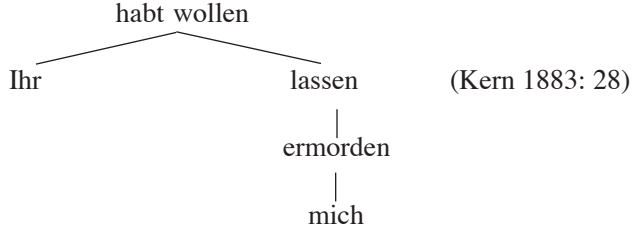
The Status of Function Words in Dependency Grammar

Timothy Osborne
Zheijiang University

1. Introduction

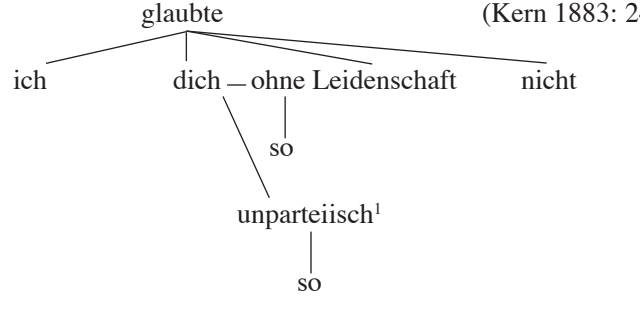
Since some of the earliest works on syntax and grammar that can be definitively acknowledged as at least partially dependency-based (e.g., Clark 1847, Reed and Kellogg 1877, Kern 1883—see Imrényi and Mazziotta's 2020 volume of essays on the history of DG), the hierarchical status of many function words has been unclear. Franz Kern (1883), for instance, produced the following structural analyses of the German sentences *Ihr habt mich ermorden lassen wolle*n 'You wanted to have me murdered' and *So ohne Leidenschaft, so unparteiisch glaubt' ich dich nicht* 'So without passion, so impartial I did not believe you to be':

(1)



(Kern 1883: 28)

(2)



(Kern 1883: 24)

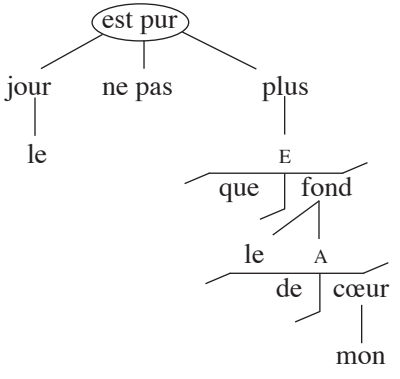
[1] The two edges reaching from *dich* to *so ohne Leidenschaft* and *so unparteiisch* indicate, presumably, that the latter two are coordinated predicative expressions, predicated over the former. Kern does not comment directly on this aspect of the tree he produces, though.

Linguistic Analysis, 43 3-4
2024 Linguistic Analysis
8832 SW Dilworth Rd, Vashon, WA 98070

Kern did not establish a hierarchical relationship between the auxiliary verb *habt* 'have' and modal verb *wollen* 'want' in (1), but rather he positioned the two together as the root node of the structure. Similarly, he did not establish a hierarchical relationship between the preposition *ohne* 'without' and the noun *Leidenschaft* 'passion' in (2), but rather he again positioned the two together in a single node.

Jumping forward 76 years to Tesnière's seminal work (1959), he too did not establish a clear hierarchical relationship between auxiliary and content verb or between preposition and noun, as can be seen from his structural analysis of Racine's sentence *Le jour n'est pas plus pur que le fond de mon coeur* 'The day is no more pure than the bottom of my heart':

(3) (Tesnière 1969: 636)



The tree shows several aspects of Tesnière's theory: the tree is a *stemma*; the bubble encloses the words of a *nucleus*; the horizontal edges with angled ends indicate the presence of *transfer* (cf. Mazziotta, this volume, pages 58, whereby the function words *que* and *de* transfer the content words *fond* and *cœur* from nouns to an adverb (E) and an adjective (A), respectively. In the current context of the analysis of function words, what is important in the stemma is its view of *est* 'is', *que* 'than', and *de* 'of'; these function words are placed equi-level with the content word with which they co-occur. Notice, however, that Tesnière did subordinate the definite article *le* 'the' to its noun each time.

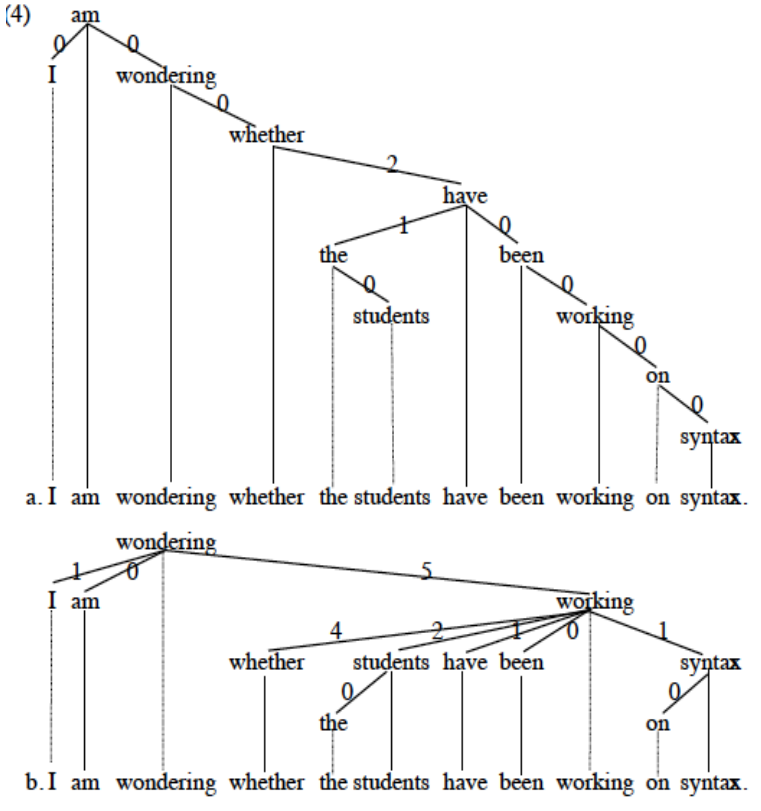
Examining the more modern history of DG, there have also been diverging views about the hierarchical status of many function words. Matthews (1981: 155-6), for instance, positioned auxiliary verbs as dependents of content verbs, whereas many grammarians at the time

assumed the opposite, placing auxiliary verbs as *heads* over content verbs (cf. Ross 1969; Hudson 1976: 150; Pullum and Wilson 1977).[2] Similarly, most DGs position determiners as dependents of their nouns, whereas Word Grammar started placing them as heads in the 1980s (cf. Hudson 1984: 90-2). Interestingly, however, by the early 2000s a consensus had mostly been reached concerning the status of auxiliary verbs and adpositions (cf. Osborne and Maxwell 2015), both of which were being consistently analyzed as heads over their co-occurring content words. Auxiliary verbs were positioned as heads over content verbs and other predicative expressions, prepositions were seen as the roots of their phrases.

Certainly from a linguistic point of view, the hierarchical status of function words influences how dependency-based approaches to the syntax of natural languages understand sentence structures. Their status can, for instance, have a major impact on dependency distance (cf. Hudson 2000, 2003; Temperley 2007; Liu 2008) and on dependency orientation, i.e., head-initial or head-final (cf. Liu 2010). The next two trees demonstrate how far reaching the choices in these areas can be. Tree (4a) consistently takes function words to be heads over content words, and the opposite is true of tree (4b). The numbers on the edges in these trees give dependency distance, that is, the distance measured in intervening words between the dependent word and its head word:[3]

[2] I use the term *head* where others may use the term *governor*. There are important issues at stake concerning the use of terminology, but to delve into such matters here would be a diversion from the core concern, which is the hierarchical status of function words.

[3] Note that there are two variations in the literature on how dependency distance is calculated. The method used here follows Hudson's (1995) original proposal in this regard: one counts the number of words that intervene between a dependent word and its head word. The other method gives each word a numerical linear index value and then subtracts the index of the dependent word from that of the head word, taking the absolute value of the result. The two methods are statistically equivalent. The advantage that the method employed here has is that the discrepancies in values are greater percentage-wise, thus allowing for more vivid comparisons of competing hierarchical analyses.



The (a)-tree is much taller than the (b)-tree, 8 levels vs. 4 levels. The taller tree has dependents appearing closer to their heads than the flatter tree. The mean dependency distance (i.e., the average distance between head and dependent) in tree (4a) is .30 (=[0+0+0+2+1+0+0+0+0+0]/10), whereas in (4b) it is 1.40 (=[1+0+5+4+2+1+0+1+0+0]/10).

Concerning the orientation of dependencies (head-initial or head-final), tree (4a) contains eight head-initial dependencies and just two head-final dependencies. In contrast, these numbers are reversed in tree (4b), there being just two head-initial dependencies there and eight head-final dependencies. It should hence be apparent that the hierarchical status of function words can have a major impact on the typological analysis of languages in terms of head-position. Tree (4a)

suggests, namely, that English as a language is more head-initial than head-final, whereas tree (4b) suggests the opposite.

Furthermore, the constituent structure of (4a) is quite distinct from that of (4b). We see, for instance, that *working on syntax, been working on syntax*, and *wondering whether the students have been working on syntax* are complete subtrees in (4a) but not in (4b).[4] Assuming that complete subtrees identify phrases, sentence (4b) indeed presents a much different take on "phrase structure" than tree (4a). In addition, the understanding of subject-verb agreement is significantly impacted. Tree (4a) has the subjects *I* and *the students* as immediate dependents of the finite verbs *am* and *have*, respectively, whereas there are no direct dependencies linking these words in (4b).

2. Functional Elements

Distinguishing function words (≈ empty words, grammatical words) from content words (≈ full lexical words, full words) is of course difficult at times and a debated matter. The next dictionary definition of "function" can serve as orientation:

> The job done by a linguistic element. The term is most often applied to the effect of a grammatical morpheme or a grammatical word such as the connecting word *of* and the *-ing* which goes on to verbs in English. These things chiefly provide the grammatical "glue" which gives structure to sentences, but they do not have meanings in the sense that lexical morphemes like *green* and *woman* have meanings.
>
> (*A Student's Dictionary of Language and Linguistics*, Trask 1997)

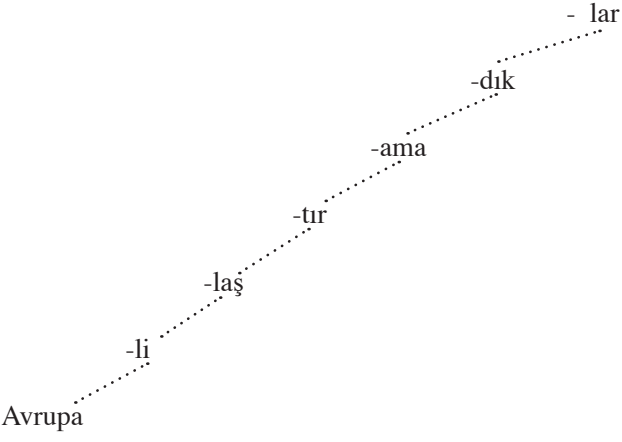
There are a couple of noteworthy points here. The first is the metaphor: function words and functional elements in general serve as "glue"; they are the grammatical mortar between the bricks of the sentence structure, the content elements being the bricks. This sort of metaphor characterizing function words has existed since the earliest works on grammar—see Mazziotta's contributions in this volume,

[4] The term *complete subtree* denotes a subtree that consists of all the words that the root word in the subtree dominates. The notion is closely similar to the constituent unit of phrase structure grammars, that is, constituents are complete subtrees and complete subtrees are constituents.

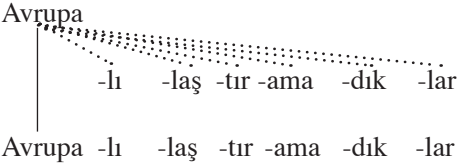
pages 359-394. The second is that functional elements can be words or parts of words (i.e., morphs). This fact is hence relevant to dependency-based approaches to word structure, and it is particularly important for the dependency analyses of the complex words in agglutinative languages such as Japanese, Finnish, and Turkish.

Examine the next dependency structures of the complex word *Avrupalılaştıramadıklar* 'unable to be Europeanised ones' from Turkish. One might assume one of the following two dependency analyses of this word (ADJ = adjective, IVRB = intransitive verb, TVRB = transitive verb, FVRB = finite verb, NEG = negation, PL = plural):[5]

(5)



(a) Avrupa -lı -las -tır -ama -dık -lar
Eruope -ADJ -IVRB -TVRB -NEG -FVRB -PL
'unable to be Europeanised ones'



(b) Avrupa -lı -laş -tır -ama -dık -lar

[5] The example is taken from the Wikipedia article on Turkish grammar and is actually abbreviated. The full word given in the article is *Avrupalılaştıramadıkları mızdanmışsınızcasına* 'as if you were reportedly of those whom we couldn't manage to Europeanise'.

The hyphens, the dotted dependency edges, and the presence of just one vertical projection line each time are the conventions Groß (2014) uses to indicate that the dependency analysis exists within the word (as opposed to between words). The analysis in (5a) positions the six function morphs, the suffixes, as heads. The opposite situation is true of (5b) where the six function morphs are all positioned as immediate dependents of the one content morph *Avrupa* 'Europe'. These two structures are of course drastically distinct, the one being tall consisting of seven levels, the other being quite flat having just two levels. If one calculates dependency distance for these structures, the mean dependency distance for (5a) is 0 (=[0+0+0+0+0+0]/6), whereas for structure (5b) it is 2.5 (=[0+1+2+3+4+5]/6).

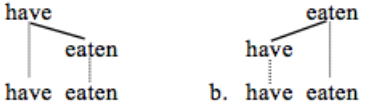
In practical terms, the debate concerning the hierarchical status of function words revolves primarily around the analysis of auxiliary verbs, adpositions, subordinators, determiners, and coordinators. These word categories occur often in many languages, and there is broad agreement that they qualify as function words. An obvious observation in this area is that these categories are closed classes, whereas typical content words belong to open classes (verbs, nouns, adjectives, and adverbs).

3. The Debates

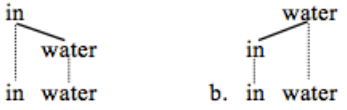
The primary point of contention concerning the hierarchical positions of function words in dependency grammar (DG) circles is easily summarized by considering the status of the five categories just mentioned, that is, of auxiliary verbs, adpositions (prepositions, postpositions), subordinating conjunctions (called *subordinators* here), determiners, and coordinating conjunctions (called *coordinators* here). There are certainly many additional types of function words and functional elements, such as particles and pleonastic elements, but given their high frequency of appearance, the primary debate revolves around the five word categories just mentioned.

The competing hierarchical analyses of function words are illustrated with the next tree pairs of English word combinations:

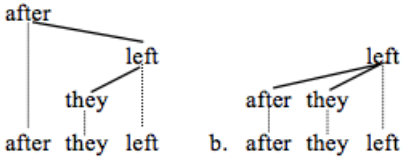
(6)



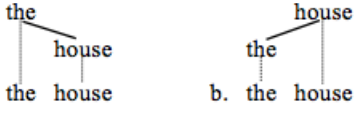
(7)



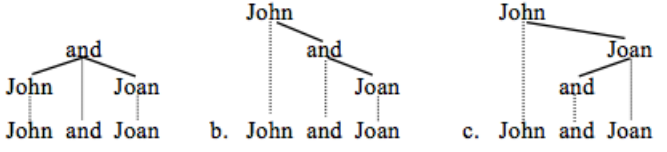
(8)



(9)



(10)



The function word is head over the content word in the (a)-trees but dependent underneath it in the (b)-trees. When one looks to the more modern literature in this area, one finds that each of these analyses is represented, although tendencies in the one direction or the other can be acknowledged. The hierarchical account of coordinators is particularly fraught with difficult decisions, in part because there are more than just two potential analyses, as illustrated in (10a-c).

A look back at developments since Tesnière (1959) reveals that a consensus of a sort concerning the analysis of auxiliary verbs, adpositions, and subordinators had been reached by about the turn of the millennium. Most DGs were positioning these words as heads over the co-occurring content words as in the (a)-trees in (6-10). Concerning determiners, though, most DGs were positioning and still do position them as dependents of their nouns as in (9b), the one major exception being Word Grammar, which does the opposite as mentioned above. Concerning the status of coordinators, there has been and continues to be much variation in how they are positioned, as suggested by trees (10a-c).

With the appearance of the Stanford Parser in the early 2000s, the situation began to change. A prominent alternative annotation scheme had emerged. The Stanford Parser originally positioned auxiliary verbs as dependents of content verbs, but it placed adpositions over their nouns. Further developments in this tradition switched the analysis, though. The grammatical framework now known as *Universal Dependencies* (UD) pursues a consistent analysis of function words: they are subordinated to the co-occurring content words. Hence the UD annotation scheme assumes the hierarchical analyses shown in the (b)-trees in (6-10), and in so doing, its understanding of dependency structures is contrary to much of the DG tradition that reaches back to the late 1950s when Igor Mel'čuk's works started to appear.

The UD annotation scheme is influential, there being hundreds of linguists and NLP practitioners who have chosen to follow its annotation guidelines. The UD webpage currently provides access to the treebank corpora of well over 100 languages. At the same time, an alternative annotation scheme has arisen in just the past few years, *Surface-syntactic Universal Dependencies* (SUD). This annotation scheme assumes the more traditional analysis of function words, the one associated with Mel'čuk's MTT framework—see Section 4.3 below. The corpora available at the UD website have all been converted to the SUD format, so one can now access parallel corpora for a given language, the one annotated according to the UD scheme along the lines of the (b)-trees in (6-9) and the other annotated according to the SUD scheme along the lines of the (a)-trees in (6-9). UD and SUD agree concerning their analyses of coordinate conjunctions, opting for (10c) in this regard. Both of these annotation schemes are represented below—see Sections 4.6 and 4.7.

4. The Contributions

The eight contributions in this volume cover a variety of issues concerning the status of function words and morphs in dependency-based syntax and grammar. The next sections summarize some of the main ideas presented in each of the eight contributions.

4.1. History

Mazziotta's contribution (this volume, p. 359) examines the status of function words in grammatical treatises from the European tradition, starting with the works of Greek grammarians from antiquity (4th century BC) and progressing through to Tesnière (1959). To limit the scope of his investigation, Mazziotta concentrates just on the grammatical status of prepositions and conjunctions, having much less to say about the other categories of function words. Since his perspective is of course a DG one, he begins with some orientation about how one identifies dependency-based structural analyses; he posits four attributes as definitional: *word-to-node mapping, connection, binarity,* and *headedness*. Mazziotta also gives a word of caution concerning the interpretation of historical material with respect to modern grammatical debates; he states that "it would be anachronistic to extrapolate genuine dependencies from the semantic and grammatical relations highlighted in ancient grammars."

Concerning the tradition itself, Mazziotta provides the next chronology as an outline of the ideas and grammarians he considers:

Greek philosophy and grammar	4th c. BC to 2nd c. AD
Priscian	fl. 500AD
Modists	12th - 13th c.
Humanists	15th - 16th c.
General grammar and Encyclopedia	17th - 18th c.
School grammar and diagrams	19th c
Tesnière	20th c.

Some prominent grammarians whose ideas Mazziotta considers are listed next: Aristotle, Aristarchus of Samothrace (217–144 BC), Dionysios Thrax (c. 170–c. 90 BC), Apollonius Dyscolus (c. 110–175 AD), Varro (c. 110–175 AD), Quintilian's (35–96 AD), Priscian (fl.

500 AD), Peter Helias's (c. 1110–a. 1166), Peter Abelard's (1079–1142 Despauterius (c. 1480–1520), Sanctius (1523-1601), Linacre (c. 1460–1524), Wallis (1616-1703), Port-Royal (1660), Buffier (1661–1737), Nicolas Beauzeé (1717-1789), Stephen W. Clark (1810-1901), Franz Kern (1830–1894). Mazziotta ends his survey with a summary of Tesnière's (1959) understanding of function words.

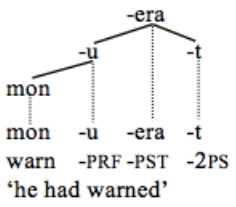
The main takeaway from Mazziotta's survey is, I believe, as follows. The status of adpositions as heads over their co-occurring nouns or as dependents underneath them has been debated throughout the millennia, and it is difficult to draw conclusions one way or the other from the historical record. Concerning conjunctions, the main difficulties are a lack of clarity about the actual conjuncts in the case of coordination and the very late recognition of the special status of complementizers and subordinating conjunctions in general.

4.2. Dependency Construction Grammar (DCxG)

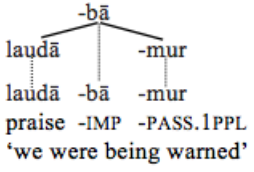
Maxwell's contribution (this volume, p. 395) explores the dependency structure of words using his framework of syntax and grammar, called *Dependency Construction Grammar* (DCxG) (Maxwell 2018). In order to illustrate the machinery of DCxG, Maxwell provides morphological analyses of Latin verbs. Maxwell begins with some discussion of the varying approaches to morphology given in Hockett (1954), mentioning Hockett's distinctions between Item and Arrangement (IA), Item and Processing (IP), and Word and Paradigm (WP). Maxwell states that these approaches are "linear models of phonology, which means there are no hierarchical relationships between the segments inside words". Maxwell also states that most traditional approaches to morphology do, though, acknowledge these hierarchical relationships, and certainly the account of Latin verbs that he provides clearly does so.

The first part of Maxwell's contribution is primarily concerned with the introduction DCxG and acknowledges a couple of challenges that the approaches face that acknowledge hierarchical structure within words. Maxwell then turns to the Latin system of verbal morphology, providing much descriptive guidance, this guidance being necessary given the complexity of the system. Latin is a synthetic language in which the morphological complexity of the verbs is particularly pronounced. The basic DCxG structural analyses of two Latin verbs are given next :

(11)



(12)



These structural analyses overlook much of the machinery of DCxG system and focus instead just on the hierarchy of segments that DCxG recognizes. The morphs that express aspect and tense are heads over the content morph, whereas the person, number and passive-voice morphs are positioned as dependents of the co-occurring tense or aspect morph.

The main takeaways from Maxwell's contribution are that dependency structure can be acknowledged inside words and that the distinction between function morphs and content morphs can also be acknowledged. There is a split in the status of the function morphs, though. Aspect and tense morphs are positioned as heads over the content morph, whereas person, number, and passive voice morphs are positioned as dependents of other co-occurring function morphs.

4.3. Meaning-Text Theory (MTT)

Mel'čuk's contribution (this volume, p. 433) considers the status of function words in Meaning-Text-Theory (MTT), his dependency-based framework for an integral description, or modelling, of natural languages (cf. Mel'čuk 1988, 2012-2015, 2021). The first works in this tradition appeared in the late 1950s and early 1960s. There is thus more than 65 years of accumulated literature. Mel'čuk begins his contribution with an important point; he states that the linguist

should not postulate a given function word as governor or dependent of some other word, but rather it simply is governor or dependent of that other word;[6] the linguist's job is to discover which it is by using appropriate methods such as rigorous definitions of the notions involved. Accordingly, the first part of Mel'čuk's contribution is devoted to producing the definitions and illustrations that enable the linguist to distinguish governor from dependent. He then examines function words in this context, providing examples from many languages, with an emphasis on Russian.

Mel'čuk presents the conditions and criteria for identifying the presence of syntactic dependencies between two lexical items, the direction of these dependencies, and the types of surface-syntactic relations of these dependencies. In the current context of the hierarchical status of function words, his criteria for identifying the directionality of syntactic dependencies are most relevant. He lists three such criteria: 1) passive surface-syntactic valence of the resulting phrase, 2) morphological link of the phrase to external context, and 3) the semantic content of the phrase (cf. Kahane 2022 for a critical assessment of these criteria). Based on the examples Mel'čuk produces in this area, it is evident that these criteria identify auxiliary verbs as governors over content verbs and adpositions as governors of their nouns.

Mel'čuk continues with an extensive example, whereby he gives both the *Deep-Syntactic Structure* (DSyntS) and *Surface-Syntactic Structure* (SSyntS) of the Russian sentence *Čelovek, kotoryj sdelal dlja spasenija traktata ot gibeli bol'še, čem kto by to ni bylo, byl nazvan pri roždenii Evstoxiem* 'The person who did more than anyone else for the rescue of the treatise from destruction was named at birth Eustochius'. He then goes on with a discussion of the value of common terminology. He rejects the term *function word* as "hackneyed," replacing it with *grammatical lexeme*; he also replaces the term *content word* with *lexical lexeme*. Grammatical lexemes are not present in DSyntS; they appear first in the transition from DSyntS to SSyntS.

Concerning grammatical lexemes (≈ function words), Mel'čuk acknowledges three distinct types: structural, grammemic, and pronominal. Adpositions, subordinators, expletives, and correlative elements of various sorts tend to be structural grammatical lexemes;

[6] Note that the term *governor* now appears (instead of *head*) in a manner that is consistent with Mel'čuk's use of terminology.

auxiliary verbs, articles, and certain particles tend to be grammemic grammatical lexemes; and pronouns are of course pronominal grammatical lexemes. The inventory of grammatical lexemes can vary a lot from one language to the next. Some types of grammatical lexemes exist in some languages but not in others.

Two take-aways from Mel'čuk's contribution are that MTT views many grammatical lexemes as governors over the co-occurring content words (e.g., auxiliary verbs, adpositions, subordinate conjunctions) and others as dependents thereof (e.g., articles, degree adverbs, comparative and superlative markers, various particles). The conjunctions and disjunctions of coordination are intermediate in a sense insofar as they are dependents of the initial conjunct(s) but governors over the final conjunct. Most importantly in these areas is that MTT proposes definitions for all linguistic notions it uses, and according to these definitions, it arrives at the hierarchies just sketched.

4.4. Functional Generative Description (FGD)

Hajičová, Panevová, Mikulová, and Hajič's contribution (this volume, p. 465) presents *Functional Generative Description* (FGD), a dependency-based framework of syntax and grammar associated with linguists in Prague (cf. Sgall et al. 1986; Bejček et al. 2017; Hajič et al. 2017). The first works in this tradition were by Petr Sgall and appeared in the late 1960s. The authors of this contribution begin with some history of FGD and sketch the organization of the framework. Like MTT, FGD is a multistratal system. There are two levels of syntax, a deep level, called the *tectogrammatical* level, and a surface level, called the *analytical* level; both of these levels assume dependency structures. Function words (*synsemantic lexical items*) do not enjoy node status at the tectogrammatical level, only content words (*autosemantic lexical items*) do that. Both function words and content words are present as nodes in the analytical level, though—and punctuation marks as well! The *Prague Dependency Treebank* (PDT), which first appeared in the early 1990s, is based on FGD.

The discussion first examines the status of function words at the analytical level, progressing then to their status at the tectogrammatical level. The examples used for illustration are all from Czech and translated to English, and the trees themselves are from the PDT. At the analytical level, there are on the one hand function words that are

parts of nominal groups (prepositions) or that connect clauses (or, as the case may be, parts of clauses) into one whole (conjunctions); there are also function words functioning within verbal complexes and containing information about the morpho-syntactic properties of verbs (i.e., auxiliaries). To be specific, adpositions appear above their nouns, subordinate conjunctions above their verbs, and coordinate conjunctions above the root words of the coordinated strings. Concerning verbs, modal verbs are positioned above their co-occurring full verbs. Most noteworthy, however, is the analysis of the Czech *být* 'be'. When *být* occurs together with a full verb, FGD views it as an auxiliary and subordinates it to the co-occurring content verb. When it appears with a predicative expression (e.g., predicative nominal, predicative adjective), however, it is understood to be copular *být*, which is not viewed as an auxiliary, and is hence positioned as head over the predicative expression. The discussion of the motivating factors behind the structural analysis of *být* are of particular interest (this volume, page 475-476).

After discussing the analytical level and its tree structures, the contribution turns to the tectogrammatical level. As stated above, function words are not present as nodes in the tectogrammatical level. Their contribution is indicated, though, as functors in the labels of dependency edges and as parts of the complex symbols attached to the nodes of the tree. These contributions are called *grammatemes* and express tense, modality, mood, voice, etc. The contribution of prepositions and subordinate conjunctions are understood as labels on the dependency edges, but are indicated as part of the dependent node. Interestingly, coordinate conjunctions do enjoy node status at the tectogrammatical level; the analysis for them is similar to their analysis at the analytical level. The discussion also considers the status of focalizers such as *jenom* 'only', positioning it as a dependent of the expression that it focuses.

To summarize, FGD is similar to MTT in its multistratal approach to grammar. It is also similar to MTT insofar as many function words are positioned as heads over their co-occurring content words (adpositions, subordinate and coordinate conjunctions, modal verbs, copular *be*), but it is unlike MTT and other dependency grammars in its analysis of the Czech auxiliary *být*, which it positions as a dependent of the co-occurring content verb.

4.5. Word Grammar (WG)

Hudson's contribution (this volume, p. 513) presents the Word Grammar account of function words, Word Grammar being a dependency-based framework of syntax and grammar developed by Richard Hudson (cf. Hudson 1984, 2007, 2010). The first publications in this framework appeared in the early 1980s and have continued over the decades since. Hudson begins his contribution by drawing attention to the Universal Dependencies (UD) analysis of function words and to the fact that UD is "a very widely used and well-known version of dependency analysis." He states that the UD decision to position function words as dependents of their co-occurring content words is not representative of how most DGs view the hierarchical status of function words. Most of his contribution is then devoted to establishing that the opposite situation, that is, that function words should be seen as immediately dominating their co-occurring content words, is the case.

Hudson first provides some historical background information concerning the debate for orientation. He discusses some works of prominent grammarians from previous centuries, e.g., John Wallis (1616-1703), Nicolas Beauzée (1717-1789), Joseph Priestly (1733-1804), Franz Kern (1830-1894), Brainerd Kellogg (1834-1920). Hudson writes that part of the reason for the lack of clarity concerning the status of function words stems from the influence of Greek and Latin grammars on the development of grammatical theory over the centuries. Greek and Latin are morphologically rich, highly inflected languages. Markers of tense, case, aspect, etc., tend to appear as affixes within words. The tendency among some DG people to subordinate function words to content words can be evaluated in this light, function words being akin to these affixed markers.

After this historical background information, Hudson moves to the *Word Grammar* framework and its account of function words. He discusses the Word Grammar analyses of English prepositions and subordinators, auxiliary verbs, and determiners. He gives many linguistic arguments motivating the Word Grammar position regarding the hierarchical status of each of these word categories. Word Grammar positions prepositions over their co-occurring nouns, subordinators over their co-occurring verbs, and determiners over their co-occurring nouns. Hudson motivates each of these decisions

with multiple linguistic insights. Interestingly, however, Hudson does not discuss the Word Grammar analysis of coordinate conjunctions, probably due to the fact that such a discussion would require more space than is available.

Of particular note is the Word Grammar stance that determiners are heads over their co-occurring nouns. The Word Grammar position in this area is contrary to most work in the DG tradition, including most of the works mentioned and the contributions appearing in this volume (Tesnière, DCxG, MTT, FGD, UD, SUD). The Word Grammar analysis does, however, receive support from phrase structure grammars, the DP analysis of nominal groups having become the default in phrase structure circles decades ago. Osborne (2020) scrutinizes this debate from a DG perspective, i.e., NP vs. DP, examining closely many of Hudson's arguments in favor DP produced here below and in the Word Grammar literature.

4.6. Universal Dependencies (UD)

De Marneffe, Nivre, and Zeman's contribution (this volume, page 549) presents the dependency-based *Universal Dependencies* (UD) famework, stating that UD "has developed primarily as a framework for cross-linguistically consistent, morphosyntactic annotation" (this volume, p. 549) (cf. Nivre et al. 2016; De Marneffe et al. 2020; Nivre et al. 2020). The first relevant publication in this tradition appeared in 2006 and is associated with the Stanford Parser. The UD annotation scheme has been applied to at least 122 languages, the treebank corpora of these languages being freely available at the UD website (https://universaldependencies.org/). The popularity of the UD approach and the Stanford Parser make UD the poster child for DG in general.

As mentioned above, though, the UD decision to position function words as dependents of content words is a controversial matter; it contradicts much of the DG tradition since Tesnière (1959) (cf. Osborne and Maxwell 2015 and Osborne and Gerdes 2019). It is important in this respect to consider the motivation for the UD decisions in this area. The next passage from the De Marneffe et al.'s contribution below gives this motivation:

> UD consistently chooses the lexical or content word as the head, and makes function words dependents of the head with

> special functional relations to indicate their status as nucleus elements. This choice follows naturally from the decision to prioritize predicate-argument and modifier relations in the syntactic structure. It also makes cross-linguistic similarities more transparent, since direct relations between the semantic cores of nuclei are more likely to be parallel across languages, whereas function words frequently correspond to morphological inflection (or nothing at all) in other languages. Thus, what counts as the head of a nucleus is likely to be more parallel across languages (and sometimes also within languages) if the content word is consistently analyzed as the head.
>
> (De Marneffe et al., this volume, page 549)

The term *nucleus* appearing in this passage is appealing to Tesnière's concept—see tree (3) above. A nucleus fulfills both a semantic and syntactic role. By positioning function words as dependents of their co-occurring content words, UD foregrounds the importance of semantic units, i.e., units of meaning, and backgrounds that of syntactic units, i.e., units of structure. In this sense, UD annotation is in fact akin to the Deep Syntactic Structure (DSynS) of MTT and the tectogrammatical level of FDG where dependency relations are also stated between content words, function words being relegated to secondary status.

De Marneffe et al. first provide basic information about UD and the specifics of its design. UD acknowledges 17 parts of speech and 37 syntactic relations of which seven directly concern the hierarchical analysis of function words: copula (cop), auxiliary (aux), marker (mark), case (case), classifiers (clf), determiner (det), and coordinating conjunction (cc). The discussion examines each of these, the associated hierarchical analyses thereof being those shown in the (b)-trees above in (6-10). In addition, the copula is subordinated to the predicative expression with which it co-occurs and classifiers, which are prominent in languages like Chinese, are subordinated to the co-occurring pre-modifier of the noun.

De Marneffe et al. then turn to the discussion of the usefulness of the UD treebanks for exploring the status of function words in 72 languages of the world. The discussion first documents the frequency of case nodes (i.e. adpositions) in the 72 languages; it then does the same for the frequency of marker nodes (i.e., subordinators), and then for the frequency of cc nodes (i.e., coordinate conjunctions),

and finally for the frequency of aux nodes (i.e., auxiliaries). After considering these frequencies, it turns to the linear order of heads and dependents in the 72 languages. Crucially in this area, however, the account first does the calculation for content words only, excluding function words. The resulting numbers thus deliver numbers that are broadly in line with typological studies in the area. English, for instance, is identified as more head-initial than head-final, and Japanese is revealed to be almost entirely head-final.

The account then continues with the calculation of function-word dependencies involving adpositions (case markers), subordinators, and auxiliaries, correlated with the dependencies among content words. It is, given UD annotation, not surprising to learn that the dependencies reaching from adpositions, subordinators, and auxiliaries to their heads, are almost entirely head-final in English, whereas these same dependencies are almost entirely head-initial in Japanese. These results should give one pause, though, since a language such as Japanese—which is widely acknowledged to be almost entirely head-final in syntax—ends up on UD annotation with a significant number of head-initial dependencies, namely concerning those dependencies that connect a function word into the structure.

4.7. Surface-Syntactic Universal Dependencies (SUD)

Gerdes, Guillaum, Kahane, and Perrier's[7] contribution (this volume, p. 589) presents the *Surface-syntactic Universal Dependencies* (SUD) approach, SUD being a dependency-based annotation scheme that takes part of its name from the Universal Dependencies (UD) framework (cf. Gerdes et al. 2018, 2019, 2021). SUD piggybacks onto UD in a sense, but stands as an alternative to UD annotation deci-

[7] Guy Perrier passed away in June 2023. The other authors of the contribution have provided the next statements in his honor:

We were deeply saddened by the loss of Guy Perrier. As Professor Emeritus at the University of Lorraine, he not only left an indelible mark on the academic world but continued to contribute actively to his field. He was especially passionate about enhancing the coherence of French treebanks and was in the midst of preparing an extracted grammar of French from these very treebanks. Guy was more than just a colleague to us; he was a vibrant spirit whose enthusiasm for formalism and grammar was contagious. Our lively debates and discussions with him added a spark to our work that we will forever miss.

It is in his memory that we dedicate this fourth joint article on the SUD annotation scheme. Guy's invaluable contributions have shaped and enriched this scheme, and we believe his legacy will live on through it. This dedication is but a small token of our gratitude and admiration for a brilliant scholar and a cherished friend.

sions, this alternative pursuing surface syntactic analyses of sentence structures as opposed to the deeper, more semantically motivated analyses of UD. The SUD project is relatively new, the first publications presenting this approach appearing in 2018. The claim that SUD piggybacks onto UD is based on the interconvertibility—UD corpora can be and have been converted to the UD annotation format, and vice versa. The treebank corpora that are freely available at the UD website are also available at the SUD website, but converted to SUD annotation (https://surfacesyntacticud.github.io/).

As the name suggests, the emphasis in SUD is on syntactic plausibility of the dependency structures assumed. To this end, the authors devote the first part of their contribution to establishing the syntactic criteria that allow one to discern the presence and direction of dependencies. They emphasize that if one has clear syntactic criteria for distinguishing between head and dependent, the difficulty of knowing where to draw the line between content and function word does not arise. The authors present three groups of criteria that allow them to distinguish the presence, direction, and syntactic function of dependencies. The approach and emphasis are similar to those of Mel'čuk's contribution, an algorithm of a sort being proposed to distinguish heads from dependents.

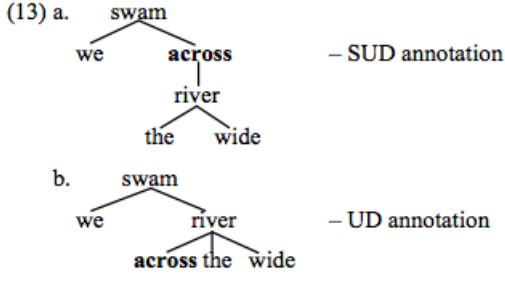
The contribution proceeds to the particular SUD analysis of the word categories that most DG theoreticians acknowledge to be function words. The discussion in this area begins wtih Tesnière's category of *translative*—adpositions, subordinators, and auxiliaries being typical translatives. The discussion presents numerous linguistic considerations that identify adpositions as heads over the co-occuring nouns, subordinators as heads over the co-occurring verbs, and auxiliaries as heads over the co-occurring content verbs. The account in this area is similar to Hudson's approach; many of the same linguistic observations are presented and taken as motivation for decisions in the area. Concerning determiners, the authors acknowledge that their criteria do not allow for an easy decision concerning the direction of the dependency. The decision in this area that is nevertheless reached takes the noun to be head over the determiner due to the semantics, a noun phrase being the semantic type of the noun, not of the determiner. Concerning coordinators, UD and SUD agree; both annotation schemes position the non-initial conjunct as a dependent of the initial conjunct, and the coordinator as a dependent of the non-initial conjunct.

The discussion also touches on additional tough cases, for instance on multi-word adopositions, and also on relative and interrogative clauses. The decisions reached in these areas are fraught with difficulty, there being competing factors making clear decisions less than obvious. What is perhaps most noteworthy, though, is that SUD is like MTT, FGD, and UD in that it allows non-projective dependencies, making it a non-projective DG, as opposed to one that is entirely projective (cf. Groß and Osborne 2009).[8]

4.8. Treebank metrics

Yan and Liu's contribution (this volume, p. 629) examines the hierarchical status of adpositions in treebank corpora with respect to three quantitative metrics: *absolute hierarchical distance* (AHD), *relative hierarchical distance* (RHD), and *probabilistic valency pattern* (PVP) (Liu 2009). It compares the status of adpositions across 11 parallel UD and SUD corpora with respect to these metrics. The key insight in this area is that on UD annotation, adpositions are leaf nodes in the structure, whereas on SUD annotation, they are intermediate in the structure. The differences in the numbers for adpositions extracted from the parallel corpora are traceable back to this hierarchical distinction.

The metrics are illustrated here briefly using the key trees that Yan and Liu provide, trees of the sentence *We swam across the wide river*:



AHD is the depth in the tree of the given word, measured starting from the root node at 0 and reaching down to that word. On the SUD

[8] Note that projectivity as a principle of word order is operational only on the surface. At deeper levels of representation, the structures may be entirely projective (as they are at the tectogrammatical level of FGD) or word order may be absent entirely as is the case at the semantic and syntactic levels of MTT.

annotation in (13a), the AHD of the preposition *across* is 1, whereas it is 2 on the UD annotation in (13b). RHD is the depth in the tree of the given word subtracted from the depth of the deepest word that word dominates in the tree. The RHD of *across* on the SUD annotation in (13a) is hence 2 (=3-1), and on the UD annotation in (13b), it is 0 (=2-2).

The PVP metric is broader; it is concerned with the so-called *centripetal* and *centrifugal* forces that connect the given word into the structure. Centripetal force concerns the governor of the given word, and centrifugal force its dependent(s). The sum influence of these forces on a given node add up to 100%. It is predictable in this area that on UD annotation, the centripetal force exerted on a given adposition is much greater than the centrifugal force, since UD annotation views adpositions as leaf nodes, meaning that they lack dependents entirely. Conversely, it is not surprising that on SUD annotation, the centripetal and centrifugal forces exerted on a given adposition tend to be more balanced, since adpositions are very often intermediate in the structure, having both a head and at least one dependent.

Based on their exploration of adpositions with respect to the three metrics, Yan and Liu conclude that SUD annotation may be better suited for theoretical studies, especially for those of a typological nature. At the same time, they state that since UD annotation is more semantically motivated, it is consistent with Tesnière's stemmatic emphasis on semantic roles.

5. Concluding comment

The goal behind putting together this collection of essays on function words in DG is to push the field forward. By raising awareness about the importance and impact of decisions in this area, the desire is to see dependency grammarians focus their interest and reach linguistically well-motivated conclusions about the dependency hierarchies they put forward.

Works Cited

1. Bejček, E., Eva Hajičová, Marie Mikulová and Jarmila Panevová. 2017. The Relation of Form and function in linguistic theory and in a multi-layer treebank . In Jan Hajič (ed.), *Proceedings of the 16th International Workshop on Treebanks and Linguistic Theories*. Praha: Karlova univerzita, pp. 56–63.
2. Clark, Stephen Watkins. 1847. *A practical grammar: In which words, phrases and sentences are classified according to their offices and their various relationships to each another*. Cincinnati: H. W. Barnes and Company.
3. De Marneffe, Marie-Catherine, Christopher D. Manning, Joakim Nivre and Daniel Zeman. 2021. Universal Dependencies. *Computational Linguistics*, 47(2): 255–308.
4. Gerdes Kim, Bruno Guillaume, Kahane Sylvain and Guy Perrier. 2018. SUD or Surface-Syntactic Universal Dependencies: An annotation scheme near-isomorphic to UD, *Proceedings of the Universal Dependencies Workshop (UDW)*, EMNLP, Bruxelles.
4. Gerdes Kim, Bruno Guillaume, Kahane Sylvain and Guy Perrier. 2019. Improving Surface-syntactic Universal Dependencies (SUD): surface-syntactic functions and deep-syntactic features, *Proceedings of the 17th international conference on Treebanks and Linguistic Theories (TLT),* SyntaxFest, Paris.
5. —. 2021. Starting a new treebank? Go SUD!, *Proceedings of 6th international conference on Dependency Linguistics (DepLing)*, SyntaxFest, ACL, 11 p.
6. Groß, Thomas. 2014. Clitics in dependency morphology. In Linguistics Today Vol. 215: *Dependency Linguistics*, ed by E. Hajičova et al. 229-252. Amsterdam: John Benjamins Publishing.
7. Groß, Thomas and Timothy Osborne. 2009. Toward a practical dependency grammar theory of discontinuities. *SKY Journal of Linguistics* 22, 159-190.
8. Hajič, Jan, Eva Hajičová, Marie Mikulová and Jiří Mírovský. 2017. Prague Dependency Treebank. In N. Ide et al. (eds.), *Handbook of Linguistic Annotation*, pp. 555–594. Dordrecth: Springer.
9. Hudson, Richard. 1976. *Arguments for a non-transformational grammar*. Chicago: University of Chicago Press.

10. —. 1984. *Word Grammar*. Oxford, UK: Basil Blackwell.
11. —. 1995. Measuring syntactic difficulty. Draft of manuscript. Available at: http://dickhudson.com/wp-content/uploads/2013/07/Difficulty.pdf.
12. —. 2000. Language as a cognitive network. In Hanne Gram Simonsen and Rolf Theil Endresen (eds.), *A Cognitive Approach to the Verb. Morphological and Constructional Perspectives*, 49–72. Berlin: Mouton de Gruyter. http://www.phon.ucl.ac.uk/home/dick/net.htm.
13. —. 2003. The psychological reality of syntactic dependency relations. In Sylvain Kahane and Alexis Nasr (eds.), Proceedings of the First International Conference on Meaning-Text Theory, 181–192. Paris: Ecole Normale Supérieure.
14. —. 2007. *Language networks: The new Word Grammar*. Oxford: Oxford University Press.
15. —. 2010. *An Introduction to Word Grammar*. Cambridge: Cambridge University Press.
16. Imrényi, András and Nicolas Mazziotta. 2020. Chapters of Dependency Grammar: A historical survey from Antiquity to Tesnière (*Studies in Language Companion Series*, 212). Amsterdam/Philadelphia: John Benjamins.
17. Kahane, Sylvain. 2022. Toward an inductive approach of the Mel'čukian model – Illustration by the criteria for the syntactic head, in J. Milićević, A. Polguère (eds), *Festschrift for 90th birthday of Igor Mel'čuk*, 201-212.
18. Kern, Franz. 1883. *Zur Methodik des deutschen Unterrichts*. Berlin: Nicolaische Verlags-Buchhandlung.
19. Liu, Haitao. 2008. Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science* (Korea) 9. 159–191.
20. —. 2009. *Dependency grammar from theory to practice*. Beijing: The science press.
21. —. 2010. Dependency direction as a means of word-order typology: A method based on dependency treebanks. *Lingua* 120(6), 1567–1578
22. Matthews, P. H. 1981. *Syntax*. Cambridge, UK: Cambridge University Press.
23. Maxwell, Daniel. 2018. Dependency Construction Grammar. Ms. Available at: Academia.edu

24. Mel'čuk, Igor. 1988. *Dependency Syntax: Theory and Practice*, Albany, N.Y.: The SUNY Press
25. —. 2012-2015. *Semantics: From Meaning to Text*. Vols. 1–3. Amsterdam/Philadelphia: John Benjamins.
26. —. 2021. *Ten Studies in Dependency Syntax*. Berlin: De Gruyter Mouton.
27. Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Jan Hajič, Christopher D. Manning, Sampo Pyysalo, Sebastian Schuster, Francis Tyers and Daniel Zeman. 2020. Universal Dependencies v2: An evergrowing multilingual treebank collection. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 4034 – 4043. European Language Resources Association.
28. Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajič, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty and Daniel Zeman. 2016. Universal Dependencies v1: A multilingual treebank collection. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pp. 1659–1666.
29. Osborne, Timothy and Dan Maxwell. 2015. A historical overview of the status of function words in dependency grammar. Depling 2015, Uppsala, Sweden. Available at https://aclweb.org/anthology/W/W15/W15-2127.pdf.
30. Osborne, Timothy and Kim Gerdes. 2019. The Status of Function Words in Dependency Grammar: A Critique of Universal Dependencies. *Glossa* 4(1): 1-28.
31. Pullum, Geoffrey and Deirdre Wilson. 1977. Autonomous syntax and the analysis of auxiliaries. *Language* 53: 741–788.
32. Reed, Alonzo and Brainerd Kellogg. 1877. *Higher Lessons in English*. New York: Clark & Maynard.
33. Ross, John. 1969. Auxiliaries as main verbs. In *Studies in Philosophical Linguistics*, Series One., 77–102. Evanston, Illinois: Great Expectations Press.
34. Sgall Petr, Eva Hajičová and Jarmila Panevová. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Prague-Dordrecht: Academia-Reidel.
35. Temperley, David. 2007. Minimization of dependency length in written English. *Cognition* 105. 300–333.

36. Tesnière, Lucien. 1959. *Éléments de syntax structural*. Paris: Klincksieck.
37. —. 1969. *Éléments de syntax structural*. Deuxième edition revue et corrigée. Paris: Klincksieck.
38. Trask, R. L. 1997. *A Student's Dictionary of Language and Linguistics*. London: Arnold.