Oliver Calder

10 June 2019

Multiple Auxiliaries in English

§ 1. Introduction

The phenomenon of multiple auxiliaries occurs frequently in English syntax, and refers to the string of one or more auxiliary verbs which may precede the main verb in a verb phrase. These auxiliary verbs modify the meaning of the main verb, and in turn, the verb phrase as a whole. The tense, aspect, and mood of the verb phrase can all be affected by the auxiliaries. The combinations of auxiliaries needed to form specific tenses, aspects, and moods follow very strict rules. For example, the sentence "*I will have been running*" is grammatical, while the sentence *"*I being will had run*" is ungrammatical. The goal of this paper is to examine possible combinations of auxiliaries in order to define rules for syntactically correct auxiliary strings.

For the purposes of this paper, the so-called auxiliary strings, along with the main verb they modify, will be referred to collectively as a verb cluster (VC). This distinguishes auxiliaries from other elements of verb phrases, such as noun phrases or prepositional phrases, and shows that auxiliaries are more closely connected to the main verbs of verb phrases than the other objects or clauses which may be constituents. When all other parts of the verb phrase have been stripped away, what remains is the verb cluster as a single unit, rather than the verb alone.

We shall see that tense and aspect heavily influence the syntactic structure of verb clusters, and that individual auxiliary verbs play both semantic and syntactic roles. The formation of the infinitive and moods other than indicative have very specific syntactic and semantic rules which differ significantly from the indicative, so we will consider these to be

beyond the scope of the paper. Overall, the goal is to establish sufficiently strict rules so as to limit both the number of grammatical sentences which our rules cannot account for and the number of ungrammatical sentences which our rules permit as syntactic.

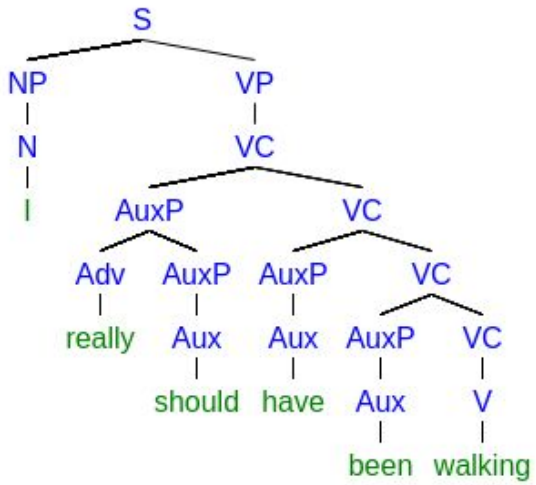
§ 2. Critique of Current Rules

We have a system in place to account for multiple auxiliaries, with rules as follows:

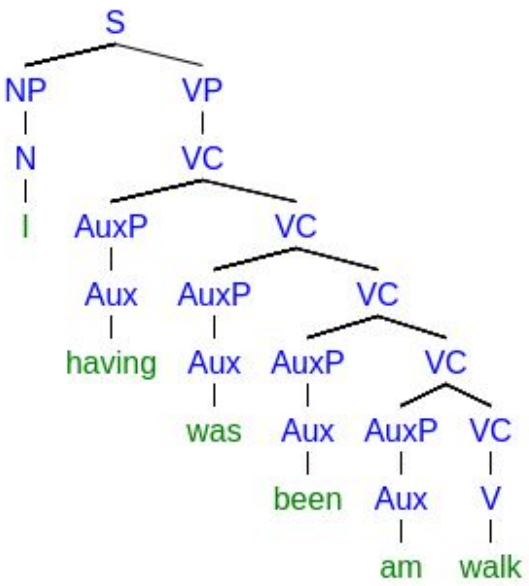
VC → AuxP VC
AuxP → Adv AuxP
AuxP → AuxP Adv
AuxP → Aux
VC → V

These rules have several positive characteristics. Firstly, they successfully account for the vast majority of grammatical verb clusters within the scope of the investigation (i.e. not passives, infinitives, etc.). They are also very concise and straightforward, as a result of their recursive property. However, this property allows for tacking on infinitely many auxiliaries in any order, which thus vastly overgenerates ungrammatical sentences.

"*I really should have been walking*"



*"*I having was been am walk*"



The current rules treat both sentences above as correct, and allow for the creation of syntax trees for their structures. However, the second sentence is ungrammatical, and the goal is to develop rules which to not allow for such ungrammatical sentences to be treed. Thus, it is necessary to examine the possible combinations of auxiliaries which yield syntactically correct verb clusters.

§ 3. Initial Presentation of Data

All English sentences in the indicative (most common) mood can be reduced to a finite set of possible verb clusters through elimination of external constituents, such as direct or indirect objects and subordinate clauses. Doing so is essentially equivalent to reducing all verbs to intransitive forms. For example, the sentence "*I was walking my dog to the store*" can be reduced to the verb cluster "*was walking*," with the understanding that it is possible to re-append noun phrases, prepositional phrases, and other constructions onto the outside of the verb cluster, in accordance with the predefined rules, in order to create a large number of possible grammatical sentences. For the verb "*to walk*," it can thus be seen that there are 12 different combinations of auxiliaries, based upon the tense and aspect of the phrase.

"*to walk*"	**Present**	**Past**	**Future**
Simple	walk	walked	will walk
Perfect	have walked	had walked	will have walked
Progressive	am walking	was walking	will be walking
Perfect Progressive	have been walking	had been walking	will have been walking

In the chart above, the columns and rows correspond to the tense and aspect, respectively. Any other verb can be substituted in for "*to walk*" to create any grammatical verb cluster.

§ 4. Tense-Based Analysis of Data

From the chart of verb clusters in § 3, several patterns can be identified. Firstly, all verbs in each of the verb clusters individually take either the present, past or progressive form. The exception to this is the word "*will*," which, as will be discussed, is not a verb in the same sense as the other verb cluster components. It shall be termed an auxiliary, or Aux. Given these assumptions, the chart may be color coded in order to reveal deeper patterns within verb clusters.

"*to walk*"	**Present**	**Past**	**Future**
Simple	walk	walked	will walk
Perfect	have walked	had walked	will have walked
Progressive	am walking	was walking	will be walking
Perfect Progressive	have been walking	had been walking	will have been walking

In the chart above, all present verbs are colored blue, all past verbs are colored red, all progressive verbs are colored green, and the auxiliary "*will*" is colored purple. Now, four patterns are clear: (1) all present VCs begin with a present tense verb, all past VCs begin with a past tense verb, and all future VCs begin with the auxiliary "*will*"; (2) the future tense is identical to the present tense, with the addition of "*will*" at the beginning; (3) the first verb, excluding auxiliaries such as "*will*," must be in either the present or the past form; (4) the order of verb tenses in the VCs is always Aux (optional), Verb (present or past), past tense Verb (optional), progressive tense Verb (optional).

It is important to clarify what is implied by the distinction between auxiliaries and verbs as they appear in verb clusters. As was stated above, verbs have three inherent forms (present, past, and progressive) which are differentiated within each verb itself, rather than being indicated

through the presence of some other verb. For example, the verb "*to ask*" has the present "*ask*," the past "*asked*," and the progressive "*asking*" forms. Also, verbs can stand alone in sentences such as "*I sit*." In contrast, auxiliaries, such as "*will*," do not have multiple forms, and cannot act as the singular verb in a sentence. The phrase "*I will*" is only valid as a responds to some other sentence such as "*Will you x?*" where "*x*" is some verb phrase. Thus, the response actually implies "*I will [x]*," where the silent verb phrase "*x*" is implied through context.

The auxiliary "*will*" appears in the chart in order to trigger the future tense. However, other auxiliaries can take its place to form a variety of different meanings. The verb clusters "*should be walking*," "*shall walk*," "*might have walked*," "*would be walking*," and "*can have been walking*" are all syntactically and semantically acceptable. Importantly, as was the case with "*will*," each auxiliary must be followed by a verb in present tense, as clusters such as *"*may walked"* and *"*could walking*" are ungrammatical. Additionally, There may only be one auxiliary in any verb cluster, as clusters such as *"*shall would walk*" are ungrammatical. An exception to this occurs in some dialects of English spoken in the southern United States, where constructions such as "*might could*" are common. However, for the purposes of this paper, this exception will be ignored in favor of broader applicability to common patterns in English.

§ 4.1 Initial Tense-Based Rules

From the patterns discussed in § 4, some initial rules can be created to specify options for the first components of the verb cluster. If there is an auxiliary, it must be followed by a verb in present tense. Otherwise, it may begin with a verb in either past or present tense. Since there is commonly more than one verb in the VC, these take the form of PresVC or PastVC, which are constituent verb clusters to the primary VC, and which must begin with a verb in the tense

corresponding to the tense marker. Thus, PresVC means a verb cluster which begins with a verb in the present tense. These rules are as follows:

VC → Aux PresVC
VC → PresVC
VC → PastVC

Now, within a PresVC, there are four options corresponding to the four entries on the present tense column: present, present+past, present+progressive, and present+past+progressive. The same applies for PastVCs, corresponding to the four entries on the past tense column: past, past+past, past+progressive, and past+past+progressive. Since it is preferable to avoid tri-branching, the last option must be grouped into a verb followed by a verb cluster. However, both PresVC and PastVC follow directly from the initial VC and, to avoid infinite recursion and the possibility of having more than three verbs in the VC, should not go to each other. Thus, while still accounting for the present/past+past+progressive structure, both must go to PastVC', which is a PastVC which can only go to a single past tense verb or a past tense verb followed by a progressive tense verb. Thus, these patterns can be compiled into the following rules:

PresVC → PresV
PresVC → PresV ProgV
PresVC → PresV PastVC'

PastVC → PastV
PastVC → PastV ProgV
PastVC → PastV PastVC'

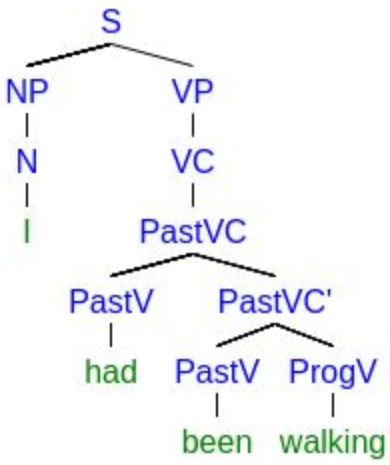
PastVC' → PastV
PastVC' → PastV ProgV

§ 4.2 Critique of Tense-Based Rules

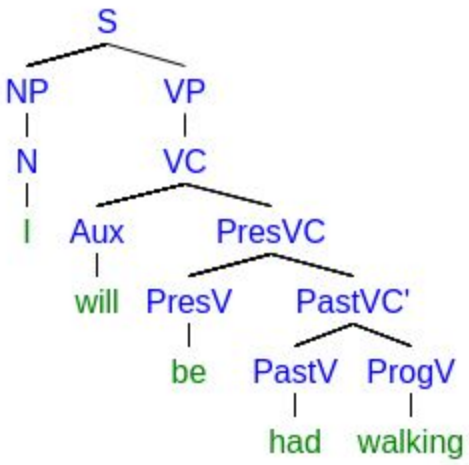
These rules would appear to work fairly well, but there are a few immediate complaints with them as well. Firstly, while PresVC and PastVC exist, there is no ProgVC, which is

inconsistent, given that verbs can take all three forms, and despite that there appears to be no need for a ProgVC construction. Secondly, there only exists a PastVC' construction, and there is no primed constructions for PresVC, which again appears to be inconsistent.

"*I had been walking*"



*"*I will be had walking*"



The rules account for both these sentences, and thus trees are able to be created. However, the latter sentence is ungrammatical, despite the fact that it follows an acceptable pattern of tenses within the verbs. The only difference between "*I have been walking*" and *"*I am had walking*" is that the verbs "*to be*" and "*to have*" are reversed in order. Thus, each has a specific location within the verb cluster where they must occur in order to create a grammatical sentence. This location is related to the aspect of the sentence. Therefore, "*to be*" and "*to have*" play both semantic and syntactic roles in the formation of aspect in verb clusters.

§ 5. Aspect-Based Analysis of Data

With the understanding that the verbs "*to be*" and "*to have*" play both semantic and syntactic roles with regard to aspect, we may revisit the table of verb clusters, this time marking "*to have*" in *italics* and "**to be**" in **bold**, and we may group by verb on the table:

"*to walk*"	**Present**	**Past**	**Future**
Simple	walk	walked	will walk
Perfect	*have* walked	*had* walked	will *have* walked
Progressive	**am** walking	**was** walking	will **be** walking
Perfect Progressive	*have* **been** walking	*had* **been** walking	will *have* **been** walking

From this annotated table, several new patterns emerge. Firstly, the verb "*to have*" is always present if and only if the verb cluster is perfective, and "*to have*" is always followed by a past tense verb or verb cluster. Secondly, the verb "*to be*" is always present if and only if the verb cluster is progressive, and "*to be*" is always followed by a verb in the progressive form, which is also the only situation in which a progressive verb can occur.

§ 5.1 Aspect-Based Rules

With those patterns in mind, recall the components of each aspect: simple is [nothing] plus the main verb; perfect is "*to have*" plus the main verb; progressive is "*to be*" plus the main verb; perfect progressive is "*to have*" plus "*to be*" plus the main verb. Simple verb clusters are already accounted for with our tense-based rules, so we need only create new rules for each of the other three aspects. These rules are as follows:

PerfVC → HaveV PastV
ProgVC → BeV ProgV
PerfProgVC → HaveV PastProgVC

These rules do not account for the tense of the verb cluster as a whole, and thus will need to be combined with the previous tense-based rules in order to fully account for possible combinations of tense and aspect. This is foreshadowed by the presence of the PastProgVC,

which is a ProgVC in which the first verb must be in the past tense. This will be further elaborated upon in § 6. Note that the ProgVC now exists, which was one criticism of the tense-based rules. Additionally, the PerfProgVC correctly treats the ProgVC as a constituent, establishing a useful pattern where each identifier is stripped off from the VC from left to right until only a single verb is left.

§ 6. Synthesis of Tense-Based and Aspect-Based Rules

It is now clear that the tense-based rules detailed in § 4.1 and the aspect-based rules detailed in § 5.1 both identify important patterns in the syntactic rules of verb cluster formation. However, neither can stand independently. Thus, they must be combined to account for all possible combinations of tense and aspect in verb clusters. From the initial VC, these combinations can be shown as follows:

VC → PresV	VC → PastV	VC → Aux PresV
VC → PresPerfVC	VC → PastPerfVC	VC → Aux PresPerfVC
VC → PresProgVC	VC → PastProgVC	VC → Aux PresProgVC
VC → PresPerfProgVC	VC → PastPerfProgVC	VC → Aux PresPerfProgVC

Notably, these combinations correspond directly to the verb cluster tables found in § 3, 4, and 5. Notice that the verb clusters with non-simple aspects all take the same form after their initial tense indicator. Similarly, the first two columns (those without an Aux) are identical except for their difference in the initial tense indicator. After this first stage of divergence from the initial VC, the rules are as follows:

PresPerfVC → PresHaveV PastV	PastPerfVC → PastHaveV PastV
PresProgVC → PresBeV ProgV	PastProgVC → PastBeV ProgV
PresPerfProgVC → PresHaveV PastProgVC	PastPerfProgVC → PastHaveV PastProgVC

The above rules are quite verbose and repetitive in their use of tense and aspect indicators. As such, both the tense and aspect indicators for each verb cluster can be abbreviated

and modularized as T and A, respectively. This allows for any combination of tense and aspect to be formed with few rules. The modularized rules are below:

VC → TV	TPerfVC → THaveV PastV
VC → TAVC	TProgVC → TBeV ProgV
VC → Aux PresV	TPerfProgVC → THaveV PastProgVC
VC → Aux PresAVC	
T ∈ {Pres, Past}	A ∈ {Perf, Prog, PerfProg

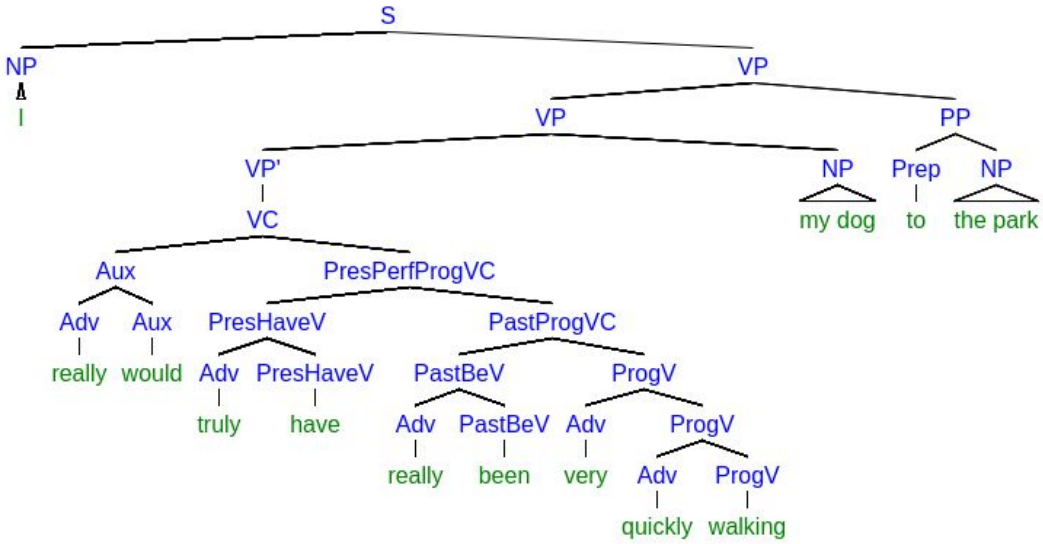
One important omission thus far has been adverbs occurring in verb clusters, such as "*I really would truly have really been very quickly walking*." Despite the absurdity of this sentence, it is grammatical, and illustrates that one or more adverbs can apply at any stage of the verb cluster. Thus, modularized rules with # as a wildcard for adverbs in VCs are as follows:

Aux → Aux Adv	#V → #V Adv	#VC → #VC Adv
Aux → Adv Aux	#V → Adv #V	#VC → Adv #VC

§ 6.1 Test and Critique of Rules

These rules can account for any standard combination of verb clusters. The lengthiest grammatical verb cluster, excluding chaining infinitely many adverbs together, is as follows:

"*I really would truly have really been very quickly walking my dog to the park*"



While the rules effectively account for grammatical and ungrammatical structures in verb clusters, they have several issues. Firstly, while the rules are concise, they use variables T and A to represent many possible structures using one line of the rules, which makes them difficult to read. Additionally, the trees themselves are very verbose and repetitive. While this helps ensure that only grammatical verb clusters can be formed, it is far more complicated than the initial recursive AuxP rules. Another issue is that the tense and aspect of the sentence must be known before the tree is begun, such that it can be correctly labeled at the first stage of the verb cluster. The recursive adverb rules also do not show true constituency, as in "*very quickly walking*," where "*very quickly*" should modify "*walking*," rather than "*very*" modifying "*quickly walking*."

§ 7. Conclusion and Future Research

Overall, the proposed rules successfully account for all standard (non-passive, non-infinitive) verb cluster combinations. However, they are not without flaw, and require substantial initial understanding of the classification of the verb cluster in order to properly tree the structure. Regardless, there are two key takeaways from this paper: (1) The system of, for a TAVC with many indicators, splitting off the first two indicators to form the first constituent, and then applying the aspect-dependent tense onto the remainder of the AVC to form the second constituent; (2) The treatment of tense and aspect as both semantic and syntactic indicators. From these, it is much easier to approach constructions such as passives, which behave very similarly to progressive verb clusters in that the final verb is replaced with two verbs, the first of which is "*to be*" in the tense of the original verb, and the second of which is the original final verb, either in the progressive tense or the past tense, for progressive and passive respectively. It is likely that a similar approach could be used to explore subjunctives and infinitives.