A Historical Overview of the Status of Function Words in Dependency Grammar

Timothy Osborne
Zhejiang University
Hangzhou
China
tjo3ya@yahoo.com

Daniel Maxwell
US Chess Center
Washington DC
USA
dan.maxwell@gmail.com

Abstract

This contribution provides a historical overview of the analysis of function words in surface syntactic dependency hierarchies. Starting with Tesnière (1959), the overview progresses through some prominent voices in the history of DG (Mel'čuk 1958, 1963, Hays 1964, Matthews 1981, Schubert 1987, Maxwell and Schubert 1989, Hudson 1976, 1984, etc.). The overview establishes that the analysis of prepositions has been almost unanimous: they are positioned as heads over their nouns. There has been more variation concerning the status of auxiliary verbs, although most DG grammarians have viewed them as heads over their content verbs. Concerning determiners, the dominant position is that they are dependents under their nouns, although there are a couple of prominent voices that assume the opposite stance.
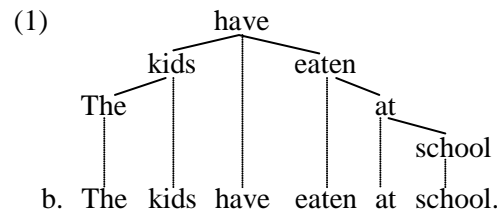
1 The dependency status of function words

The distinction between function words and content words has been made by linguists of various backgrounds. A rough definition of the distinction might be that content words can be understood without any supporting context whereas function words cannot. The discussion in this article takes place in terms of specific syntactic categories which happen to be function words according to this definition.

A recent proposal for surface analyses of dependencies categorically subordinates function words to content words. Universal Stanford Dependencies (USD) (de Marneffe et al. 2014) advocate an annotation scheme that positions auxiliary verbs, adpositions, subordinators, and determiners as dependents of the content words with which they co-occur. Thus according to this scheme, the DG analysis of the sentence *The kids have eaten at school* would be as follows:

(1)

a. The kids have eaten at school.
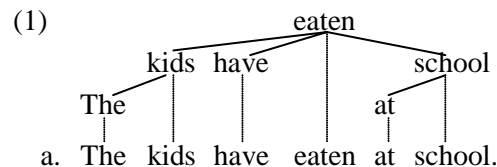
The USD analysis shown with (1a) stands in contrast to more conventional analyses, which position auxiliaries as heads over content verbs and prepositions as heads over their nouns:

(1)

b. The kids have eaten at school.

In pursuing the analysis in (1a), USD is advocating an understanding of surface dependencies that is generally contrary to the views about function words that have crystallized over the decades in support of the analysis in (1b).

This contribution surveys some prominent voices in the DG tradition, in order to determine the extent to which the analysis in (1a) has been advocated over the decades since Tesnière (1959) and the period in which computational linguistics has become influential. Space does not allow us to try to evaluate these analyses in detail, but when the authors of these analyses attempt to justify their decisions, we occasionally make comments about their argumentation.

2 Tesnière (1959/2015)

Tesnière's analysis of function words has not survived into most modern DGs. The reason it has not done so is tied to the fact that Tesnière's subtheory of transfer has also not survived into most modern DGs. Tesnière viewed most function words as *translatives* (auxiliary verbs, prepositions, subordinators, many determiners). As such, they were not granted autonomy in the syntactic representations, but rather they appeared together with a content word, the two

forming a *dissociated nucleus*. What this means in the current context is that Tesniere's *Éléments* (1959) did not provide much direct guidance concerning the dependency analysis of function words.

While Tesnière is widely credited as the father of modern DGs, he himself was not aware of the dependency vs. constituency distinction. That distinction was first established a few years after his death. Hays (1964) is generally credited with establishing the term *dependency grammar* (as opposed to *phrase structure grammar*). Thus the fact that Tesnière's account of function words left much room for debate about the actual dependency status of function words should not be so surprising. Tesnière never intended to produce an account of function words that would be consistent with the purely dependency-based theories of syntax that followed him.

The relevant aspect of Tesnière's grammar can be seen with the sentence *Bernard est frappé par Alfred* 'Bernard is hit by Alfred', the structure of which Tesnière showed with his Stemma 95:

(2)   est frappé

Bernard   par Alfred
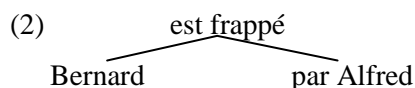
The noteworthy aspect of this stemma is the manner in which *est frappé* and *par Alfred* together occupy a single node each time. Neither can the function word be viewed as head/parent over the content word nor can the content word be viewed as head over the function word.[1] For Tesnière, the two words *est* and *frappé*, and the two words *par* and *Alfred*, formed a single nucleus together each time.

When Tesnière wanted to draw attention to the fact that a given function word forms a nucleus together with a content word, he put the two in a bubble. His stemmas of the sentences *Alfred est arrivé* 'Alfred has arrived' (Stemma 27) and *Alfred est grand* 'Alfred is big' (Stemma 28) are given here:

(3)   est arrivé   est grand[2]

    a.   Alfred   b.   Alfred

---
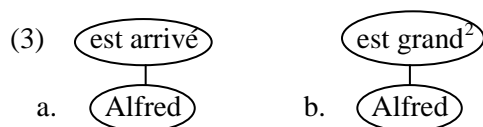
[1]One of us coauthors prefers the term *parent* to *head* in this context, reserving the latter term for one specific kind of parent, namely those which are content words rather than function words.

[2]The use of *est* 'is' in what is often considered two different ways is taken up in section 3.2.

Each of the top bubbles encloses the words of a single nucleus. Neither word in a bubble is head over the other.

This greater point established, one can nevertheless examine Tesnière's analysis of dissociated nuclei more closely. When one does so, one sees that he actually did provide some indirect guidance concerning the dependency analysis of function words. He drew a distinction between the structural and the semantic function of a nucleus (Chapter 29). In a disassociated nucleus, one of the words guarantees the structural integrity of the nucleus, and the other its semantic integrity. He also comments (Chapter 29, §18) that from an etymological point of view, the subsidiary word in a dissociated nucleus was once dependent on the constitutive word.

Tesnière later (Chapter 38, §19) states that auxiliary verbs, despite the fact that they lack semantic content, are constitutive of the nucleus in which they appear. Thus given these statements, one can extrapolate that, had Tesnière been forced to choose, he would have viewed the auxiliary verb as head over the content verb. Using his conventions, he might have produced an analysis of the sentence *Alfred est grand* 'Alfred is big' along the following lines:

(4)   est
         grand
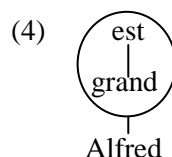
    Alfred

The two words *est* and *grand* together still form a single nucleus, but now the auxiliary verb is shown as head over the adjective, and at the same time, the two together form a single head over the noun *Alfred*.

One can also extend this extrapolated analysis to prepositions. Tesnière viewed a preposition as subsidiary to the noun with which it forms a dissociated nucleus (Chapter 29, §4). Thus for the phrase *livre d'Alfred*, lit. 'book of Alfred', the following analysis reflects the distinction between constitutive and subsidiary words inside the nucleus:

(5)   livre

    Alfred
       d'
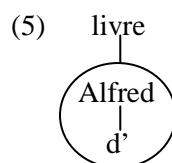
In other words, Tesnière would probably have preferred an analysis that views prepositions as dependents of their nouns, if forced to choose.

Concerning the other two relevant types of function words, i.e. subordinators and articles, Tesnière viewed subordinators as **translatives** that were essentially the same as prepositions in how they function. Translatives are part of Tesnière's theory of transfer, developed in the second half of his book. The purpose of a translative is to transfer a content word of one category to a content word of another category, e.g. a noun to an adjective, an adjective to an adverb, a verb to noun, etc. Thus he probably would have favored subordinating them to the verb with which they co-occur inside the nucleus. And concerning articles (definite and indefinite), he took a varied stance, viewing them as translatives when they perform a translative function, but as mere dependents of their noun when they do not perform a translative function.

3 Some early works

3.1 Mel'čuk (1958, 1963, 1974, etc.)

Unlike Tesnière, Igor Mel'čuk has been clear and consistent about the dependency status of function words in surface syntax. Mel'čuk and his collaborators have consistently subordinated content verbs to auxiliary verbs, nouns to adpositions, verbs to subordinators, and determiners to nouns in surface syntax in their prolific dependency-based works on syntax and grammar in the MTT (Meaning to Text Theory) framework and otherwise. Mel'čuk and his collaborators have been doing this since his earliest works, starting in 1958 (Mel'čuk 1958: 252–4, 1963: 492–3, 1974: 221–4). These early works are in Russian, but the approach is consistent with the prominent MTT works in English from the 1980s (e.g. Mel'čuk and Pertsov 1987, Mel'čuk 1988).

Judging by the dates of these early publications, it seems likely that Mel'čuk's works earn the honor of being the pathfinder in this area, since the majority of DGs that have come later have done the same, as will become increasingly clear below. Mel'čuk's position concerning the status of function words has been and is particularly firm.

3.2 Hays (1964)

Hays (1964) is considered a milestone in the development of dependency theory, in part because the article appeared in such a prominent journal, *Language*. Hays' article does not, however, say much about the dependency status of function words. The article concentrates instead on the formalization of dependency-based rewrite rules, as inspired directly by Chomsky (1957), and on the extent to which dependency and constituency formalisms are weakly or strongly equivalent. One can merely glean a sense of Hays' understanding of function words from the article.

In particular, Hays discusses the ambiguous sentence *They are flying planes*, and he captures the ambiguity with the following trees:

(6)

a. They are flying planes.

b. They are flying planes.
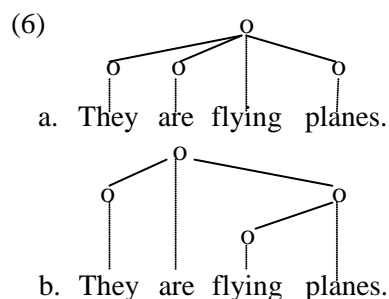
Tree (6a) reflects the meaning 'They are making the planes fly', whereas tree (6b) reflects the meaning 'They are planes that are flying'. Hays does not discuss the varying status of the auxiliary/copula *are* in these cases. But from the trees the reader can see that *are* is taken to be an auxiliary verb in tree (6a) and a copular verb in tree (6b).

Hays' interpretation of the auxiliary/copula *be* is representative of how it was widely viewed at the time. Chomsky (1957: 38f.) viewed auxiliaries as a separate class (Aux), meaning he did not classify them as verbs. Thus when a form of *be* appears with a main verb, it was deemed an auxiliary, but when it appeared in the absence of a main verb, it was deemed a copula. For Hays then, *are* in (6a) was an auxiliary, whereas *are* in (6b) was a copular verb.

From a modern perspective, the distinction Hays was building on cannot be maintained. Diagnostics for identifying auxiliary verbs reveal that the two putative types of *be* behave the same in important ways:

(7) a. Are they flying planes?
    b. They are not flying planes.
    c. They are flying planes, and they are, too.

The words *flying planes* show the same ambiguity in all three of these sentences, just as in Hays' example. Thus the putative distinction between auxiliary *be* and copular *be* lacks an empirical basis, since the two show the same syntactic behavior with respect to subject-auxiliary inversion (7a), sentence negation (7b), and VP-ellipsis (7c). The two are in fact the same type of verb; they are both auxiliary *be*.

3.3 Hudson (1976, 1984, 1990)

Richard Hudson's dependency-based framework Word Grammar (1984, 1990) has consistently taken auxiliary verbs as heads over content verbs, prepositions as heads over nouns, and subordinators as heads over verbs. And concerning determiners, Hudson has mostly preferred an analysis that positions determiners as heads over nouns (Hudson 1984: 90–2). Thus Hudson's approach concerning function words is consistent insofar as function words are heads over the content words with which they co-occur.

Concerning the latter issue, i.e. the status of determiners, the determiner phrase (DP) vs. noun phrase (NP) debate has been an ongoing dispute since the 1980s. Interestingly in this regard, Hudson's position is a minority one in the DG community in general, but it certainly finds much support among generative grammarians, the majority of whom presently pursue a DP analysis of nominal groups. This issue is not addressed here. The discussion focuses instead on auxiliary verbs.

By the time of Hudson's 1976 book, he had apparently become convinced that auxiliary verbs are heads over content verbs (p. 149–51), and in his 1984 book *Word Grammar*, Hudson writes in this regard:

> "It is now widely accepted that a main verb is syntactically subordinate to its auxiliary verb (Pullum and Wilson 1977 is particularly important collection of evidence), and I have accepted this analysis in all my dependency analyses." (p. 91)

The reference here to Pullum and Wilson (1977) is pointing to a significant debate that took place in the 1970s concerning the status of auxiliary verbs (Ross 1969, Chomsky 1972, Huddleston 1974, Pullum and Wilson 1977). The question concerned the extent to which auxiliaries should be viewed as verbs at all.

As mentioned above, Chomsky (1957) took auxiliaries to be a syntactic class that was to a large extent distinct from that of content verbs, labeling this class simply Aux (p. 39). This view of auxiliaries led to a ternary-branching analysis of basic clause structure in which an auxiliary is present, as in (8a):

(8)

a. They have succeeded.
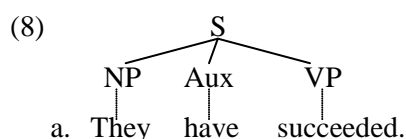
This basic constituency-based analysis is present in syntax textbooks from the 1970s (e.g. Bach 1974, Emonds 1976, Baker 1978).

Based in part on Pullum and Wilson's 1977 article, a stance took hold in the 1980s that viewed auxiliaries of every sort (aspect, voice, and modality) to be syntactically like full verbs. For example, they distinguish between present and past tense and agree with the subject. In European languages other than English, both full verbs and auxiliaries behave the same way, as already discussed in section 3.2. For this reason, auxiliaries should be granted the status of verbs in the hierarchy. This led to analyses like the following one, which shows the auxiliary as head over the content verb, e.g.

(8)

b. They have succeeded.
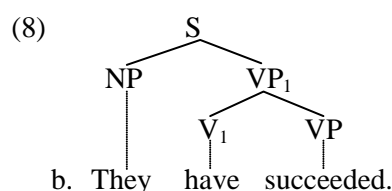
Analyses along these lines can be found in many textbooks of the era (e.g. Haegeman 1991, Napoli 1993, Ouhalla 1994), and despite the addition of numerous varied functional categories, the basic hierarchy of verbs shown with (8b) remains intact in many recent constituency grammars (e.g. Culicover 2009, Carnie 2013).

The interesting aspect of this trend in constituency grammars, i.e. the trend toward auxiliary verbs as heads over content verbs, is the fact that Hudson's dependency-based system (and Mel'čuk's) was ahead of its time. On a dependency-based analysis, there are just two basic possibilities for the hierarchical analysis of auxiliaries that must be considered: either the auxiliary is head over the content verb, or vice versa:

(9)

a. They have understood.

b. They have understood.
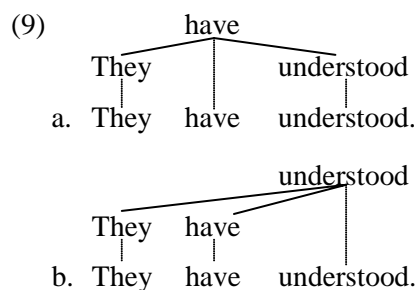
Apparently Hudson had already decided firmly in favor of the analysis in (9a) by 1976. In this regard, the trend in constituency grammars was lagging significantly behind the stance that Hudson and Mel'čuk had adopted early in the development of the Word Grammar and the Meaning-Text frameworks.

3.4 Matthews (1981)

Matthews (1981) discusses the distinction between dependency- and constituency-based syntax at length, and in this regard his book *Syntax* was a major contribution to our developing understanding of the distinction between dependency- and constituency-based systems. In the book, Matthews took content verbs to be heads over auxiliaries, prepositions to be heads over their nouns, and nouns to be heads over their determiners.

Two examples from the book illustrate Matthews' positions:

(10) a. No animal shall wear clothes. (p. 155)

b. leave the meat in the kitchen (p. 87)

The important observations here are that the modal auxiliary *shall* is dependent on the content verb *wear* in (10a), that the noun *kitchen* is dependent on the preposition *in* in (10b), and that the determiners *no*, *the*, and *the* are dependent on their nouns.

Concerning prepositions as heads over their nouns, Matthews did not motivate his position empirically, but rather appealed to traditional case government. He wrote "Grammarians also talk of prepositions having objects…, or having complements…" (p. 146). Every dependency hierarchy in Matthews' book containing a preposition shows the preposition as head over its noun.

Concerning the status of determiners, Matthews did not produce specific empirical evidence in favor of determiners as dependents of their nouns, but rather he appealed to the fact that they form a closed class. His assumption was that closed class categories are more appropriately viewed as dependents than as heads. The position Matthews was taking concerning determiners was not controversial at that time, so there would have been little reason for him to justify his decision in the area with further empirical observations.

Concerning the status of auxiliary verbs, however, Matthews had a bit more to say. He motivated their status as dependents of content verbs in two ways. The first was to point to their status as a closed class, and as a closed class, they were like determiners and thus should be viewed as dependents. He was drawing an analogy: just like determiners determine their head nouns, auxiliaries determine their head verbs. To him, this also meant that the hierarchical relationships between determiner and determined should be the same in both constructions.

The second motivation Matthews produced in favor of auxiliaries as dependents of their content verbs was subcategorization. He briefly discussed the example *has appeared* (p. 63). According to Matthews, *appeared* influences the syntactic category and semantic content of the noun with which it appears, whereas *has* lacks this ability. Matthews wrote:

"As a form of APPEAR it can take just a subject (*He has appeared*) but not both a subject and an object (**He has appeared the speech* or **He has appeared Cicero*). For other lexemes it can be the reverse: *He has distributed the speech* or *He has visited Cicero*, but not **He has distributed* or **He has visited*. A relation is thus established between *appeared*, or the morpheme APPEAR, and a subject element. But at that level the relation of *appeared* to *has*, or of the morpheme APPEAR to the discontinuous HAVE...past participle, is quite incidental." (p. 63)

To restate Matthews' point in other words, content verbs influence their linguistic environment in a way that auxiliary verbs do not, and for this reason, auxiliary verbs should be subordinated to content verbs.

A noteworthy aspect of Matthews' reasoning in this area concerns its lexico-semantic nature. Matthews overlooked the fact that from a purely syntactic point of view, it is the finite verb (i.e. the auxiliary verb) that licenses the appearance of the subject, not the nonfinite verb, e.g. *He has gone home*, **He gone home*, *He goes home*; *She has eaten a lot*, **She eaten a lot*, *She eats a lot*. From this point of view, there should be a direct dependency linking the subject to the finite verb.

3.5 Schubert(1987)/Maxwell & Schubert(1989)

Using the dependency relations of Schubert (1987), Maxwell and Schubert (1989) gathered annotation schemes from a number of authors for machine translation of a number of languages (Bangla/Bengali, Danish, English, Esperanto, Finnish, French, German, Hungarian, Japanese, Polish). The project, based in the Netherlands, was known as *Distributed Language Translation* (DLT, 1984–1990). DG was used to provide

syntactic representations of sentences in a source language, in Esperanto (the intermediate language), and in a target language. Grammars written for this project aimed to show the structural relationships to be derived by automatic parsing.

In the current context, how function words were dealt with in the various languages is the point of interest. In Schubert's analysis of Esperanto, auxiliary verbs are heads over content verbs, common nouns are heads over determiners, and subordinators are heads over verbs. Schubert (1987: 45) states that his understanding of DG was influenced by the Mannheim school of DG (Engel 1982).

These patterns were followed for all other grammars with two exceptions. In Danish, subordinators are positioned as dependents of the following finite verb rather than as the head of the subordinate clause. Here is Ingrid Schubert's (1989: 58) statement on this matter:

> "These clause introducers may under certain circumstances be omitted in Danish. I have not decided to let them be governed by a subordinating conjunction, but to consider the verb of the subordinate construction a direct dependent of the verb in the superordinate sentence."

Perhaps these cases are something like the omission of the complementizer *that* in English, which makes no contribution to meaning and accordingly can often be left out, as in *Say (that) it's true*. It is arguably the only subordinator which has this property. If so, it seems wrong to base the analysis on this one instance. The alternative of an empty node could be considered.

The other exception is in Lobin's (1989) analysis of German. The determiner rather than its noun is taken as the head of the nominal group. Lobin justifies his analysis in this area by pointing to cases like the following one:

(11) unsere Fahrt an die See und eure
     our   trip to the sea and yours

    in die Berge
    in the mountains

    'our trip to the sea and yours to the mountains'

The absence of *Fahrt* in the second part of the coordinate structure forces one to view *eure* as the head of the nominal group (or to posit an empty nominal node).

While the two exceptions just noted provide insight into the difficult choices that had to be made by the authors who participated in the project, the greater point is that there was a large measure of agreement concerning the status of most function words. With the exception of determiners, most function words were taken to be heads of the content words with which they co-occur.

4 The German schools

DG has enjoyed particular favor in the German speaking world. German grammarians recognized the potential of dependency-based syntax early on. This early recognition may have been due to the particular compatibility of dependency-based syntax, which emphasizes verb centrality, with the verb second (V2) principle of German (and other Germanic languages). The finite verb is anchored in second position in German declarative clauses, thematic material tending to precede this position and rhematic material tending to follow it.

The interesting and noteworthy point about the German schools is the unanimity that one encounters among the leading voices. DG grammarians (Kunze 1975, Lobin 1993, Engel 1982, 1994, among others) are mostly unanimous in the basic hierarchical analyses of function words that they assume: auxiliary verbs are heads over content verbs; prepositions are heads over nouns; and subordinators are heads over verbs. The only area where one encounters some variation among these experts concerns determiners. The majority stance is certainly that nouns are heads over determiners, but Lobin (1993) takes the opposite stance, as he does in Lobin (1989), already mentioned in section 3.5, and Eroms (2000) has argued for interdependence between article (definite or indefinite) and noun.

Due to the large measure of agreement concerning the hierarchical status of most function words, the German-language DG world can be viewed as speaking with a single voice, and this voice is particularly loud by virtue of the fact DG enjoys a prominence at schools and universities that is not generally encountered outside of the German-speaking world. A point of interest, perhaps, is the reasoning that one finds in the German-language DG literature about the sentence root. In a two-verb combination such as *hat gelegt* 'has layed', the German schools of DG unanimously view *hat* as head over *gelegt*. It is worth considering briefly why they do so.

Engel (1994:107–109) points to facts about subcategorization. He sees the lexical stem *hab* of the auxiliary *hat* determining the form of the nonfinite verb *gelegt* as a participle. This reasoning does not work in the opposite direction, that is, one cannot view the lexical stem *leg-* as determining the syntactic form of the auxiliary *hat*. The insight can then be extended to all manner of auxiliary verbs. For instance in a combination such as *wird wollen* 'will want', the modal auxiliary *wird* subcategorizes for the infinitive form of the stem *woll-*, but not vice versa, that is, the lexical stem *woll-* does not determine the syntactic subcategory of the auxiliary *wird*.

A related issue concerns the motivation for positioning the finite verb as the root of the clause. Eroms (2000: 129ff.) motivates the hierarchical status of auxiliary verbs in another way. He appeals to the fact that when an auxiliary verb and full verb co-occur, it is the auxiliary that is finite, not the full verb. The auxiliary verb then bears the functional information of person, number, tense, and mood. The nonfinite verb typically does not express this information. Thus in the example from the previous paragraph, i.e. *hat gelegt* 'has layed', the finite auxiliary verb *hat* expresses number (singular), person (3rd person), tense (present), and mood (indicative). The participle form *gelegt*, in contrast, can be construed as helping to convey perfect aspect only. This functional load that the auxiliary verb bears motivates its status as the root of the clause, i.e. as head over the content verb.

5 The Prague school

The Prague school of DG, as associated with the Prague Dependency Treebank (PDT), agrees with most of the other DG mentioned in this contribution concerning the hierarchical status of adpositions; they are heads over their nouns. However, the annotation scheme for the PDT (Hajič 1998) began subordinating auxiliary verbs to content verbs in 1996. This aspect of the PDT remains anchored today in the analytical level (surface syntax) of the PDT. Due to the prominence of the PDT in the development of DG theory in general, the linguistic motivation for its choice to subordinate auxiliary verbs to content verbs in surface syntax is worth considering, however briefly.

There has been a difficulty in this area, though. Attempts to locate the linguistic motivation behind this aspect of the PDT annotation scheme have not turned up anything concrete in published works. For this reason, the two linguistic observations produced next are based on personal communication with Jarmila Panevová, one of the founding members of the Prague school of DG.

Worth noting first, though, is that the PDT annotation scheme subordinates only non-modal auxiliary verbs to content verbs. Modal verbs, in contrast, are heads over their content verbs. What this means is that the PDT annotation scheme for surface syntax deviates from the majority position only regarding a single auxiliary verb, namely *být* 'be' (in all its forms).

There are two linguistic motivations for subordinating the forms of *být* to the content verb. One of these concerns a general aspect of the verb 'be' in Slavic languages in general. Many Slavic languages lack or omit the finite form of this verb in certain environments. Czech omits a form of this verb in the 3rd person of the compound past, but a 1st and 2nd person form of this verb appear in such environments:

(12) a. Já jsemspal.   (masculine)
        I   am.slept
        'I slept.'

    b. Já jsemspala.   (feminine)
        I   am.slept
        'I slept.'

    c. John   spal.
        'John slept.'

If the clitic auxiliary *jsem-* is subordinated to *spal/spala*, then there is a direct dependency that connects the subject to *spal* in each of these three cases. But if *spal/spala* is subordinated to *jsem-*, then an asymmetry appears: the subject is at times (in the 1st and 2nd person) an immediate dependent of the auxiliary verb, and at other times (in the 3rd person), it is an immediate dependent of the participle. One thing that is different about this construction from others discussed earlier is that here the auxiliary and main verb form a single word. The question therefore concerns the extent to which this asymmetry should influence choices about the syntactic hierarchy.

A possible drawback of this choice is that it forces the PDT to draw a distinction between auxiliary *být* and copular *být*, since when a form of *být* is the only verb present, e.g. *Mary je velká* 'Mary is tall', the PDT positions that verb as head over the predicative expression (here *je* 'is' is head over*velká* 'tall'). Examples (7a–c) above illustrate that there is no syntactic motivation for

distinguishing between an auxiliary *be* and a copular *be* in English.

Another linguistic argument for subordinating forms of the auxiliary *být* to its content verb is seen when multiple forms of *být* appear in one and the same clause:

(13) John by   byl   býval   spal.
      John be   been been   slept
      'John would have slept.'

This sentence is an example of the past conditional. Three distinct forms of *být* appear together. By subordinating all three of them directly to *spal*, a relatively flat structure obtains, and a flat structure has the advantage of avoiding projectivity violations (cf. Mel'cuk and Pertsov 1987:181 for a discussion of such a violation in terms of the arcs used in some forms of DG), which would occur, assuming that certain other sequences of these verb forms are possible.

Gruet-Skrabalova (2012) offers a different account of the auxiliary *být* in the course of a Minimalist analysis of some kinds of ellipsis. She shows that there is a certain degree of freedom in the word order, but the only alternative order shown by her does not produce projectivity violations. In (13), *by* is the 3rd person (singular or plural) form of the conditional mood and the only one of the three which is finite. The other two are both past participles, the first in the perfect aspect and the second in the imperfect aspect. The first of them is subcategorized for by the preceding conditional verb. Gruet-Skrabovala does not discuss any sentence in which the two participles co-occur. However, sentences in this article suggest that both past tense forms of this verb subcategorize for a participial complement. From this, we judge that the first participle in (13) subcategorizes for the second, which in turn subcategorizes for the participial form of the main verb. Gruet-Skrabovala states that the final participle of the auxiliary must directly precede the main verb, although the finite conditional form can follow the main verb, just like in subordinate clauses in German.

The PDT decision to subordinate forms of *být* to the co-occurring content verb constitutes a narrow exception to the majority position concerning function words. The PDT is otherwise consistent with the majority position regarding the status of modal verbs, copular verbs, adpositions, subordinators, and determiners.

6   Concluding discussion

This survey has revealed that there is little support in the sources examined above for the Universal Stanford Dependencies' (USD) decision to categorically subordinate function words to content words. Not one of the sources surveyed clearly supports the USD analysis of adpositions, and only two of the sources provide support for the USD decision to subordinate auxiliaries to content words. The dominant position, which has crystallized through the decades, is that auxiliaries are heads over content verbs and prepositions are heads over their nouns. And concerning determiners, they are more widely viewed as dependents of their nouns – although their status has been the focus of more debate.

The survey has turned up three published arguments in support of the USD position and two unpublished arguments that partially support the USD position. The published ones are Matthews' argument concerning English auxiliaries in section 3.5, and the arguments concerning Danish subordinators and German determiners in section 3.6. The unpublished ones concerned the Czech auxiliary *být* in section 5.

De Marneffe et al. do give a few indications of the supposed linguistic superiority of USD. The choice of having nouns as heads over adpositions allows parallelism between prepositional phrases and morphological case-marking (p. 4585) and also between adpositions and adverbial clauses (p. 4587). However, it ignores the fact that adpositions assign case to their complement nouns, not vice versa. Hence what one achieves in the way of more parallelism across the structures of distinct languages, one pays for with the unorthodox stipulation that adpositions assign case up the syntactic hierarchy to their nouns.

The choice of making predicates heads over auxiliaries allows parallelism between constructions which in some languages omit the copula and those which do not (p. 4586). This is true, but alternative solutions such as an empty node should be considered. Also, if there are several linked auxiliaries, as in *might have been dreaming*, they must all have *dreaming* as their head, so the subcategorization relationship between any two consecutive auxiliaries cannot be shown by the dependency links (cf. the discussion of subcategorization by Engel in section 4).

De Marneffe et al. note that the choices discussed in their article have a negative effect on parsing:

"It is now fairly well-known that … dependency parsing numbers are higher if you make auxiliary verbs heads … and if you make prepositions the head of prepositional phrases… Under the proposed USD, SD would be making the 'wrong' choice in each case." (p. 4589)

Parsing accuracy is not the most important criteria, however, as the following statements concerning the importance of linguistic quality and downstream applications document:

"…it seems wrongheaded to choose a linguistic representation to maximize parser performance rather than based on the linguistic quality of the representation and its usefulness for applications that build further processing on top of it." (p. 4589)

For this reason, de Marneffe et al. propose transforming the USD system to provide two other results, one for **parsing** and one called **enhanced**.

Thus by de Marneffe et al.'s own admission, parsing accuracy tends to be higher if function words are heads over content words, and given the analysis and discussion above, the DG tradition agrees to a large extent that linguistic considerations support most function words as heads over the content words with which they co-occur, contrary to USD's stance.

If USD wants to claim that linguistic considerations support its unorthodox approach to surface dependencies, it of course has every right to do so in the clash of ideas. But the point we hope to have established in this contribution is that the DG tradition does not support this claim. Quite to the contrary, the DG tradition has crystallized over the decades to a position that contradicts the USD approach. Thus if USD wants to maintain its claim to "linguistic quality", the burden of proof rests firmly on its shoulders; it needs to produce the linguistic reasoning that supports its position in part by discussing and refuting the observations and reasoning that have coalesced over the decades to the opposite position.

References

Emmon Bach. 1974. *Syntactic Theory*. Holt, Rinehart and Winston, Inc., New York.

Carl Lee Baker. 1978. *Introduction to Generative Transformational Syntax*. Prentice Hall, Englewood Cliffs, NJ.

Andrew Carnie. 2013. *Syntax: A Generative Introduction*. Wiley-Blackwell, Malden, MA.

Noam Chomsky. 1957. *Syntactic Structures*. Mouton, The Hague.

Noam Chomsky. 1972. Some emperical issues in the theory of Transformational Grammar. In S. Peters (ed.), *Goals of Linguistic Thoery*, 63–130. Prentice Hall, Englewood Cliffs, NJ.

Peter Culicover. 2009. *Natural Language Syntax*. Oxford University Press, Oxford UK.

Joseph Emonds. 1976. *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations*. Academic Press, New York.

Ulrich Engel. 1982. *Syntax der deutschen Gegenwarts- sprache*, 2nd edition. Erich Schmidt, Berlin.

Ulrich Engel. 1994. *Syntax der deutschen Gegenwarts-sprache*, 3rd fully revised edition. Erich Schmidt, Berlin.

Hans-Werner Eroms. 2000. *Syntax der deutschen Sprache*. Walter de Gruyter, Berlin.

Hana Gruet-Skrabalova. 2012. VP-ellipsis and the Czech auxiliary *být* ('to be'). *Xlinguae*, 2012, 5, 3–15.

Liliane Haegeman. 1991. *Introduction to Government and Binding Theory*. Blackwell, Oxford, UK.

Jan Hajič. 1998. Building a syntactically annotated corpus: The Prague Dependency Treebank. In Eva Hajičcová (ed.), *Issues of Valency and Meaning: Studies in Honor of Jarmila Panevová.*106–32. Karolinum, Charles University Press, Prague.

David Hays. 1964. Dependency theory: A formalism and some observations. *Language* 40, 511–25.

Rodney Huddleston. 1974. Further remarks on the analysis of auxiliaries as main verbs. *Foundations of Language* 11, 215–29.

Richard Hudson. 1976. *Arguments for a non-Transformational Gramma*r. University of Chicago Press, Chicago.

Richard Hudson. 1984. *Word Grammar*. Basil Blackwell, New York.

Richard Hudson. 1990. *An English Word Grammar*. Basil Blackwell, Oxford.

Ray Jackendoff. 1977. *X-bar Syntax: A Study of Phrase Structure*. The MIT Press, Cambridge, MA.

Jürgen Kunze.1975.*Abhängigkeitsgrammatik. Studia Grammatica* 12. Akademie Verlag, Berlin.

Henning Lobin. 1989. *A dependency syntax of German*. 17–39. In Maxwell and Schubert.

Henning Lobin. 1993. *Koordinationssyntaxalsproze-durales Phänomen. Studien zur deutschen Grammatik* 46. Gunter NarrVerlag, Tübingen.

Marie-Catherine de Marneffe, Timothy Dozat, Natalia Silveira, Katri Havarinen, Filip Ginter, Joakim Nivre, Christopher Manning. 2014. Universal Stanford Dependencies: A crosslinguistic typology. LREC, 4585–92

Peter Matthews. 1981. *Syntax*. Cambridge University Press, Cambridge, UK.

Dan Maxwell and Klaus Schubert (eds.). 1989. *Metataxis in Practice: Dependency Syntax for Multilingual Machine Translation*. Foris Publications, Dordrecht.

Igor Melʹčuk. 1958. Mašinnomperevode s vengersko-gojazykanarusskij — On the automated translation of Hungarian to Russian, *Problemykibernetiki*, v. 1, 222–64.

Igor Melʹčuk. 1963. Avtomatičeskijanaliztekstov (namaterialerusskogojazyka) — Automated analysis of texts (on the basis of Russian). In *Slavjanskoejazykoznanie — Linguistique slave*. Academy of Sciences of the U.S.S.R., Moscow, 477–509.

Igor Melʹčuk. 1974. *Opytteoriilingvisti
českixmodelej "Smysl - Tekst". Semantika, Sintaksis*. — Sketch of a linguistic model of the sort Meaning – Text. Semantics, Syntax. Nauka, Moscow.

Igor Mel'čuk and Nikolaj Pertsov. 1987. *Surface Syntax of English: A Formal Model within the Meaning-Text Framework*. John Benjamins, Amsterdam.

Igor Meľčuk. 1988. *Dependency Syntax: Theory and Practice*. State University of New York Press, Albany.

Donna Jo Napoli. 1993. *Syntax: Theory and Problems*. Oxford University Press, New York.

Jamal Ouhalla. 1994. *Transformational Grammar: From rules to Principles and Parameters*. Edward Arnold, London.

Geoffrey Pullum and Deirdre Wilson. 1977. Autonomous syntax and the analysis of auxiliaries. *Language* 53, 4, 741–88.

John Robert Ross. 1969. Auxiliaries as main verbs. *Studies in Philosophical Linguistics* 1, edited by W. Todd, 77–102. Great Expectations Press, Evanston.

Ingrid Schubert. 1989. *A dependency syntax of Danish*. 39-68, in Maxwell and Schubert.

Klaus Schubert. 1987. *Metataxis: Contrastive dependency syntax for machine translation*. Foris Publications, Dordrecht.

Klaus Schubert. 1989. *A dependency syntax of Esperanto*. 207–32, in Maxwell and Schubert.

Lucien Tesnière. 1959. *Éléments de syntaxe structural*. Klincksieck, Paris.

Lucien Tesnière. 2015 (1959). *Elements of Structural Syntax*. Translated by Timothy Osborne and Sylvain Kahane. John Benjamins, Amsterdam.

Diagnostics for Constituents: Dependency, Constituency, and the Status of Function Words

Timothy Osborne
Zhejiang University
Hangzhou
China
`tjo3ya@yahoo.com`

Abstract

This contribution delivers two messages: 1) that the tests for constituents that are widely employed in linguistics and syntax textbooks are more congruent with dependency-based syntax than with constituency-based syntax and 2) that these same tests support the conventional analysis of function words, that is, the analysis that takes most function words (auxiliary verbs, adpositions, subordinators) to be heads over the content words with which they cooccur. The latter issue is important at present, since a recent annotation scheme is choosing to subordinate all function words to the content words with which they coocur.

1 Two messages

Most English language textbooks on syntax and linguistics rely on tests for constituents to introduce the concept of syntactic structure. Tests such as coordination, proform substitution, topicalization, answer fragments, clefting, VP-ellipsis, pseudoclefting, etc. are used to demonstrate the presence of constituents, and thus, the presence of sentence structure. The tests show that words are being grouped together into phrases, and smaller phrases are grouped into ever larger phrases, until the largest phrase, the sentence, is reached. The tests are very widely employed, so widely that they enjoy a prominent spot in the syntactician's toolbox; they are basic tools with which the syntactician works.

An interesting aspect of most tests for constituents, however, is that they identify much less syntactic structure than most constituency grammars assume. In this respect the data delivered by the tests are relatively congruent with dependency grammars (DGs), since by its very nature dependency-based syntax posits much less syntactic structure than constituency-based syntax. Interestingly, the DGs currently in existence rarely draw attention to this fact, that is, they rarely draw attention to the fact that the dependency-based understanding of syntactic structures is strongly supported by the basic tests that are, ironically, so widely employed by constituency grammars.

Tests for constituents can also be employed to shed light on debates about the best hierarchical analysis of various syntactic structures, for instance concerning the hierarchical status of function words. The tests are consistent with the traditional DG analysis of function words, namely that auxiliary verbs are heads over content verbs and prepositions are heads over their nouns.

This contribution draws attention to the two points just mentioned. It delivers two messages: 1) most commonly used tests for constituents are much more consistent with dependency-based syntax than with constituency-based syntax and DGs can and should draw attention to this fact, and 2) the tests reveal that auxiliary verbs are heads over content verbs and prepositions are heads over their nouns.

The data examined in this contribution are limited to English. This is due mainly to the fact that the most widely employed tests for constituents are employed in English language textbooks, applied to the syntactic structures of English. Section 6 below reflects on this aspect of the tests, considering the extent to which they can be employed in other languages.

2 Constituents

The term *constituent* is associated with *constituency* grammars, the morphological relatedness of the two words, *constituent* and *constituency*, being obvious. In this respect the first message delivered in this manuscript might seem contrary to basic terminology, this terminology suggesting that dependency and constituency are distinct principles of syntactic organization and that the constituent unit is not compatible with depen-