How a Construction Grammar account solves the auxiliary controversy

Remi van Trijp
Sony Computer Science Laboratories Paris

The English auxiliaries have been a matter of dispute for decades with two opposing views: one analysis treats them as main verbs that take a VP complement; the other considers them as feature carriers. Proponents of both approaches have convincingly pointed out each other's weaknesses without however settling the debate and without accounting for the fact that the English VP is still evolving today. The goal of this paper is to show that Construction Grammar offers a way out of the current status quo. This claim is substantiated by a computational formalization of the English verb phrase in Fluid Construction Grammar that includes a bi-directional processing model for formulation and comprehension available for online testing.

Keywords: verb phrase, Construction Grammar, auxiliaries, tense, aspect, modality, grammar formalization

1. The auxiliary controversy

English auxiliaries are among the most hotly disputed topics in the history of linguistics. In the mid-eighties, when the debate had already been raging for almost two decades, Schmerling (1983) carefully approached the topic as follows:

> Writing a paper on English auxiliaries these days is a little like proposing a solution to the problems of the American car industry: one has the feeling that all conceivable approaches have probably been proposed, that the best we can hope for is a more cleverly refined version of one of these approaches, and that the reception one's proposal will get is determined essentially by the politics of one's audience. It is an unfortunate fact of life that it is impossible to propose a solution either for Detroit or for auxiliaries that does not step on somebody's toes.
>
> (Schmerling 1983: 1)

Constructions and Frames 9:2 (2017), **250–275**. DOI 10.1075/cf.00004.van
ISSN 1876–1933 / E-ISSN 1876–1941 © John Benjamins Publishing Company

We are now more than thirty years later and the debate still has not been settled. There are roughly two opposing views. Consider Example (1a) and the two possible analyses in (1b) and (1c), quoted from Falk (2008: 861).

(1) a. *The children will take syntax.*
b. *will* is the head of the sentence and *[take syntax]* is a complement of *will*;
c. *take* is the head of the sentence and *will* is a "modifier" or morphological marker expressing or realizing future tense.

The analysis in (1b) treats auxiliaries as main verbs that take a VP as their complement. The analysis in (1c), on the other hand, is closer to the traditional notion of auxiliaries as *helper verbs* that express functions such as tense, aspect, modality, and voice.

1.1 The main-verb analysis and its problems

The main-verb analysis of auxiliaries was pioneered within the context of transformational linguistics by Ross (1969), who argued that auxiliaries and lexical verbs are two *surface* manifestations of the same *deep* category *verb*. That claim has since lost its support because transformational grammars are no longer in vogue and because there is abundant empirical evidence to support the distinction between auxiliaries and lexical verbs (see for example Palmer 1979).

Despite losing its original raison d'être, however, the analysis survives in contemporary theories such as Head-Driven Phrase Structure Grammar (HPSG; Pollard & Sag 1994), where auxiliaries are raising verbs that share their subject with their VP complement. One motivation for maintaining the main-verb analysis is that it allows for expressing straightforwardly the constraints that auxiliaries impose on other verbal forms through head-complement rules. However, this convenience comes at a cost: if auxiliaries are the head of a phrase, then sentences such as *The children will have been taking syntax* require a multi-clausal analysis of four embedded clauses, as shown in Figure 1.

There are two problems with a multi-clausal analysis. First, the notion of *raising* (also a term from transformational linguistics) has always been somewhat problematic; this is because while the so-called raising verb acts as the syntactic head, it is the embedded lexical verb that governs the sentence's argument structure: Figure 1 involves a transitive clause not because of the constraints imposed by the auxiliaries that occur in it, but because *take* is a transitive verb. Moreover, descriptive linguists have provided ample evidence for treating VPs that contain only a single lexical verb as a simple clause (Palmer 1979).

Secondly, a multi-clausal analysis has also been criticized for its processing complexity in comprehension because it creates spurious ambiguities when

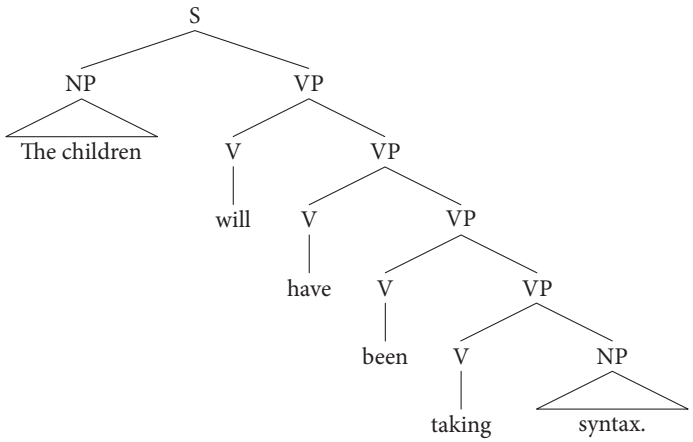


Figure 1. Even for simple sentences such as *The children will have been taking syntax*, a multi-clausal analysis is needed if auxiliaries are treated as the head of a phrase.

retrieving a sentence's dependency relations (Butt et al. 1996). Butt et al. (1996) also show that the main-verb analysis works well for a strongly configurational language such as English, but that it rapidly becomes costly even for family-related languages, such as German, in which verb phrases may be discontinuous. Obviously, data from other languages cannot be directly used as evidence against a main-verb treatment of English auxiliaries, but it is nevertheless important to investigate whether the strategies found in other languages can also account for the English auxiliary system.

1.2 The feature-carrier analysis and its problems

The feature-carrier analysis proposes a rather flat, single-clause analysis of verb phrases such as *will have been taking*. In principle, this approach solves both problems of the main-verb analysis: it matches the observations from empirical linguists and it avoids costly multi-clausal representations. Moreover, the different behaviors of auxiliaries and lexical verbs are functionally motivated, whereas the main-verb analysis offers no explanation for this distinction. Unfortunately, the main problem of the feature-carrier analysis is that so far, no one has been able to propose a satisfactory formalization. The most explicit proposals have been made in Lexical Functional Grammar (LFG; Butt et al. 1996; Frank & Zaenen 2004). Instead of taking phrase structure as the main representation device from which all information can be "read off", the key idea in LFG is to separate grammatical functions from how they are encoded by the grammar. LFG, therefore, represents linguistic information as correspondences between two separate structures: the *f(unctional)-structure* and *c(constituent)-structure*. The grammatical functions

expressed by the English auxiliaries, such as tense and aspect, are hypothesized to belong to the f-structure.

However, the LFG approach creates a different problem, namely how to decide which information should be represented in which structure. This problem is not immediately apparent for languages such as English in which grammatical functions are often expressed in a periphrastic way (in other words, where the correspondences between f- and c-structures are relatively straightforward), but it comes to the fore when applying the analysis to languages that use morphological inflection instead (such as French). Since LFG assumes a distinction between syntax and morphology, Butt et al. (1996) have proposed an additional structure (called *m-structure*) for representing morphological forms.

The m-structure is however controversial within the LFG community; as it turns out, it is impossible to cleanly separate the relevant linguistic information into three different structures (Dyvik 1999; Frank & Zaenen 2004; Falk 2008). For instance, Frank & Zaenen (2004) show that the m-structure analysis needs to repeat most of the information from the f-structure in order to account for agreement of number, person, gender, and case.

1.3 Coping with language change

An additional issue for both analyses is that they assume that the English verb phrase is fully systematic and well-structured. Indeed, at first sight, the VP has a clear canonical structure that can be roughly schematized as shown in (2):

(2) (Modal) (Perfect) (Progressive) (Passive) Lexical-Verb

The English VP has five slots of which only the last one is obligatory: the lexical verb, which governs a sentence's argument structure. All the other slots are syntactically optional and can be filled independently of each other. Whenever that happens, its fillers appear in a fixed relative order and each filled slot has an impact on the next slot that is filled. Example (3) only presents a few possible realizations of the VP:

(3) a. *The children take syntax.*
b. *The children are taking syntax.*
c. *The children have taken syntax.*
d. *The children have been taking syntax.*
e. *The children will be taking syntax.*

The empirical evidence also justifies the hypothesis that there is a categorial difference between auxiliaries and lexical verbs. For example, auxiliaries have an anomalous conjugation scheme, and they are the only verb forms that can express the so-called 'NICE'-properties (i.e. Negation, Inversion, Code, and Emphasis;

Huddleston 1976). In the absence of an auxiliary, English calls on DO-support for expressing those properties, as illustrated in Example (4):

(4) *The children didn't take syntax.*

But looks can be deceiving and functional and historical linguists have been pointing out for decades that the English VP is not the perfect system that it appears to be: it has actually been in constant flux for centuries (Denison 2000) and it continues to evolve today (Bolinger 1983). For instance, diagnostic tests show that the distinction between auxiliaries and lexical verbs is a matter of gradience (Palmer 1979: 4–5) and English is in fact still in the process of developing new auxiliaries (Bolinger 1983; Denison 2000). Example (5) illustrates the use of *passive-GET* as an alternative to *passive-BE*, while Example (6) illustrates how *progressive-BE* can be replaced by several verbs to modify the main verb's aspectual contour.

(5) a. *I was pulled over by the police for speeding.*
b. *I got pulled over by the police for speeding.*

(6) a. *I was reading a book.*
b. *I started reading a book.*
c. *I continued reading a book.*
d. *I stopped reading a book.*

However, all of these alternatives remain different from prototypical auxiliaries because they cannot take any of the NICE-properties, as shown in Examples (7) and (8):

(7) *Were you pulled over by the police for speeding?*
**Got you pulled over by the police for speeding?*

(8) a. *Were you reading a book?*
b. **Started you reading a book?*
c. **Continued you reading a book?*
d. **Stopped you reading a book?*

So rather than trying to pour these intricate data into a rigid mold, any adequate model of the English VP must be able to capture its systematic behavior while at the same time staying fluid enough for handling the ongoing changes in the system.

1.4 Construction Grammar makes a synthesis possible

Manning (1995) makes the interesting observation that HPSG and LFG have grown closer to each other in the past decades in order to allow for more flexibility in their analyses. On the one hand, HPSG has moved away from a strict

phrase-structural approach. For instance, *linearization-based* HPSG introduces a new virtual layer of *ordering domains* that allows a sentence's word order to be separated from its constituent structure (Reape 1994). On the other hand, LFG researchers have argued for tighter connections between LFG's different structures. For instance, Frank & Zaenen (2004) and Falk (2008) have suggested that the categories of c-structure could be made more fine-grained in order to incorporate morphological inflection instead of pushing that information to a separate m-structure.

However, there is an alternative linguistic approach that is already able to combine the best of both worlds: Construction Grammar. In a construction grammar, *constructions* are considered to be the primitives of linguistic analysis. On a technical level, this means that there is only one structure for linguistic representation – in contrast to LFG's separate structures – so the problem of deciding which information goes in which structure never occurs. Moreover, constructions are able to cut across all levels of linguistic information (Fillmore 1988) so there is no need to shoehorn an analysis into phrase-structural relations. Instead, a construction may simultaneously describe an utterance's functional, structural and/or morphological features (van Trijp 2016).

2. Modeling the English verb phrase

This section presents how the English verb phrase can be operationalized in Fluid Construction Grammar (FCG; Steels 2011; see also Steels in this issue for a comprehensive introduction), an open source grammar formalism for bidirectional language processing and learning. More specifically, I will focus on how the English VP expresses distinctions in the domains of tense, aspect, and modality (TAM). This paper's linguistic account is firmly grounded in cognitive-functional analyses and some familiarity with such approaches is assumed (see for example Radden & Dirven 2007: Chapter 3 for an introduction). The implementation is part of recent efforts to operationalize key ideas from Construction Grammar in a robust broad-coverage processing model for English (van Trijp 2017), which can be tested online at http://www.fcg-net.org/fcg-interactive/ and which shows that the proposal in this paper is scalable. While this paper only investigates the English VP, similar approaches are explored by other contributions in this special issue for Dutch (Van Eecke), Spanish (Beuls), and Russian (Beuls et al.).

2.1 Overview of the analysis

The model presented in this paper assumes a flat structure (i.e. single-clause analysis) for the English verb phrase. In other words, the verb that controls the utterance's argument structure is also considered to be the verb phrase's syntactic head. Auxiliaries are thus not main verbs but rather *helper verbs* that indicate that the lexical semantics of the main verb should be modulated, as shown in Figure 2.

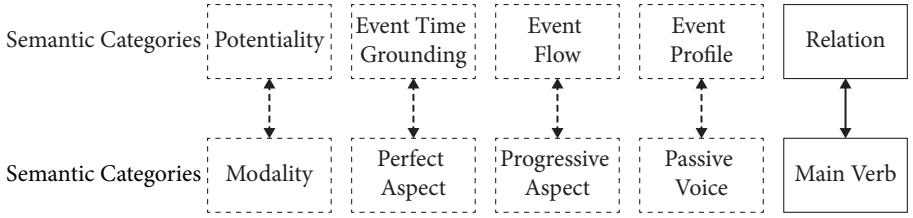


Figure 2. The model presented in this paper posits a flat structure in which the lexical verb acts as the head of the verb phrase. Grammatical constructions may modulate the main verb's lexical semantics in terms of its potentiality, temporal grounding, situation type and event profile.

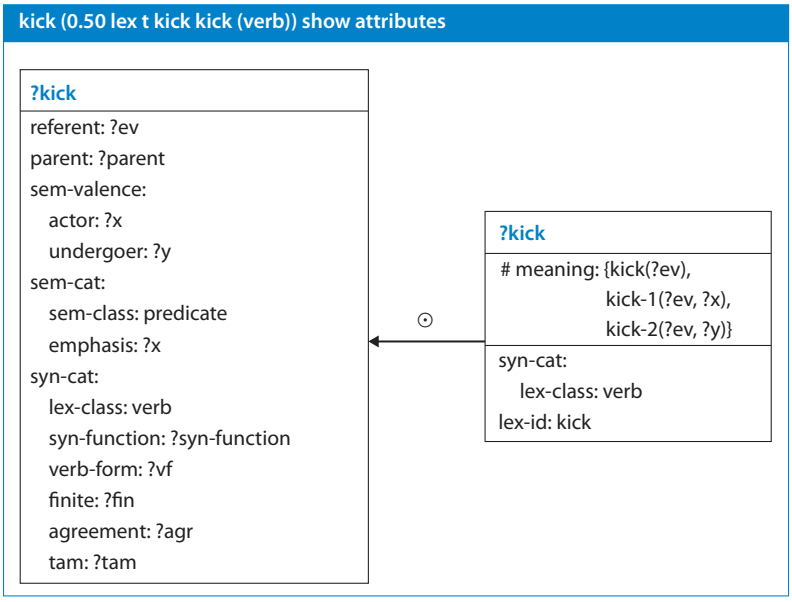
Note that Figure 2 only shows the schematic mapping between semantic and syntactic *categories* instead of directly referring to specific auxiliaries. The reason is that the modulation of the main verb's lexical semantics is achieved by *grammatical constructions* rather than directly by the auxiliaries themselves, which stands in stark contrast to traditional lexicalist approaches. For instance, the Progressive construction has a double-verb generalization that is not dependent on auxiliary-*be* for modulating the aspectual contour of a lexical verb. Instead, it requires a double-verb pattern in which the second verb is an *-ing* form. Typically, the first verb is auxiliary-*be*, but depending on the semantic distinctions that need to be expressed, the construction will also allow other verbs such as *continue*, *stop* and *start* to play the role of 'helper verb', as shown before in Example (6), even though these verbs are not strictly speaking auxiliaries. Throughout this paper, I will use the following metaphor from Fillmore (1988) for clarifying this approach:

> It may be useful to think of the positions within a construction as **offices** […]. The obligatory features associated with positions in the description of the construction can be thought of as the **qualifications** for the office, and the role indicator identifies the **function** of the office. […] A candidate which does not satisfy the qualifications of the office cannot fill the office. When a particular **incumbent** occupies the office, that incumbent has properties of its own, not only the properties which allowed it to occupy the office, but also properties which cause it to make its own demands. (Fillmore 1988: 41)

2.2 Verbal lexical constructions

Constructions are often informally defined as mappings between meaning and form. In Fluid Construction Grammar, such mappings are implemented as *units* that contain *feature-value pairs*. Each unit can be thought of as an office in the aforementioned metaphor, and the feature-value pairs specify the qualifications, functions and constraints of the office. The following example shows the formal notation for the lexical *kick* Construction (see Steels in this issue for an introduction to the notation), taken from the FCG web interface. The right-hand side of the construction specifies the *conditions* that need to be satisfied before the construction is applied. The left-hand side contains all the information that the construction contributes itself.

(9) **Kick Construction:**



As can be seen, the construction concerns a single office: a lexical unit for the verb *kick*. The right-hand side specifies the qualifications of the office, divided into a *formulation-lock* (above the single line) and a *comprehension-lock* (below the line). This distinction captures the fact that the qualifications for the office change depending on whether the construction has been evoked during formulation (speaking) or comprehension (listening).

The formulation-lock only contains the feature MEANING. Its feature name is marked by the symbol #, which indicates that the construction should look for this feature in the utterance's *input* (in the case of formulation, the input consists

of the meanings that the speaker conceptualizes). None of the features in the comprehension-lock are marked by this symbol, which means that the construction expects all those qualifications to be supplied by other constructions. In the following paragraphs, I will discuss some of these features in more detail because they play an important role in the English Verb Phrase.

Meaning
The model assumes a frame semantics approach for event structure (Baker et al. 1998), so the lexical meaning of a verb includes an explicit representation of an event frame and its frame elements. For example, the meaning of the verb *to kick* includes the kick-frame and its frame elements kick-1 (i.e. the one who is doing the kicking), kick-2 (i.e. the entity that is being kicked), and so on. In the implementation, a verb's semantic frame is formalized as *an unordered, flat list of logic predicates and their arguments* as shown in Example (10).

(10) {kick(?ev), kick-1(?ev, ?x), kick-2(?ev, ?y)}

Every symbol that starts with a question mark is a variable. Here, the variable ?ev needs to be bound to the specific kick-event that the speaker wishes to talk about, ?x needs to be bound to the participant that performs the kicking (e.g. *a boy*), and ?y needs to be bound to the participant that undergoes the kicking (e.g. *a ball*).

SEM-VALENCE
While the meaning of a verb specifies an *event-specific* semantic frame, the feature SEM-VALENCE recategorizes that frame in terms of more abstract *semantic roles* (see also van Trijp 2011, 2015 for a detailed account). A verb such as *kick* includes two highly abstract roles that I will call *Actor* and *Undergoer*.[1] These semantic roles are linked to the frame elements that are expressed as the subject and direct object of active transitive and ditransitive clauses.[2] For instance, the kicker (kick-1) is categorized as the Actor and the kicked object (kick-2) is categorized as Undergoer.

The value of SEM-VALENCE is an unordered list of features (e.g. Actor and Undergoer) whose values repeat one of the variables used in the meaning of the construction. For instance, the variable ?x (which was bound to the kick-1) is reused for the Actor. This **variable equality** indicates that the kicker is mapped onto

1. Please note that the labels *Actor* and *Undergoer* have been chosen for mnemonic reasons only, and that these abstract roles actually cover many more functions than their names suggest. For example, the Actor role can not only be used for categorizing Agents, but also Experiences, States, and so on.

2. The presence of the Undergoer role also makes the verb compatible with the Passive Voice Construction.

the actor.[3] Besides the highly abstract roles of Actor and Undergoer, the value of SEM-VALENCE may also include less abstract semantic roles such as Instrument or Location. The value of SEM-VALENCE is open-ended and does not impose any maximum number of semantic roles, nor does it assume a fixed set of roles. Secondly, the SEM-VALENCE of a verb only influences its combinatorial potential but does not specify which role should be explicitly expressed and how. For example, there is no direct link between the Actor role and the grammatical function of Subject. The choice of which roles are expressed and which form they are realized in depends on the grammatical constructions that the verb interacts with. Interested readers can also find more details in van Trijp (2011, 2015).

SEM-CAT

The feature SEM-CAT contains two subfeatures: sem-class and emphasis. Both subfeatures fall beyond the scope of this paper, which focuses on verb phrases, but interested readers can nevertheless check the online FCG grammar for English to see the consequences of these features.

First, the sem-class feature categorizes the verb as a *predicate*, which is also how adjectives are categorized in the online grammar. Essentially, this allows verbs to function as adjectives in nominal phrases, as in *the motivated student*. Secondly, the emphasis feature repeats one of the variables from the meaning of the verb to indicate whether the verb takes accusative or ergative alignment, which regulates how verbs interact with many other grammatical constructions. For example, the Intransitive Construction will map the Actor of an accusative verb onto the Subject role (e.g. *the horse kicked*), but the Undergoer of an ergative verb (e.g. *the window broke*).

SYN-CAT

Lexical constructions do not directly specify a word's surface form, but instead introduce an abstract feature called SYN-CAT. The value of this feature is an *unordered list* of feature-value pairs. One important subfeature is lex-class, which categorizes *kick* as a verb. For auxiliaries, the value is aux. The distinction between verb and aux is used for controlling which verbs are compatible with the NICE properties of English, and which are not. However, verbs that are classified as lexical verbs are still able to function as auxiliaries, as I will illustrate further below.

It is useful to think of the other subfeatures of SYN-CAT as a blackboard on which other constructions can write information. A lexical construction such as *to*

3. For readers who are more familiar with the terminology of Goldbergian Construction Grammar (Goldberg 1995): the variable equality formalizes the *fusion* between the "participant role" *kicker* and the "argument role" *Actor*.

kick will add some features to this blackboard, but will typically leave a lot of values *underspecified*. For example, a verbal construction already adds the features verb-form, TAM and agreement to its SYN-CAT feature, but it does not yet specify any value for them because these values will depend on the combination of the verb with other constructions. Underspecification is again represented using a *variable* as shown in the construction in Example (10).

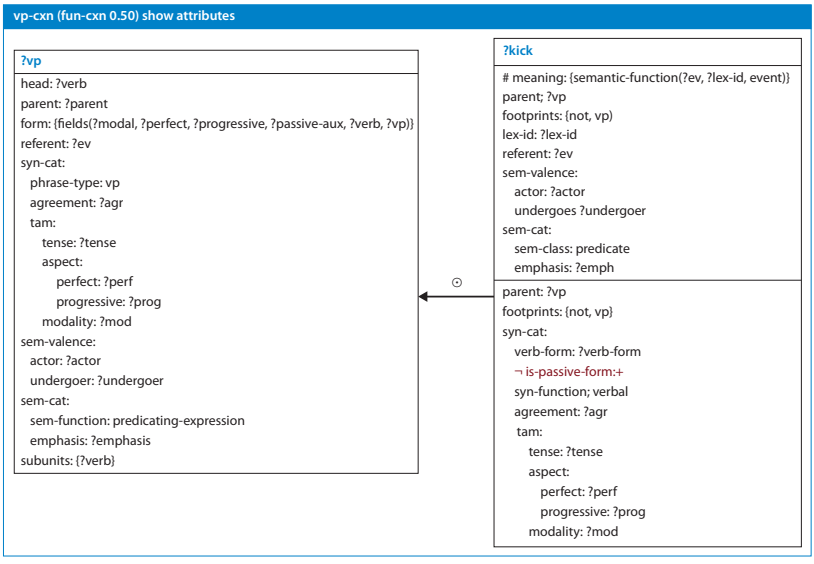
LEX-ID

The ~~,ex-id~~ feature takes a symbol as its value and can be considered as the *stem* of the verb. The choice of using lex-id rather than the linguistic notion of stem was made to avoid commitment to any particular morphological theory. The actual form of the construction will be contributed by morphological constructions depending on the values of the subfeatures of SYN-CAT: *kick*, *kicks*, *kicked*, or *kicking*. In comprehension, the same morphological constructions are able to supply the requested qualifications so the lexical construction is immediately activated.

2.3 Verb Phrase Construction

The Verb Phrase Construction (shown in Example (11) is a construction that takes a verb and makes it the head of a VP. This VP can be built by the construction itself, or it may already exist if it has already been built by one of the TAM-constructions discussed further below. To continue the metaphor of offices and incumbents, the

(11) **Verb Phrase Construction:**



verb is a suitable candidate for filling the (main) verb office of the VP unless it already fills an office in a different phrase (e.g. when it is used as an adjective in a NP). The qualifications for the main-verb office are shown on the right-hand side of the construction. The formulation-lock specifies that the office's incumbent must be a predicate. The comprehension-lock states that the incumbent must be a verb.

Let us illustrate how the Kick Construction and the Verb Phrase Construction interact with each other in language comprehension. When processing an utterance, FCG keeps all of the obtained information about the utterance in a *transient structure* (Steels, this volume). After applying the morphological construction that recognizes *kicks*, the lexical construction shown in Example (9) is immediately triggered. These two constructions add a lexical unit to the transient structure, as shown in Example (12). As can be seen, the verb's features contain more specific information than specified in the lexical construction. These values were contributed by the morphological construction, which I will discuss later in this paper.

Now that the transient structure contains a lexical unit, the Verb Phrase Construction checks whether that unit would make a suitable filler for its (main) verb office. This is indeed the case: *kicks* has all the required qualifications, so the Verb Phrase construction adds a new unit to the transient structure.

Here, the left-hand side of the construction in Example (11) becomes important. Just like its main verb, the verb phrase contains features such as referent, agreement, and TAM, whose values are underspecified through a variable. Note however that the variable names in the ?VP-unit are the same ones as those of the main verb. These variable equalities have the effect that the values of these features are copied from the main verb and repeated in the verb phrase. Example (13) shows the expanded transient structure after applying the Verb Phrase Construction.

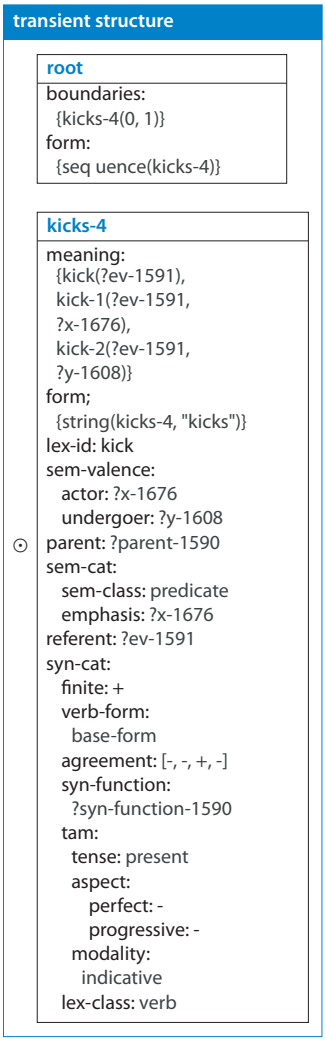
The subsequent sections explain how other grammatical constructions for tense, aspect and modality are able to open new office spaces in the verb phrase.

2.4 Present- and Past-Tense-Indicative Constructions

TAM systems relate to a far more complex domain than a "linear time concept familiar from high school physics" (Moens & Steedman 1988: 15). Building on the tense logics of Reichenbach (1947), cognitive linguists typically use three time points for analyzing temporal expressions: the moment of speaking (henceforth S or *Speech Time*), the time when the event described by the speaker takes place (E or *Event Time*), and R or *Reference Point*.

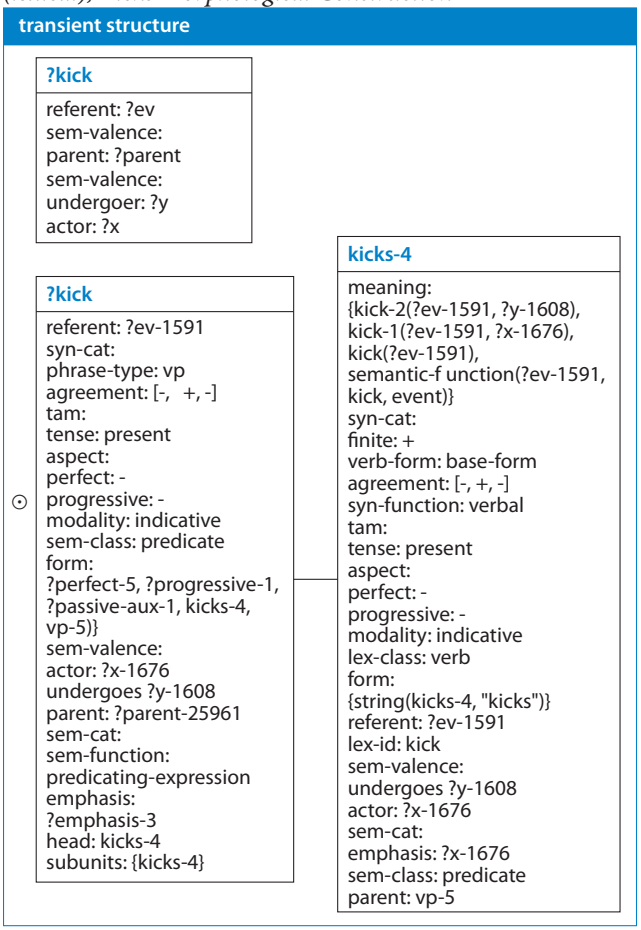
When locating events in time, the speaker does not directly relate Event Time to Speech Time. Instead, the speaker first establishes the Reference Point of the event and relates it to the Speech Time. If R precedes S, it is a *Recalled Point* and situated in the past time sphere. If R co-occurs with S, it is a *Present Point* and

(12) **Transient Structure**
Constructions applied so far: *Kick Construction (lexical), Kicks Morphological Construction*



situated in the present time sphere. R can also be an *Anticipated Point*. For the time being, I will say that an Anticipated Point is situated in the future time sphere, though we will see later that this statement is not entirely truthful. Example (14) shows how these three relations are represented using a diagrammatic notation on the left and a first-order logic notation on the right.

(13) **Transient Structure**
Constructions applied so far: *Verb Phrase Construction, Kick Construction (lexical), Kicks Morphological Construction*



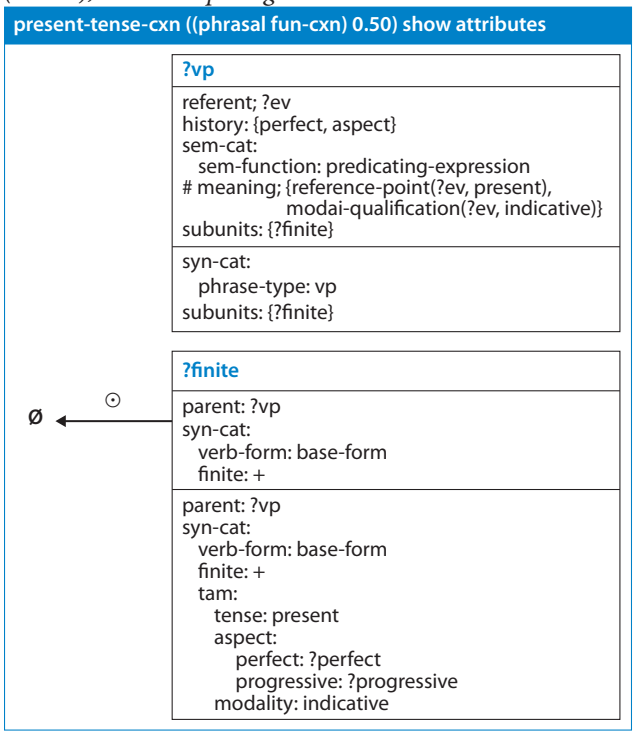
(14) a. R – S *reference-point(?event, recalled)*
b. S,R *reference-point(?event, present)*
c. S – R *reference-point(?event, anticipated)*

Morpho-syntactically speaking, English only distinguishes two tenses: past and present. That is, English has not grammaticalized a future tense as found in languages such as Spanish or French. Instead, it expresses futurity using modal auxiliaries. The proposed model therefore only has two tense constructions: a Past-Tense-Indicative Construction and a Present-Tense-Indicative Construction.

Present-Tense-Indicative Construction

The Present-Tense Construction – shown in Example (15) – assigns the meaning of (14b) to all temporal expressions that use a verb form in the present tense and that are in the *indicative mood* (i.e. there are no modal auxiliaries), such as *I read a book*, *I am reading a book*, *I have read a book*, *I have been reading a book*, and so on. In other words, the construction specifies that the Reference Point of the event is in the present time sphere. The constraints of the construction are straightforward and basically stipulate that there must be a VP that contains a finite verb.

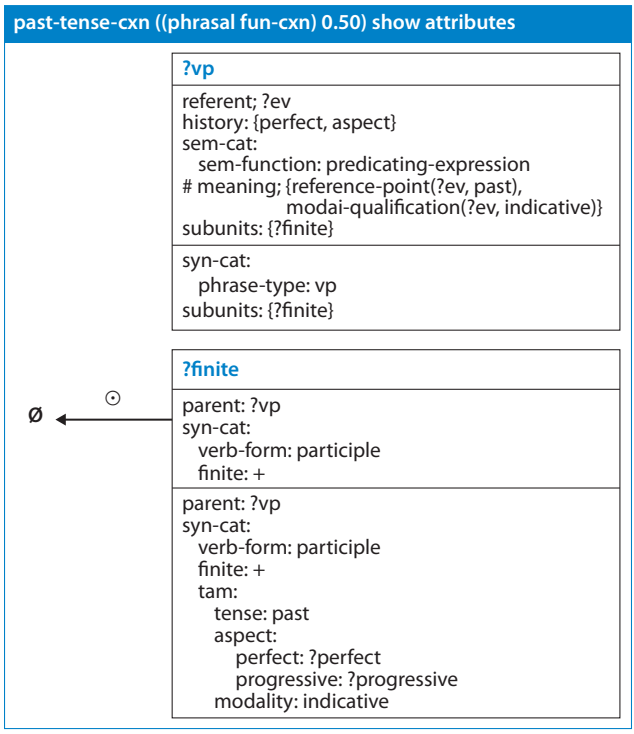
(15) **Present-Tense Construction:**
~~Constructions applied so far: *Verb Phrase Construction, Kick Construction (lexical), Kicks Morphological Construction*~~



Past-Tense-Indicative Construction

The Past-Tense-Indicative construction is almost the same as its Present-Tense counterpart. There are only two differences. First, its meaning feature indicates that the event's Reference Point is a Recalled point, as shown in (14a). Secondly, it specifies that the value for the feature tense must be past. Example (16) shows the construction in full detail.

(16) **Past-Tense Construction:**



2.5 Perfect and Non-Perfect Constructions

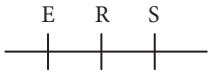
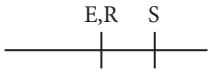
After establishing a Reference Point, the speaker needs to *ground* Event Time in the Reference Point. Event Time can be anterior to or simultaneous with the Reference Point.[4] Example (17) shows an informal representation of these two relations on the left, followed by their formal representation on the right.

(17) a. E – R relation(?E,?R,anterior)
b. E,R relation(?E,?R,simultaneous)

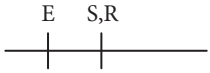
Example (18) illustrates how English expresses these three relations between Event Time and the Reference Point if the Reference Point is situated in the past time sphere. Example (19) does the same for Reference Points situated in the present time sphere.

4. In Reichenbach's (1947) system and most cognitive linguistics approaches (e.g. Radden & Dirven 2007), E can also be posterior to R. I did not include this possibility in my model because English does not seem to express such relations.

(18)

Semantics	Syntax	Example
	Past Perfect	*He had just finished his book.*
	Simple Past	*He read his book.*

(19)

Semantics	Syntax	Example
	Present Perfect	*They have reached the top.*
	Simple Present	*He passes the ball.*

Perfect Construction

Whenever Event Time is anterior to the Reference Point, English employs the Perfect Construction. Linguists disagree on whether the English perfect should be treated as a tense or an aspect marker (Radden & Dirven 2007: 206). From an implementation point of view, however, it is clear that the Perfect Construction is not a tense marker: if the perfect were a tense construction, the model would require separate constructions for the simple past, past perfect, simple present, present perfect, and so on. By treating the perfect as an aspect construction, it can be combined with either the Present, Past or Modal constructions to form different temporal expressions in a compositional fashion.

Also from a purely descriptive point of view, it makes sense to treat the perfect as an aspect construction. Consider Example (20):

(20) *Jennifer has cut her hair shorter than shoulder length.*

The Event Time is clearly in the past, yet English uses present tense morphology. The reason is that the English Perfect Construction profiles a contingency relation between E and R, which makes it more of an aspect construction. In Example (20), the result of the haircutting event is still visible at the Reference Point and at Speech Time. If Jennifer's hair had already grown to her waist, the sentence would have been inappropriate to native speakers. Likewise, the utterance *They have reached the summit of Mount Everest* implies that the mountaineers are still at the top of the mountain. Examples (21–24) show how the contingency relation expressed by the perfect can be instantiated in numerous ways.

(21) *I have lived in France for nine years.*
(i.e. I moved to France nine years ago and am still living there now)

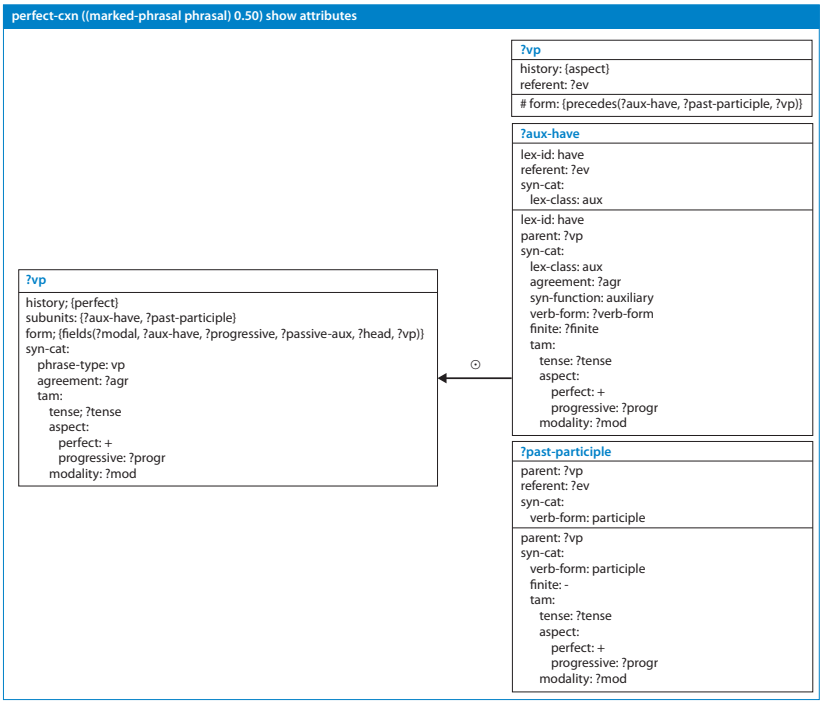
(22) *I have already bought the album.*
(i.e. the album is still in my possession)

(23) *I have never seen a real whale.*
(i.e. this fact was true in the past and is still true now)

(24) *I have visited Japan twice.*
(i.e. the visits were in the past but the statement is still true now)

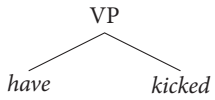
Example (25) shows the formal details of the Perfect Construction. On the right-hand side, three offices are specified for a Verb Phrase, Auxiliary Have, and a Past Participle. The qualifications for the auxiliary specifically require the verb *to have* (in any of its forms). The auxiliary must also agree with the verb-phrase with respect to its TAM and agreement features, and it needs to precede the past participle ~~(indicated by the symbol <)~~. Note that the word order constraint (specified in the form feature) of the ?VP-unit is preceded by a #, which means that the order must have been observed in the input. Since there are no other qualifications for the VP in the comprehension lock of the unit, the construction does not require the VP to have already been built by other constructions. Indeed, if the VP does not exist yet, the Perfect Construction will build it and add it to the transient structure.

(25) **Perfect Construction:**



As mentioned before, the model presented in this paper involves a rather flat structure for the verb phrase. Example (26) shows the structure as a tree diagram, where all verbal forms are immediate children of the VP, as opposed to a multi-clausal analysis in which the auxiliary functions as the syntactic head.

(26)



All TAM constructions operate according to the same principle: they check for candidate elements in the transient structure to fill their offices, and if so, they may admit auxiliaries as new members of an existing verb phrase; alternatively, they build the verb phrase if it does not exist yet. In the case of the Perfect Construction, there is also a Non-Perfect Construction that applies only in the absence of auxiliary-have. In this case, the Non-Perfect Construction adds the meaning to the transient structure that Event Time is simultaneous with the Reference Point.

2.6 Progressive and Non-Progressive Constructions

Besides the perfect, English has another aspect category that is often called the *progressive* or *continuous*. In its most prototypical form, the progressive consists of auxiliary-BE followed by an *ing*-form, as in *I am writing*.

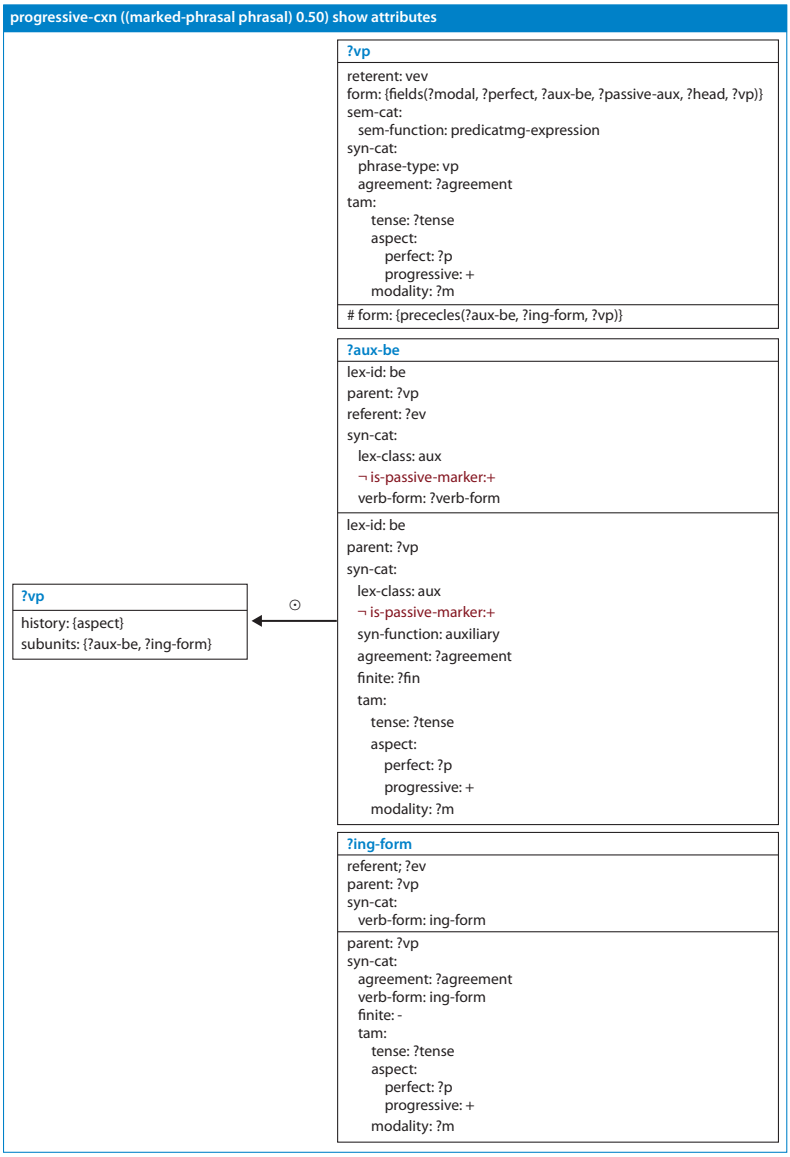
Progressive Construction
Just like the Perfect Construction, the Progressive Construction needs suitable candidates for filling three offices: a Verb Phrase, an Auxiliary, and an *ing*-form, as shown on the right-hand side in Example (28). The syntactic constraints on these offices (specified in the comprehension-locks) are relatively straightforward: variable equalities indicate that the auxiliary must agree with the verb phrase in terms of the features agreement and TAM. The auxiliary must also precede the *ing*-form.

The Progressive Construction can be combined with the Perfect Construction and the Tense Constructions in a compositional way, as illustrated in Example (27):

(27)

S ↔ R	**E ↔ R**	**Progressive**	**Example**
S,R	E – R	–	*He has read the book.*
S,R	E – R	+	*He has been reading the book.*
S,R	E,R	–	*He reads a book.*
S,R	E,R	+	*He is reading a book.*
R – S	E – R	–	*He had read the book.*
R – S	E – R	+	*He had been reading the book.*
R – S	E,R	–	*He read a book.*
R – S	E,R	+	*He was reading a book.*
R – S	R – E	–	*He would read a book.*
R – S	R – E	+	*He would be reading a book.*

(28) **Progressive Construction:**



2.7 Modal Constructions

We now turn to future time expressions. Following Reichenbach's (1947) system, most cognitive linguists assume that a simple future tense expression involves the meaning presented in (29):

(29) **Diagram**

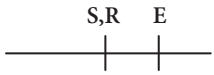


Formal Notation

{ *reference-point(?e,posterior),*
relation(?e, ?r, simultaneous) }

However, as already said, there is no future tense inflection in English, so there is no way of discriminating the meaning of (29), in which the Reference Point is set in the future, from the meaning in (30), where the Reference Point is in the present time sphere.

(30) **Diagram**



Formal Notation

{ *reference-point(?e,present),*
relation(?e, ?r, posterior) }

Another approach is to circumvent this inconsistency by stating that "simple" temporal expressions do not require an additional Reference Point and instead directly relate Event Time to Speech Time (cf. Radden & Dirven 2007). While this solution has intuitive appeal, it is undesirable from a processing perspective because it leads to inconsistencies in the grammar and loss of compositionality.

The solution presented here is to stick closer to the facts of English, which expresses futurity not through tense but through *modals*. Why modal auxiliaries? The difficulty about the future is that it is not set in stone. Speakers, therefore, have to add all kinds of qualifications to future time expressions, such as the degree of likelihood that an event will indeed happen. Since modality concerns "the speaker's assessment of, or attitude towards, the *potentiality* of a state of affairs" (Radden &Dirven 2007: 233, my emphasis), modal auxiliaries offer a natural way of expressing potential events in the future. Example (31) illustrates this point through three future time expressions. Sentence (a) announces a scheduled sports game in the near future almost as if it were a fact (using the simple present) because it will only be cancelled in extraordinary circumstances. Utterances (b) and (c), on the other hand, include the modal auxiliaries *will* and *might* for indicating how certain the speaker is of his/her prediction.

(31) a. *FC Barcelona plays Real Madrid tonight.*
b. *FC Barcelona will beat Real Madrid tonight.*
c. *Real Madrid might win against FC Barcelona tonight.*

Modality is typically formalized using *possible world semantics* (Kripke 1963). A *possible world* simply means any possible situation, which can be contrasted to an actual situation. The kind of relation between the possible world and the actual situation (e.g. deontic, epistemic, and so on) has an impact on which auxiliaries can be used. For example, when the relation is *deontic*, the speaker expresses his/her

desire for the realization of a possible world (e.g. *you MUST leave now* expresses a stronger desire than *you CAN leave now*).

A comprehensive implementation of modality is still in progress for the English FCG grammar (van Trijp 2017), so unfortunately the remainder of this paper can only sketch a prototype implementation that offers a first step towards a deep, linguistically motivated analysis. I will restrict myself to the auxiliaries *will* and *would*. In both cases, there is a generic Marked-Modality Construction that establishes a relation between a modal auxiliary and the remainder of the verb phrase. The nature of the link is specified in the meaning of the modal auxiliary itself.

Modal-Will Construction

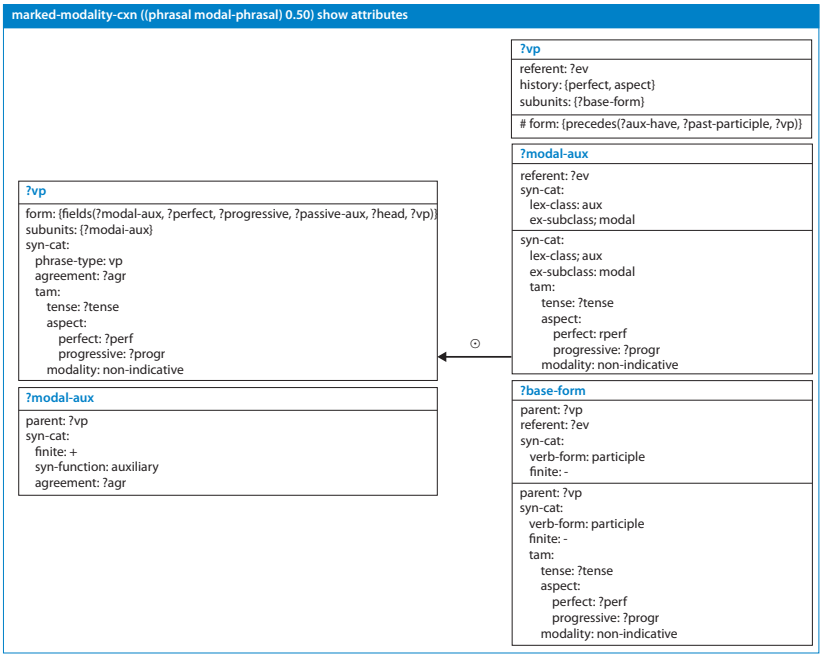
The modal auxiliary *will* expresses the fact that a speaker expresses some volition or prediction about a possible world. The MEANING of the construction is shown in (32):

(32) *{ reference-point(?ev, possible-world), modal-qualification(?ev, volition-or-predication }*

Marked-Modality Construction

The marked modality construction specifies three offices: a VP, a modal auxiliary, and some verb in its "base form" (i.e. any form except an *ing*-form or past-participle):

(33)



The past tense counterpart of the simple future expressions (e.g. *I would go*) involve the same Marked-Modality Construction, but the auxiliary *would* contributes a different modal meaning to the sentence. The constructions can again be combined with all of the previously defined constructions in a fully compositional manner, except from the Present-Tense-Indicative and Past-Tense-Indicative constructions (because these conflict in terms of the utterance's modality). Some combinations for future time expressions are shown in (34):

(34)

R	E	Progr.	Example
Anticipated	Anterior	–	*In 2020, I will have lived in France for 15 years.*
Anticipated	Anterior	+	*In 2020, I will have been doing research for 15 years.*
Anticipated	Simultaneous	–	*In 2020, I will retire from football.*
Anticipated	Simultaneous	+	*In 2020, I will be traveling.*

3. Conclusions

This paper opened with a citation from Schmerling (1983), suggesting that no satisfying solution could be found for the English auxiliaries, just like no solution could be found for the American car industry. More than 30 years later, the debate on whether auxiliaries should be treated as main verbs or feature carriers still has not been settled; and as a grim foreboding the city of Detroit filed for bankruptcy on 18 July 2013.

This paper proposes that a Construction Grammar approach allows us to move out of this deadlock situation. Through a fully implemented, bidirectional model of the English auxiliaries and verb phrase in Fluid Construction Grammar, the paper has demonstrated that a constructional approach solves the empirical inadequacies of the auxiliaries-as-raising-verbs account, avoids the formalization problems of the horizontal/modular traditional analyses, and is able to cope with the less conventionalized usage patterns of English verbal expressions.

Acknowledgements

This paper and the model it describes have greatly benefited from intense discussions with Luc Steels, Katrien Beuls, Miquel Cornudella and Paul Van Eecke. It was further improved thanks to the constructive feedback of Emília Garcia Casademont, Yana Knight, Johan Loeckx, Tânia Marques, Dirk Pijpops, Michael Spranger and Freek Van de Velde at the FCG Summer Meeting 2014 in Orval (Belgium).

References

Baker, C. F., Fillmore, C. J., & Lowe, J. B. (1998). The Berkeley FrameNet project. In *Proceedings of the 17th International Conference on Computational linguistics*, Morristown, NJ, USA. Association for Computational Linguistics. doi: 10.3115/980451.980860

Bolinger, D. (1983). The go-progressive and auxiliary formation. In F. B. Agard & C. F. Hockett (Eds.), *Essays in honor of Charles F. Hockett* (pp. 153–167). Leiden: E. J. Brill.

Butt, M., Niño, M. -E., & Segond, F. (1996). Multilingual processing of auxiliaries in LFG. In D. Gibbon (Ed.), *Natural language processing and speech technology: Results of the 3d KONVENS conference* (pp. 111–122). Berlin: Mouton De Gruyter.

Denison, D. (2000). Combining English auxiliaries. In O. Fischer, A. Rosenbach, & D. Stein (Eds.), *Pathways of change* (pp. 111–147). Amsterdam: John Benjamins. doi: 10.1075/slcs.53.07den

Dyvik, H. (1999). The universality of f-structure: Discovery or stipulation? The case of modals. In M. Butt (Ed.), *Proceedings of the LFG '99 Conference*, Manchester. Stanford: CSLI Publications. Available at http://web.stanford.edu/group/cslipublications/cslipublications/LFG/4/lfg99dyvik.pdf

Falk, Y. N. (2008). Functional relations in the English auxiliary system. *Linguistics*, 46(5), 861–889. doi: 10.1515/LING.2008.028

Fillmore, C. J. (1988). The mechanisms of Construction Grammar. In S. Axmaker & H. Singmaster (Eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society* (pp. 35–55). Berkeley, CA: Berkeley Linguistics Society.

Frank, A., & Zaenen, A. (2004). Tense in LFG: Syntax and morphology. In L. Sadler & A. Spencer (Eds.), *Projecting morphology* (pp. 23–65). Stanford: CSLI Publications.

Goldberg, A. E. (1995). *A Construction Grammar approach to argument structure*. Chicago: Chicago University Press.

Huddleston, R. D. (1976). Some theoretical issues in the description of the English verb. *Lingua*, 40(4), 331–383. doi: 10.1016/0024-3841(76)90084-X

Kripke, S. (1963). Semantical analysis of modal logic. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 9, 67–96. doi: 10.1002/malq.19630090502

Manning, C. D. (1995). Dissociating functor-argument structure from surface phrase structure: The relationship of HPSG order domains to LFG. Unpublished Manuscript. URL http://nlp.stanford.edu/manning/papers/hpsglfg1.pdf.

Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational Linguistics*, 14(2), 15–28.

Palmer, F. R. (1979). Why auxiliaries are not main verbs. *Lingua*, 47(1), 1–25. doi: 10.1016/0024-3841(79)90064-0

Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.

Radden, G., & Dirven, R. (2007). *Cognitive English Grammar, volume 2 of Cognitive Linguistics in practice*. Amsterdam: John Benjamins.

Reape, M. (1994). Domain union and word order variation in German. In *German in Head-Driven Phrase Structure Grammar* (pp. 151–197). Stanford: CSLI Publications.

Reichenbach, H. (1947). *Elements of symbolic logic*. New York: Macmillan.

Ross, J. R. (1969). Auxiliaries are main verbs. In W. Todd (Ed.), *Studies in philosophical linguistics* (pp. 77–102). Evanston, IL: Great Expectations Press.

Schmerling, S. F. (1983). A new theory of English auxiliaries. In F. Heny (Ed.), *Linguistic categories: Auxiliaries and related puzzles, volume two: The scope, order, and distribution of English auxiliary verbs* (pp. 1–53). Dordrecht: D. Reidel Publishing Company. doi: 10.1007/978-94-009-6992-6_1

Steels, L. (Ed.). (2011). *Design patterns in Fluid Construction Grammar*. Amsterdam: John Benjamins. doi: 10.1075/cal.11

van Trijp, R. (2011). A design pattern for argument structure constructions. In L. Steels (Ed.), *Design patterns in Fluid Construction Grammar* (pp. 115–145). Amsterdam: John Benjamins. doi: 10.1075/cal.11.07tri

van Trijp, R. (2015). Cognitive vs. generative construction grammar: The case of argument structure and coercion. *Cognitive Linguistics*, 26(4), 613–632. doi: 10.1515/cog-2014-0074

van Trijp, R. (2016). Chopping down the syntax tree: What constructions can do instead. *Belgian Journal of Linguistics*, 30, 15–38. doi: 10.1075/bjl.30.02van

van Trijp, R. (2017). A constructional language processing model for English in Fluid Construction Grammar. In *Proceedings of The AAAI 2017 Spring Symposium on Computational Construction Grammar and Natural Language Understanding, Technical Report SS-17-02* (pp. 266–273). Palo Alto, CA: AAAI Press.

Author's address

Remi van Trijp
Sony Computer Science Laboratories Paris
6, rue Amyot
F-75005 Paris
France

remi@csl.sony.fr