J. Linguistics **47** (2011), 407–446. © Cambridge University Press 2010
doi:10.1017/S0022226710000381 First published online 13 December 2010

Reconstructing last week's weather: Syntactic reconstruction and Brythonic free relatives[1]

DAVID WILLIS

Department of Linguistics, University of Cambridge

(Received 8 April 2010; revised 3 November 2010)

Lightfoot (2002) argues that syntactic reconstruction is rendered impossible by the lack of any analogue in syntax to the traditional notion of the phonological 'correspondence set' of the Comparative Method and by the radical discontinuity caused by reanalysis between successive grammars. Alice Harris and Lyle Campbell, in various works, have defended the notion of 'syntactic pattern' as the analogue of the correspondence set, arguing that patterns can be compared across languages, with innovations being stripped away to reveal aspects of the protolanguage. In this article, I argue that syntactic reconstruction can be carried out while maintaining and indeed utilizing core notions in generative approaches to syntactic change such as the central role of reanalysis and child language acquisition and the distinction between the abstract grammatical system and the surface output of that system. Reanalysis itself is constrained by the fact that both pre- and post-reanalysis grammars must be acquirable on the basis of the same primary linguistic data. This imposes limits on the possible hypotheses that can be entertained ('local directionality') even in the absence of any crosslinguistic generalizations about patterns of change ('universal directionality'). This approach is then applied to aspects of the syntax of free relative clauses and negation in the early Brythonic Celtic languages (Welsh, Breton and Cornish), showing that non-trivial reconstructions can be achieved even where the daughter languages manifest significant differences.

1. Introduction

David Lightfoot has recently argued that the prospects for realist syntactic reconstructions are very limited because of the 'chaotic' nature of syntactic

[1] The research presented here developed as part of the research project 'The development of negation in the languages of Europe', funded by the Arts and Humanities Research Council. Earlier versions were presented to audiences at the British Academy Research Project 'The development of the Welsh language', the Seventeenth Welsh Syntax Seminar and the Linguistics Association of Great Britain Meeting 2010. I am grateful to those audiences and to three anonymous *Journal of Linguistics* referees for their perceptive comments. I am also grateful to various other people for comments and discussion on the research presented here: Gareth Bevan, Bob Borsley, Torsten Meissner, Ian Roberts, Peter Schrijver and George Walkden. All remaining errors or omissions are my own.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

change. Outside of very narrowly defined limits (essentially where daughter languages show identity), he suggests that 'one can no more reconstruct the syntax of a proto-language than one can reconstruct last week's weather, and for the same reason: both reflect chaotic systems' (Lightfoot 2002: 135). Lightfoot makes the important observation that the central mechanism of syntactic change, abductive reanalysis, replaces one grammatical analysis with another one that may differ from its predecessor without limit. Consequently, he argues, the new grammar can tell us nothing about its immediate predecessor.

In this article, I will argue that, while the radical discontinuity of abductive reanalysis is an important aspect of syntactic change and reconstruction, it does not have the disastrous consequences for reconstruction that Lightfoot claims. It is accepted that children cannot compare their grammars with those of the adults around them and that, therefore, there can be no theory of the direct relationship between successive grammars as proposed in, for instance, Longobardi's (2001) Inertial Theory (Roberts 2007: 231–234; Walkden 2010). However, it does not follow from this that a grammar can differ in unlimited ways from its immediate predecessor. An innovative abduced grammar must be as good a fit for the primary linguistic data (PLD) of language acquisition as its predecessor, and this fact sharply constrains what hypotheses the historical linguist can entertain about an earlier grammar. I propose that this fact can be used in syntactic reconstruction: the earlier reconstructed grammar must have produced an output very similar to that which must have led to the attested systems. Reanalysis proceeds via the availability of acquisitionally ambiguous sentence types and does not substantially alter the surface output of the ambiguous sentence type that was subject to reanalysis. Whether or not a particular change manifests a significant trend towards unidirectionality across different language histories ('universal directionality'), these facts will further limit possible reconstructions across instances of historical reanalysis ('local directionality'). I will demonstrate the usefulness of this approach using data from free relatives in the early Brythonic languages (Middle Welsh, Middle Breton and Middle Cornish).

1.1 *Historical background*

A series of attempts were made in the 1970s to reconstruct Indo-European word order on the basis of typological generalizations. For instance, Lehmann (1974) showed that early Indo-European languages all had features typologically characteristic of SOV languages (modifier–head orderings such as adjective–noun order, genitive–noun order, standard–pivot–comparative order etc.). Tracing a development away from these features over the attested histories of the early languages, he reconstructed Proto-Indo-European as a typologically consistent SOV language. Other

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

researchers (Friedrich 1975, Miller 1975) reached radically different conclusions, but made reference to similar techniques and assumptions. These approaches, which assumed the protolanguage had typologically consistent syntax, were widely criticized at the time. Watkins (1976), for instance, criticized them for reducing syntax essentially to word order, and for then reducing word-order reconstruction to asserting implicational consistency of head–modifier order in a small number of head–modifier pairs in the protolanguage.

The legacy of the 1970s typological work has tarred all research in syntactic reconstruction with the same brush. However, more recent work is radically different in approach, and has focused not on broad-brush typological generalizations about languages, but rather on the development of the syntax of individual functional items and narrowly defined syntactic patterns, witness for instance the fine-grained reconstructions of various syntactic patterns across Finno-Ugric in Campbell (1990) or internal reconstruction of the syntactic origins of various verbal forms in Swahili in Givón (1999). It is this second approach that I pursue in this article. I argue that, even accepting the validity of several of Lightfoot's basic premises (namely the radical discontinuity in abductive reanalysis and the need to reconstruct abstract grammatical systems), progress can still be made. Successful reconstruction can be achieved by investigating microvariation in closely related varieties. Reanalyses are not completely unconstrained: a successful reanalysis must have had some basis in the earlier grammar. Some sentences must have manifested acquisitional ambiguity, the possibility of two different structural analyses at the point of transition and the output of that earlier grammar cannot have been radically different from its immediate successor. These facts can be used to 'reverse' reanalysis without any appeal to universal directionality of change. Directionality can be assessed at a purely local level: often, in a given case, a plausible reanalysis can be proposed for one possible historical scenario but not for another. This approach also has the advantage that it does not require us to give up the internalized abstract grammar as the object of investigation: we investigate both surface patterns (E-language) and the abstract grammatical systems that generate them (I-language), reconstructing the diachronic interplay between the two.

This article is structured as follows. Traditional problems associated with syntactic reconstruction are outlined in Section 2. Section 3 addresses possible solutions to these problems and methods for engaging in syntactic reconstruction, outlining the approach adopted here which focuses on reconstructing fragments of grammars and reanalyses rather than surface outputs alone. Section 4 applies this method to a data set consisting of free relatives in the three attested medieval Brythonic languages, Welsh, Breton and Cornish, leading to a reconstruction of the relevant area of the Common Brythonic grammar.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

	Italian	Spanish	Portuguese	French	
(A)	k	k	k	k	
	correre	correr	correr	courir	'run'
	costare	costar	costar	coûter	'cost'
(B)	k	k	k	ʃ	
	caro	caro	caro	cher	'dear'
	capo	cabo	cabo	chef	'head, top'

Table 1
Simplified correspondence sets for Proto-Romance */k/

2. Problems

The problems in applying the Comparative Method to syntax are well-known and have been used to deny the possibility of non-trivial syntactic reconstruction in its entirety. This section sets out the four main problems, dubbed here the CORRESPONDENCE-SET PROBLEM, the DIRECTIONALITY PROBLEM, the RADICAL-REANALYSIS PROBLEM and the TRANSFER PROBLEM.

Phonological reconstruction relies on establishing correspondence sets: series of lexical items containing a cognate sound in a particular phonological environment in a set of languages hypothesized to be related. This is illustrated, in simplified form, adapted from Campbell (1998: 111, 122), for two correspondence sets in Romance in Table 1. We find a set of cognate lexical items containing a /k/ before a back vowel in correspondence set (A) in all the modern Romance languages listed. In correspondence set (B), before a front vowel, we find /k/ in Italian, Spanish and Portuguese, but /ʃ/ in French.

We assume that any sound change that gave rise to this variation across Romance applied to all lexical items in a given phonological environment (in this case, after a front vowel) (Regularity Hypothesis; Osthoff & Brugmann 1878, Labov 1981). We hypothesize a sound change, namely /k/ > /ʃ/ before a front vowel in French, and therefore reconstruct each of the items in correspondence set (B) as originally containing an initial /k/. In establishing the historical development, we appeal to two factors: (i) economy: reconstruct a history with as few sound changes as possible; and (ii) universal directionality: /k/ > /ʃ/ before a front vowel is a commonly attested and, above all, physiologically motivated change and hence 'natural', while /ʃ/ > /k/ before a front vowel is not. These considerations are crucial to the method since, otherwise, we would not be able to determine whether French or the other Romance languages best reflected the ancestral situation. The method would be able to demonstrate

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

relatedness but would not reconstruct the shape of the given lexical items in the parent language.

This procedure cannot be straightforwardly applied to syntax. First, it is hard to know what the analogue of the correspondence set is in syntax (the CORRESPONDENCE-SET PROBLEM). The obvious answer is that it is collections of 'cognate' sentences containing a particular feature rather than sets of lexical items. However, sentences are not transmitted as whole units from generation to generation; rather an entire grammatical system is transmitted, and, if a particular sentence survives as a possible sentence of a language from one generation to the next, this is because the relevant aspect of the grammar has been transmitted intact, not because the sentence itself has. In this respect, sentences are different from lexical items, which are transmitted directly from generation to generation. It has therefore been concluded that there are no correspondence sets in syntax (Lightfoot 2002: 120–121).

Secondly, while research on grammaticalization appears to offer an analogue to universal directionality in sound change, this too has been challenged: Lightfoot, in particular, argues that grammaticalization changes have purely local motivation and doubts the relevance of any general theory of grammaticalization to explain them. From this, he concludes that 'we have no well-founded basis for claiming that languages or grammars change in one direction but not in another, no basis for postulating algorithms mapping one kind of grammar into another kind' (Lightfoot 2002: 125–126). The result is that, even when presented with a series of alternative grammatical systems, it is hard or impossible to know which one to posit for the parent language (the DIRECTIONALITY PROBLEM).

Linked to the directionality problem is another problem relating to the central mechanisms of change. On many interpretations, both generative and traditional, syntactic change is often mediated by radical reanalyses of structure, in which children abduce an entirely new grammatical structure for a particular construction, organizing other aspects of the language system (e.g. lexicon) in line with the analysis they adopt. Reconstruction amounts to working out what structure the language had before the reanalysis, and there is no trace of that structure in the new language: if children had been able to work out what the earlier structure was, they would not have introduced the new structure. Radical reanalysis breaks a language's relationship with its past (in some particular area of the grammar), obscuring its genetic connections (the RADICAL REANALYSIS PROBLEM). This seems to present formidable obstacles to reconstruction, although the suggestion that reanalysis leaves no trace of the former structure is not entirely true (see Section 3.2.1 below).

Discussion of a fourth problem, the TRANSFER PROBLEM, will be delayed until further discussion of correspondence sets in syntactic reconstruction, in Section 3.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

	Subject of transitive	Subject of unergative	Subject of unaccusative	Direct object
Laz				
Series I	NAR	NAR	NOM	NOM
Series II	NAR	NAR	NOM	NOM
Mingrelian				
Series I	NOM	NOM	NOM	DAT
Series II	NAR	NAR	NAR	NOM
Svan/Old Georgian				
Series I	NOM	NOM	NOM	DAT
Series II	NAR	NAR	NOM	NOM

Table 2
Case-marking patterns in various Kartvelian languages

3. Solutions

3.1 *Patterns as elements of a correspondence set*

One possible solution to the correspondence-set problem is to propose that 'syntactic patterns' can usefully be treated as the elements of the correspondence set. In various works (Harris 1985, 2008; Campbell 1990, 2003; Harris & Campbell 1995; Campbell & Harris 2002), Alice Harris and Lyle Campbell have defended the possibility of non-trivial syntactic reconstruction using this method. That is, they claim that the relevant type–token relationship in syntax is between syntactic patterns (types) and sentences (tokens), just as the relevant relationship in phonology is between the phoneme in a given phonological environment (e.g. /t/ in intervocalic position) and instances of that phoneme in particular lexical items (e.g. the /t/ in *butter*, *better*, *eating*). In this section, I outline and review this approach, arguing that, while comparing grammatical systems is a legitimate way of pursuing syntactic reconstruction, it does not solve the correspondence-set problem, and we should not pretend that syntactic patterns are real analogues to correspondence sets.

Harris and Campbell's general approach can be illustrated by Harris's reconstruction of case-marking patterns in Kartvelian. From observation of actual sentence patterns, Harris (1985: 37–58) abstracts the patterns of case marking for various Kartvelian languages in Table 2, where 'Series I' and 'Series II' refer to particular verbal categories (tense–mood–aspect combinations). Narrative case (NAR) corresponds to what is termed ergative case in other grammatical traditions; on the terminology used, see Harris (1985: 38).

Laz has an active-ergative case-marking system (distinguishing agentive subjects from non-agentive, thematic subjects and objects), with identical

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

patterning in both series: agentive subjects are marked using the narrative case, while non-agentive subjects and objects are marked using the nominative. Mingrelian has an accusative case-marking system in both series (distinguishing subjects from objects), but uses different morphological cases to express that system: in Series I contexts, subjects are marked nominative and objects dative, while, in Series II contexts, subjects are marked with the narrative and objects with the nominative. Svan and Old Georgian have an accusative case-marking system in Series I, but an active-ergative one in Series II: in Series I contexts, subjects are marked nominative and objects dative (as in Mingrelian), while, in Series II contexts, the distinction is between narrative-marked agentive subjects and nominative-marked non-agentive subjects and objects. The languages can be compared within this subsystem, and we can ask what type of case system the parent language could have had that would have allowed all daughter systems to emerge. For instance, Harris proposes that, for Series I, the system common to Mingrelian, Svan and Old Georgian is the ancestral one, with Laz having extended nominative case marking from direct objects to subjects of unaccusatives by analogy with the active-ergative system that it uses in Series II. Note that there is no appeal here to (universal) directionality: the claim that Laz has been innovative in Series I is not made on the basis of any claim that accusative systems tend to turn into active-ergative systems, but rather on the basis of very local factors, namely the existence of a possible extension of the pattern found in Laz in Series II (local directionality).

This procedure produces plausible reconstructions in general, and the particular claim about the innovative nature of the Laz system above seems entirely convincing. However, it is hard to see that syntactic patterns really are correspondence sets in syntax. Phonological correspondence sets emerge historically because a given sound change is manifested in many individual lexical items, each of which is learned in its new form during language acquisition by each new generation. While it is true that a syntactic change is manifested in many individual sentences, those sentences are generated anew and not learned as units by each new generation. Thus, while in phonology, each affected lexical item is independent evidence of a prior sound change, in syntax, there is only really a single observation, whether this is conceived of as a regular syntactic pattern or as a more abstract grammatical rule.

We could, nevertheless, operate with the syntactic pattern as the approximation in syntax to the correspondence set in phonology. However, this understanding of the correspondence set would endanger the two crucial functions of correspondence sets in sound change, namely to eliminate chance or typological resemblance and to protect against and identify language contact.

Elimination of the possibility that resemblances could be due to contact and chance is important to the Comparative Method in phonology for two reasons. The first concerns establishment of relatedness: similarity due to

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

contact may offer false evidence that two languages are related. Secondly, the fact that an item is common to two related languages A and B would normally be treated as evidence to postulate its presence in their common ancestor; however, if the item was loaned from A into B, we cannot be confident of its presence in the ancestor. The first of these two concerns is not really relevant to syntax: syntactic reconstruction is normally carried out among languages whose genetic relatedness is already known. While some attempts have been made to demonstrate relatedness and perform subgrouping using syntactic parameters (Guardiano & Longobardi 2005; Longobardi & Guardiano 2009: 368–375; Roberts 2007), this is not the usual situation. The second concern is, however, a real impediment. If a grammatical rule is present in two languages, this could be because there is a continuous line of transmission from an ancestor grammar where that rule was present or alternatively the rule could have been transferred via contact from one to the other. Furthermore, where the number of possible patterns is limited, two languages may share the same general pattern by chance (that is, both have innovated the same pattern independently, perhaps favoured by the typological similarity likely to be manifested by two closely related languages).

Correspondence sets in phonological reconstruction address this, being used to identify and eliminate lexical borrowings: a loaned item will only participate in sound changes that got underway after it was introduced into the language, and will therefore show irregular correspondences in cases where the expected correspondences are due to sound changes that took place before the item in question was borrowed. However, syntactic cross-linguistic comparisons will often contain only one item (pattern or rule) in each language. Treating patterns as elements of a correspondence set in syntax means that we cannot eliminate the possibility that a particular syntactic pattern was transferred from one language to another (the TRANSFER PROBLEM). We therefore need to be aware of the need to identify transfer by other means. Even if we suspect transfer, syntactic comparisons will provide no particular clue to the direction of transfer, which will have to be determined on external grounds. On the positive side, however, the risks of failing to identify contact can be reduced if the two languages share details of the patterns (e.g. exceptional language-specific rules) or if the crucial functional items in the relevant patterns are shared (cognate).

Harris and Campbell's approach creates a general analogue between syntax and phonology: in phonology, we use the lexical item as basis for forming correspondence sets, and not the phoneme alone, for the simple reason that instances (tokens) of phonemes are not transmitted from generation to generation via child language acquisition; lexical items are. The phoneme is already one level of abstraction away from the observed data. In syntax, Harris and Campbell abstract patterns from sentences in the same way. This procedure makes sense in a theory of grammar in which

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

syntactic patterns are stored as units of the grammar (e.g. Construction Grammar), and Harris and Campbell seem to be assuming this general type of model of grammar. Note also that phonological reconstruction also operates at a fairly limited level of abstraction: the deepest abstraction normally used is the phoneme in a particular environment in the daughter languages. The procedure does not reconstruct the phonemic status of an item in the parent language, only a speech sound which would have to be subjected to further analysis to establish its phonemic status within the phonological system of that language. Furthermore, historical linguists undertaking phonological reconstruction do not normally attempt to reconstruct a phonological grammar (e.g. within Lexical Phonology or Optimality Theory), but rather are concerned with establishing a broadly phonemic system understood in structuralist terms. Harris and Campbell's relatively surface-oriented approach is therefore not so different from standard practice in phonological reconstruction.

This procedure is, in fact, comparing grammatical systems, not correspondence sets, and we should be honest about this. Comparing grammatical systems may nevertheless be a useful enterprise, and this is a procedure that will be developed in the next section. While I am not unsympathetic to Harris and Campbell's general approach, I shall argue that it is not necessary to abandon central tenets of generative grammar (namely the existence of an abstract grammatical system that projects syntactic structure from individual lexical items) to pursue syntactic reconstruction fruitfully.

3.2 *Techniques for reconstructing from correspondence sets*

In this section I turn to setting out a general methodological toolbox for syntactic reconstruction to be applied to the Brythonic data in Section 4. The approach taken here, while grounded in traditional approaches to syntactic reconstruction, including in particular the work of Harris and Campbell, is integrated into an acquisition-based model of syntactic change. If we take seriously the notion that major structural innovations in a language tend to arise during the acquisition process, then reconstruction is not simply about comparing patterns, but compares abstract grammatical systems, asking what reanalyses during the acquisition process in the past could have given rise to subsequent differences between systems.

The approach adopted here shares with pattern-based and construction-based approaches the idea that comparison of syntactic units (whether patterns, constructions, rules or the more abstract features of lexical items within generative grammar) can yield fruitful reconstructions. It also shares a number of specific methodological tools (for instance, use of archaism and directionality) with Harris and Cambpell's approach. However, the aim is to reconstruct grammars rather than patterns, an aim specifically rejected by Harris & Campbell (1995: 371–372). In doing so, I proceed on the assumption

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

that syntactic reconstruction should aim to reach, where possible, the same level of analysis that one would expect for an attested language, including therefore a proposed analysis for how structures are generated.

In this sense it also contrasts with recent construction-based approaches to syntactic reconstruction that have produced explicit characterizations of constructions in various protolanguages. Barðdal & Eythórsson (2010) and Barðdal (in press) reconstruct aspects of patterns of case marking of arguments in Germanic, positing, for instance, a dative experiencer subject construction for verbs such as 'like' in the ancestor language because of its presence in all the daughter languages considered (Old Norse, Old English, Old High German and Gothic). In all the cases that Barðdal and Eythórsson discuss, there is either identity or strong similarity in all the daughter languages and they are never obliged to reconstruct across a reanalysis. In fact, they emphasize that even reconstruction of continuity is a useful exercise. Since they do not make reference to language acquisition, it is unclear how their approach could generalize to a case, such as the Brythonic one below, where a structural reanalysis is involved and where there are significant, although not extreme, differences between the daughter languages. One would need to incorporate an account of how constructions are reanalysed, innovated or dissolved, such as those proposed by Trousdale (2008, 2010), to achieve this. The approach adopted here is thus more able to deal with reconstruction across a reanalysis than construction-based approaches have done to date.

3.2.1 *Reconstruction of abstract grammatical systems*

Syntactic reconstruction involves both syntactic patterns and abstract mental grammars that produce those patterns as their output. A given syntactic pattern may be produced by more than one possible grammar, so the relationship between the two is not trivial. At all stages of the enterprise, we must ask what grammatical system generated the outputs that we observe or postulate. As a methodology, we can observe patterns in the daughter languages, analyse those patterns as syntactic structures and grammars to generate those structures, and ask what earlier outputs could have led ultimately to the historical emergence of those grammars (via reanalysis and actualization/extension). From these reconstructed outputs we infer a reconstructed abstract grammatical system. This basic approach follows Pires & Thomason (2008) in requiring the process of syntactic reconstruction to involve the following three elements:

(i) hypotheses about the grammars that generate the outputs of the daughter languages;
(ii) hypotheses about the grammar internalized by speakers of the protolanguage;

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

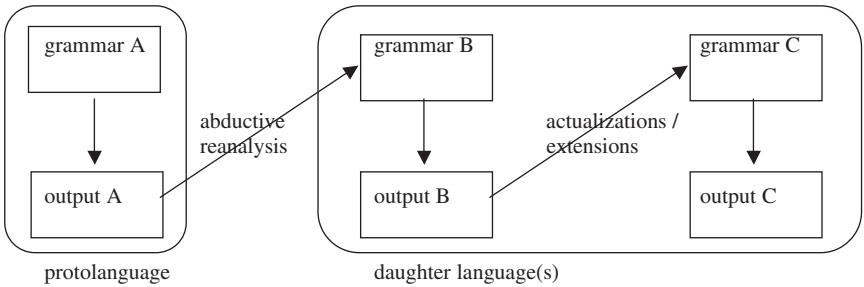


Figure 1
Model of syntactic change for reconstruction.

(iii) hypotheses about how the different grammars of the daughter languages could have developed from exposure to the output of the proposed protogrammar. (Pires & Thomason 2008: 45)

The model of change assumed here follows a standard model of reanalysis and actualization/extension (Andersen 1973, Timberlake 1977, Harris & Campbell 1995). The basic schema is given in Figure 1. The language state that we are aiming to reconstruct is grammar A, which generates a set of sentences (I-language A) on the basis of which a particular set of linguistic data (output A) was produced as the input to a subsequent acquisition process. The daughter languages may either replicate grammar A successfully for a given aspect of syntax or undergo reanalysis, yielding grammar B, and subsequent actualizations or extensions on the basis of the new system, yielding various possible versions of grammar C.

Reanalysis occurs when acquirers assign a new abstract structure to a particular (type of) string, and construct a new grammar to generate that structure. According to the classical definition, reanalysis is a 'change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation' (Langacker 1977). However, linguists have differed in their understanding of the extent to which reanalysis leads to immediate change in the set of sentences generated by the grammar and hence in the output. In the tradition stemming from Andersen (1973) just discussed, speakers initially develop adaptive rules to 'patch up' their grammar, thereby ensuring that, for a while, the language remains identical to its predecessor. This 'patch-up' operation is necessary when acquirers encounter new data that falsifies their initial reanalysis: instead of starting again from scratch and formulating another analysis leading potentially to successful acquisition of the ancestral grammar, they add other rules that lead to the correct surface forms. A change in the syntactic patterns found occurs subsequently, as the new grammatical system is more open to certain types of innovations (generalizations/syntactic analogies) than the

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

old one was. As adaptive rules fail to be acquired by subsequent generations, the full effects of the reanalysis are played out in the process of actualization (Timberlake 1977; Andersen 2001a, b, 2006). On this approach, since there is no change in the surface manifestation of the string undergoing reanalysis, output B does not differ systematically from output A with respect to the structure at issue. However, this could nevertheless leave open the possibility that the output might differ immediately in other ways (for instance, through the creation or obsolescence of some related type of sentence). The empirical consequences of a reanalysis may not be immediately apparent, but may be masked by 'an innovated "usage rule" that enables the speaker to conform to received usage' (Andersen 2006: 72). Significant observed innovations, then, occur only in output C. Actualization is therefore distinct from and subsequent to reanalysis iteslf. The reanalysis only becomes apparent to the linguist once some of these adaptive rules have fallen away and output C appears in the historical record.

In a rather different tradition, Lightfoot operates with a notion of reanalysis that allows for abrupt and even substantial shifts in the set of sentences generated directly after the reanalysis (normally coupled with a parameter shift):

> The emergence of a grammar in a child is sensitive to the initial conditions of the primary linguistic data. If those data shift a little, changing the distribution of the cues, there may be significant consequences for the abstract system. A new system may be triggered, which generates a very different set of sentences and structures. There is nothing principled to be said about why the cues should shift a little; those shifts often represent chance, contingent factors. (Lightfoot 2002: 133)

This follows from Lightfoot's cue-based approach to parameter setting and from degree-0 learnability, the idea that children fail to consider most data from embedded clauses during acquisition (Lightfoot 1991, 1999). Both of these principles provide an explanation for why children should produce a grammar that is evidently and blatantly falsified by their own linguistic experience, that is, why they do not revise their (evidently historically incorrect) hypotheses once they encounter further evidence. On this view, they fail to notice counterevidence either because it does not provide a robust enough cue or because it is found in embedded structures that they do not use for acquisition. This is also inherent in Lightfoot's notion of 'catastrophic' parametric change in which parametric change is associated with reanalyses and the abrupt working out of the consequences in the language. One advantage of the Lightfoot-style approach is that it provides an immediate answer to the question why a reanalysis cannot be reversed: if there is no change to the set of sentences generated by the language and therefore no change to the Primary Linguistic Data used by other children for acquisition, why should the reanalysis spread? If the reanalysis does lead to immediate

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

changes in the I-language and the output, other children will be biased in their own acquisition towards the new grammar since they will now have some positive evidence in its favour. On this approach, the set of sentences produced by grammar B may differ from that produced by grammar A, and hence output B may differ systematically from output A.

The choice between these two approaches to reanalysis is not crucial to the current discussion, provided the changes observed directly after the reanalysis are not too extreme. However, I shall broadly assume the second approach, but maintaining the concept of extension/actualization (Timberlake 1977, Andersen 2001a, b), partly because it is closer to Lightfoot's own conception of change, allowing us to see more easily that syntactic reconstruction is possible even adopting Lightfoot's own general model, and partly because the Brythonic data discussed in Section 4 seem to motivate it at one point (see Section 4.3.2, examples (32)–(34) below). I therefore assume that, while the grammar after reanalysis produces an output that approximates to that found before the reanalysis, we may expect some changes to be manifested immediately, with others only becoming apparent in the full course of actualization.

Finally, reanalysis requires acquisitional ambiguity: some subset of the examples of a particular construction must be amenable to two possible analyses by children acquiring the language (Timberlake 1977; Harris & Campbell 1995: 70–72; Willis 1998: 41). That is, output A must have been amenable in principle to two different analyses, namely grammar A and grammar B. If we can establish grammar B, then we can ask what instances of acquisitional ambiguity there may have been, thereby limiting the possible forms that grammar A could have taken.

These two observations (the output of the grammar after reanalysis approximates to that of the grammar before reanalysis, and the need to posit plausible and motivated reanalyses) guide our methodology in carrying out syntactic reconstructions and impose restrictions on possible hypotheses. While there is a single grammar A and output A, this representing the protolanguage to be reconstructed, any given daughter language may successfully replicate this system or else undergo changes of the type given in Figure 1, leading to the emergence of output C. Different daughter languages may undergo different changes, hence we can have multiple versions of output C, plus, potentially, some languages with output A. For any given daughter language, we have direct access to a version of output C in Figure 1, which we can analyse to establish grammar C. Output A may be present in some other language in the group, or it may not, having disappeared in all languages. The different versions of grammar C will present features that can only be explained as being a relic of the language once having manifested grammar A. That is, elements of grammar C in some languages could only have arisen in a language with grammar A. Even if grammar A is not present in any daughter language, this allows us to reconstruct its properties. This approach

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

therefore incorporates traditional criteria of reconstructing on relics and archaisms, reversal of extension and reversal of grammaticalization.

As a concrete (and simplified) example, consider Brythonic word order. In main clauses, Modern Breton is a verb-second (V2) language (Borsley & Kathol 2000), while Modern Welsh is VSO. On the basis of the modern languages, should we reconstruct Common Brythonic as V2 or as VSO in main clauses? If we look at the outputs in the daughter languages (versions of output C) and construct grammars for them (versions of grammar C), we see various features of the Welsh grammar that are puzzling in a VSO language. Welsh allows initial verbs in main clauses to be preceded by *mi* or *fe*, particles that function as affirmative complementizers. These are identical in form to pronouns (first person singular *mi* and third person masculine singular *fe*), but clearly do not function as such. In (1), the verb is third person singular and the subject is *Steffan*, not *mi*.

(1) Mi welodd Steffan gath.
PRT see.PAST.3S Steffan cat
'Steffan saw a cat.' (Present-day Welsh)

These will have to be listed in the lexicon.

In some southeastern varieties, subject pronouns may precede the verb, as in (2), although lexical subjects must follow the postverbal pattern already seen in (1).

(2) ti ˈgwe·läsd iː
you see.PAST.2S her
'You saw her.' (Phillips 1955: 298)

The precise analysis of these structures is not relevant for current purposes: some additional feature of the grammar must be posited to account for them, since they will not be generated by the ordinary VSO word-order mechanism. The most obvious addition is to propose a feature triggering fronting of pronouns to preverbal position, although there are reasons in fact to suppose that the preverbal 'pronouns' are actually affirmative complementizers manifesting subject agreement (Willis 2007). The crucial point is that whatever feature of the grammar accounts for these, it will be unmotivated in terms of the VSO grammar; essentially, it will be a 'language-specific' operation.

Finally, the lexicon of Modern Welsh contains the lexical item *efallai* 'perhaps'. While this is etymologically transparent, it is nevertheless curious. It self-evidently derives from *ef* 'it, he' and (*g*)*allai* 'could' (3SG conditional of *gallu* 'be able'). However, it manifests SV word order, not generated by the ordinary rules of Welsh grammar; so how did it arise? (Regarding these last two points, see also Comrie 1980 on the possibilities for using pronominal position for word-order reconstruction and some of the pitfalls therein.)

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

These features are all irregularities in the grammar of Modern Welsh (grammar C). Some are patently archaisms; others are more difficult to interpret. They lead us to the conclusion that there must have existed an earlier grammatical system (a grammar A) in which those features did not need special mechanisms to account for them. Either a V2 or an SVO grammar would suffice. In the current instance, since a closely related language, Breton, manifests V2, we posit an earlier common V2 grammar, with a reanalysis (on which, see further below) giving rise to a VSO grammar in Welsh.

This is the basic approach that will be adopted in the reconstruction of Brythonic free relatives shortly. It will be supplemented by local directionality, to which we now turn (Section 3.2.2.), along with a series of traditional considerations – economy, archaisms and extension (Section 3.2.3).

3.2.2 *Use of directionality and 'possible change' in syntactic reconstruction*

Research in grammaticalization has apparently made syntactic reconstruction easier by allowing us to turn back the clock on grammaticalization paths. Given a language with a lexical verb 'want' and another with a similar-looking future auxiliary, grammaticalization tells us that the emergence of future markers from verbs of volition is a unidirectional pathway, and hence, that we can confidently reconstruct the lexical verb rather than the auxiliary for the parent language (Bybee & Pagliuca 1987, Bybee, Pagliuca & Perkins 1991, Bybee, Perkins & Pagliuca 1994). Such research supplies us with general pathways of change. I shall refer to this criterion as UNIVERSAL DIRECTIONALITY (reversal of grammaticalization): a change must have proceeded in a particular direction because it is a general property of language change that it proceeds only in this direction and not the reverse.

Grammaticalization is, however, not absolutely unidirectional (Campbell 2001, Newmeyer 2001, Norde 2009), and this may pose problems; see, for instance, Willis (2010) for an example of the potential problems posed for reconstruction by counterdirectional changes in the Slavonic conditional. However, recognition that grammaticalization is not unidirectional does not preclude us from using statistical tendencies about directionality or particular instances of change that do seem to be unidirectional. Universal directionality, then, is a useful tool in syntactic reconstruction, but not an infallible one. The same in fact applies in phonology: universal pathways of sound change vary according to how unidirectional they are, but we must always keep open the possibility that a particular constellation of circumstances led to an unexpected direction of change in a particular instance.

Lightfoot argues that 'a distinction between possible and impossible changes is in principle a necessary prerequisite for reconstruction' (Lightfoot 1979: 154), and concludes that syntactic reconstruction is impossible because

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

there is no such distinction in syntax and therefore no directionality: 'in syntax there are no formal constraints on possible changes independent of those which follow from a definition of a possible grammar' (Lightfoot 1979: 155).

This approach confuses two things. On the one hand, Lightfoot is correct to argue that there can be no constraints imposed by the acquisition algorithm on possible changes. Children cannot limit the extent or nature of the difference between their grammar and that of the adults around them, because, if they could, that would mean that they knew the exact form of the adult grammar, and, if children knew the exact form of the adult grammar, change would not arise in the first place. Hence any possible grammar can change into any other possible grammar. However, change is constrained in another way: both the adult grammar and children's hypotheses about their language are constrained by the fact that they must be consistent with the Primary Linguistic Data upon which they were based. In reversing a reanalysis we are therefore looking for two analyses, both of which account for the data, one of which was adopted by children in preference to an earlier one used by the adults from whom they acquired their language. This fact imposes quite narrow limitations on transitions, essentially requiring us to state the nature of the acquisitional ambiguity at each point of reanalysis in our reconstruction. The upshot of this is that, at the macrolevel, Lightfoot's claims may well be true, at least for a significant amount of the domain we are investigating: an SVO language may become a V2 language or a V2 language may become an SVO language, or, to put it another way, the value of the V2 parameter may shift in apparently random fashion. However, at the microlevel, things are much more constrained.

Consider again the example of Brythonic word order, this time with reference to directionality. At the macrolevel of argumentation, we have no basis for deciding whether it is more likely that V2 in the parent language became VSO in Welsh or that VSO in the parent language became V2 in Breton. We have no evidence that the historical relationship between the two word-order types manifests any kind of universal directionality, so either is a possible reconstruction. However, at the microlevel, we ask how Welsh could have undergone a transition from V2 to VSO and how Breton could have undergone a transition from VSO to V2: is there local directionality? Examination of Welsh phrase structure reveals a means by which Welsh could have developed VSO word order from earlier V2. We have already seen above (Section 3.2.1) that Welsh has affirmative preverbal particles *mi* and *fe*, homophonous with pronouns. Clauses containing them, such as (1) above, are today analysed as verb-initial (VSO) since the particles signal only that the clause is affirmative and are not arguments of the verb. These particles have no analogue in Breton. If Welsh main clauses were formerly verb-second, then the availability of sentences such as (1) can be understood as the result of a change in which *mi*, a preverbal subject pronoun 'I' was

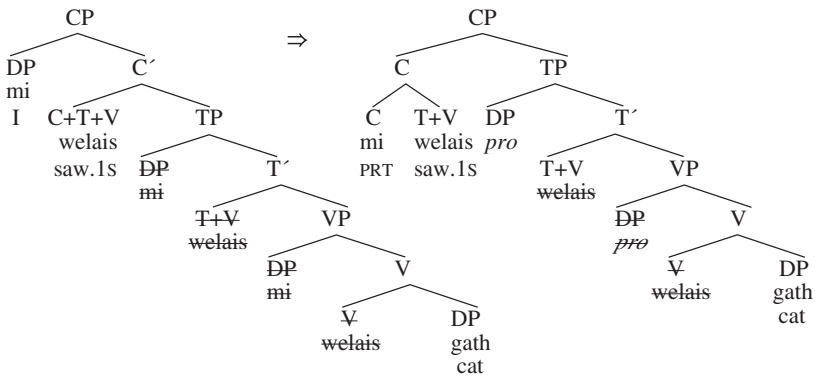
https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

reanalysed as an affirmative complementizer, glossed as PRT below (Willis 1998) (henceforth, the earlier position of moved elements is indicated by strikethrough):

(3) Mi welais gath.
I/PRT see.PAST.1S cat
'I saw a cat.'

(4) [CP [Spec mi] [C welais] [TP ~~mi~~ [T ~~welais~~] [VP ~~welais~~ gath]]]
⇒ [CP [Spec ø] [C mi + welais] [TP *pro* [T ~~welais~~ [VP ~~welais~~ gath]]]]

(5)



To the extent that the reconstruction given in (3)–(5) is plausible, we can therefore reconstruct a V2 system for the parent language, a conclusion borne out by the actual historical records of the languages in question. We do not need to assume that pronoun > complementizer is a universal unidirectional path of development to conclude this. Even if we permitted the reverse reanalysis to have occurred in Breton (reconstructing the particle *mi* in the parent language), it would not explain the creation of a generalized system of V2 in Breton. This approach, then, agrees with the observation that we need to 'tell a plausible story about how grammatical objects in different languages developed from a single antecedent grammatical object' (Harrison 2003: 225), elements of which are not guided by a narrow interpretation of the Comparative Method.

To conclude, it is consistent to accept that differences between successive grammars are not subject to any universal diachrony-specific constraints and to believe that the unidirectionality of grammaticalization is not exceptionless (a belief which in fact follows from the first claim), while at the same time believing that a local form of directionality has a role to play in reconstruction. For any given reconstruction problem, we can evaluate the relative likelihood or plausibility that the development took place in one direction rather than the other. This includes an assessment of whether factors known to play a role in acquisition, such as simplicity or perhaps markedness, as

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

highlighted by Roberts & Roussou (2003) and van Gelderen (2004), suggest that a particular direction of change is more likely than another.

3.2.3 *Other techniques*

Within such an approach, some familiar techniques of reconstruction play their usual role. The first of these is economy. As in phonological reconstruction, all other things being equal, we posit the reconstructed form which requires the smallest number of innovations in the daughter languages. Any reconstruction which requires the same development to occur in more than one daughter language or in more than one branch of the daughter languages (independent parallel development) is immediately suspect. A related concept is 'majority rules': the option found in the largest number of languages is reconstructed. However, this can only be of use where the subgrouping is uncertain. Once subgrouping is known, the more sophisticated concept of economy must be brought into play. This takes into consideration the fact that an option found in two distantly related parts of a family is more likely to be inherited and that a single innovation may lead to an option appearing in many daughter languages if one part of the family tree is more densely packed with languages than other parts.

In line with Antoine Meillet's observation that we reconstruct on exceptions, not on rules (Meillet 1931), identification of archaisms allows us to evaluate hypotheses about the ancestral grammar against one another: it is crucial to establishing a plausible history of reanalyses, as we saw with the Brythonic word-order example above. In this context, archaisms are features (exceptional grammatical rules) that are motivated (generated by the regular pattern) only within an earlier grammar. Exceptions can often be explained as having once been motivated and therefore having once been part of a regular and therefore more widely applied grammatical rule. On the other hand, it is generally difficult to construct a convincing account of how an isolated exception could arise from nothing. Campbell (1990: 82–86) cites the example of Finnish infinitives. Finnish normally uses the 'third infinitive' after a verb of motion ('Father went to cut hay'), but when the infinitive is 'lie down', we unexpectedly find the 'first infinitive' being used. This can be explained as a relic: the third infinitive is an innovation and replaced the first infinitive after verbs of motion, but some fixed environments retain the earlier pattern.

Finally, analogical extension of a pattern in one language to contexts where it is not found in another language must be eliminated. Finnish uses what looks like the negative of the perfect to express the negative of the preterite, while some other Balto-Finnic-Lapp languages use something that looks straightforwardly like the negative of the preterite (Campbell 1990: 70–74). Finnish has extended use of the negative of the perfect to the preterite. The problem here though is identifying that this is a case of

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

analogical extension of a pattern, rather than an archaism, which would lead to the opposite reconstruction.

4. The reconstruction of Brythonic free relatives

In the second half of this article, this general approach to syntactic reconstruction is applied to a problem concerning the reconstruction of patterns for free relatives in Brythonic. While all modern and medieval Brythonic languages have a free-relative marker cognate with Welsh *bynnag*, corresponding more or less to *-ever* in English *whoever* etc., the details of the syntactic rules for its use differ significantly from language to language. I will argue that the patterns of the daughter languages only make sense if a reanalysis is reconstructed for late Brythonic, and that both the details of that reanalysis (that is, both the ancestral grammar and the grammar resulting from that reanalysis) can be reconstructed along with a series of extensions that led to a better motivated system in some of the daughter languages. In doing so, I follow the basic model outlined in Section 3.2.1 above, including reference to other tools, such as identification of archaisms, plausible patterns of extension and identification of language contact, as discussed in Sections 3.1 and 3.2.3.

4.1 *The Brythonic subfamily*

The Brythonic subfamily of Celtic consists of three languages with extensive attestation in the medieval period, namely Middle Welsh, Middle Breton and Middle Cornish. All are descended from a Common Brythonic language, itself unattested except via names in Latin sources and possibly in one inscription discovered in Bath (Tomlin 1987). While once spoken in a single dialect continuum across the whole of what is now England, Wales and southern Scotland, most varieties of the language were replaced by English and Scots from the fifth to the eighth centuries. A western variety gave rise to Welsh, while in the southwest, Cornish emerged, with migrants from and via this region creating a new language, Breton, in Brittany. In the northwest, another Brythonic language, Cumbric, survived into the Middle Ages, but remnants of it are too poor to be of any use for reconstruction. These relations are traditionally summarized in the family tree given in Figure 2. It is worth bearing in mind, however, that these languages originally formed a dialect continuum and essentially remained in this relationship (via sea routes) for a long time after they became physically separated from one another by land. Cumbric was most likely closer to Welsh than to the other varieties, and could be considered to form a western subgrouping with Welsh. Again, this makes little difference for current purposes.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

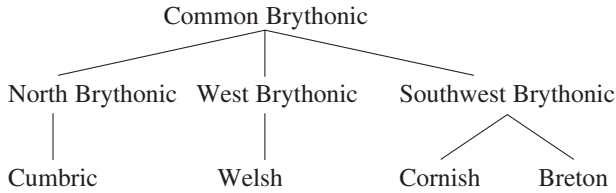


Figure 2
Subgrouping for the Brythonic branch of Celtic.

4.2 *The data*

Brythonic languages have a cognate set comprising Welsh *bynnag*, Breton *bennak* and Cornish *penag*, used to form free relatives.[2] All the languages allow this item to follow a *wh*-word, giving Welsh *pwy bynnag*, Breton *piv bennak* and Cornish *pyw penag* 'whoever':

(6) **pwy bynnac** a 'm metrei i yuelly
who ever REL IS.ACC strike.IMPF.SUBJ.3S me thus
'whoever could strike me thus'
(*Pedeir keinc y Mabinogi* 87.2) (Middle Welsh)

(7) **Piou pennac** no cred nendeu guir
who ever NEG + 3P believe.PRES.3S NEG + be.PRES.3S true
seruicher da doue
servant to God
'Whoever does not believe them is not a true servant of God'
(*Middle-Breton Hours* 4) (Middle Breton)

(8) **pyv penagh** a 'm gwellha vy
who ever REL IS.ACC see.PRES.SUBJ.3S me
'whoever may see me'
(*Resurrexio Domini Nostri Jhesu Christi* 2384) (Middle Cornish)

[2] The following descriptive outline of the data for the medieval Brythonic languages is based on published grammars and exhaustive searching of a range of medieval texts. The grammars used are Evans (1964) for Middle Welsh, Hemon (1975) for Middle Breton and Lewis (1946) for Middle Cornish. The texts used are: for Middle Welsh, *Brut Dingestow*, *Pedeir keinc y Mabinogi*, *Peredur* and *Ystoryaeu Seint Greal*; for Middle Breton, *Le breton de Gilles de Keranpuil*, *Le dialogue entre Arthur et Guinclaff*, *Doctrin an christenien*, *Le mystère de sainte Barbe*, *La vie de sainte Catherine*, and extracts from *Le mirouer de la mort* (lines 1–1200); for Middle Cornish, *Origo mundi*, *Passio Christi*, *Life of Saint Meriasek*, *Bewnans Ke* and *Gwreans an bys*. While the relevant items are attested in Old Welsh and Old Breton sources (Fleuriot 1964: 269; Falileyev 2000: 131), the evidence is too sparse to add anything significant to the picture given by the much more voluminous Middle Welsh and Middle Breton sources. Full bibliographic information for all textual sources is given at the end of this article, before the list of references.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

Wh-phrases other than 'who' may precede:[3]

(9) Canys **pa beth bynnac** ry gollych ti eno, minheu
for which thing ever PERF lose.PRES.SUBJ.2S you there I
a'e henillaf yty yma.
PRT + 3S.ACC win.PRES.1S to.you here
'For whatever you might lose there, I shall win it [back] for you here.'
(*Brut Dingestow* 77.1–2) (Middle Welsh)

(10) A **pha borthua bynnac** y delhei y bydei
and which harbour ever REL come.IMPF.SUBJ.3S PRT be.COND.3S
Edwin a llu ganthav yn y ludyas.
Edwin and force with.3MS PROG 3MS.GEN impede.INF
'And whichever harbour he came to, Edwin would be there with a force with him impeding him.'
(*Brut Dingestow* 195.21–22) (Middle Welsh)

(11) **pé quen bras pennac** vé an offancz
how big ever be.PRES.SUBJ.3S the offence
'however big the offence may be'
(*Le breton de Gilles de Keranpuil* 240.1–2) (Middle Breton)

(12) **py le penag** y's kyffyn
what place ever REL + 3S.ACC find.IMPF.1S
'wherever I find it' (*Passio Christi* 1551) (Middle Cornish)

All Celtic languages once had both a strong (independent) *wh*-element and a weak (dependent) one (Lewis & Pedersen 1937: 226–229). The strong one survives as Middle Welsh *pwy*, Middle Breton *piu* and Middle Cornish *pyw* 'who'. The weak one, *pa* or *py*, is largely obsolete on its own in Middle Welsh, but there are a few examples where it is used alone to mean 'what' or 'why':

(13) **Pa** derw ytti?
what happen.PERF.3S to.you
'What has happened to you?'
(*Pedeir keinc y Mabinogi* 67.21) (Middle Welsh)

(14) Ha uab, **py** liuy ti?
VOC son why blush.PRES.2S you
'Son, why are you blushing?'
(*Culhwch ac Olwen* 54) (Middle Welsh)

Otherwise it survives only in a range of phrases: *pa* X 'which X', *paham* 'why' (<*pa* 'what' + *am* 'for'), *pyr* 'why' (<*pa* 'what' + *yr* 'for'),

[3] In the Middle Welsh example in (10), the base form is *pa borthua bynnac* 'whichever harbour', with the shift from *pa* /pa/ to *pha* /fa/ (aspirate mutation) being triggered by the presence of the co-ordinator *a* 'and'.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

padiw 'to whom' (<*pa* 'who'+*di* 'to'), *pieu* 'whose' (<*pi* 'to whom'+*eu* 'is'), *pwy enw* (<*py* 'what'+*yw* 'is'+*enw* 'name') 'what is the name of' etc.[4]

The weak *wh*-element is not found before *bynnag* in Middle Welsh or in Middle Breton, but it does occur in this environment in Middle Cornish, giving *pepenag*, which may mean 'whoever', 'whatever' or 'wherever':

(15) **pepenag** vo a 'n barth wyr
whoever be.PRES.SUBJ.3S of the part true
'whoever is of the true part' (*Passio Christi* 2025) (Middle Cornish)

(16) **pe-penag** vo
what-ever be.PRES.SUBJ.3S
'[It is necessary to follow his will,] whatever it may be.'
(*Origo mundi* 662) (Middle Cornish)

Cornish is also alone in allowing *penag* to occur without any preceding element at all to mean either 'whoever' or 'whatever':

(17) **penag** a worthya ken du
whoever REL worship.PRES.3S other god
'whoever worships another god'
(*Life of Saint Meriasek* 764) (Middle Cornish)

(18) **pynak** vo lettrys py lek
whoever be.PRES.SUBJ.3S lettered or lay
'whoever he may be, lettered or lay'
(*Passio Christi* 681) (Middle Cornish)

(19) **panak** vo age deses
whatever be.PRES.SUBJ.3S their disease
'whatever their disease may be'
(*Life of Saint Meriasek* 3104) (Middle Cornish)

Both Cornish and Breton allow a reinforcing element to follow. In Cornish this is *-ol* or *-el*, the latter illustrated in (20), transparently derived from *ol* 'all':

(20) **penagel** nath car
who-ever-all NEG+2S love.PRES.3S
'whoever does not love you' (*Bewnans Ke* 1792) (Middle Cornish)

In Breton, the suffix *-et* is added, but this is not attested until after the Middle Breton period (Hemon 1975: 144, 1976–present: 222).

[4] The final *-w-* of *padiw* is unexpected. It may reflect a conjugated form of the preposition – Lewis & Pedersen (1937: 228) and Evans (1964) suggest *-diw* 'to him'.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

Middle Breton alone combines *bennak* with the indefinite article *un* to form indefinite noun phrases, such as *un lech pennac* 'some place':

(21) en **un lech pennac** … / ez crethenn ef-fe em
in a place some PRT believe.COND.1S PRT + be.COND.3S REFL
tennet de hem cuzet
pull.PASTPART to + 3SF.GEN REFL hide.INF
'I should think that she has slunk away somewhere to hide.'
(*Le mystère de sainte Barbe* 364) (Middle Breton)

(22) Ret eu diff gouzout … / Diouz **un re pennac** … /
necessary be.PRES.3S to.me know.INF from a one some
Vn tra …
a thing
'I must learn a thing from somebody.'
(*Le mystère de sainte Barbe* 107) (Middle Breton)

(23) Rac na couezhemp **en vn fæczon pennac** dindan an temptation
lest NEG fall.COND.1P in a way some into the temptation
'lest we should in some way fall into temptation'
(*Le breton de Gilles de Keranpuil* 240.8) (Middle Breton)

(24) A palamour … ma-z dleont **vn guez bennac** bezaff
because COMP must.3P one time some be.INF
reuniet ouz ho eneffou / glorius
reunite.PASTPART to 3P.GEN souls glorious
'because … they must at some time be reunited with their glorious souls'
(*Doctrin an christenien* §24) (Middle Breton)

In Modern Breton, an entire series of indefinite pronouns is based on this pattern:

(25) *un … bennak*-series
person *unan bennak/un den bennak* 'someone' (plur. *ur re bennak*)
thing *un dra bennak* 'something'
quantity *un … bennak* 'some'
place *ul lec'h bennak/un tu bennak* 'somewhere'

Breton *bennak* is stressed on the final syllable, which is unexpected given the usual penultimate stress rule of the language. Welsh *bynnag* receives regular penultimate stress.

Finally, there are two dialect innovations in Welsh that are transparently recent. Specifically, recently (in the twentieth century) many speakers have begun to prepose *bynnag* to the head noun, giving, for instance, *pa bynnag fath* 'whichever kind' for more traditional *pa fath bynnag*. This is clearly on the model of English, which does not split *whichever* and keeps both elements in prenominal position. Another innovation is somewhat earlier (having its roots in the seventeenth century): in this innovation, *bynnag* is preposed and

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

reanalysed as *gan nad* 'since that-not', giving *beth bynnag* 'whatever' > *bynnag beth* > *gan nad beth*. Finally, the negative *nad* in *gan nad beth* is replaced by its affirmative counterpart *taw* '(complementizer) *that*', giving *gan taw beth* 'whatever' and then reduced phonologically to *ta beth* 'whatever', the usual form across much of the south (Thomas 1950–2002: 1379). To go further into the details of these interesting but complex developments would take us beyond the topic of the current article, and they will therefore be disregarded in the following discussion of reconstruction.

4.3 *The reconstruction*

The etymology of *bynnag* is understood in its general outline. It is clear that two elements are involved in its emergence: *pa*, the weak, unstressed form of the general interrogative pronoun 'who, what', plus a negative element *na*(*g*) (Lewis & Pedersen 1937: 231).[5] But what grammar exactly should we posit for the parent language in this area?

4.3.1 *Preliminaries to reconstruction*

First, we can strip away the Cornish suffix *-ol*. This is clearly an innovation to the form of the lexical item. We reach this conclusion because:

(i) it is only found in Cornish, which occupies a middle position in the dialect continuum, hence it is more economical to posit that it is an innovation (economy);
(ii) some instances of Cornish *-ol* violate the metre of the texts, suggesting that the texts were originally composed without *-ol*; this is the case with example (20) above, where the metre requires a four-syllable line rather than the five syllables found in the manuscript, suggesting that the original form here was *penag*, to which the editor emends; and
(iii) universal-type elements are frequently added to free relatives; compare also their use in free-choice indefinite pronouns, where 'always, ever' is frequently added (Haspelmath 1997: 137–138) (universal directionality).

Similarly, the historical record shows that Breton *-et* is a recent innovation, along with the two dialectal Welsh innovations described above, so these can be stripped away easily too.

[5] Morris-Jones (1913: 294) suggests a connection with *pan* 'when' and *a*(*c*) 'and', but the logic of this seems weak. Hemon (1975: 144) describes the origin of Breton *bennak* as 'obscure'. I shall therefore pursue Lewis & Pedersen's (1937) suggestion as the only coherent proposal worth considering. Its coherence is reinforced by the possibility of basing detailed reconstructed reanalyses around it.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

Next, we can strip away the Breton *un … bennak* construction. This is reasonably well attested in Middle Breton, so we cannot dismiss it out of hand. However, we need a story for how it arose, whether in the parent language or later. *Un … bennak* can be treated as an innovation specific to Breton in which Welsh and Cornish have not participated, because: (i) the development of a free relative ('Bring whoever you like') to a free-choice indefinite ('Bring whoever' with omission of the relative clause) to a non-specific pronoun 'someone (unknown)' and finally to a specific pronoun 'someone' is known crosslinguistically, but not the reverse (Haspelmath 1997) (universal directionality); and (ii) precisely this development occurred in French, slightly earlier than the earliest Middle Breton texts (transfer). Old French *quel … que* 'which(ever) X that' developed from free-relative marker 'whatever, whichever' along the same pathway to free-choice indefinite marker, attested in the fourteenth and fifteenth centuries, illustrated in its former use in (26), to the Modern French *quelque*-series (*quelque* 'some', *quelqu'un* 'someone', *quelque chose* 'something' etc.) (Foulet 1919).

(26) Qui femme prend, de quelque taille, / Il ne puet faillir a
who wife takes of any size he NEG can lack at
bataille.
battle
'Anyone who takes a wife, of whatever/any size, he cannot be short of battles.'
(Jean le Fèvre, *Les lamentations de Matheolus* l. ii, v. 3817–3818)
(Old French, *c.* 1371)
(example from Foulet 1919: 227)

The Breton pattern seems likely, then, to be a case of transfer from French, in which case we can eliminate it from our reconstruction. This view is reinforced by the general pattern of heavy transfer from French in many aspects of Breton syntax, including, for instance, the innovation of a verb 'have' and even the indefinite article *un* (from the numeral *un* 'one'), which itself features as part of the *un … bennak* indefinite marker.

4.3.2 *Reconstructing the reanalysis*

This leaves us with three types of free relative:

(a) free relatives with a strong *wh*-pronoun preceding *bynnag* in all the Brythonic languages (the Welsh *pwy bynnag*-type);
(b) free relatives with a weak *wh*-pronoun preceding in Middle Cornish (the *pepenag*-type);
(c) free relatives with *pynag* alone in Middle Cornish.

Since patterns (a) and (b) are essentially variants of one another, differing only in whether a strong or weak form of the *wh*-pronoun is used, we are

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

reduced to considering two basic hypotheses with regard to the surface outputs of the grammar:

(i) pattern (a)/(b) is original and Cornish innovated pattern (c), dropping the *wh*-pronoun;
(ii) pattern (c) is original, but pattern (a)/(b) was innovated, adding the *wh*-pronoun; pattern (c) was lost everywhere except Cornish.

'Majority rules' would suggest treating the Cornish patterns in (c) as an innovation (although both hypotheses are equally economical provided pattern (a)/(b) is innovated in the latter stages of the parent language in hypothesis (ii)). However, the basis for the Cornish innovation, for instance, in terms of reanalysis, would be quite unclear: why would Cornish speakers innovate free relatives without a *wh*-pronoun on the basis of primary linguistic data that contained a *wh*-pronoun?

Hypothesis (ii) raises the following analogous question: why would speakers of Brythonic innovate the *pwy bynnag*-type if their language allowed only the *bynnag*-type? The obvious answer is that it seemed natural for a free-relative pronoun to contain a *wh*-pronoun and, crucially, speakers did not perceive *bynnag* to be a *wh*-pronoun, that is, it was not a *wh*-pronoun in their grammar. To put it more concretely, non-specific free relatives with *bynnag* had the same structure as specific free relatives, such as that in (27), but lacked the *wh*-pronoun found in the latter. This could be remedied by introducing a *wh*-pronoun into *bynnag* free relatives. Using the unmotivated element *bynnag* in free relatives seemed odd, and the degree of motivation of the grammar could be increased by introducing *wh*-pronouns into this context.

(27) A sef a wnaeth y uorvyn ryuedu
and FOCUS PRT do.PAST.3S the maiden marvel.INF
[pvy a'e galwassei].
who REL + 3FS.ACC call.PLUPERF.3S
'And the maiden marvelled at who had called her.'
(*Brut Dingestow* 199.29) (Middle Welsh)

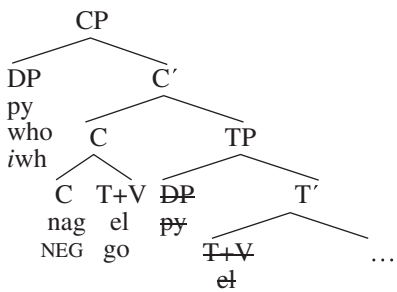
However, this suggestion requires the language to have a poorly motivated system which is made more transparent via an innovation (an extension rather than a reanalysis). It also raises another question, namely, how had this unmotivated system arisen in the first place? Why use a completely non-transparent free-relative marker like *bynnag* in the first place? Reanalysis often leads to the loss of transparency in a system (loss of structure), and in this case we can suggest that the lack of motivation was the result of a reanalysis that changed an earlier grammatical system in which *bynnag* made sense, that is, a system in which *bynnag* was treated as having internal syntactic structure and as containing a *wh*-pronoun.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

We can therefore posit a series of changes in structure as follows, using a hypothetical example clause *py nag el* 'whoever may go' with Middle Welsh forms for the purposes of exemplification.[6] In the ancestral grammar, *py* is a weak *wh*-pronoun that moves from subject position to [Spec, CP]. Free relatives are just relative clauses headed by a *wh*-pronoun (identified as such by bearing an interpretable *wh*-feature) and with pleonastic negation. The ancestral structure is therefore that reflected in (28) and (29). The free relative clause is in the subjunctive (but see evidence in Section 4.3.4 below that this was not compulsory). As is common in irrealis contexts in many languages, it contains pleonastic negation, cf. Russian *kto by to ni byl* 'anyone'<'whoever it might (not) be' (Lewis & Pedersen 1937: 231; Haspelmath 1997: 136). Compare also pleonastic negation in other contexts in various languages, such as French *Je doute qu'il ne soit là* 'I doubt that he'll be there' (Rowlett 1998: 26–27) or German *Was es nicht alles gibt!* '(Look) what is(n't) there'.

(28) py nag el
WH REL.NEG go.PRES.SUBJ.3S
'whoever goes' (stage 1 = Common Brythonic)

(29)



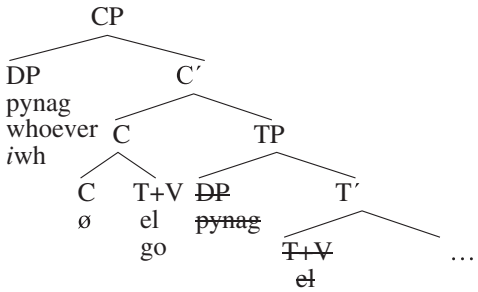
Acquirers had to establish whether the sequence *py* + *nag* was one word or two, both hypotheses being in principle available (acquisitional ambiguity). A reanalysis occurs in which the word boundary between the *wh*-pronoun and the pleonastic negative relative marker fails to be acquired. The resulting single item *pynag* is listed in the lexicon as a *wh*-pronoun, yielding stage 2:

(30) pynag el
WH go.PRES.SUBJ.3S
'whoever goes'
(stage 2 = reanalysis of *py nag* as a single item, Common Brythonic)

[6] It is assumed that clause-inital particles such as the negative *nag* are complementizers (heads of CP) and that inflected verbs raise to adjoin to these particles passing through T on the way. For discussion, see Borsley & Roberts (1996a: 25–29) and Tallerman (1996). The position of moved elements is indicated by strikethrough on the original copy of the element.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

(31)



Since the effects of this reanalysis are evident in all the daughter languages, it dates to Common Brythonic.

In all Brythonic languages there are separate relative particles in the affirmative and in the negative. Since the ancestral construction involves pleonastic negation, we do not expect to find an affirmative relative particle in it, only the negative one. As a result of the reanalysis, *nag* is no longer analysed as a negative relative marker, but is instead an integral part of *bynnag*. This leaves us with a structure containing a null C in a relative clause, a situation otherwise unavailable in Brythonic: relative clauses with extraction from subject position feature the particle *a* in all the daughter languages. We therefore posit that the particle *a* was extended to this environment, giving us the pattern with the particle *a* that is actually found in all the medieval Brythonic languages:

(32) A phwy bynhac **a** vynho ennill clot
and who ever REL want.PRES.SUBJ.3S win.INF praise
'and whoever wants to earn praise' (*Peredur* 57.28) (Middle Welsh)

(33) piou pennac **a**n exerczo
who ever REL + 3MS.ACC practise.FUT.3S
'whoever practises it' (*Middle-Breton Hours* 12) (Middle Breton)

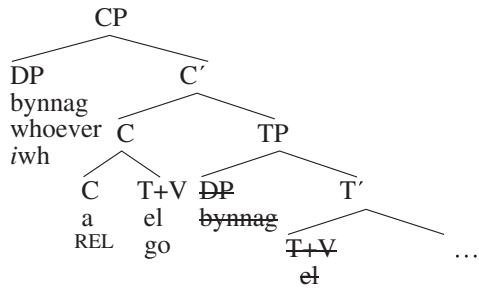
(34) penagel **a** gows er ow fyn
whoever REL speak.PRES.3S against.me
'whoever speaks against me' (*Bewnans Ke* 3077) (Middle Cornish)

This extension would be expected to be more or less automatic, and hence, while stage 2 is a theoretical possibility, in practice an immediate shift from stage 1 to stage 3 would be expected. Stage 3 is shown in (35) and (36). From stage 3, we enter the realm of the daughter languages, so the Welsh form *bynnag* will be used as the standard representation of the word.

(35) bynnag a el
FREE.REL REL go.PRES.SUBJ.3S
'whoever goes'
(stage 3 = addition of relative particle, Common Brythonic)

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

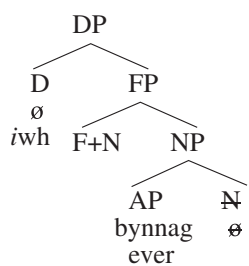
(36)



Again, this is reflected in all the daughter languages, so dates to Common Brythonic. At stage 3, acquirers still successfully identify *bynnag* as a *wh*-pronoun. However, the merger of the two items *py* and *nag* coupled with the likely declining frequency of independent uses of the weak *wh*-pronoun *py* may have made it more difficult for acquirers to establish that *bynnag* should be a *wh*-pronoun (that is, bear an interpretable *wh*-feature). In terms of acquisitional ambiguity, acquirers had to decide whether the *wh*-dependency in the structure was contributed by the element *bynnag* or not. In the transition to stage 4, they fail to do so, but must instead attribute the *wh*-feature of phrases containing *bynnag* to the overall structure. This is a reanalysis, since the structure of the noun phrase containing *bynnag* changes and the feature specification of *bynnag* changes. This is implemented in (38) by positing that the *wh*-feature of the phrase is no longer contributed by *bynnag*, but by a null determiner (D) head. The tree in (38) shows only the structure of the noun phrase *bynnag*; the rest of the structure remains as at stage 3.

(37) bynnag a el
FREE.REL REL go.PRES.SUBJ.3S
'whoever goes'
(stage 4 = failure to acquire *wh*-feature of *bynnag*, Middle Cornish)

(38)



This stage seems to be attested in Middle Cornish, where the surface output in (37) is found; given that Middle Cornish also allows stage 5 constructions, it seems likely that Middle Cornish speakers no longer analysed *bynnag* as a

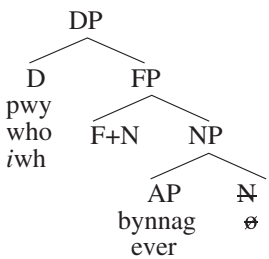
https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

wh-element (otherwise there would have been no motivation to introduce stage 5 constructions, even marginally).

The structure in (38) features a null D providing the structure with a *wh*-feature. A null head noun undergoes raising to an intermediate functional head position (given the neutral label F in (38)), as can be assumed for all noun phrases in Brythonic (given noun–adjective word order in these languages). In specific free relatives, as in (27) above, this feature was associated with an overt *wh*-pronoun. Acquirers fail to notice that an overt *wh*-pronoun is allowed in specific free relatives, but not in non-specific free relatives, and begin to insert them as the source of the *wh*-feature in the latter context too. This gives us the forms which appear as Middle Welsh *pwy bynnac*, Middle Breton *piu pennac* and Middle Cornish *pyw penag* 'whoever' etc.:

(39) pwy bynnag a el
who ever REL go.PRES.SUBJ.3S
'whoever goes'
(stage 5 = addition of *wh*-pronoun, all medieval daughter languages)

(40)



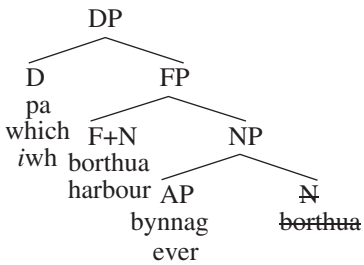
As a motivation for the extension at stage 5, it is also worth noting that, like the early version of *bynnag*, *py* was neutral with respect to ontological category (it could mean 'who(m)', 'what', 'where' or 'why'), but its strong pronoun successors increasingly distinguished category (e.g. Middle Welsh *pwy* 'who', (*pa*) *beth* 'what', *pa le* 'where', *paham* 'why'). The failure of *bynnag* to distinguish category may have promoted this extension, since the creation of sequences such as *pwy bynnag* 'whoever' allows different categories to be distinguished. Since these were increasingly distinguished in the *wh*-system as a whole, speakers would expect them to be distinguished within free relatives too.

A second extension follows on from this reanalysis: the grammar expands by generating *bynnag*-phrases with a lexical element within the *wh*-phrase, such as *pa borthua bynnac* 'whichever harbour' ('which harbour ever') in

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

Middle Welsh in (10) above, and parallel forms in the other daughter languages:

(41)



Such structures could not be generated by a grammar in which *bynnag* was a *wh*-pronoun (stage 3), so we can reconstruct the absence of this structure in the parent language, expecting instead something like (42) (using Middle Welsh forms).

(42) **py borthua nag …
which harbour NEG.REL
'whichever harbour'

In the scenario presented above, we have two reanalyses: one reanalysis involves the loss of the boundary between *py* and *nag* (stage 2), followed by an extension which introduces an affirmative relative particle into the structure (stage 3); a second reanalysis involves the loss of the *wh*-feature on *bynnag* (stage 4), followed by an extension which introduces a new overt *wh*-pronoun into the structure (stage 5).

This series of reconstructed changes is summarized in terms of Figure 1 above, adapted to the current situation in Figure 3. This shows two rounds of reanalysis and extension, with the final stage of the first round acting as the input first stage of the second round.

Let us now consider this reconstructed history in terms of its use of the methodology developed in Section 3. With respect to the aspects of the data discussed in this section, the grammars of Middle Welsh and Middle Breton are essentially identical: both require free relatives of the form *wh*-pronoun *bynnag* + relative clause, a grammatical system which is synchronically straightforward. The Middle Cornish system, however, is more complex and difficult to understand in synchronic terms, since Cornish free relatives are found both with and without the *wh*-pronoun. While Cornish allows the type with the *wh*-pronoun (*pwy bynnag*) (output C2, and hence grammar C2), it incorporates the option of allowing the *wh*-pronoun to be null in free relatives: in the type with the *wh*-pronoun, *bynnag* appears to be a generalizing morpheme, with *pwy* as the *wh*-element providing an interpretable *wh*-feature. If so, some additional statement must be postulated to allow the pattern without *pwy* to be generated (hence the null *wh*-bearing D). The existence of

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

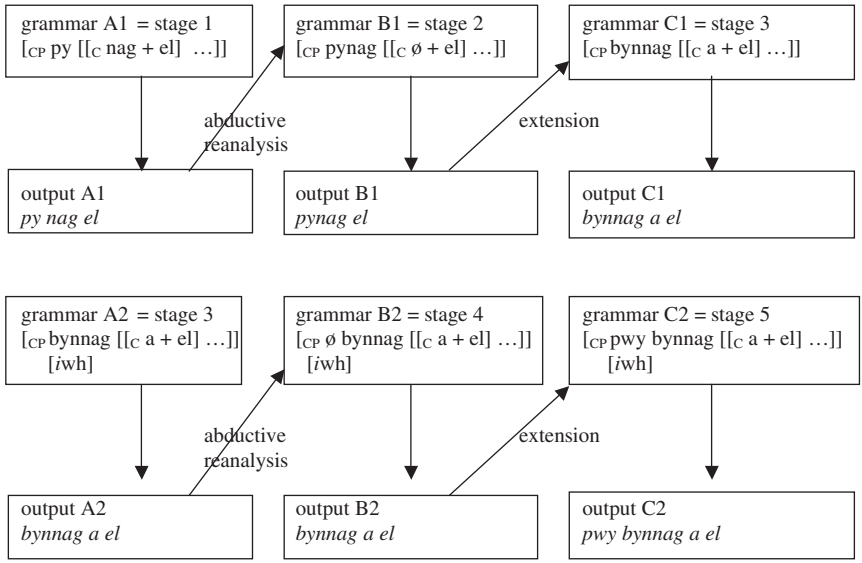


Figure 3
Reanalysis and extension of Brythonic free relatives.

this second pattern is a complexity within the grammar of Middle Cornish, since it requires us to posit an additional null *wh*-element specially for free relatives in that language. This irregularity in the grammar of Middle Cornish could not have emerged from a language with a grammar like Middle Welsh or Middle Breton where *wh*-pronouns are compulsory in free relatives. This therefore leads us to posit an earlier grammar in which the type with *bynnag* alone was not exceptional, but generated by ordinary mechanisms of the grammar. This would be the case if the language at the time treated *bynnag* itself as the *wh*-element in the free relative, a possibility which is supported by the fact that *bynnag* etymologically appears to contain a *wh*-element *py*. This leads us to posit the stage 3 grammar as the reconstructed ancestral grammar.

Further reconstruction from stage 3 back to stage 1 is essentially an application of reversal of grammaticalization (cf. Section 2 above): if *bynnag* is a *wh*-pronoun at stage 3, and its parts recur as independent words in the language, we posit an earlier stage where those parts were distinct, and posit a reconstructed grammar accordingly via internal reconstruction.

This is the central part of the reconstruction. However, some related issues remain, namely the variable use of weak and strong *wh*-pronouns in the different languages and the form of the negative particle. We need to resolve these aspects of the reconstruction too.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

4.3.3 *The status of weak and strong* wh-*pronouns in the protolanguage*

While Welsh and Breton only ever use stressbearing *wh*-pronouns to create new free-relative pronouns, Cornish uses both stressbearing *wh*-pronouns and unstressed *py*- (also *pe*-) to do this, thereby creating both *pepenag* and *pyw penag* for 'whoever'. The latter option is, however, rather rare. How does this impact on our reconstruction?

We know that unstressed *wh*-forms became restricted to adnominal use in all the Brythonic languages (Welsh *pa dŷ*, Breton *pe di* 'which house') and became frozen in other contexts where they were found (e.g. with a preposition such as Welsh *pam* < *paham* 'why' < *pa am* 'for what' or Breton *perak* 'why' < *py rak* 'on account of what' etc.) (Lewis & Pedersen 1937: 226–229). We must suppose that their distribution was wider in Common Brythonic and contracted in late Common Brythonic and in the daughter languages. This leaves us with two possible hypotheses for reconstruction:

(i) the complex free-relative pronouns (stage 5 above) were innovated while weak *wh*-pronouns were the norm in Cornish, but only once strong ones had generalized in Welsh or Breton; this may indicate that they were innovated earlier in Cornish or (more likely, given that Middle Cornish seems to retain the weak *wh*-pronouns better than Welsh or Breton) that Cornish use of weak *wh*-pronouns was simply more conservative;
(ii) the complex free-relative pronouns were innovated in late Common Brythonic and both strong and weak *wh*-pronouns were possible; only strong ones survived in Welsh and Breton; predominantly weak ones survived in Cornish.

According to hypothesis (i), the absence of a form such as ***pybynnac* in Middle Welsh and Middle Breton is explained by the fact that the weak *wh*-pronoun *py* was already obsolete as a free pronoun in those languages when *wh*-pronouns were introduced into *bynnag*-phrases. According to hypothesis (ii), these developments, although they show some degree of differentiation between the daughter languages, are ascribed to the period of the parent language. That is, both the ancestor of Middle Welsh *pwy bynnac*, Middle Breton *piu pennac* and Middle Cornish *pyw penag*, and the ancestor of Middle Cornish *pepenag* arose in the parent language. The latter remained morphologically transparent as a sequence of *pe*- '*wh*-pronoun' and *penag* '-ever' in the early stages of the daughter languages, and hence fell out of use in Welsh and Breton as the unstressed form of the *wh*-pronoun fell out of use. In Cornish, the loss of the unstressed form of the *wh*-pronoun fell out of use more slowly. Consequently, *pepenag* grammaticalized as a single unit before the loss of the *wh*-pronoun *pe*, thereby surviving into the attested period.

It is hard to evaluate these hypotheses against one another. Universal directionality ('strong forms replace weak forms') points against hypothesis

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

(ii), since it requires Cornish to increase the frequency of a weak form (*pepenag*) over that of a strong form (*pyw penag*) over time. On the other hand, economy favours hypothesis (ii), since it requires a single innovation in late Common Brythonic, while hypothesis (i) posits independent parallel development across the three languages. Neither argument is conclusive.

4.3.4 *Negative particles in the protolanguage*

There is another question that we need to resolve in our reconstruction: why is there a word-final velar, either /k/ or /g/, in all of the languages? The reconstruction suggests the negative marker used was *na*(*c*). This makes no sense in Middle Welsh. Middle Welsh has three negative markers: *ny*(*t*) in main clauses and in relative clauses; *na*(*t*) in non-relative embedded clauses; and *na*(*c*) in certain modal clauses, specifically imperatives and optatives, and also in responses to questions.[7] All are clause-initial, immediately preceding the verb. The consonant in parentheses appears when the particle appears before a vowel. *Wh*-words are followed by relative-clause syntax, which requires *ny*(*t*), orthographically *nyd* in example (43), in Middle Welsh:

(43) ny wydynt pwy a oed yn ev herbyn. na phwy
NEG know.IMPF.3P who REL be.IMPF.3S against.3P nor who
nyd oed.
NEG.REL be.IMPF.3S
'they didn't know who was against them and who was not.'
(*Brut y brenhinedd* Cotton Cleopatra B.v. 91.6) (Middle Welsh)

On the basis of Middle Welsh, then, we would expect *bynnag* to be based on *ny*(*t*), not on *na*(*c*). The appearance of *na*(*c*) is therefore inexplicable with respect to Middle Welsh grammar, leading us to suggest that things were not always this way, that is, that there was an earlier grammatical system in which *na*(*c*) would have been generated straightforwardly in this context.

In Breton and Cornish, on the other hand, the appearance of *na*(*c*) is expected. Middle Breton has two negative particles: *ne*(*nd*) (Old Breton *nit*) in main clauses and *na*(*c*) in all embedded clauses and in imperatives and optatives (Hemon 1975: 282). In relative clauses and after *wh*-words, we find *na*(*c*):

(44) an nep **nac** eu discret
the anyone NEG be.PRES.3S discreet
'anyone who is not discreet'
(*Mirouer de la mort* 1200) (Middle Breton)

[7] Negative interrogatives are excluded from discussion here as they differ in more fundamental ways in their syntax across the early Brythonic languages.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

	Middle Welsh	Middle Breton	Middle Cornish
main clauses	*ny*(*t*)	*ne*(*nd*)	*ny*(*ns*)
embedded clauses	*na*(*t*)	*na*(*c*)	*na*(*g*)
relative clauses	*ny*(*t*)	*na*(*c*)	*na*(*g*)
imperatives	*na*(*c*)	*na*(*c*)	*na*(*g*)

Table 3
Clause-initial negative particles in the medieval Brythonic languages

Middle Cornish is parallel, using *ny*(*ns*) in main clauses and *na*(*g*) in embedded clauses and imperatives (Lewis 1946). Again, the relevant form in a relative clause and after a *wh*-word is *na*(*g*):

(45) cusyll **nag** o vas
advice NEG be.IMPF.3S good
'advice that was not good'
(*Pascon agan Arluth* 31.3) (Middle Cornish)

These patterns are summarized in Table 3.

We can conclude that *bynnag* must have grammaticalized in a language which used *na*(*c*) in relative clauses. This has the side effect of leading us to reconstruct *na*(*c*) in negative relative clauses in the parent language, with Middle Welsh extending *ny*(*t*) into relative clauses from main clauses (local directionality). The irregularity found in Middle Welsh (unexpected appearance of *-nag* within *bynnag*) is thereby resolved. Note that, in this case, we are reconstructing the extension of a pattern, rather than a reanalysis, but the basic pattern of argumentation remains the same.

We also need to explain why the form with a velar consonant generalizes irrespective of whether a vowel or consonant follows, that is, why we do not end up with *bynna*(*g*) instead of *bynnag*. This would suggest that, in its formative period, *bynnag* was used habitually preceding a verb that began with a vowel. The only realistic prospects are the indicative forms of the verb 'be' in the present and imperfect tenses (Middle Welsh *yw* and *oed*, respectively). Some Welsh dialects are reported with forms of *bynnag* without a final velar, namely *benna* or *bynna* (Morris-Jones 1913: 293; Thomas 1950–2002: 364). Morris-Jones treats these as innovative, since they are attested late, but we cannot exclude the possibility that they are archaisms reflecting an earlier ancestral situation with an alternating form *bynna*(*g*).

4.3.5 *Stress patterns*

Finally, this reconstruction explains the irregular stress pattern of Breton *bennak*: exceptional word-final stress is found because *bennak* was originally

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

two words. This reinforces our strategy of reversing the grammaticalization (back from stage 3 to stage 1). Stress was on the negative element as in some other cases of these patterns, cf. the stress pattern of Russian *któ by to ní byl* 'whoever' with stress on the negative marker *ni*.

5. Conclusion

In this article I have argued that the absence of any direct counterpart of correspondence sets in syntax and the abrupt break in transmission caused by abductive reanalysis need not be insuperable barriers to syntactic reconstruction. Maintaining a distinction between an abstract grammatical system (I-language) and the surface manifestations of that system (E-language), we can analyse the grammatical systems of the daughter languages and investigate what sequences of reanalysis and extension could have given rise to that microvariation, focusing in particular on those aspects of the grammar that are exceptional and therefore unlikely to have arisen within the current grammatical system. Directionality exists at two levels: the more reliable 'local directionality', which constrains reconstructions to those hypotheses which include plausible reanalyses consistent with detailed aspects of the known daughter language systems; and the less reliable 'universal directionality', which guides syntactic reconstruction through broad-brush rules of thumb, just as it does in phonological reconstruction. This approach to syntactic reconstruction can be embedded within a standard generative model of syntactic change without the need to abandon central distinctions such as that between the grammatical system and the surface output. The focus on grammars rather than syntactic patterns or constructions means than this approach is particularly suited to situations where we need to postulate historical reanalyses, rather than stability or extension of existing patterns. We are led to ask what grammatical system could have given risen to a particular feature that appears exceptional in the daugher languages' grammars.

We noted at the outset that Lightfoot rejects syntactic reconstruction because 'in syntax there are no formal constraints on possible changes independent of those which follow from a definition of a possible grammar' (Lightfoot 1979: 155) and hence, in his view, nothing can be inferred about the state of a grammar immediately before a reanalysis from the grammar immediately after the reanalysis. I have attempted to show in the Brythonic example that the range of hypotheses that need to be considered about the pre-reanalysis grammar is often quite limited. The previous output must have been susceptible to reanalysis; the subsequent output often does not differ radically or catastrophically from the previous output; and frequenly the reanalysis leaves behind synchronically exceptional features that can only be explained in terms of an earlier, reconstructed grammatical system. All these factors limit the space of possible reconstruction hypotheses. Under

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

such conditions, as with Brythonic free relatives, syntactic reconstruction is an entirely feasible enterprise. Reconstruction therefore does not meet a brick wall at historical points where reanalysis cuts us off from the past. While language change is indeed chaotic in the sense that minor changes in the Primary Linguistic Data may lead children to posit grammars that show radical discontinuities with those of their ancestors, much of the time we are dealing with much smaller discontinuities constrained by the relative stability of the data. Except in truly 'catastrophic' scenarios, this allows us to make progress in syntactic reconstruction.

Coupled with other, more traditional notions such as economy, use of archaism and reversal of analogy, this method has proven fruitful in reconstructing parts of the grammar of the Brythonic ancestral language. While these reconstructions do not touch on large-scale word-order parameters such as those favoured in the 1970s, the reconstructions that they provide are not trivial, nor are they limited to cases of identity or cases where minor distorting data can be pruned away to reach identity.

TEXTS CONSULTED

Middle Breton

Ernault, Émile (ed.). 1887. La vie de sainte Catherine. *Revue Celtique* 8, 76–95.
Ernault, Émile (ed.). 1888. *Le mystère de sainte Barbe*. Paris: Libraire du Collège de France.
Ernault, Émile (ed.). 1914. *Mirouer de la mort*. Paris: H. Champion.
Ernault, Émile (ed.). 1928–30. Le breton de Gilles de Keranpuil. *Revue Celtique* 45, 202–271; 47, 72–159.
Hemon, Roparz (ed.). 1977. *Doctrin an christenien*. Dublin: Dublin Institute for Advanced Studies.
Largillière, René (ed.). 1929. Le dialogue entre Arthur et Guinclaff. *Annales de Bretagne* 38, 627–674.
Stokes, Whitley (ed.). 1876. *Middle-Breton hours*. Calcutta: [n.p.].

Middle Cornish

Norris, Edwin (ed.). 1859. *The ancient Cornish drama*. Oxford: Oxford University Press. [Contains *Origo mundi*, *Passio Christi* and *Resurrexio domini nostri Jhesu Christi*.]
Stokes, Whitley (ed.). 1860–1. Pascon agan Arluth. Appendix to the *Transactions of the Philological Society* 1860–1.
Stokes, Whitley (ed.). 1864. *Gwreans an bys. The creation of the world, A Cornish mystery*. London: Williams and Norgate.
Stokes, Whitley (ed.). 1872. *The life of Saint Meriasek, bishop and confessor*. London: Trübner and Co.
Thomas, Graham C. G. & Nicholas Williams (eds.). 2007. *Bewnans Ke: The life of St Kea*. Exeter: University of Exeter Press.

Middle Welsh

Bromwich, Rachel & D. Simon Evans (eds.). 1992. *Culhwch ac Olwen. An edition and study of the oldest Arthurian tale*. Cardiff: University of Wales Press.
Goetinck, Glenys W. (ed.). 1976. *Historia Peredur vab Efrawc*. Cardiff: Gwasg Prifysgol Cymru.
Jones, Thomas (ed.). 1992. *Ystoryaeu Seint Greal*. Cardiff: Gwasg Prifysgol Cymru.
Lewis, Henry (ed.). 1942. *Brut Dingestow*. Llandysul: J. D. Lewis a'i Feibion.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

Parry, John Jay (ed.). 1937. *Brut y brenhinedd: Cotton Cleopatra version*. Cambridge, MA: Mediaeval Academy of America.
Williams, Ifor (ed.). 1930. *Pedeir keinc y Mabinogi*. Cardiff: Gwasg Prifysgol Cymru.

REFERENCES

Andersen, Henning. 1973. Abductive and deductive change. *Language* 49, 765–793.
Andersen, Henning. 2001a. Introduction. In Andersen (ed.), 1–20.
Andersen, Henning. 2001b. Markedness and the theory of linguistic change. In Andersen (ed.), 21–57.
Andersen, Henning (ed.). 2001c. *Actualization: Linguistic change in progress*. Amsterdam: John Benjamins.
Andersen, Henning. 2006. Synchrony, diachrony and evolution. In Ole Nedergaard Thomsen (ed.), *Competing models of linguistic change*, 59–90. Amsterdam: John Benjamins.
Barðdal, Jóhanna. In press. Construction-based historical-comparative reconstruction. In Graeme Trousdale & Thomas Hoffmann (eds.), *Oxford handbook of Construction Grammar*. Oxford: Oxford University Press.
Barðdal, Jóhanna & Thórhallur Eythórsson. 2010. Reconstructing syntax: Construction Grammar and the Comparative Method. In Hans C. Boas & Ivan A. Sag (eds.), *Sign-based Construction Grammar*, 161–216. Stanford, CA: CSLI Publications.
Borsley, Robert D. & Andreas Kathol. 2000. Breton as a V2 language. *Linguistics* 38, 665–710.
Borsley, Robert D. & Ian Roberts. 1996a. Introduction. In Borsley & Roberts (eds.), 1–51.
Borsley, Robert D. & Ian Roberts (eds.). 1996b. *The syntax of the Celtic languages*. Cambridge: Cambridge University Press.
Bybee, Joan L. & William Pagliuca. 1987. The evolution of future meaning. *7th International Conference on Historical Linguistics*, 109–122. Amsterdam: John Benjamins.
Bybee, Joan L., William Pagliuca & Revere D. Perkins. 1991. Back to the future. In Elizabeth Closs Traugott & Bernd Heine (eds.), *Approaches to grammaticalization*, vol. II, 17–58. Amsterdam: John Benjamins.
Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
Campbell, Lyle. 1990. Syntactic reconstruction and Finno-Ugric. In Henning Andersen & E. F. Konrad Koerner (eds.), *Historical linguistics 1987: Papers from the 8th International Conference on Historical Linguistics, Lille, August 30–September 4, 1987*), 51–94. Amsterdam: John Benjamins.
Campbell, Lyle. 1998. *Historical linguistics: An introduction*. Edinburgh: Edinburgh University Press.
Campbell, Lyle. 2001. What's wrong with grammaticalization? *Language Sciences* 23, 113–161.
Campbell, Lyle. 2003. Beyond the Comparative Method? In Barry Blake & Kate Burridge (eds.), *Historical linguistics 2001: Selected papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001*, 33–57. Amsterdam: John Benjamins.
Campbell, Lyle & Alice C. Harris. 2002. Syntactic reconstruction and demythologizing 'Myths and the prehistory of grammars'. *Journal of Linguistics* 38, 599–618.
Comrie, Bernard. 1980. Morphology and word order reconstruction: Problems and prospects. In Jacek Fisiak (ed.), *Historical morphology* (Trends in Linguistics Studies and Monographs 17), 83–96. The Hague: Mouton.
Evans, D. Simon. 1964. *Grammar of Middle Welsh*. Dublin: Dublin Institute for Advanced Studies.
Falileyev, Alexander. 2000. *Etymological glossary of Old Welsh*. Tübingen: Max Niemeyer.
Ferraresi, Gisella & Maria Goldbach (eds.). 2008. *Principles of syntactic reconstruction*. Amsterdam: John Benjamins.
Fleuriot, Léon. 1964. *Le vieux breton: Éléments d'une grammaire* (Collection Linguistique Publiée par la Société de Linguistique de Paris 63). Paris: Librairie C. Klincksieck.
Foulet, Lucien. 1919. Études de syntaxe française: I Quelque. *Romania* 45, 220–249.
Friedrich, Paul. 1975. *Proto-Indo-European syntax* (Journal of Indo-European Studies, Monograph 1). Butte, MT: Montana College of Mineral Science and Technology.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

Givón, Talmy. 1999. Internal reconstruction: As method, as theory. In Spike Glidea (ed.), *Reconstructing grammar: Comparative linguistics and grammaticalization*, 107–159. Amsterdam: John Benjamins.
Guardiano, Cristina & Giuseppe Longobardi. 2005. Parametric comparison and language taxonomy. In Montse Batllori, Maria-Lluïsa Hernanz, Carme Picallo & Francesc Roca (eds.), *Grammaticalization and parametric variation*, 149–174. Oxford: Oxford University Press.
Harris, Alice C. 1985. *Diachronic syntax: The Kartvelian case*. Orlando, FL: Academic Press.
Harris, Alice C. 2008. Reconstruction in syntax: Reconstruction of patterns. In Ferraresi & Goldbach (eds.), 73–95.
Harris, Alice C. & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
Harrison, Shelly. 2003. On the limits of the Comparative Method. In Brian D. Joseph & Richard D. Janda (eds.), *Handbook of historical linguistics*, 213–243. Oxford: Blackwell.
Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Clarendon Press.
Hemon, Roparz. 1975. *A historical morphology and syntax of Breton*. Dublin: Institute for Advanced Studies.
Hemon, Roparz. 1976–present. *Geriadur istorel ar brezhoneg*. Quimper: Preder.
Labov, William. 1981. Resolving the Neogrammarian controversy. *Language* 57, 267–308.
Langacker, Ronald W. 1977. Syntactic reanalysis. In Li (ed.), 57–139.
Lehmann, Winfred P. 1974. *Proto-Indo-European syntax*. Austin, TX: University of Texas Press.
Lewis, Henry. 1946. *Llawlyfr Cernyweg Canol*. Cardiff: Gwasg Prifysgol Cymru.
Lewis, Henry & Holger Pedersen. 1937. A concise comparative Celtic grammar. Göttingen: Vandenhoeck & Ruprecht.
Li, Charles N. (ed.). 1977. *Mechanisms of syntactic change*. Austin, TX: University of Texas Press.
Lightfoot, David W. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.
Lightfoot, David W. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.
Lightfoot, David W. 1999. *The development of language: Acquisition, change and evolution*. Oxford: Blackwell.
Lightfoot, David W. 2002. Myths and the prehistory of grammars. *Journal of Linguistics* 38, 113–136.
Longobardi, Giuseppe. 2001. Formal syntax, diachronic minimalism, and etymology: The history of French *chez*. *Linguistic Inquiry* 32, 275–302.
Longobardi, Giuseppe & Cristina Guardiano. 2009. Evidence for syntax as a signal of historical relatedness. *Lingua* 119, 1679–1706.
Meillet, Antoine. 1931. Caractère secondaire du type thématique indo-européen. *Bulletin de la Société de linguistique de Paris* 32, 194–203.
Miller, D. Gary. 1975. Indo-European: VSO, SOV, SVO, or all three? *Lingua* 37, 31–52.
Morris-Jones, John. 1913. *A Welsh grammar*. Oxford: Oxford University Press.
Newmeyer, Frederick J. 2001. Deconstructing grammaticalization. *Language Sciences* 23, 187–220.
Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.
Osthoff, Hermann & Karl Brugmann. 1878. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen*, vol. I. Leipzig: S. Hirzel.
Phillips, Vincent Howell. 1955. Astudiaeth o Gymraeg llafar Dyffryn Elái a'r cyffiniau [A study of the spoken Welsh of the Ely Valley and surrounding areas]. MA dissertation, University of Wales Cardiff.
Pires, Acrisio & Sarah G. Thomason. 2008. How much syntactic reconstruction is possible? In Ferraresi & Goldbach (eds.), 27–72.
Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.
Roberts, Ian & Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.
Rowlett, Paul. 1998. *Sentential negation in French*. Cambridge: Cambridge University Press.
Stathi, Katerina, Elke Gehweiler & Ekkehard König (eds.). 2010. *Grammaticalization: Current views and issues*. Amsterdam: John Benjamins.

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press

Tallerman, Maggie. 1996. Fronting constructions in Welsh. In Borsley & Roberts (eds.), 97–124.
Thomas, R. J. (ed.). 1950–2002. *Geiriadur Prifysgol Cymru: A dictionary of the Welsh language*. Cardiff: Gwasg Prifysgol Cymru.
Timberlake, Alan. 1977. Reanalysis and actualization in syntactic change. In Li (ed.), 141–177.
Tomlin, R. S. O. 1987. Was ancient British Celtic ever a written language? Two texts from Roman Bath. *Bulletin of the Board of Celtic Studies* 34, 18–25.
Trousdale, Graeme. 2008. A constructional approach to lexicalization processes in the history of English: Evidence from possessive constructions. *Word Structure* 1, 156–177.
Trousdale, Graeme. 2010. Issues in constructional approaches to grammaticalization in English. In Stathi et al. (eds.), 51–71.
Van Gelderen, Elly. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins.
Walkden, George. 2010. Against inertia. Ms., University of Cambridge. http://www.dspace.cam.ac.uk/handle/1810/227574 (retrieved 30 November 2010).
Watkins, Calvert. 1976. Towards Proto-Indo-European syntax: Problems and pseudo-problems. In Sanford B. Steever, Carol A. Walker & Salikoko S. Mufwene (eds.), *Chicago Linguistics Society: Papers from the Parasession on Diachronic Syntax*, 305–326. Chicago: Chicago Linguistics Society.
Willis, David. 1998. *Syntactic change in Welsh: A study of the loss of verb-second*. Oxford: Clarendon Press.
Willis, David. 2007. Specifier-to-head reanalyses in the complementizer domain: Evidence from Welsh. *Transactions of the Philological Society* 105, 432–480.
Willis, David. 2010. Degrammaticalization and obsolescent morphology: Evidence from Slavonic. In Stathi et al. (eds.), 151–177.

Author's address: *Department of Linguistics, University of Cambridge,*
Sidgwick Avenue, Cambridge CB3 9DA, UK
dwew2@cam.ac.uk

https://doi.org/10.1017/S0022226710000381 Published online by Cambridge University Press