the correspondence problem of double cognacy (see amongst others, Willis (2007b) and Walkden (2014)), I focussed on the reconstruction of the functional particles in the C-domain that can still be found in the Brythonic languages. I then described the further developments of reanalysis of hanging topics and relative clauses (the 'Mixed Sentence') and the extension of IS functions leading to the postulation of a generalised Edge Feature on the C-head. On the basis of further possible sentence types like the periphrastic construction with the auxiliary *bod* 'to do' in Middle Welsh, I further argued that this Edge Feature must be on a lower C-head, $C_{\texttt{Fin}}$. The phonological erosion of the C-particles in the Early Modern Welsh period finally resulted again in the loss of V2.

I then put these diachronic developments in a wider cross-linguistic context and sketched a tentative feature hierarchy for word order patterns including V2. Finally, I returned to the role of information-structural features in diachronic syntax.