

Syntactic lexicalization as a new type of degrammaticalization

David Willis

Department of Linguistics, University of Cambridge

Grammaticalization, the historical emergence of new items with grammatical function from earlier lexical items, is generally considered to be a unidirectional process. Much recent interest has, however, focused on degrammaticalization changes that run counter to this general direction. This paper considers three cases of degrammaticalization from Bulgarian and Welsh, involving shifts from pronoun to noun, and from preposition to verb. These cases exhibit a common set of properties, such as the central role played by syntactic reanalysis and pragmatic inferencing, that justify viewing them as examples of a new type of degrammaticalization. Degrammaticalization via syntactic reanalysis appears to be cross-linguistically rare, because it is constrained by two factors: the requirement that the item undergoing degrammaticalization should have become grammatically or semantically isolated; and the requirement that it should match a possible morphological pattern for the lexical category that it is to join.

1 GRAMMATICALIZATION AND DEGRAMMATICALIZATION

Central to the standard account of grammaticalization is the idea that it is a unidirectional process. Lexical items may over time acquire a grammatical function, and items with a less grammatical function may acquire a more grammatical function, but not the reverse. However, much recent research has been concerned with challenging this orthodoxy, both by claiming the existence of extensive counterexamples to unidirectionality (Janda 2001), and by claiming that grammaticalization itself is not a unified or explanatory process, but rather a frequent constellation of independent processes (Campbell 2001; Newmeyer 2001). This paper considers the existing typology of degrammaticalizations. It begins by asking what a convincing example of degrammaticalization would look like, and proposing a typology of the existing case of degrammaticalization discussed in the literature. In the process, it concludes that at least one, morphological lexicalization, is of little interest to historical linguists. Rather the interesting cases of degrammaticalization all involve reanalysis of some sort or another. Of the existing types of degrammaticalization proposed in the literature, the most convincing category is that where former clitics or bound morphemes acquire greater positional freedom ('deflexion'). This paper suggests that another type can be identified at the syntactic level. In this type, grammatical items undergo syntactic reanalysis as lexical items.

1.1 Defining unidirectionality

Before looking at individual types of degrammaticalization, we need to decide exactly what sort of cases would count as degrammaticalization. It has often been noted that,

under the strictest definition, grammaticalization is unidirectional by definition, rather in the way that Neogrammarian sound change is by definition regular. A change of an item from lexical to grammatical is grammaticalization, and a change from grammatical to lexical is not, hence grammaticalization always proceeds in the direction lexical to grammatical (Campbell 2001: 124–127). The only falsifiable claim is not that grammaticalization itself is unidirectional, but rather that grammaticalization exists in the absence of a parallel reverse phenomenon, degrammaticalization, which, if it were attested, would occur when items with a formerly exclusively grammatical function changed into items with a (more) lexical function.

Even this second hypothesis is not as testable as might first appear, because there is relatively little agreement about what conditions a change has to fulfil in order to count as a convincing example of degrammaticalization. This paper takes the view that, in order to be theoretically interesting, degrammaticalization must be parallel to and linked to grammaticalization. That is, the nature of the mechanisms involved must, in some sense, be the same in both cases, but they must lead to opposite results. The existing cases of degrammaticalization fall into a small number of broad categories. These cases are considered individually below. Some cases involve mechanisms that are so completely unconnected with any of the processes that contribute to grammaticalization that they have little relevance for issues in grammaticalization. Others have a better claim to involve the reversal of grammaticalization.

Grammaticalization has been argued to operate at phonological, morphological, syntactic and semantic-pragmatic levels. At a phonological level, grammaticalized items may undergo phonological erosion. At a morphosyntactic level, there is reanalysis, typically either reanalysis of boundaries (Hopper & Traugott 1993: 40–42, 48–50), for instance, word boundaries become stem-affix boundaries, or reanalysis of category (Hopper & Traugott 1993: 103–113), for instance, from verb to preposition.¹ The semantic-pragmatic processes that have been identified as forming part of grammaticalization are metaphor and metonymy, and, in particular, pragmatic inferencing (Traugott & König 1991). For degrammaticalization to constitute a challenge to grammaticalization theory, it needs to be ‘grammaticalization in reverse’. That is, some (or all) of the processes that contribute to grammaticalization must be at work, but must lead to the opposite result from that expected in grammaticalization. The following five putative processes, which lead to the reverse outcome from that normally found in grammaticalization, might be posited as being involved:

- (i) phonological ‘strengthening’;
- (ii) boundary reanalysis rightwards along the cline: affix > clitic > independent word;
- (iii) category reanalysis from grammatical (preposition, pronoun, article etc.) to (more) lexical (noun, verb, adjective);
- (iv) metaphorical shift from abstract to concrete (‘antimetaphor’);
- (v) pragmatic inferencing from abstract to concrete.

¹ Some authors (for instance, Haspelmath 1998) have been sceptical of the role of reanalysis in grammaticalization. The orthodox position, however, seems to remain that “reanalysis is the most important mechanism for grammaticalization” (Hopper & Traugott 1993: 32).

Processes (ii), (iii) and (v), reanalysis and pragmatic inferencing are identical to those proposed in grammaticalization, but, in the proposed version of degrammaticalization, lead to the reverse outcome. The two reanalysis processes, (ii) and (iii), would be the ones most central to the idea of degrammaticalization, since they involve a change of status of some element from a more grammatical status (more bound) or category (more functional) to a less grammatical status (less bound) or category (less functional). Processes (i) and (iv) cannot, by definition, be the same in grammaticalization and degrammaticalization: phonological weakening, which is proposed for grammaticalization, cannot, by definition, lead to an increase in phonological material; and metaphor, since it is, by definition, the expression of the abstract using linguistic material used to express the concrete, cannot, strictly speaking, lead to a change from abstract to concrete. If phonological weakening and metaphor were to result in the reverse outcome, they would have different names. For this reason, the list suggests ‘phonological strengthening’ as a process of phonological change that would lead to the reverse outcome from that found in grammaticalization; and ‘antimetaphor’ as a process of semantic change that would lead to the same type of reverse outcome. All of these outcomes have been judged impossible, although existing proposals for degrammaticalization have involved some of them.

The discussion below will show that, even given these strict requirements, several types of degrammaticalization can be identified. The two most important seem to be deflexion and syntactic lexicalization. The next section considers a number of major proposed types of degrammaticalization, measuring them against the criteria for being ‘grammaticalization in reverse’.

1.2 Typologies of degrammaticalization

Norde (2001: 231–232) notes that most proposed instances of reversals of grammaticalization fall into one of two types, which she refers to as lexicalization of grammatical items and degrammaticalization / deflexion. Since the term degrammaticalization is already well-established in the literature for any change posited to be counter to the general direction of grammaticalization, I shall depart from her terminology in order to keep degrammaticalization as an overarching and essentially pretheoretical term, but her basic distinction is kept in what follows. Much the same typology of degrammaticalization is effectively implicit in other work in the field (for instance, Newmeyer 2001, Börjars 2003).² I first consider these two types in terms of the criteria for being ‘grammaticalization in reverse’, accepting the first, but rejecting the second, before looking at two other possible types explicit or implicit in the literature, degrammaticalization of modals and syntactic reanalysis.

1.2.1 Deflexion

Perhaps most disturbing for the claim of unidirectionality are cases where bound inflectional affixes or clitics become free morphemes of some kind (either clitics or

² Haspelmath (2002) identifies seven types of degrammaticalization, of which he rejects all but one as examples of grammaticalization in reverse (‘antigrammaticalization’, in his terms). I ignore most of his rejected types in the current discussion, and focus on the types that have the best case to be considered real cases of degrammaticalization.

independent phonological words), gaining a greater degree of positional freedom in the process. That is, whereas grammaticalization represents change of an item's status towards the right on the cline in (1), cases of degrammaticalization represent change towards the left:

(1) free morpheme > clitic > affix

As such, they represent a case of boundary reanalysis ((ii) above), and, in fact, Janda (2001: 303) refers to this type of case as 'upgrading via reanalysis'. Examples include the English possessive clitic *'s* from an earlier genitive case ending,³ the Irish first-person-plural pronoun *muid* from an earlier verbal inflection (Bybee, Perkins & Pagliuca 1994: 13–14; Doyle 2002), and the (older) Estonian question marker *es* from an earlier bound clitic (Campbell 1991). If the newly degrammaticalized item has the potential to bear stress, then examples of this type will also be cases of phonological strengthening ((i) above), on the assumption that a stress-bearing or optionally stress-bearing item is phonologically 'stronger' than one that can never bear stress. Although in some cases the empirical evidence in support of the change has been the subject of debate, the scenario itself seems clearly to amount in principle to degrammaticalization. Indeed, Haspelmath (2002) treats this as the only true type of degrammaticalization.

1.2.2 Lexicalization of grammatical items

A second group of proposed cases of degrammaticalization involves the formation of a lexical item directly from an independent grammatical item. Proposed cases of such a development include the shift of *down* from preposition (or adverb / particle) to verb (as in *He downed the beer in one*). These cases are sometimes referred to under the term lexicalization (Hopper & Traugott 1993: 127) or lexicalization of grammatical items (Norde 2001: 232). It seems reasonable to characterize lexicalization of grammatical items as essentially a process of derivational morphology, and therefore as inadmissible for being considered the 'reverse of grammaticalization'. The English verb *to down* is derived by a process of zero-derivation (conversion) from a base that happens to be a preposition (or adverb / particle). In essence, this process is no different from the way in which verbs like *to boot (out)*, *to sock (it)*, *to head* or *to mother* are derived from bases which just happen to be nouns. In fact, the syntactic category of the base is irrelevant in present-day English. If it is correct that most of the degrammaticalized lexicalized verbs in English are of fairly recent provenance, then, historically, a productive morphological rule has been extended to a new environment, that is, rule extension has occurred: a morphological rule that once converted nouns and adjectives into verbs is now applied also to prepositions.

Lexicalization of grammatical items in English is paralleled by overt derivational morphological processes in other languages. For instance, in French and German, verbs derived from pronouns, French *tutoyer* < *tu, toi*, German *duzen* < *du* 'to address using the familiar pronoun'; and French *vouvoyer* < *vous* + ending analogical to *tutoyer*, German *siezen* < *Sie* 'to address using the polite pronoun', have explicit morphology that distinguishes them from the corresponding pronouns. This demonstrates beyond any doubt that processes involved in grammaticalization (phonological erosion, reanalysis, and pragmatic inferencing) play no role at all here. In

³ Norde (2001) argues for a very similar change having taken place in Swedish, but this is open to some dispute (see Börjars 2003).

particular, there is no plausible way in which reanalysis can have played a role, and there is no suggestion that any current or past instance of the form *down* is potentially ambiguous between a prepositional and a verbal analysis.⁴ Therefore, lexicalization of grammatical items is not ‘grammaticalization in reverse’ and is therefore not degrammaticalization. None of the five processes of degrammaticalization posited above is involved. It is rather an independent process of derivational morphology, hence purely of morphological change.

1.2.3 *Degrammaticalization of modals*

Another group of well-known examples does not fit into either of the types proposed so far. These are the cases of modal verbs becoming more lexical over time, either by replacing modal meanings with lexical ones or by adopting syntactic patterns associated with lexical verbs rather than auxiliaries. A number of examples of this have been proposed, for instance, the shift of Pennsylvania German *wotte* from auxiliary to lexical verb (Burrige 1998); the Swedish modal *må*; Latin *posse*; and English *need* and *dare* (Beths 1999). It is unclear whether to treat these as genuine examples of degrammaticalization. A number of authors (for instance, Börjars, Eythórsson & Vincent 2003) have suggested that, in some of these cases, the shift towards lexical uses has its basis in the survival of an earlier pre-grammaticalization lexical use. That is, a lexical item A develops a grammaticalized form B (A > A and B); the lexical form survives, and, at some later point, the grammaticalized form becomes obsolete (A and B > A) (cf. Haspelmath’s ‘retraction’, Haspelmath 2002). Clearly the fact that functional items sometimes become obsolete is not part of any theory of grammaticalization. To the extent that this scenario is the correct one for these cases, they do not represent true degrammaticalization. I therefore leave open the question of whether this type of drift backwards needs to be recognized as a distinct typed of degrammaticalization.

1.2.4 *Syntactic lexicalization: A new type of degrammaticalization?*

There is another type of degrammaticalization that has received little attention in the literature, but which seems to involve category reanalysis and pragmatic inferencing (processes (iii) and (v) above). It is to these cases that much of the rest of this article is devoted. I shall argue that these cases really are ‘grammaticalization in reverse’. This is not to be understood in the sense that a particular linguistic item with a grammatical function returns to the SAME form and function that it formerly had as a lexical item. This is clearly either impossible – unless languages have memories – or else likely to arise only by pure chance. Just as with lexicalization of grammatical items, these are cases where a lexical category (noun or verb) derives historically from an earlier grammatical category (preposition or pronoun). However, in sharp contrast to lexicalization of grammatical items, processes of derivational morphology, which are not involved in grammaticalization, are not involved in these cases of degrammaticalization either. The three main examples to be discussed in more detail below are:

- (i) Bulgarian *nešto* ‘thing (noun)’ < ‘something, anything (indefinite pronoun)’

⁴ Hopper & Traugott (1993: 49) do view examples of this kind as instances of reanalysis, but this ignores the cross-linguistic evidence, and the possibility of making a distinction between morphological zero-derivation and reanalysis.

- (ii) Welsh *eiddo* ‘property (noun)’ < ‘his (possessive pronoun)’
- (iii) Welsh *nôl* ‘fetch, bring (verb)’ < ‘after (preposition)’

Superficially these look like the same type of example as lexicalization above. However, although they involve the same sorts of category shift, for instance preposition to lexical verb, the cases considered here have a very different flavour, in that they involve a demonstrable continuity between the old and the new function, with the new function arising out of reanalysis of ambiguous instances of the old function. For this reason, I shall refer to them as ‘syntactic lexicalization’.

Of the existing cases of degrammaticalization described in the literature, one group has a good claim to belong in this category. In a number of Germanic languages, there are cases of the degrammaticalization of particles in separable verb construction into adjectives. Examples are the degrammaticalization of German *zu* from preposition ‘to’ to become an (attributive) adjective ‘closed’ (Janda 2001: 299–300), or the degrammaticalization of Dutch *bij* from a preposition ‘by’ to an adjective ‘smart, tough, alert’ (Perridon 2003). These involve syntactic lexicalization provided that they arose through a category reanalysis based on potentially ambiguous environments. Janda suggests that the first example arose through successive stages beginning with sentences of the kind *die Tür ist zugemacht* ‘the door is closed’ with elision of the verb to *die Tür ist zu* ‘the door is closed (lit. to, preposition)’. This is open to reanalysis since the syntactic frame is the same as that for an adjective *die Tür ist blau* ‘the door is blue’. It might be added that even the elision does not need to be postulated because sentences like *Er machte die Tür zu* ‘He closed the door’ could be construed as having a structure with *zu* as an adjective, parallel to the adjective *kürzer* ‘shorter’ in such sentences as *Er machte die Hose kürzer* ‘He made the trousers shorter; he shortened the trousers’. Either way, what is actually a preposition can plausibly be interpreted as an adjective in certain syntactic environments, and is reanalyzed as such, giving rise to fully adjectival uses, for instance, in attributive position, as in *die zu((e)n)e Tür* ‘the closed (adjective) door’. In as much as this scenario reflects the actual emergence of these adjectives, the process involved is indeed syntactic lexicalization.⁵

Notice that in all of these cases, one form of the degrammaticalized lexical item is identical to the grammatical item from which it arose. This seems to be a requirement of syntactic lexicalization, since it is only if this requirement is met that the syntactic ambiguity underlying the reanalysis, and necessary for it, can arise. Contrast this with morphological lexicalization, where this condition need not be met. For instance, the German verb *duzen* has no morphological form that is identical to the pronoun *du*. Note that from this it follows that Germanic-style reanalysis of prepositions as adjectives could only occur as syntactic lexicalization in a language where adjectives lack an overt ending in at least one form or syntactic environment.

1.2.5 Evaluation of the typology of degrammaticalization

In this section, I have divided the most plausible types of degrammaticalization into four types: (i) deflexion; (ii) lexicalization of grammatical items; (iii) degrammaticalization of modals; and (iv) syntactic lexicalization. I have taken it as self-

⁵ If this sketch of the historical development is correct, then *zu* and parallel cases belong in a distinct category from cases of preposition to verb in English (such as *to down*, *to out* or *to off*), contra Janda (2001: 299). The former are syntactic lexicalizations, whereas the latter are lexicalizations of grammatical items.

evident that the most interesting cases of counterdirectional changes, and the ones for which the term ‘degrammaticalization’ seems most appropriate, are those which involve the same sorts of processes as core examples of grammaticalization, but with reverse outcomes. If so, then (ii) lexicalization of grammatical items can be dismissed. Degrammaticalization of modals is possible, but dubious, mainly for empirical reasons. This leaves deflexion and syntactic lexicalization as the central phenomena in degrammaticalization, and the ones to which historical linguists should pay most attention.

In the rest of this article, I demonstrate that syntactic lexicalization is indeed a credible scenario for degrammaticalization. If counterdirectional changes (degrammaticalizations) do exist, then the task of historical linguists is to account both for the existence of the two types of change, and for their relative (and unequal) frequencies. In discussing the various cases of syntactic lexicalization, I attempt to consider what unusual properties of the scenario lead to counterdirectional change.

2 CASES OF SYNTACTIC REANALYSIS AS DEGRAMMATICALIZATION

2.1 Bulgarian *nešto* ‘thing’ < ‘something (indefinite pronoun)’

Generic nouns such as ‘thing’ or ‘person’ frequently give rise historically to indefinite (unknown-specific) pronouns such as ‘someone’, ‘something’ or ‘somewhere’. Heine & Kuteva (2002: 208–209, 232–233, 295–296) recognize the following grammaticalization paths of this type:

- (2) THING > INDEFINITE PRONOUN e.g. Nahuatl *itlaa* ‘thing’ > *tlaa* ‘something’;
 MAN > INDEFINITE PRONOUN e.g. Latin *homo* ‘man’ > French *on* ‘indefinite pronoun’;
 PERSON > INDEFINITE PRONOUN e.g. Albanian *njeri* ‘person’ > ‘somebody, anybody’.

This process seems to be a very common one. Indeed, Haspelmath (1997: 182) notes that, in a sample of 100 languages, 42 had indefinite pronouns that can be derived from generic nouns. The diachronic process is quite straightforward. First, the generic noun is used without modifiers as a noun phrase to convey ‘someone’ or ‘something’ in sentences of the type ‘I saw a person’. Then it acquires phonological, morphological and syntactic features distinct from other nouns, in effect, becoming grammaticalized, with the nominal and pronominal uses diverging. Finally, the original generic noun may become obsolete in its former meaning, as with English *body* in *somebody* (no longer possible with the meaning ‘person’) or French *rien* ‘nothing’ (formerly ‘thing’) (Haspelmath 1997: 182–183).

In the light of this typical pattern, the history of Bulgarian *nešto* is particularly interesting, since it seems to instantiate exactly the reverse development, namely

INDEFINITE PRONOUN > THING. In Old Church Slavonic (Old Bulgarian) *něčito* (genitive *něčeso*) is used as an indefinite pronoun meaning ‘something’.⁶

- (3) a. Tŭgda pristŭpi kŭ nemumati snovu zevedeovu sŭ snĭma
 then came.3S to him mother son.DAT Zebedee’s.DAT with sons.INST
 svoima klanějŭšti se i prosešti něčeso otŭ
 her.INST bow.PRES.PART REFL and ask.PRES.PART something.GEN from
 nego.
 him
 ‘Then the mother of the son of Zebedee came to him with her (two) sons,
 bowing down and asking something from him.’ (Mar. Matt. 20.20)
- b. Simone imamŭ ti něčŭto rešti...
 Simon have.PRES.1S you.DAT something say.INF
 ‘Simon, I have something to say to you...’ (Mar. Luke 7.40)

The semantic interpretation of examples such as those in (3) could perhaps be open to debate, but the morphosyntax of *něčito* is clear and points unambiguously to a status distinct from that of a lexical noun. It is unlike nouns in having a highly irregular genitive form *něčeso* (exemplified in (3a)), not exactly paralleled by any noun or pronoun, except for the transparently related interrogative pronoun *čĭto* ‘what’. Its syntactic behaviour is typical of indefinite pronouns in other languages. In particular, modification is highly constrained, essentially limited to modification by adjectives in the short (indefinite) form. This is illustrated in (4), and paralleled in the English translations ‘something else’ and ‘something evil’.

- (4) a. mněaxŭ bo ŭ zlata plŭny sošte ili něč’to ino.
 think.IMPF.3P PRT it gold.GEN full be.PRES.PART or something else.NEUT
 ‘...for they thought it was full of gold or something else.’ (Supr. i.26v.19)
- b. ...živŭštii vŭ g’radě nadě’axŭ se zŭlo něč’to prieti
 living in city were-sure.3P REFL evil.NEUT something
 receive.INF
 otŭ neju...
 from them.DUAL
 ‘...those living in the city expected to receive something evil from them...’
 (Supr. i.106r.27)

Modification by the long (definite / specific) form of the adjective is not found in the canonical Old Church Slavonic texts.⁷

Haspelmath (1991: 107, 1997: 131–132), following Miklosich (1886: 214), derives Old Church Slavonic (Old Bulgarian) indefinite pronouns such as *někŭto* ‘someone’ and *něčĭto* ‘something’ historically from clauses such as *ne vě kŭto / čĭto* ‘I don’t know who / what’ (sic, presumably for *ne vě / věstĭ kŭto / čĭto* ‘he / she doesn’t

⁶ The relevant data are: Mar. Matt 20.20, Luke 7.40, 11.53, John 13.29; Supr. i.17v.20, i.26v.19, i.104r.12, i.105v.21, i.106r.27, iii.9v.29, iii.24v.3, iii.46r.25, iii.63r.7, iii.77v.29, iii.80v.12, iii.82r.9, plus examples in Kurz (1958–66: ii.453).

⁷ Kurz (1958–66: ii.453) gives an example of *drugŭje něčto* ‘another thing, something else’ from a thirteenth-century Russian Church Slavonic text ‘The Homilies of St. Gregory the Great’, with a definite form of an adjective preceding *něčĭto*.

know who / what’ or *ne věmī / vědě kūto / čīto* ‘I don’t know who / what’ etc.) via grammaticalization paths of the kind ‘She told him I don’t know what’ > ‘She told him something’ with corresponding extreme phonological erosion (Haspelmath 1997: 143). Alternatively, it has been seen merely as a combination of the negative *ne* plus the interrogative pronoun *čīto* ‘what’, with (somewhat mysterious) ‘affective’ secondary lengthening of the first vowel (Trubačev 1997: xxiv.92–93; see also Vasmer 1953: 209). Either way, it is clear that *něčīto* is not related historically to a noun, and its use as an indefinite pronoun meaning ‘something’ is etymologically primary, and well established as its only function by the time of the earliest textual attestation.

In Modern Bulgarian, on the other hand, *nešto* functions both as a noun and as a pronoun. It has gained a number of morphosyntactic properties that identify it also as a neuter noun. It has a regular plural inflection in *-a*, and hosts forms of the clitic definite article, in both cases using forms identical to those of a neuter noun such as *mjasto* ‘place’ (including a shift of stress to the second syllable in the plural). The relevant forms are given in Table 1.

	‘thing’	‘place’
sing. indef.	<i>nešto</i>	<i>mjasto</i>
plur. indef.	<i>nešta</i>	<i>mesta</i>
sing. def.	<i>neštoto</i>	<i>mjastoto</i>
plur. def.	<i>neštata</i>	<i>mestata</i>

Table 1. Forms of *nešto* ‘thing’ and *mjasto* ‘place’ in Modern Bulgarian.

These morphosyntactic innovations are illustrated in the examples in (5).⁸

- (5) a. *Vratata beše razbita, no neštata v stajata bjaxa po mestata si.*
 the.door was broken but the.things in the.room were in the.places REFL
 ‘The door had been broken in, but the things in the room were in their
 (respective) places.’ *(Rečnik na bālgarskija ezik 1123)*
- b. *Točno tova e neštoto, koeto naj-mnogo drazni.*
 exactly that is the.thing REL.NEUT most-of-all irritates.
 ‘That is exactly the thing that irritates the most.’

It occurs in syntactic frames typical of nouns, for instance, following a numeral, as in (6a), a quantifier, as in (6b), or a demonstrative, as in (6c).

- (6) a. *Stava vāpros, če imapet nešta, koito trjabva da se napravjat.*
 arise.3S question that has five things RELis-necessary to REFL do.3P
 ‘The issue is that there are five things that need to be done.’

⁸ Unless otherwise indicated, example sentences are from the website of Bālgarska Nacionalna Televizija (www.bnt.bg).

- b. Za vsički tezi nešta sa nužni vekove, a ne desetiletija.
for all those things are necessary centuries and not decades
'Centuries are needed for all those things, not decades.'
- c. ...graždanite trjabva da znajat tova nešto.
the.citizens is-necessary to know.3P that thing
'...the citizens need to know that thing.'

These properties are all innovations not found in Old Church Slavonic. It is also worth noting that there are a number of nouns derived from *nešto* in Modern Bulgarian, such as diminutives *neštičko*, *neštinko*, and *neštice*. It would be odd if such forms had been derived from a pronoun rather than from a fully lexical noun.

Alongside its use as a noun, *nešto* maintains its use as an indefinite pronoun, for instance, *nešto drugo* 'anything else, something else' or *nešto novo* 'something new, anything new', in which case any adjective follows. However, when used as a pronoun, *nešto* does not have a plural. Adding plural morphology to the latter phrases results in the formation of the ungrammatical **nešta drugi* 'somethings else' and **nešta novi* 'somethings new', with unacceptable ordering of adjectives. With adjectives, the distinction between the indefinite pronoun use of *nešto* and its use as a common lexical noun is particularly clear, contrast (7), with nominal *nešto*, and (8), with pronominal *nešto*.

- (7) ...vsjako novo nešto e dobre zabraveno staro.
every new thing is well forgotten old
'...every new thing is a well-forgotten old one.'
- (8) Predi dve godini xorata glasuvaxa sigurno za promjanata,
before two years the-people voted.3P decisively for the.change,
iskajki nešto po-dobro.
want.GER something better
'Two years ago people voted decisively for change, wanting something better.'

Nešto, therefore, has split into two items in Modern Bulgarian. One, a noun meaning 'thing', has a full range of nominal morphosyntactic properties, and is an innovation. The other, an indefinite pronoun 'something', continues Old Church Slavonic *něčito*, and is much more restricted morphosyntactically. Furthermore, the emergence of a lexical noun from an indefinite pronoun is a change that runs counter to the general expectations of grammaticalization.

How has this happened? A scenario can be sketched out that is relatively similar to that proposed for grammaticalization. In some contexts, 'something', that is, some thing known to exist, but whose actual identity is not known to the speaker or hearer, may be interpreted as 'a thing', the identity of which is known to one or both participants. This ambiguity is already present in the examples in (3) and (4), in particular (4b), and is facilitated by the relatively free word order of Old Church Slavonic. Semantically, then, the shift is relatively easy to envisage at any time.

Morphosyntactically things are more changeable. Each new generation of speakers has to establish whether *nešto* is morphosyntactically a generic noun or an indefinite pronoun. Three facts seem relevant in making this decision. First, the nominative form *nešto* could be the nominal form of a neuter noun (with the ending *-o*), and, in fact, happens to parallel very closely the form of a real generic neuter noun *mjasto* 'place'. Second, in Old Church Slavonic, *něčito* is transparently synchronically

related to the interrogative pronoun *čito* ‘what’. Third, the irregular genitive *něčeso* alerts the learner to the fact that this is no ordinary noun.⁹

The first fact leads learners to posit that *nešto* is a generic noun. The second and third both lead learners to posit that *nešto* is an indefinite pronoun. In Old Church Slavonic, all three facts are relevant, and the indefinite pronoun analysis therefore ‘wins’. Independent developments in Bulgarian remove the second and third facts. The interrogative pronoun *čito* ‘what’ becomes obsolete and is replaced by *kakvo*, formerly ‘which one (neuter)’ (for details, see Ivanova-Mirčeva & Xaralampiev 1999: 114–115). In time, the erosion of the Bulgarian case system, in particular, the loss of genitive forms for nouns, makes the third fact redundant too. With these two gone, the hypothesis that *nešto* is a generic neuter noun rather than a pronoun becomes a highly attractive one. There is some (morphological) evidence for it, and no longer any non-negative evidence against it. Degrammaticalization of *nešto* as a generic pronoun results. This involves a counterdirectional metonymic semantic change, plus category reanalysis from pronoun to noun. Both of these are processes of types typically associated with the change of GENERIC NOUN > INDEFINITE PRONOUN, but lead to the reverse outcome.

2.2 Welsh *eiddo* ‘property’ < ‘his (possessive pronoun)’

A second case of this type of degrammaticalization comes from the history of Welsh, and again involves the emergence of a generic noun. In Modern Welsh, the word *eiddo* may be a noun, meaning ‘property’ in most varieties,¹⁰ or it may function as a masculine third-person singular possessive pronoun, meaning ‘his (one)’, where it forms part of a paradigm along with other possessive pronouns inflected for the person and number of the possessor, and, in the third person singular, the gender of the possessor. Use of the *eiddo*-paradigm as a possessive pronoun in Modern Welsh is rather formal, and it is replaced by various circumlocutions in spoken and less formal varieties. The nominal use is neutral with respect to register. The paradigm of the possessive series of pronouns is given for Middle and Modern Welsh in Table 2.

	Middle Welsh		Modern Welsh	
	singular	plural	singular	plural
1st person	meu	einym	eiddof	eiddom
2nd person	teu	einwch	eiddot	eiddoch
3rd person	eidaw (masc.) eidi (fem.)	eidu(nt)	eiddo (masc.) eiddi (fem.)	eiddynt

Table 2. Paradigm of possessive pronoun *eidaw* / *eiddo* in Middle and Modern Welsh.

⁹ There is also of course a considerable amount of negative evidence, that is, the absence of various types of syntactic modification, and the absence of various morphological forms (especially plural ones). However, it is widely assumed that such negative evidence is only marginally taken into consideration in language acquisition.

¹⁰ The meaning ‘furniture’ is reported for the dialect of northern Ceredigion (R. J. Thomas 1950–2002: 1189).

The pronoun occurs with a preceding definite article if used in an argument position, and is optionally followed by a ‘reinforcing’ pronoun:

- (9) 'Rwy 'n hoffi... ei gwmni a 'i wlad yn well na 'r
 am PROG like.VN his company and his country PRD better than the
 eiddot ti.
 yours you
 ‘I like... his company and his country better than yours.’ (Thorne 1993: 169)

The dual nominal-pronominal nature of *eiddo* in Modern Welsh is most clearly confirmed by the fact that it may be modified by adjectives, demonstratives and quantifiers, whereas other forms of the pronoun, such as the feminine third-person singular *eiddi*, cannot. This is shown in (10)–(12).

- (10) a. eiddo lledrad / coll b. *eiddi ledrad / goll
 EIDDO stolenlost EIDDI stolen lost
 ‘stolen / lost property’ ‘stolen / lost things of hers’
- (11) a. yr eiddo hwn b. *yr eiddi hwn / hon / hyn
 the EIDDO this.MASC the EIDDI this.MASC / FEM / PLUR
 ‘this property, these belongings’ ‘these things of hers’
- (12) a. ei holl eiddo ef b. *ei holl eiddi hi
 3SM.GEN all EIDDO 3SM 3SF.GEN all EIDDI 3SF
 ‘all his belongings’ ‘all her belongings’

If the nominal lexical use of *eiddo* can be shown to be a secondary innovation, then this amounts to degrammaticalization. The word’s etymology, taken together with syntactic and semantic evidence from Middle Welsh, can be used to piece together a historical development that leads via a number of stages away from a pronominal status to that of a lexical noun.

Etymologically *eiddo* is clearly pronominal in origin. It derives historically from the stressed form of the Common Celtic possessive adjective **/esjo/* ‘his’, cognate with Sanskrit *asya*. Stressed **/esjo/* gives **/'eið/* by regular phonological changes. Subsequently, Middle Welsh *eidaw* /'eiðaw/ arises as a result of the addition of the ending /aw/ by analogy with the masculine third-person singular ending of inflected prepositions (see below). Unstressed **/esjo/* gives rise to the Middle Welsh genitive pronominal clitic *y /i/* ‘his, him’ (Modern Welsh *ei*) (Lewis & Pedersen 1937: 216).

If we compare the syntax of the *eiddo*-paradigm in Middle Welsh and Modern Welsh, we also observe a development in the direction of more lexical uses. Broadly speaking, *eiddo*-pronouns appear in three syntactic environments in Middle Welsh. In predicative positions, they occur alone, as in (13a), or with a following ‘reinforcing’ pronoun (glossed as 3SM in (13b)), or a noun phrase, as in (13c):

- (13) a. A mi a weleis y sarff heuytyn treissyaw y llew
 andI PRT saw.1S the snake also PROG attack.VN the lion
 o 'r hynn a oed eidaw...
 of the that REL was his
 'And I saw the snake too forcibly taking from the lion what was his...'
 (YSG 2120)
- b. ef... a gymyrth y dynas yn eidaw ef ehvn.
 he PRT took the kingdom PRD his 3SM himself
 'he... took the kingdom as his own.'
 (BB 194.15-16)
- c. ...yr enys... a uu eidu uy ryeni inheu...
 the island PRT was theirs my parents 1S
 '...the island... was my parents'...'
 (BD 85.27-8)

When they appear in argument positions, they must further be preceded by a definite article, but are otherwise parallel:

- (14) a. Y guyr hynny... a ouynyssant idaw, pa darpar oed yr eidaw...
 the men these PRT ask.PAST.3P to-him what intent was the his
 'These men ... asked him what his intent was.'
 (PKM 32.22)
- b. ...ac y dodet emelltith Duw a 'r eidaw ynteu a
 and PRT put.PAST.IMPERS curse God andthe his 3SM and
 'r hon Gymry oll ar y neb nys katwei...
 the that Wales all on the anyone REL.NEG+3S.ACC keep.IMPF.SUBJ.3S
 '...and the curse of God and his own and that of the whole of Wales was put
 on anyone who might not keep it [the law]...'
 (LIB 2.10-12)

Given the pronominal etymology of the *eiddo*-paradigm, the patterns in (13a) and (14a), with no material following the pronoun, must be historically primary, since we would not want to reconstruct a pronoun that could be followed by another pronoun that expressed the same function.

The pattern in (13b) and (14b), with reinforcing pronoun, then arises from treating the *eiddo*-paradigm syntactically as prepositions, which in Welsh inflect for person, number, and, in the third person singular, gender. In fact, throughout its history, the *eiddo*-paradigm has been strongly influenced, both morphologically and syntactically, by prepositions. This is not as surprising as it first appears, since, in some contexts, the *eiddo*-paradigm occupies the same syntactic environment as a preposition. For instance, in (13a), the form *eidaw* could be replaced by the preposition *gandaw* 'with him', the usual periphrastic way of expressing 'have' in Welsh, which has no lexical verb 'have'. Hence it is not unreasonable to suppose that, during parts of its history, *eiddo* was actually treated as a preposition.¹¹

Morphologically too, the entire *eiddo*-paradigm has been reformed as though it were a preposition, specifically one of the second conjugation (using the traditional labelling of D. S. Evans 1964: 58-60; and S. J. Williams 1980: 127-129), such as *yn* 'in' or *heb* 'without'. Paradigms for *heb* 'without' are given in Table 3. Traditionally, prepositions like these insert /ð/ (orthographic Middle Welsh <d>, Modern Welsh <dd>) in the third person forms. As a result Middle Welsh *eidaw* /'eiðaw/ resembles a third person singular masculine form such as *hebdaw* /'hebðaw/ 'without (him)'. In

¹¹ My thanks to Mair Parry for pointing this fact out to me.

most varieties of modern spoken Welsh, /ð/ is inserted in all the inflected forms of the preposition, making the resemblance with the reformed *eiddo*-paradigm even stronger.

	Modern literary Welsh		Modern spoken Welsh (King 1993: 280)	
	singular	plural	singular	plural
1st person	hebof	hebom	hebdda i	hebddon ni
2nd person	hebot	heboch	hebdat ti	hebdoch chi
3rd person	hebdo (m.) hebdi (f.)	hebdynt	hebdo fe (m.) hebdi hi (f.)	hebbyn nhw

Table 3. Paradigm of second conjugation inflected preposition *heb* ‘without’ in Modern Welsh.

Old and Middle Welsh inflected prepositions allowed either overt or null objects. If the object was null, the ending of the preposition was sufficient to identify the person and number of the object unambiguously. The parallelism between inflected prepositions and possessive pronouns led to the pronouns being treated in the same way, as in (15). Rather than being the possessor itself, the pronoun was reanalyzed as merely a head that agreed with the possessor, which, being unambiguously identified by the ending of the pronoun, could be omitted.

(15) *eiddo* > *eiddo* (*ef*) on the analogy of *hebdo* (*ef*)

It is also worth pointing out that this brought *eiddo* more into line with the syntax of nominal possessor noun phrases such as *llys Arthur* ‘Arthur’s court’, where the possessor is overt and follows the head noun.¹² This development is already complete by the time of the earliest records of Welsh, but can be inferred from the etymology of the *eiddo*-paradigm and the comparative evidence of the absence of overt possessors accompanying possessive pronouns in most other Indo-European languages.

Once pronominal possessors are permitted, the appearance of nominal possessors, that is, the pattern in (13c), is inevitable, but a difficulty arises with agreement. Inflected prepositions have an uninflected form for use with nominal objects (the form *heb* in the case of the example in Table 3). The possessive pronoun lacks such a form. Instead, we find the pronoun agreeing with its possessor. That is, in (13c), we find *eidu uy ryeni inheu* with third-person-plural possessive pronoun *eidu* agreeing with the third person plural possessor *uy ryeni inheu* ‘my parents’. This agreement runs counter to the way the syntax of prepositions works in Middle Welsh, since prepositions never agree with nominal objects. This may account for the unexpected rarity of the pattern in (13c) in Middle Welsh texts.

The scene is now set for degrammaticalization. The *eiddo*-paradigm is syntactically and semantically ambiguous in two contexts. First, when there is no overt possessor, the sequence definite article + *eiddo* is sometimes contextually vague, and some instances, such as that in (16), are open to being interpreted as referring to physical objects (‘the property’) rather than denoting the identity of the possessor

¹² Morphological case plays no role in either instance, since neither Middle nor Modern Welsh has morphological case marking on nouns or independent pronouns.

(‘his’). Both semantic widening and narrowing are involved. In many cases, if something is ‘his’ it will also be a physical entity, hence also ‘property’ (narrowing). On the other hand, widening occurs if the notion of masculine third person singular is viewed as being only incidental, and contextually motivated.

- (16) *yaun yu caffael o perchennauc e da er eydau ket*
 right is get.VN of owner the property the EIDDO although
roder mach arnau
 give.PRES.SUBJ.IMPERS surety on-it
 ‘[if a man gives someone a surety of something which he does not own...] it is
 right for the owner of the property to have what is his [=‘the property’?] even
 though there is a surety on it...’ (LII 39.16–17)

Secondly, constructions of the type in (13c), if the pronoun is third person singular masculine, will almost always be open to reanalysis of *eiddo* as a lexical noun. For instance, in (17), the phrase *eydav vn tewyssav[c]* is syntactically ambiguous. Welsh possessive noun phrases may never begin with a definite article even if they are definite (a ‘construct state’ effect), so, even if *eydav* were a noun, with the whole phrase meaning ‘the property of a single prince’, there would be no preceding definite article. Semantically, the distinction between ‘X’s’ and ‘X’s property’ will only be relevant if a physical object is possessed, and, even then, may be slight, as is also the case in (17).

- (17) ...nyt eydav vn tewyssav[c] e vudvgolyaeth namyn e gwyr
 NEG EIDDO one prince the victory but the men
 a emlado trostvnt.
 REL fight.PRES.SUBJ.3S for-them
 ‘...the victory is not a single prince’s [property?] but the men who fight for
 them.’ (*Rhyddiaith Gymraeg o Lawysgrifau’r 13eg Ganrif*, Peniarth 44, 55.9–12)

The creation of a new lexical item *eiddo* ‘property’ is the result of both of these factors.¹³ The emergence of this new item is best confirmed by the appearance of sequences such as *y eidaw (ef)* ‘his property’, with a preceding possessive clitic *y* ‘his’ in the pattern found with other lexical nouns (cf. *y lys (ef)* ‘his court’). This pattern is illustrated in (18). Dictionaries treat this as an innovation dating from the fourteenth century (R. J. Thomas 1950–2002: 1189; and Lloyd-Jones 1931–63: i.454).

¹³ Both Lloyd-Jones (1931–63) and R. J. Thomas (1950–2002) take the second factor to be responsible for the change, but there are two reasons to doubt this. First, the pattern in (13c) and (17) is quite rare in Middle Welsh texts, so rare that it seems unlikely that it alone could lead to doubt over the categorial status of *eiddo*. Secondly, two other forms of the pronoun, namely first person singular *meu* and second person singular *teu*, also develop for a while into nouns meaning ‘property’, before becoming obsolete. With these, the first factor plays a role, but ambiguous sentences of the type found in (17) cannot be constructed.

- (18) Pan vo marw righill, yn trugared yr arglwyd y byd
 when be.PRES.SUBJ.3S dead sergeant in mercy the lord PRT be.FUT.3S
 y eidaw.
 his property
 ‘Whenever a sergeant dies, his property is at the mercy of the lord.’ (LIB 29.1–2)

Notice also that, with (18), the task of conveying the possessor, previously encoded primarily in the pronoun *eiddo*, is ultimately transferred to the accompanying pronoun or lexical noun phrase. This transfer of function from one element to another element with which it habitually co-occurs is a phenomenon also found in grammaticalization (cf. the transfer of the value of negation to *pas* in French *ne ... pas* constructions).

To sum up, a lexical item meaning ‘property’ has an undisputed etymology based on a pronoun. The historical development from pronoun to noun can be partly reconstructed and partly documented historically as a series of semantic and analogical syntactic changes that lead to the gradual development of lexical properties. This is extremely surprising given a unidirectional view of grammaticalization.

2.3 Welsh *nôl* ‘bring’ < *yn ôl* ‘after (preposition)’

Our final case study is also from Welsh, but this time involves the emergence of a lexical verb from a preposition. Many varieties of present-day Welsh have a verb meaning ‘fetch’ that derives historically from a form of the Middle Welsh complex preposition *yn ol* ‘after’. The most widely distributed form of the verbnoun,¹⁴ found across the whole of north Wales, is *nôl*. Many southern varieties have another variant, *hól* or *ôl*, while in the midlands ‘fetch’ is normally expressed using some other verb, such as *moyn*, or using a periphrasis. For the precise geographical distribution of the forms in the traditional dialects, see A. R. Thomas (1977: 534–535).

For this to be a case of degrammaticalization, it is necessary to demonstrate that *yn ol* in Middle Welsh does indeed behave like a preposition, and that *nôl* in the relevant varieties of present-day Welsh is a verb. We also need to demonstrate that the lexical verbal use and meaning arose from the preposition, and not, say, some historical remnant of an earlier lexical meaning. In fact, for the claim that this is ‘grammaticalization in reverse’, and therefore degrammaticalization, the details of the transition are crucial.

2.3.1 The development of Middle Welsh *yn ol* as grammaticalization

First, consider Middle Welsh *yn ol*. Its etymology is not in dispute. It derives from the preposition *yn* ‘in’ plus a noun *ol* meaning ‘track(s), path, trail’, and is therefore itself an instance of grammaticalization. Together these form a complex preposition of a type that is widespread in Middle and Modern Welsh and in other Celtic languages. In fact, in a number of cases such complex prepositions have completely ousted earlier simple

¹⁴ Celtic languages have nonfinite verbal forms, traditionally referred to as verbnouns, which fulfil a number of infinitive-like and gerund-like functions in other languages. Welsh verbnouns are typically either identical to the verbal stem of inflected forms (as is the case with *nôl*) or related to it by the addition of a lexically idiosyncratic suffix (as in *rhed-* ‘run’ > vn. *rhedeg* or *clyw-* ‘hear’ > vn. *clywed*). Citation forms of verbs throughout are verbnoun forms.

prepositions in concrete uses, for instance, Modern Welsh *o flaen* ‘in front of’ for earlier *rhag*, and Modern Welsh *yn ystod* ‘during, in the course of’ for earlier *er*. The noun *ol* itself is attested in Middle Welsh, as in (19), and survives into present-day Welsh, usually with the slightly more abstract meaning ‘trace’, often in the plural form *olion* ‘traces, remains’.

- (19) Ac ol y march a welei.
and tracks the horse PRT saw.IMPF.3S
‘And he saw the horse’s tracks.’ (P 11.9)

This lexical use of *ol* is relatively infrequent in Middle Welsh texts.¹⁵ The overwhelmingly most frequently attested context for *ol* is as part of complex prepositions, either in *yn ol* or *ar ol*.¹⁶ *Yn ol* is used in two major senses, first of all, a spatial meaning ‘behind, after’, usually directional, illustrated in (20); and, secondly, a temporal meaning ‘after’, illustrated in (21).

- (20) Ac yn ol y baed y kerdassant...
and after the boar PRT walk.PAST.3P
‘And they went after the boar...’ (PKM 55.20)
- (21) A hi a glywei lef corn, ac yn ol llef y corn
and she PRT hear.IMPF.3S sound horn and after sound the horn
llyma hyd blin yn mynet heibaw...
there stag tired PROG go.VN past
‘And she heard the sound of a horn, and after the sound of the horn an exhausted stag went past...’ (PKM 84.14)

2.3.1.1 The grammaticalization path FOOTPRINT > BEHIND

These meanings fit in neatly with a hypothesized historical development fully consistent with the existing results of research on grammaticalization. The spatial meaning arose from grammaticalization of the phrase [*yn ol*] + [noun phrase], literally ‘in the track(s) of’, where [*ol* + noun phrase] once formed a larger possessive noun phrase. Grammaticalization of spatial prepositions from nouns is well attested. For instance, Heine & Kuteva (2002: 141) cite FOOTPRINT > BEHIND as a grammaticalization path, attested, for instance, in Zande, where the preposition *fuo* ‘after’ derives from the noun *fuo* ‘footprints, trace’:

¹⁵ Comments on Middle Welsh are based on an exhaustive search for (*yn / ar*) *ol* in the Middle Welsh texts, *Llyfr Blegywryd (LIB)*, *Pedeir Keinc y Mabinogi (PKM)*, *Peredur (P)*, and *Ystoriaeu Seint Greal (YSG)*.

¹⁶ *Ar ol* occurs once in *Pedeir Keinc y Mabinogi (PKM 33.26)*, with the meaning ‘(spatial) after’. It is more frequent in later Middle Welsh texts, for instance *Ystoriaeu Seint Greal*, where it is also found in a temporal use, and has ultimately come to replace *yn ol* as the primary preposition for both spatial and temporal ‘after’. Since it is not central to the development of the verb *nôl*, further discussion of it will be omitted.

- (22) a. Fuo bahu) du erE) .
 ‘A lion’s footprints are here.’
 b. Mi nandu fuo ko.
 ‘I am going after him.’ (Heine & Kuteva 2002: 141)

A number of syntactic and semantic facts show that grammaticalization of *yn ol* as a preposition was already complete in Middle Welsh. The object of *yn ol* is no longer restricted to noun phrases that denote entities that can plausibly have (foot)prints or tracks. For instance, in (23), weapons can hardly leave tracks, so the function of *yn ol* must be prepositional.

- (23) Ac ar hynny, yn ol yr arueu yd aeth hi.
 and on this after the weapons PRT went she
 ‘And at this, she went after the weapons.’ (PKM 82.22)

There are also cases where, although still spatial and directional, a meaning such as ‘on someone’s trail’ or ‘along someone’s track’ is barely appropriate. In these cases, a generalization of meaning, such as that generally observed in the semantic bleaching often associated with grammaticalization, is apparent (cf. PKM 12.2, 33.14). Both types of evidence suggest that a form of semantic bleaching, specifically a widening of the semantic contexts where the use of the item is appropriate, had already taken place. A further generalization of meaning is noticeable in cases where *yn ol* is used in expressions that mean ‘seek, look for, demand, go after’, where the idea of motion is either only vaguely present, as in (24a), or where it is subordinate to the sense of ‘searching for’, as in (24b).

- (24) a. Mi a baraf iawn y ti yn gyntaf, ac yn ol
 I PRT cause.1S compensation for you PRD first, and after
 uy iawn y bydaf inheu.
 my compensation PRT be.FUT.1S I-too
 ‘I shall get compensation for you first, and [then] I will be after
 compensation for myself.’ (PKM 74.16)
- b. Ac ar hynny nachaf varchawc yn dyuot, ac amouyn a
 and on this lo knight PROG come.VN and ask.VN with
 Pheredur a welsei y kyfryw varchawc yd oed ef
 Peredur Q see.PLUPERF.3S the such knight REL was he
 yn y ol.
 after.3SM
 ‘And thereupon a knight arrived, and asked Peredur whether he had seen a
 knight like the one he was after / he was following.’ (P 28.24)

Note also, that syntactically no verb of motion is present in either of the examples in (24). The example in (24b) might even be considered, in and of itself, as an instance of a degrammaticalization path BEHIND > FOLLOW (cf. Heine & Kuteva’s grammaticalization path FOLLOW > BEHIND).¹⁷ For these reasons, it seems safe to

¹⁷ Clearly, to establish degrammaticalization here would require two extra conditions to be met. Firstly, the meaning ‘follow’ would have to stem historically from the meaning ‘behind’, rather than be the residue of the original meaning of *ol* ‘trail,

conclude that the grammaticalization of *yn ol* from prepositional phrase to preposition had been completed by the time of these texts.

2.3.1.2 The grammaticalization path BEHIND (SPATIAL) > AFTER

Yn ol in Middle Welsh is also used in a temporal meaning ‘after (in time)’. This meaning is secondary. This can be inferred (weakly) from general facts about grammaticalization, namely the fact that, cross-linguistically, temporal meanings typically derive historically from spatial meanings, and, specifically, BEHIND (SPATIAL) > AFTER is a recognized grammaticalization path (Heine & Kuteva 2002: 52). Synchronically in Middle Welsh, evidence for the grammaticalization (although of course not its directionality) comes from cases that are ambiguous between a spatial and temporal meaning, and which therefore instantiate the sort of context in which this scenario for grammaticalization could have arisen:

- (25) a. Ac rac diruawr wres y kyrchwys y bleit a ’e
 and for great heat PRT charge.PAST.3S the wall with 3SM
 yscwyd a ’y tharaw gantaw allan, ac yn y ol ynteu y
 shoulder and 3SF hit.VN by.3SM out and after/behind him his
 wreic.
 wife
 ‘And because of the great heat, he charged the wall with his shoulder and hit it out, and after / behind him his wife [went] too.’ (PKM 36.21)
- b. ...a ’r baed yn kyrchu yr gaer yn uuan, a ’r
 and the boar PROG head.for to.the fortress PRD swift and the
 cwn yn y ol.
 dogs after/behind him
 ‘...and the boar was heading for the fortress swiftly, with the dogs after it / behind it.’ (PKM 55.23)

2.3.1.3 The path AFTER > ACCORDING TO

The two uses discussed so far account for the majority of cases of *yn ol* in Middle Welsh texts, but some other uses are worth noting. One is its use after nouns and verbs of emotion to encode the cause or source of the emotion, for instance, *hiraeth yn ol* ‘longing for, after’ (P 40.18), or *galar yn ol* ‘mourning for, after’ (PKM 26.24). This use is again consistent with the view that *yn ol* undergoes grammaticalization, since the acquisition of more abstract grammatical prepositional uses in fact represents a further stage of the process.

For the future development of the preposition, one particular sense for *yn ol*, namely ‘according to’, is of great importance, even though it is rare, albeit attested, in Middle Welsh. *Geiriadur Prifysgol Cymru* (the University of Wales dictionary, R. J. Thomas 1950–2002) cites the following example from a medieval law text as the

track’. Secondly, in addition to grammaticalization of semantics, grammaticalization of syntax, in the form of a category change from preposition to verb, would be necessary. The first condition may well not be met in this instance. The second condition is clearly not met with the meaning of ‘follow’, but it is hard to see why degrammaticalization of BEHIND > FOLLOW should be excluded in principle.

earliest example (the discussion concerns whether an unborn child should be treated as male or female for the purposes of calculating blood-money):

- (26) E keureyth a deweyt bot en yaunhaf barnu
 the law PRT say.PRES.3S be.VN PRD more.correct judge.VN
 en ol e peth pennaf, a bot galanas gur arnau
 according-to the thing main and be.VN blood.money man on.3SM
 a henne ene uedydhyer.
 and that until baptize.PRES.SUBJ.IMPERS
 ‘The law says that it is more correct to judge according to the more important
 thing, and that it [the unborn child] has the blood-money value of a man until it
 is baptized.’ (LII 65.9–11)

It is natural to suppose that this meaning arises out of the temporal meaning of *yn ol*, from contexts where to do something later than someone else is to do something in the same way, following their pattern. The example in (27) illustrates the sort of context that is envisaged.

- (27) A gwedy hynny Lawnslot a dyngawd a Gwalchmei
 and after that Lawnslot PRT swear.PAST.3S and Gwalchmai
 a Pheredur a Bwrt a Lionel a chwbyl o’r milwyr
 and Peredur and Bwrt and Lionel and all of.the warriors
 ereill pob un yn ol y gilyd.
 other every one after each.other
 ‘And after that, Lawnslot swore an oath, and Gwalchmai and Peredur and
 Bwrt and Lionel and all the other warriors, each one after the others / each
 one in the same way as the others.’ (YSG 451)

Again, this is paralleled in other languages, for instance German *nach* ‘after’ > ‘according to’, although present-day Welsh has perhaps gone further than most in abandoning the spatial and temporal meanings of Middle Welsh *yn ol* entirely, in favour of the meaning ‘according to’, which in present-day Welsh is the only usual meaning of *yn ôl*. *Yn ôl* has become obsolete as a preposition meaning ‘after’, although the spatial meaning survives in some related adverbial uses, in particular, the adverb *yn ôl* ‘behind’ in contexts such as ‘to leave something behind’.¹⁸

2.3.2 Early Modern Welsh *nôl* as *degrammaticalization*

So far, the story of *yn ol* has been one of successive processes of grammaticalization. It seems clear that by late Middle Welsh, *yn ol* was fully grammaticalized as a spatial and temporal preposition, having moved away from its origins as a prepositional phrase, and, alongside these meanings, had even acquired the meaning ‘according to’, the ultimate end of its grammaticalization path. It is therefore perhaps surprising that,

¹⁸ The main preposition for ‘(temporal) after’ and ‘(spatial) behind, after’ is *ar ôl* in Modern Welsh, a form that first appears (with both senses) in the fourteenth-century text *Ystoria Bown de Hamtwn* (R. J. Thomas 1959–2002: 2640). It is also well attested in the late-fourteenth-century *Ystoriaeue Seint Greal*. This is of course also a grammaticalization from *ol* ‘track, trail’.

precisely at this point in its development, it should have begun to develop a highly lexical use, which led eventually to the emergence of a new verb, (nonfinite) verbnoun *nól* ‘fetch’. This verb is found in a number of northern Welsh dialects today, alongside variants *dól* (Bangor, Fynes-Clinton 1913: 397) and (*h*)*ól* (southern). For some speakers, it has a full regular inflectional paradigm, for instance, with first person singular future *nola* ‘I will fetch’ or third person singular past *nolodd* ‘he, she fetched’.

The University of Wales dictionary cites the following example as a case of the type of sentence from which this may have arisen (R. J. Thomas 1950–2002: 2595):

- (28) Yna yd aeth y gweisson yn ol y varch a 'e arueu
 then PRT went the lads after his horse and his weapons
 y Arthur.
 for Arthur.
 ‘Then the lads went after his horse and his weapons for Arthur. / Then the lads
 went to fetch his horse and his weapons for Arthur.’ (P 41.19)

There are in fact plenty of similar examples in Middle Welsh. Further cases are given in (29).

- (29) a. Dos yn ol y marchawc a aeth odyma y'r
 weirglawd...
 go.IMPER after the knight REL went.3S from-here to.the meadow
 ‘Go after the knight who went away from here to the meadow...’ (P 13.26)
- b. drwc y medreist am dyn fol a yrreist yn [ol] y
 bad PRTbe-able.PAST.2S for man fool REL send.PAST.2Safter the
 marchawc.
 knight
 ‘you treated badly the fool that you sent after the knight’ (P 14.27)
- c. “Ie, Arglwyd,” heb wy, “anuon etwa genhadeu yn y ol.”
 yes lord saidthey send.IMPER again messengers after.him
 “‘Yes, Lord,’ they said, ‘send more messengers after him.’” (PKM 33.14)
- d. “Kyuodwch ... ac ewch yn y ol,” heb ef...
 get-up.IMPER.2P and go.IMPER.2P after-him said he
 “‘Get up ... and go after him,’ he said...’ (PKM 33.16)

Cases such as these seem entirely plausible as the input construction that led to the emergence of a new verb. Speakers and learners encountering data such as (28) are faced with an ambiguity: is *yn ol* a preposition ‘after’ or a sequence of the (pre-existing) purpose marker *y* ‘to’ plus an unknown nonfinite verb ‘fetch’?¹⁹ Such ambiguity sets the scene for the structural reanalysis in (30).²⁰

¹⁹ It is possible that speakers considered the possibility that this involved a verb with the meaning ‘fetch’ because of contact with English, which lexicalizes the concept ‘go and bring back’ as *fetch*. Previously, Welsh had expressed this with the verb *dwyn* ‘bear, carry, take’, which had a wider range of meanings, and did not necessarily imply going somewhere in order to get something. The change could therefore be characterized as contact-induced degrammaticalization (cf. on contact-induced grammaticalization, Heine & Kuteva 2003). However, this is clearly not the only factor. It is also worth noting that the verb *dwyn* had moved out of the relevant field in some

- (30) yna yd aeth y gweisson [PP [P yn ol] y varch]
 then PRT went the lads after his horse
- ⇒ yna yd aeth y gweisson [IP y [VP [V nol] y varch]]
 then PRT went the lads to fetch his horse

If this is indeed the source, then some of the properties of the process are surprising. The change from ‘go after’ to ‘go and fetch’ seems to involve pragmatic inferencing in the same way as standard examples of grammaticalization. If someone goes after something, then, on the assumption that the speaker is making a relevant contribution, it is quite likely that the purpose of going after something is to bring it back, in which case ‘go after’ may imply ‘go and fetch’. This is illustrated with English *after* in (31). In (31a), there is no implicature; whereas, in (31b), it is implied that I will bring bread back if at all possible. In English, this implicature is not yet conventionalized with *after*, and pragmatic inferencing has never led to degrammaticalization.

- (31) a. I went after him (to find out where he was going).
 b. I went after some bread because we ran out.

This seems to be the same stage as Middle Welsh has reached in (29). In (29a) and (29b), it is clear from context that Peredur (the addressee in (29a), and the ‘foolish man’ in (29b)) is not supposed to return with the knight, but rather merely to find him and challenge him. In (29c) and (29d), the messengers are given the option of bringing back the Irish king, Matholwch (‘he’ in both sentences). However, in none of these examples does it seem to be the case that use of *yn ol* means that someone *must* attempt to bring something or someone back. In that sense, ‘fetching’ may be implied or inferred but is not (yet) part of the lexical meaning. The pragmatic inference has not yet been conventionalized and has not yet led to degrammaticalization.

This case clearly involves pragmatic inferencing. If so, it demonstrates that pragmatic inferencing is a bidirectional process, in the sense that it can lead from a more grammatical to a more lexical meaning despite the fact that standard examples of it (such as English *going to*) involve the reverse direction. This inevitably leads to the question of why pragmatic inferencing should lead from the grammatical to the lexical in the particular cases under consideration, that is, what makes them different from standard cases.

Degrammaticalization of *nól* involves ‘split’, or ‘divergence’ in the sense of Hopper & Traugott (1993: 116–117), a phenomenon also typical of grammaticalization. The context of incipient degrammaticalization is only one of the environments in which *yn ol* may appear. In the other environments it survives with its existing meaning.

It is hard to be certain exactly when *nól* had fully degrammaticalized, and thereby acquired the status of a full lexical verb. The first examples of the verb *nól* cited

varieties of spoken Welsh by the sixteenth century, having developed the dominant sense ‘steal, take without permission’.

²⁰ An anonymous referee for *Linguistics* suggests an alternate bracketing for the input structure, namely [PP [P yn] [PP [P ol] y varch]]. Even this structure would be consistent with the essential point that the category status of *yn ol* is functional (P) in the input to reanalysis, but lexical (V) in the output.

in the University of Wales dictionary are from the sixteenth century (R. J. Thomas 1950–2002: 2595):

- (32) a. Da i awen aeth Duw yw nol...
 good his muse go.PAST.3S God to+him fetch.VN...
 ‘His muse [being] good, God went to fetch him...’
 (*Barddoniaeth Wiliam Llŷn* 120)
- b. ...anfon kenad I nol gwrlais a Orvc Uthr.
 send.VN messenger to fetch.VN Gwrlais PRT did Uthr
 ‘...Uthr sent a messenger to fetch Gwrlais.’ (Llanstephan 195, 153)

Some (for instance, (32b)), but by no means all, of the early ones could possibly be analyzed as instances of the preposition, although they are spelled in a way that implies the writers understood them to be distinct from the preposition. Examples with verbal inflection (for instance *nolodd* ‘fetched (third person singular past)’, and a new variant of the verbnoun, *noli*) are given for the seventeenth century, and these provide clear evidence that *nól* had developed into a verb by this time.²¹ This is illustrated in (33), with a second-person-plural imperative form, *nolwch*.

- (33) Nolwch y Brenin i ’w examnio.
 fetch.2P.IMPER the King to 3SM examine.VN
 ‘Fetch the King to be cross-examined.’ (RhC 860, late 17th century)

A parallel development occurred through reanalysis based on the form *hol*, which would appear in the first person singular, third person singular feminine, first person plural and the third person plural of the paradigm of the preposition *yn ol* (see the paradigm in Table 4).²² Plausible inputs to the reanalysis, with good contexts for pragmatic inferencing, can be found in Middle Welsh, such as (34), where ‘after me’ implies ‘to fetch me’.

- (34) “Dywet,” heb y marchawc, “a weleisti neb o ’r
 tell.IMPER said the knight Q see.PAST.2s+you anyone from the
 llys yn dyuot y’ m hol i?”
 court PROG come.VN after.me
 “‘Tell me,’ said the knight, ‘have you seen anyone from the court coming after
 me?’” (P 14.7)

Unambiguously verbal uses of *hól* are found from around 1600, and are illustrated in (35). The earliest example in the University of Wales dictionary dates from a text composed in 1687.

²¹ The gap between reanalysis of *nól* as a verb and the appearance of inflected forms appears to be a real phenomenon, and not simply a gap in textual attestation. This is suggested by the fact that, for some speakers of modern Welsh, *nól* is a defective verb, still having only a verbnoun.

²² Sources for Table 4: 1s. P 14.7; 2s. PKM 19.7, 76.23; P 9.5; 3sm. PKM 1.17, 22.24, 26.24, 32.14, 32.16, 34.16, 35.21, 54.23, 82.16; 3sf. PKM 9.27, 12.2; 1p. PKM 69.14; 2p. PKM 69.21; 3p. PKM 28.18; P 47.9.

- (35) a. ...mi a ddanfonaf yvory yddy hol ef.
 I PRTsend.1S tomorrow to+3SM fetch.VN him
 ‘...I shall send tomorrow to fetch him.’ (RhG i.124.3)
- b. ...gan erchi yddo veddwl am yr awr nodedig i vyned
 with ask.VN to.3SM think.VN about the hour set to go.VN
 i hol y daliad.
 to fetch.VN his reward.
 ‘...asking him to think about the hour set to go to collect his reward.’
 (GR 2168)
- c. ...ony ddaw hi i hvnan i ’r llong yddy hol ef.
 unless come.3S she herself to the ship to+3SM fetch.VN it
 ‘...unless she comes herself to the ship to fetch it.’ (GR 3403)

The geographical distribution of forms, *nôl* in northern dialects, *hól* in southern dialects, may at first sight seem arbitrary, but a very satisfactory historical explanation for their distribution can in fact be arrived at, provided the reanalysis account of the historical development is accepted. The form *nôl* can be reached by reanalysis of *yn ol* in virtually any context. In particular a non-pronominal context is entirely adequate; compare example (28), which could give rise to the reanalysis *yn ol y varch* ‘after his horse’ > *nol y varch* ‘to fetch his horse’.

	sing.	plur.
1st person	y’ <i>m</i> hol (i)	yn an ol
2nd person	y’ <i>th</i> ol (di)	yn ych ol (chwi)
3rd person	yn y ol (m.) yn y hol (f.)	yn eu hol

Table 4. Paradigm of *yn ol* in Middle Welsh (minor spelling variation ignored).

With *hól*, however, things are different. On the basis of the evidence in Table 4, it seems that *hól* can arise only by reanalysis of the first person singular form of *yn ol*, namely *y’*m* hol (i)*. All other forms either contain an /n/, and hence would give rise to *nol*; or else they lack the /h/, and therefore fail to account for the initial /h/ of *hól*. This means that, at first sight, the frequency of the syntactic pattern that gave rise to *nôl* was far higher than the frequency of the syntactic pattern that gave rise to *hól*.

In fact, the input pattern for *hól* must have been quite rare, so rare that the reanalysis seems almost implausible. I suggest, therefore, that this reanalysis was unlikely to take place in the ‘classical’ Middle Welsh variety, but must have taken place in a variety with slightly different linguistic properties. Sixteenth-century southern Welsh is a variety that provides rather better conditions. First of all, the loss of the /n/ of *yn ol*, found in Middle Welsh before a first- or second-person preposed genitive pronoun (*y’*m* hol* ‘after me’, *y’*th* ol* ‘after you’), spread analogically at least to the first person plural. Thus, we find *y’*n* herbryn ni* ‘against us’ (GR 585) for earlier expected *yn an herbryn ni* (and therefore presumably also *y’*n* hol ni* for earlier *yn an ol* ‘behind us’, PKM 71.13–14). Secondly, loss of or uncertainty about the phoneme /h/, a characteristic of southern varieties of Welsh today, had probably already begun to be manifested in the south by this time.

3 EXPLANATIONS FOR DEGRAMMATICALIZATION

These three case studies document fairly clearly the existence of a type of syntactic degrammaticalization that has a good claim to being described as ‘grammaticalization in reverse’. Detailed examination of the evidence for them demonstrates the involvement of syntactic reanalysis, both as category reanalysis (pronoun > noun and preposition > verb) and structural reanalysis (in the case of *nól*, [PP [p yn ol] NP] > [IP y [VP [v nol] NP]]). Pragmatic inferencing is very clearly involved in the case of *nól* (‘go after something’ implies ‘go and fetch something’). Metonymic developments are also evident in the case of *eiddo* (metonymic link from something ‘belonging to someone’ to being ‘property’), which could perhaps also be formulated in terms of pragmatic inferencing (‘X belongs to someone’ implies ‘X is property’). The semantic shift is more nebulous in the case of *nešto*, but may also be broadly characterized as metonymic.

The processes seem to be of the same type (reanalysis, pragmatic inferencing, lexical semantic change involving metonymy) as those encountered in grammaticalization, but with the outcome reversed. The conclusion then is that both reanalysis and pragmatic inferencing are bidirectional processes. This conclusion is not particularly controversial in the case of reanalysis, but is perhaps more surprising in the case of the semantic and pragmatic developments associated with grammaticalization, for instance, in the light of such statements as “Change by inference as well as by generalization appears to be unidirectional” (Bybee, Perkins & Pagliuca 1994: 13) (cf. also the discussion in Traugott & König 1991: 192–193).²³ Many linguists have considered grammaticalization to be built up out of a number of separate and conceptually independent processes (for instance, Fischer 2000: 149–153). If this view is taken, it is to be expected that each of the individual processes should have properties of its own. In fact, on such a view, unidirectionality might be expected to hold (or not hold) independently on each of a number of dimensions (Campbell 2001: 132–133). The conclusion that some of them are bidirectional leads inevitably to the conclusion that, under some circumstances, they may together build cases of grammaticalization that are counterdirectional.

The relationship between grammaticalization and degrammaticalization in this respect may be viewed as parallel to that between assimilation and dissimilation in sound change. While assimilation is regarded as the historical norm, no one denies the existence of dissimilation, which is generally regarded as being the result of language learners’ attempts to ‘discount’ or reverse processes of assimilation in their linguistic experience, a process which Ohala (1993: 249) describes as “‘correction’ erroneously implemented’.

3.1 Divergence and isolation

The similarities between grammaticalization and syntactic lexicalization do not stop here. Grammaticalization frequently leads to the split or divergence of a single item. As

²³ It is necessary to distinguish two ‘directions’ in pragmatic inferencing here: from concrete to abstract, and strengthening rather than weakening of inference. Clearly, the current example violates the former by creating a more concrete meaning from a less concrete one, but it does involve strengthening of inference.

it grammaticalizes in one context, it remains ungrammaticalized in another, and the two diverge, resulting in independent lexical items, and independent subsequent histories. This has happened, for instance, with English *one*, the weak form of which grammaticalized into the indefinite article *a(n)* in Old English (Hopper & Traugott 1993: 117). Once the two diverged, they were subject to radically different phonological developments.

Divergence is a feature of all the degrammaticalizations discussed here. In each case, the ‘undegrammaticalized’ variant survives. In the case of Bulgarian *nešto* and Welsh *eiddo*, it survives in more or less the same function and syntax as it had previously. In the case of Welsh *nôl*, divergence involves both items shifting, and, in fact, a striking aspect of the development is the large extent to which divergence had taken place *prior* to degrammaticalization, and may therefore be viewed as a cause of degrammaticalization. The degrammaticalized verb *nôl* is attested in the seventeenth century, by which time the preposition *yn ôl* was already seriously restricted in use in its former core spatial and temporal meanings. This is clearer from Figure 1, which summarizes the functions of *yn ôl* / *nôl* over time.

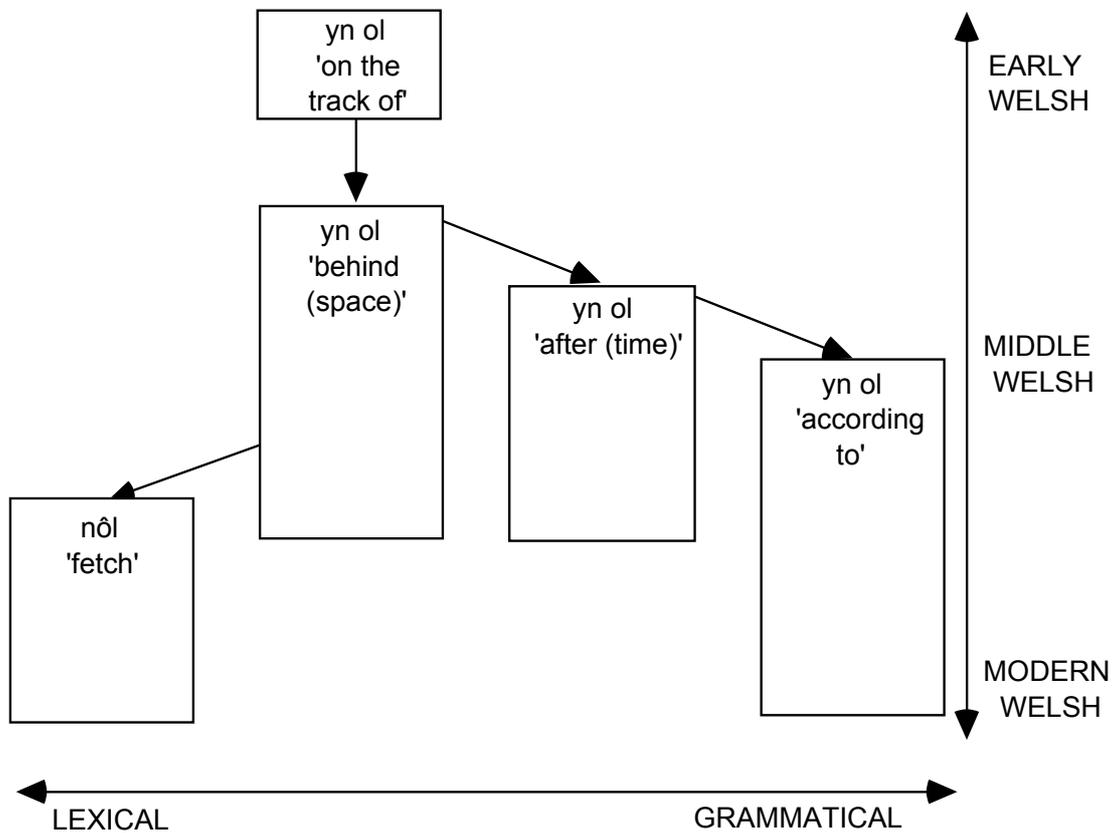


Figure 1. Summary of the development of *yn ôl* / *nôl*.

This suggests a restriction on degrammaticalization, along the lines of ‘In order to degrammaticalize, a form must have become grammatically or semantically isolated’. By ‘semantically isolated’ is meant that the lexical semantics of the item in question no longer involves a regular relationship with other items containing the same morpheme(s). Consider how this applies to the case of Welsh *nôl*. At the relevant point in the history of Welsh, *yn ôl* had developed the primary meaning ‘according to’ via the

grammaticalization path AFTER (TEMPORAL) > ACCORDING TO. It had more or less abandoned its earlier temporal and spatial meanings ‘after’ and ‘behind’, except that these survived as relics in a few frequent (perhaps idiomatic) constructions, such as ‘go after’ and ‘leave behind’. The semantic relationship between *yn ôl* meaning ‘according to’ and the cases where it meant ‘(prepositional) (spatial) after’ or ‘(adverbial) behind’ was not regular – the relationship cannot be derived by synchronic pragmatic inference, since the intervening meaning ‘(temporal) after’ had been lost entirely. Learners must therefore have posited ‘according to’ as the meaning of *yn ôl*, but were then left with the task of working out the syntax and semantics of cases where *yn ôl* clearly does not mean ‘according to’. There is no synchronic link from these cases to another item, and therefore no limit on the range of hypotheses that a learner may entertain about the syntactic category and semantic value of the item in question. My claim is that this is a possible scenario for degrammaticalization.

Isolation is also evident in the case of *eiddo* and *nešto*, albeit not as a result of grammaticalization. Isolation here is grammatical: by grammatical isolation is meant that the item does not form part of a synchronic paradigm (broadly defined to include not only paradigms of case and person-number forms, but also other paradigmatically organized parts of the grammar, including, here, paradigmatic relations within pronominal systems). *Nešto* becomes isolated as the result of two other changes: its counterpart in the system of interrogative pronouns, Old Church Slavonic *čto* ‘what’, becomes obsolete and is replaced by the unrelated *kakvo*; and its case forms become obsolete as a result of the loss of morphological case in Bulgarian nouns. With *eiddo*, the situation is less clear, but other forms of the pronoun (such as *eiddof* ‘mine’ or *eiddot* ‘yours’) are considered literary in modern Welsh, and are not used in unscripted speech.²⁴ Use of these pronouns may already have been restricted in the spoken language at the period when degrammaticalization was taking place. This would have left the third-person masculine singular form *eiddo* alone in the spoken language without a paradigm.

In some other cases of degrammaticalization (deflexion) it is noticeable that the degrammaticalized form has become isolated within the grammatical system as a result of other independent changes. With the English and Swedish possessives, it seems crucial that both languages in question were losing case marking on nouns during the period in which the genitive case ending degrammaticalized as a possessive clitic. The genitive ending could no longer be fitted into a productive grammatical system. The same applies to Irish *-muid*, which was the only person-number ending in an otherwise analytic paradigm at the point when it degrammaticalized into the pronoun *muid(e)* ‘we’ (Bybee, Perkins & Pagliuca 1994: 13–14; see also Newmeyer 2001: 208).

This leaves degrammaticalization closely related to exaptation, the reuse of marginal material (‘junk’) in a new grammatical context (Lass 1990). Traugott (2001: 12) has recently drawn attention to the special place that exaptation may have in degrammaticalization, noting that, in exaptation, “individual morphemes have become relatively unanalyzable, or have lost connectedness with other member of their class, and have opportunistically and idiosyncratically been reused”. This is effectively the same as the scenario that has been sketched for the cases of degrammaticalization discussed above. The differences are that exaptation standardly (but perhaps not

²⁴ In spoken present-day Welsh, either a circumlocution using the verb *piau* ‘belongs’ would be used, or else the pronominal *un* ‘one’ (e.g. *f’un i* ‘my one’ with *fy* ... *i* ‘my’).

necessarily) involves morphemes, not words, and that, in standard examples of exaptation, the function of the morpheme undergoing exaptation is said to have become completely obsolete before exaptation takes place.²⁵ In the examples discussed here, the items had become isolated, having lost connectedness with other members of their class, but had not become meaningless.²⁶ They were then reassigned to a different class (category). Whereas in standard cases of exaptation, the speakers are left with two options – either to create a new function for the ‘junk’ material, or else to do away with it entirely – the cases of degrammaticalization discussed here seem to allow another possibility, namely to allow the item in question to continue in its marginal function (cf. the retention of *yn ôl* as a historical relic in modern Welsh as an adverb meaning ‘(left) behind’). Traugott (2001: 13) further notes that “it is probable that some of the most important counterexamples [to unidirectionality of grammaticalization] will turn out to be instances of exaptation”. Given a somewhat modified definition of exaptation, the examples discussed here could be accommodated within it.²⁷

3.2 Morphological factors in degrammaticalization

Another factor which seems relevant in degrammaticalization, and which may represent a serious barrier to it, is the fact that the target categories for degrammaticalization are typically inflected. Target categories for degrammaticalization, nouns and verbs, typically inflect, whereas target categories for grammaticalization, such as complementizers, prepositions or grammatical markers, do not need to acquire

²⁵ It is hard to see that this can really be the case, since it would imply that it is possible in language acquisition to learn the distribution of a morpheme that has no function (whether formal or semantic).

²⁶ Heine (2003) makes a distinction between exaptation, where “grammatical forms which have lost most or all of their semantic content ... are put to new uses as semantically distinctive grammatical forms”, and adaptation, where “old taxa are adapted to new taxonomic categories ... in particular to adapt grammatical forms to new word classes or morphological paradigms”. Heine calls adaptation a process, and, if it is a process, it is presumably identical to what has been called category reanalysis in this paper. Exaptation is related, in that it is category reanalysis (primarily of morphological categories), where the input form has already become obsolete as a member of its original category. I agree with Heine that category reanalysis (adaptation) does not exhibit any exceptionless directionality, although category reanalyses do tend towards moving in the direction of more grammatical. I also agree that category reanalysis is one of the components of grammaticalization. However, it should be clear from details of the exposition in this paper that his assertion that category reanalysis / adaptation and exaptation are so different from grammaticalization that counterdirectional instances of them do not amount to counterexamples to unidirectionality is not justified.

²⁷ Kiparsky (2004) suggests that “apparent exceptions [to the unidirectionality of grammaticalization] are really instances of analogical change”. This seems to work better for the cases of deflexion that he discusses than for the cases of syntactic lexicalization discussed here. In particular, it is hard to see the reassignment of *yn ôl* to the category of verb as being analogical, unless the definition of analogy is defined so broadly as to be effectively meaningless.

inflection as they grammaticalize. For degrammaticalization to succeed, a grammatical item must be of a phonological form that could plausibly also be an inflected form of some lexical item. That is, for a grammatical item to be reanalyzed as a verb, it must look as though it has some appropriate form of person, number and tense marking. Similarly, for it to degrammaticalize as a noun, it must look as though it has some appropriate form of case and number marking. This situation must arise by chance, but in a highly inflected language this is unlikely. From the start, then, degrammaticalization will be restricted to occurring in relatively isolating languages, or, at least, subsystems of particular languages that are in the process of becoming more isolating.

This is really a subcase of the generalization that the potential for structural ambiguity (the availability of two plausible competing analyses for speakers) is central to reanalysis (Timberlake 1977), and is perhaps a necessary condition for it. If the existing grammar of the language makes available two possible morphological analyses for the item in question, the possibility of reanalysis can be entertained. If the inflectional system is such that no plausible analysis of the item as a member of a lexical category is available, then reanalysis is excluded.

Again, chance factors in the three case studies examined here happen to provide the right conditions. In the cases of Welsh *nól*, the crucial morphological factor is the fact that, although most Welsh nonfinite verbs (verbnouns) are formed from a verbal stem plus a suffix, a number of very common ones have a zero ending. Examples include *agor* ‘open’, *ateb* ‘answer’, *cau* ‘close’, *dal* ‘catch, hold’ and *deall* ‘understand’. *Nól* is therefore a perfectly plausible form for a verbnoun. Similar factors apply in the case of Bulgarian *nešto*, which just happens to be a morphologically plausible form for a nominative singular neuter noun (with ending *-o*). In the case of *eiddo*, the fact that it was a highly plausible form for a masculine third-person-singular form of an inflected preposition (despite unambiguous syntactic and semantic evidence that it was not a preposition) seems to have played an important role in its early development. Later, it was important that Welsh places no restrictions on the form of nouns, and that nouns in the language inflect only for number (not case). Furthermore, a noun meaning ‘property’ could be assumed to be a mass noun with no plural, hence even the negative evidence (the absence of an attested plural form) could play no role.

Finally, it is worth pointing out that, if it is assumed that morphosyntactic change can occur during the discontinuity of language transmission between generations, it is hard to see how such counterdirectional category reassignment could be excluded from a theory of language change. Given that children acquiring their language have no direct access to the grammar(s) that form the basis of the adult language that they hear around them (Andersen 1973, Janda 2001), they must build hypotheses about the category membership of each item in their language. As Janda (2001: 267) points out, they “cannot know ... if the global status that earlier speakers ... assigned to a particular linguistic element ... was lexical or grammatical”. This means that category reanalysis from grammatical to lexical should in principle be possible during the discontinuity of transmission between generations.

4 CONCLUSION

This paper has focused on three case studies of degrammaticalization that involve upgrading of a grammatical item to a fully lexical category. Despite first appearances, these cases turn out to have little in common with standard examples of lexicalization of

grammatical items, such as *to down (a beer)*. Rather, they involve processes of syntactic reanalysis, and in some cases also pragmatic inferencing. The continuity of development from the grammatical item to its degrammaticalized lexical form is striking. The similarity of the processes involved to those found in grammaticalization justifies the claim that these cases really are instances of ‘grammaticalization in reverse’, and therefore represent the syntactic counterpart of deflexion, which often involves upgrading from affix to clitic or phonological word via morphological reanalysis.

The existence of these examples leads inevitably to the question as to why degrammaticalization of this type should be rare. These three examples suggest two conditions that may severely limit the possibilities for this type of degrammaticalization:

- (i) the potential source of degrammaticalization must have become grammatically or semantically isolated;
- (ii) the potential source of degrammaticalization must (by chance) be morphophonologically acceptable (ambiguous) as some potential form of a lexical category.

Further instances of this type of degrammaticalization need to be examined in order to establish the extent to which these conditions apply more generally.

REFERENCES

Texts cited

- BB* *Brut y brenhinedd: Cotton Cleopatra version*, John Jay Parry (ed.). Cambridge, Mass.: Mediaeval Academy of America, 1937.
- BD* *Brut Dingestow*, Henry Lewis (ed.). Cardiff: Gwasg Prifysgol Cymru, 1942.
- GR* *Gesta romanorum*, Patricia Williams (ed.). Cardiff: Gwasg Prifysgol Cymru, 2000.
- LLB* *Cyfreithiau Hywel Dda yn ôl Llyfr Blegywryd*, Stephen J. Williams and J. Enoch Powell (eds.). Cardiff: Gwasg Prifysgol Cymru, 1942.
- LLI* *Llyfr Iorwerth*, Aled Rhys Wiliam (ed.). Cardiff: University of Wales Press, 1960.
- Mar.* *Mariinskoe četveroevangelie s primečanjami i priloženijami*, I. V. Jagić (ed.). Berlin: Weidmann, 1883 (reprinted Graz: Akademische Druck- und Verlagsanstalt, 1960).
- P* *Historia Peredur vab Efracw*, Glenys Witchard Goetinck (ed.). Cardiff: Gwasg Prifysgol Cymru, 1976.
- PKM* *Pedeir Keinc y Mabinogi*, Ifor Williams (ed.). Cardiff: Gwasg Prifysgol Cymru, 1930.
- RhC* ‘Y rhyfel cartrefol’, National Library of Wales, Cwrtmawr ms. 42.
- RhG* *Rhyddiaith Gymraeg, i. Detholion o lawysgrifau 1488–1609*, T. H. Parry-Williams (ed.). Cardiff: Gwasg Prifysgol Cymru, 1954.

Rhyddiaith Gymraeg o lawysgrifau'r 13eg ganrif, G. R. Isaac and Simon Rodway (eds.). Aberystwyth: Adrannau'r Gymraeg, Prifysgol Cymru, 2002 (CD-ROM).

Supr. Suprasălski ili Retkov sbornik, Jordan Zaimov and Mario Kapaldo (eds.). Sofia: Izdatelstvo na Bălgarskata Akademija na Naukite, 1982.

YSG Ystoryaeu Seint Greal, Thomas Jones (ed.). Cardiff: Gwasg Prifysgol Cymru, 1992.

Secondary literature

Andersen, Henning (1973). Abductive and deductive change. *Language* 49, 765–793.

Beths, Frank (1999). The history of *dare* and the status of unidirectionality. *Linguistics* 37, 1069–1110.

Börjars, Kersti (2003). Morphological status and (de)grammaticalization: The Swedish possessive. *Nordic Journal of Linguistics* 26, 133–163.

Börjars, Kersti; Eythórsson, Tolli; and Vincent, Nigel (2003). On defining degrammaticalisation. Paper presented at the Linguistics Association of Great Britain Spring Meeting 2003, University of Sheffield.

Burridge, Kate (1998). From modal auxiliary to lexical verb: The curious case of Pennsylvania German *wotte*. In *Historical linguistics 1995: Selected Papers from the 12th International Conference on Historical Linguistics, Manchester, August 1995*, Richard M. Hogg and Linda van Bergen (eds.), 19–33. Amsterdam: Benjamins.

Bybee, Joan; Perkins, Revere; and Pagliuca, William (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Campbell, Lyle (1991). Some grammaticalization changes in Estonian and their implications. In *Approaches to Grammaticalization*, Elizabeth Closs Traugott and Bernd Heine (eds.), 285–299. Amsterdam: Benjamins.

Campbell, Lyle (2001). What's wrong with grammaticalization? *Language Sciences* 23, 113–161.

Čolakova, Kristalina. (1977–2000). *Rečnik na bălgarskija ezik*. Sofia: Izdatelstvo na Bălgarskata Akademija na Naukite / Akademično izdatelstvo Prof. Marin Drinov / Emas.

Doyle, Aidan (2002). Yesterday's affixes as today's clitics. In *New Reflections on Grammaticalization*, Ilse Wischer and Gabriele Diewald (eds.), 67–81. Amsterdam: Benjamins.

Evans, D. Simon (1964). *Grammar of Middle Welsh*. Dublin: Dublin Institute for Advanced Studies.

Fischer, Olga (2000). Grammaticalization: Unidirectional, non-reversible? The case of *to* before the infinitive in English. In *Pathways of change: Grammaticalization in English*, Olga Fischer, Anette Rosenbach and Dieter Stein (eds.), 149–169. Amsterdam: Benjamins.

Fynes-Clinton, O. H. (1913). *The Welsh Vocabulary of the Bangor District*. Oxford: Oxford University Press.

Haspelmath, Martin (1991). Zur Grammatikalisierung von Indefinitpronomina. In *Sprachwandel und seine Prinzipien*, Norbert Boretzky, Werner Enninger, Benedikt Jeßing and Thomas Stolz (eds.), 103–125. Bochum: Brockmeyer.

Haspelmath, Martin (1997). *Indefinite Pronouns*. Oxford: Clarendon Press.

- Haspelmath, Martin (1998). Does grammaticalization need reanalysis? *Studies in Language* 22, 315–351.
- Haspelmath, Martin (2002). On directionality in language change with particular reference to grammaticalization. Unpublished manuscript, Max-Planck-Institut für evolutionäre Anthropologie, Leipzig.
- Heine, Bernd. 2003. On degrammaticalization. In *Historical linguistics 2001: Selected papers from the 15th International Conference on Historical Linguistics, Melbourne, 13-17 August 2001*, Barry J. Blake and Kate Burridge (eds.), 163–179. Amsterdam: Benjamins.
- Heine, Bernd; Claudi, Ulrike; and Hünemeyer, Friederike (1991). *Grammaticalization: A conceptual framework*. Chicago: Chicago University Press.
- Heine, Bernd; and Kuteva, Tania (2002). *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd; and Kuteva, Tania. (2003). On contact-induced grammaticalization. *Studies in Language* 27, 529–572.
- Hopper, Paul J.; and Traugott, Elizabeth Closs (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- Ivanova-Mirčeva, Dora; and Xaralampiev, Ivan (1999). *Istorija na bǎlgarskija ezik*. Veliko Tǎrnovo: Faber.
- Janda, Richard D. (2001). Beyond “pathways” and “unidirectionality”: On the discontinuity of language transmission and the counterability of grammaticalization. *Language Sciences* 23, 265–340.
- King, Gareth (1993). *Modern Welsh: A Comprehensive Grammar*. London: Routledge.
- Kiparsky, Paul (2004). The unidirectionality of grammaticalization: A reformulation. Paper presented at Diachronic Generative Syntax (DIGS) VIII, Yale University.
- Kurz, Josef (1958–66). *Slovník jazyka staroslověnského*. Prague: Nakladatelství Československé Akademie Věd.
- Lass, Roger (1990). How to do things with junk: Exaptation in language evolution. *Journal of Linguistics* 26, 79–102.
- Lewis, Henry; and Pedersen, Holger (1937). *A Concise Comparative Celtic Grammar*. Göttingen: Vandenhoeck & Ruprecht.
- Lloyd-Jones, J. (1931–63). *Geirfa Barddoniaeth Gynnar Gymraeg*. Cardiff: Gwasg Prifysgol Cymru.
- Miklosich, Franz (1886). *Etymologisches Wörterbuch der slavischen Sprachen*. Vienna: Braumüller.
- Newmeyer, Frederick J. (2001). Deconstructing grammaticalization. *Language Sciences* 23, 187–220.
- Norde, Muriel (2001). Deflexion as a counterdirectional factor in grammatical change. *Language Sciences* 23, 231–264.
- Ohala, John J. (1993). The phonetics of sound change. In *Historical Linguistics: Problems and Perspectives*, Charles Jones (ed.), 237–278. Longman: London.
- Perridon, Harry. (2003). Grammaticalization as exceptional semantic change. Paper presented at the Sixteenth International Conference on Historical Linguistics, Copenhagen.
- Thomas, Alan R. (1977). *The Linguistic Geography of Wales: A Contribution to Welsh Dialectology*. Cardiff: University of Wales Press.
- Thomas, R. J. (1950–2002). *Geiriadur Prifysgol Cymru: A Dictionary of the Welsh Language*. Cardiff: Gwasg Prifysgol Cymru.

- Thorne, David A. (1993). *A Comprehensive Welsh Grammar*. Oxford: Blackwell.
- Timberlake, Alan (1977). Reanalysis and actualization in syntactic change. In *Mechanisms of Syntactic Change*, Charles N. Li (ed.), 141–177. Austin: University of Texas Press.
- Traugott, Elizabeth Closs (2001). Legitimate counterexamples to unidirectionality. Paper presented at Freiburg University.
- Traugott, Elizabeth Closs; and König, Ekkehard (1991). The semantics-pragmatics of grammaticalization revisited. In *Approaches to Grammaticalization*, Elizabeth Closs Traugott and Bernd Heine (eds.), 189–218. Amsterdam: Benjamins.
- Trubačev, O. N. (1997). *Ėtimologičeskij slovar' slavjanskix jazykov*. Moscow: Nauka.
- Vasmer, Max (1953). *Russisches etymologisches Wörterbuch*. Heidelberg: Winter.
- Williams, Stephen J. (1980). *A Welsh Grammar*. Cardiff: University of Wales Press.