§1  Introduction

In this response to Anna Wärnsby’s article, I address her two sets of concerns: (a) general issues surrounding constructional approaches to language, and (b) particular issues surrounding English modal verb constructions. The former is the subject of considerable debate (see for example the debate between Aarts (2004, 2007) and Croft (2007) on gradience, and more recently, that between Adger (2013) and Goldberg (2013a) on the explanatory power of constructions in the usage-based model), and I will not address all of the possible issues. Instead, I structure my commentary by making reference to the four main issues that Wärnsby (hereafter W) sees in constructional approaches to modals in English, namely:

a. the number and complexity of constructions required for an accurate grammatical description of the modals;
b. the role of context, and the lack of formalization of context in Construction Grammar (hereafter CxG);
c. the scope of schematic constructions;
d. indeterminacy.

The balance between explicitness and economy in representation is a problem for all grammatical frameworks. The English modals are a particularly good case study to illustrate this point, especially when it comes to the issue of language change. The facts surrounding the development of (English) modals suggest a very messy picture, with a significant lack of uniformity in the development of
the individual modal verbs over time (see Warner 1993, and the commentary in Fischer 2007). Traditional generative approaches risked oversimplification by failing to treat each micro-construction with sufficient care, while more modern generative approaches typically adopt an approach whose levels of detail and fine-grainedness are similar to those of constructional approaches (see e.g. Roberts 2010). The detail is necessary because the modals are a particularly messy category (synchronically, as a result of their diachronic development).

§2  Issue 1: The number and complexity of constructions required

W's first concern derives from her analysis of a number of tokens extracted from the English-Swedish Parallel Corpus and presented in detail in Wärnsby (2002). Based on these data, W suggests that at least 13 different constructions are required to accurately capture the range of complementation patterns (e.g. where the lexical verb is in the form of the past participle, as in *it may be parked somewhere in Ipswich* [Wärnsby 2002, ex. 11], or where the verbal complement is a NP, as in *it may be the thin end of the wedge* [Wärnsby 2002, ex. 10]). W's reasons for positing these constructions appear to be connected to a principle outlined by Goldberg (1995), which W refers to as a Principle of No [Syntactic] Synonymy1.

---

1 It is not absolutely clear to me what is meant by a Principle of No [Syntactic] Synonymy, i.e. why the word *syntactic* is included and then put in brackets, and whether this is intended to mean something different from Goldberg (1995). Goldberg outlines a Principle of No Synonymy; this is concerned with the claim that if two constructions differ in syntactic form, they will be different in some way in terms of their meaning – if there is semantic synonymy, the constructions will be pragmatically distinct; if there is pragmatic synonymy, the constructions
Concerned that this involves speakers having to learn an unnecessarily high number of distinct constructional patterns, W then points out that a potential solution may be found by invoking multiple inheritance, i.e. a situation where a given token inherits from a number of different parent constructions. However, she feels that even this would result in a 'bewildering web of constructions', and suggests that, unlike the case of the Caused Motion Construction, it is much harder to suggest a motivation for positing a prototypical modal construction for English which would contribute most frequently or most centrally in such an inheritance network.

My own response to these concerns (possibly not shared by other practitioners of CxG) is as follows. First, it is not clear to me that the examples W provides are genuinely of different constructional types. Rather, they are tokens, instances of the combinatory output of multiple inheritance (to which I will return below). Thus, [NP MAYe BE VenSTATE] is not a construction of English; the corpus example *it may be parked somewhere in Ipswich* is a token that is the product of inheritance from a number of constructions, some contentful (such as *park*) and some procedural (such as the Copular Construction), but is not sanctioned by a micro-construction such as the one W suggests, because there is no need to posit a separate form-meaning pairing at that level. We can certainly find examples where *may* appears in sequences with a pronominal subject, and a passive auxiliary, and a locative adjunct, but that does not require a separate micro-construction. What speakers of English know is that inheritance from both

---

will be semantically distinct. I will assume that Wärnsby is referring to the principle as laid out in Goldberg (1995), without modification.
a modal construction and a copular construction can result in well-formed expression.

Second, the issue of multiple inheritance does indeed require knowledge of a very complex web of constructional types. In addition, it can also mean the storage of many redundancies, and in some cognitive theories, the (short-term) storage of tokens (Hudson 2007, 2010), or of exemplars (Bybee 2013). This is not a problem restricted to the modals, but rather a more general issue concerning how much of our knowledge of language involves storage, and how this links to inheritance. Consider in this regard the following example cited in Denison and Cort (2010; emphasis added):

(1)  lol well you better start staying in betternt you !!!!!!! lol

For many readers of this article, this may be the first time they have encountered a token involving the multiple inheritance of a modal construction BETTER and the tag question construction. Indeed, in such a case this particular constructional inheritance may be so unusual as to strike some readers as ungrammatical. Yet the ability to process this construct requires such a reader to go beyond what they know (i.e. their current grammar) and allow inheritance to apply to this novel exemplar. It is not conventional for such a reader to have modal BETTER and the tag question construction to combine in an inheritance network, so this particular use requires the reader to extend their network to allow processing of this particular combination. Given the relative infrequency of the use of BETTER as a modal anyway (in contrast to the sequence HAD BETTER), it may well be that this inheritance pattern does not become
conventionalized for the reader. The crucial point is that what is stored varies between language users: for those for whom the sequence is unremarkable and grammatical, all that is stored are the types (i.e. the modal construction BETTER and the tag question construction); for those for whom the sequence is ungrammatical, there may be temporary storage of the token, but this will fade if not reinforced by frequent use. In neither case, however, is it necessary to suggest that better\textit{nt you} is stored as a type. And in the former case, it is not necessary to suggest that a construction whose form is $[[\text{NEG MODAL}]$ [PRONOUN]] is stored: tokens such as \textit{won't you} and \textit{can't I?} can be processed and produced using more general constructions and multiple inheritance.

This leads to the third issue, which is to do with prototype effects. I agree with W that none of her examples (1)-(6) are more epistemic than any of the others, but as I have suggested above, I would not treat (1)-(6) as separate modal constructions. But her corpus analysis (as presented in Wärnsby 2002) does suggest that, of the various possible combinations (e.g. with different subject types, or with multiple auxiliaries), the most frequent for epistemic \textit{may} is the sequence where, among other things, there is no other auxiliary present, and the main verb is a stative. Such a combination accounts for just over 40\% of the tokens. A fuller statistical analysis would need to consider, among other things, the overall frequency of statives in the corpus, and it would be useful to carry out a collocational analysis to see the extent to which statives are indeed the preferred context for may in comparison with other epistemic modals (cf. Hilpert 2008 on the changing collocational profile of English \textit{BE going to} over time). In order to establish whether or not CxG can add something that other frameworks cannot, I think it would be necessary to re-examine the corpus data.
using the tools associated with corpus-based approaches to constructional preferences (see e.g. the discussion in Stefanowitsch 2013).

§3 Issue 2: The role of context

The role of context in CxG has been described in some detail, both synchronically and diachronically. W’s decision tree for *may* (her figure 1) is compatible with a constructional approach to context, and indeed co-text (see further Bergs and Diewald (2009) on this distinction in constructional terms), as is her insight into the subtleties of meaning variation associated with English modal verbs.

As Boogaart (2009: 231) has observed, in exploring the role of context and modal constructions, analysis has focused not on the properties of the individual model item (such as Dutch *kunnen* or English *may*), but on the properties of the construction for which the modal verb is the profile determinant (Langacker 1987), and the particular contexts in which such constructions are used. Boogaart (2009) also foregrounds the importance of semantic maps in constructional analyses of modals (see e.g. van der Auwera and Plungian 1998). How might we bring together the kind of decision tree that W proposes, with the constructional analysis of context that Boogart (2009) advocates?

One possibility is to see epistemic, deontic and dynamic modality not as properties of individual modal verbs, but as part of a set of semantic domains which come to be associated with various modal constructions (see also Boogaart 2009, Croft 2010). Sometimes these modal constructions have
modality as a feature of the profile determinant (i.e. as a clausal head), as is the case with English modal verbs, but this need not be the case (e.g. when modality is expressed by an adjunct such as the adverb perhaps). W's decision tree could then be seen as a speaker or hearer's guide towards a particular location on the map. Furthermore, if this decision tree was to be considered ‘cumulatively’, one could imagine a situation for a given instance of use, where the subject was ‘not in control of the action’, the aspect perfective, the verb stative and so on, in which we have a series of properties which mark out the conceptual core of an epistemic modal construct. In other words, the ‘decision tree’ is the result of speaker generalisations about typical and less typical epistemic constructions. This too fits in with perspectives on semantic maps, where particular semantic categories such as modality types can be seen as “generalizations over a large class of specific occurrences of [...] meanings in situation types produced in utterances” (Croft 2010: 56). W's decision tree features particular co-textual properties (e.g. person of the subject, presence of other (aspectual) auxiliaries) which can also be incorporated into an account in which we see particular constructs as the product of multiple inheritance from a number of constructional types.

One final remark on the topic of context concerns W's claim that “the role of context as well as what contextual features to be included in a modal construction is not yet formalised.” It is true that some variants of CxG have a less developed formalism, but this is not true of all variants. Some models of CxG, such as Sign-Based Construction Grammar (e.g. Sag 2012), are highly formalized. Furthermore, some research on constructional change, such as Fried (2009), has been explicitly concerned with “what types of context may enter conventional
linguistic patterning” (Fried 2009: 63). Fried (2009) suggests that three particular types of context need to be considered in relation to the shape of constructions and their evolution:

a. the pragmatic conditions in which constructions are used by speakers and hearers;

b. information structuring and the roles played by discourse participants;

c. the cotext (including collocational preferences).

Quite how such contextual properties are to be formalized is an ongoing issue in constructional research (semantic frames again feature in this formalization); but there is clear evidence in the work of Fried and others that the role of context and what contextual features should be included in a construction has been formalized in some varieties of CxG.

§4  Issue 3: The scope of schematic constructions

The third concern that W raises is one which goes beyond the question of modal constructions in particular. Essentially, it focuses on the extent of generality in the constructional network – just how schematic can constructions get? W then links this issue to the problem of recovering implicit purpose clauses which might be associated with particular uses of deontic modal constructions.

How schematic constructions can get is an empirical question. It is part of a more general question that is central in constructional approaches to language:
what are conventional symbolic units (i.e. constructions) used by a community or network of speakers? As Kay observes, a linguist cannot know whether, for an individual speaker, particular generalizations represent “psychologically real entities. The relevant obligation of the grammarian, under this view, is to abstract from the data of the language all the generalizations [...] that a speaker-hearer might extract” (Kay 2004: 697, emphasis original). However, there is, in usage-based models of CxG, the further issue of frequency: “schematic constructions are posited only when justified by sufficiently high type frequency and degree of morphosyntactic and semantic similarity; more specific constructions may also be posited if they are of sufficiently high token frequency” (Croft 2013: 223).

There are, then, two grounds for proposing a given form-meaning pairing constitutes a construction in a community. One involves idiosyncracy, in some aspect of form, meaning or the association between form and meaning (Goldberg 1995); the other involves frequency of use, where fully compositional strings may be considered constructions if they are used with sufficient frequency (Goldberg 2006). The question then is whether, in W’s terms, there is a “motivated reason for positing meaningful modal constructions”. As atomic constructions, there is clearly idiosyncracy involved, and many individual modals are highly frequent in use, of course. But a more critical issue is whether we should be focusing on the combination of modals (as atomic constructions) with other more abstract schemas. Citing Eide (2002), W suggests that (2) – her example (7) – is epistemic:

(2) The patient must have been mistreated
What determines this interpretation? W suggests that the addition of a purpose clause (e.g. *to get compensation*) would render the interpretation as deontic, but without it, the epistemic reading is preferred. Consider, however, an example like (3):

(3) Tickets must have been validated

It is not clear to me that we require some implicit purpose clause here to force a deontic reading: there clearly *is* a purposive that is implicit here (e.g. something like *in order to qualify for free parking*), but my point is that an epistemic reading is dispreferred anyway, in contrast to (3), even without invoking that purposive. This does not, however, solve the problem of W’s example (9), given here as (4):

(4) I must be brave, I must maintain my own high standards.

W observes that this example is one where a speaker is a prisoner telling herself that she needs to be brave in order to remain sane. W suggests that it is difficult in CxG to account for the deontic use here, given that speakers rarely give themselves orders; we would need to stipulate this somewhere in the construction.

My suggested solution to this problem is to consider a related context: the context when speakers engage in self-talk. As Holmberg (2010) has observed, there are cases where speakers can vary between using *I* and *you* in self-talk:
I'm an idiot!

You're an idiot!

but there are other cases where variation is not possible. For instance, if the matrix verb is a verb of cognition, you appears to be dispreferred, even if there is variation in the subject of the non-matrix verb. The following examples come from Holmberg (2010: 187), his examples (6) and (10), with his own grammaticality judgements:

I think I've had it!/I think you've had it!
*You think I've had it!/*You think you've had it!

The key issue is that in self-talk, “you can assume the role of speaker or addressee”, but in some instances (associated with assertions or questions regarding state of mind, for example), only one of the options is available (Holmberg 2010: 187-188).

How does this relate to the issue of the modal construction under discussion? Here is the example W provides, but in its wider context. It is taken from a novel by Sue Townsend, about an imagined world where the United Kingdom becomes a republic, and the British Queen is forced to live as an ordinary citizen. In this part of the novel, the Queen is prevented from leaving the street on which she now lives:

The Queen walked around the Close four times. Nobody was about apart from the odd mongrel dog. She thought, I am living in a
ghetto. I must consider myself a prisoner of war. I must be brave, I must maintain my own high standards. (Townsend 1992)

This is not an instance of self-talk, but one of direct thought. But as in the case of self-talk, the context suggests a situation where the self can be conceptualized as both ‘thinker’ and ‘thought-of’, which serves as an equivalent to the speaker or addressee in self-talk (cf. she thought, you are living in a ghetto, etc.), and either first or second person pronouns may be used. In this context, there is an obligation laid down by the ‘thinker’ on the ‘thought-of’, and while typically one might not expect the use of first person pronouns in the laying down of an obligation, speakers of English permit variation in contexts of the ‘split-self’, i.e. in self-talk or self-directed thought.²

However, even if the reader is not convinced by this possible explanation of the use of the first-person pronoun in this case, there is a more general point to be made. Some of the analytic concerns that W raises strike me as problems that are not particular to constructional approaches. The problem of the relationship between world knowledge and linguistic form is an issue for any model of language. What CxG does (and this is true of all usage-based network models) is to try to establish which routes to follow between a given usage-event (and all its contextual properties) and an abstract type which provides the best fit (Hudson 2007). From a usage-based perspective, understanding (and producing) utterances relies on our capacity to:

² I am essentially treating the linguistic representation of direct thought in literature on the one hand, and self-talk on the other, as instances of intrapersonal communication.
a. conceptualise individual items, both procedural (e.g. *must*) and contentful (e.g. *brave*);

b. understand how those items combine and are interpreted once combined (e.g. *I must be brave* is most likely intended/interpreted as deontic while *I must be stupid* is most likely epistemic, but both can be overridden in particular discourse contexts);

c. establish the relevant inheritance properties.

Repeated exposure to similar patterns of use gives rise to probabilistic associations of meaning with form. Thus one explanation for the difference *W* alludes to is our experience of characteristics we ascribe to ourselves, whether this be exhortations to a particular kind of behaviour (leading to a deontic interpretation), or a particular kind of self-assessment (linked to an epistemic interpretation).

§5 Issue 4: Indeterminacy

The matter of modals and indeterminacy links up to the issues discussed in section 3, on context, but relates more crucially to the notion of inheritance relations that are central to usage-based constructional approaches, as discussion in section 2. As a result, I will not comment extensively on this topic for risk of repeating myself, but the main issues seem to be these. Recent work on language change within a cognitive linguistics framework has focused on multiple inheritance (see De Smet *et al* 2013), and both constructional and non-
constructional usage-based linguistic frameworks (see e.g. Hudson 2007 for an example of the latter) make use of the concept of multiple inheritance in the grammatical architecture. As discussed above, in such models, a particular token utterance may be sanctioned by more than one type, which may either be associated with differences in parsing, or polysemy effects. Construct indeterminacy is a natural consequence of such inheritance networks. For further discussion of the relationship between monosemy, polysemy and semantic maps in connection with modal constructions, the reader is referred to Boogaart (2009).

§6 General criticisms of CxG

W ends her article with a number of general concerns about CxG, which I will briefly respond to here. I hope to have shown that CxG is more than just a possible tool for the analysis of English modal constructions, and that it can provide a plausible and coherent account based on a series of general principles. My view is that generalisations are possible, but they are perhaps not the kind of generalisations that researchers in linguistic modality are accustomed to. W suggests that more psycholinguistic research is needed in order to delimit the ‘bewildering web of constructions’, but I think it is fair to say that there have been many psycholinguistic studies which have investigated the status of constructions (see Bencini 2013 for a summary). I agree with W that constructions should only be posited when there is sufficient evidence so to do, but it may simply be the case that the constructional network is vast and
complex. W seems to suggest that most work in CxG is concerned with argument structure (I assume this is what is meant by “constructions localized within a clausal structure”); while this might have been true of some early work, there is now much work within CxG on information structuring (e.g. Lambrecht 2001, Pattern 2012), on filler-gap constructions (e.g. Sag 2010), and on morphological constructions (e.g. Booij 2010), to name just three other areas of inquiry.

§7 Conclusions

In her paper, W has clearly and cogently addressed a number of concerns she has with the constructional enterprise more generally, and with a constructional analysis of the English modal system in particular. In this response, I have attempted to show how, based on my personal view of constructions and CxG, it is possible to respond to those concerns. CxG is not monolithic, and while there are many areas of agreement (see Goldberg 2013b), different varieties involve different degrees of formal representation. There are some very formal variants of CxG, such as Sign-Based Construction Grammar (Sag 2012), while the more usage-based models tend to involve less formal representation.

In all cases, there is the important question of the empirical verification of particular claims. One concern of W’s is establishing verifiable measures to ascertain how schematic a particular part of the constructional network is. There is a balance to be struck between economy and explicitness, both in terms of processing and production. But the point is that this is an empirical question, working from the principle that speakers generalize as little as possible but as
much as is necessary. There appears to be corpus-based evidence that some
types of change involve more schematization (or strengthening of the cognitive
representation of a schema) than others (Hilpert 2015), and this seems to be
further supported by computational work on emergent systematicity in
simulated languages (van Trijp and Steels 2012).

References

Language*, 31(2), 431-443.
Adger, D. (2013). Constructions and grammatical explanation: comments on
Bencini, G. (2013). Psycholinguistics. In T. Hoffmann and G. Trousdale (Eds.), *The
Oxford Handbook of Construction Grammar* (pp. 379-396). New York: OUP.
Bergs, A., & G. Diewald (Eds.). (2009). *Contexts and Constructions*. Amsterdam:
John Benjamins.
case of modal verbs. In A. Bergs & G. Diewald (Eds.). (2009). *Contexts and
Constructions* (pp. 213-241). Amsterdam: John Benjamins.
Bybee, J. (2013). Usage-based theory and exemplar representations of
constructions. In T. Hoffmann and G. Trousdale (Eds.), *The Oxford Handbook of
Construction Grammar* (pp.49-69). New York: OUP.


