[Written in 2017 and heavily revised in 2018 for an HPSG handbook to be published by Language Science Press (https://hpsg.hu-berlin.de/Projects/HPSG-handbook/) edited by Anne Abeillé, Robert D. Borsley, Jean-Pierre Koenig and Stefan Müller]

HPSG and Dependency Grammar

1. Introduction

HPSG is firmly embedded, both theoretically and historically, in the phrase-structure (PS) tradition of syntactic analysis, but it also has some interesting theoretical links to the dependency-structure (DS) tradition. This is the topic of the present chapter, so after a very simple comparison of PS and DS and a glance at the development of these two traditions in the history of syntax, we consider a number of issues where the traditions interact.

The basis for PS analysis is the part-whole relation between smaller units (including words) and phrases, so the most iconic notation uses boxes (Müller 2018, 6). In contrast, the basis for DS analysis is the asymmetrical dependency relation between two words, so in this case an iconic notation inserts arrows between words. (Although the standard notation in both traditions uses trees, these are less helpful because the lines are open to different interpretations.) The two analyses of a very simple sentence are juxtaposed in Figure 1. As in

HPSG AVMs, each rectangle represents a unit of analysis.

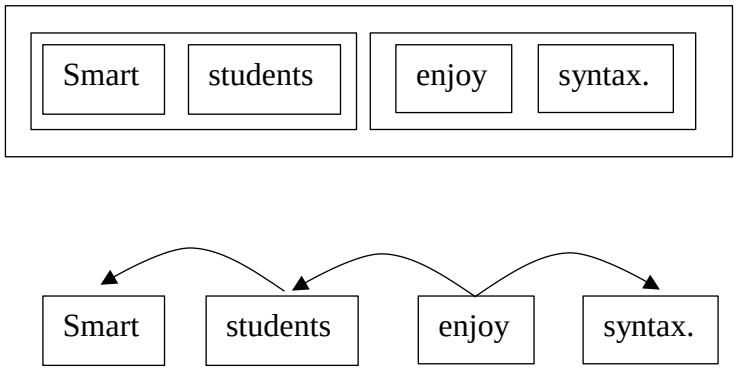


Figure 1: Phrase structure and dependency structure contrasted

In both approaches, each unit has properties such as a classification, a meaning, a form and relations to other items, but these properties may be thought of in two different ways. In PS analyses, an item contains its related items so it also contains its other properties – hence the familiar AVMs contained within the box for each item. But in DS analyses, an item's related items are outside it, sitting alongside it in the analysis, so for consistency other properties may be shown as a network in which the item concerned is just one atomic node. This isn't the only possible notation, but it is the basis for the main DS theory that I shall juxtapose with HPSG, Word Grammar.

What, then, are the distinctive characteristics of the two traditions? In the following summary I use 'item' to include any syntagmatic unit of analysis including morphemes, words and phrases (though this chapter will not discuss the possible role of morphemes).

(1) Containment: in PS, but not in DS, if two items are directly related, one must contain the other.
(2) Continuity: therefore, in PS, but not in DS, all the items contained in a larger one must be adjacent.
(3) Asymmetry: in DS, but not in PS, a direct relation between two items must be asymmetrical, with one depending on the other (the head of the relation).
(4) Functions: therefore DS, but not PS, recognises subtypes of dependency, viz the traditional grammatical functions (e.g. 'subject') as distinct relations.

These contrasts apply without reservation to 'plain vanilla' (Zwicky

1985) versions of DS and PS, but as we shall see in the history section, very few theories are plain vanilla.

As far as HPSG is concerned, the term *Head-driven* points immediately to dependency: an asymmetrical relation 'driven' by a head word. On the other hand, *Phrase Structure* clearly locates the theory in the PS tradition. This chapter reviews the relations between HPSG and the very long DS tradition of grammatical analysis. The conclusion will be that in spite of its PS roots, HPSG implicitly (and sometimes even explicitly) recognises dependencies; and it may not be a coincidence that one of the main power-bases of HPSG is Germany, where the DS tradition is also at its strongest (Müller 2018, 359).

Why, then, does HPSG use PS rather than DS? As far as I know, PS was simply default syntax in the circles where HPSG evolved, so the choice of PS isn't the result of a conscious decision by the founders, and I hope that this chapter will show that this is a serious question which deserves discussion. Indeed, I once wrote a paper (which was never published) called 'Taking the PS out of HPSG' – a title I was proud of until I noticed that *PS* was open to misreading, not least as 'Pollard and Sag'. Carl and Ivan took it well, and I think Carl may even have entertained the possibility that I might be right. But the historical roots and the general dominance of PS have so far discouraged discussion of this fundamental question.

HPSG is a theoretical package where PS is linked intimately to a collection of other assumptions; and the same is true for any theory which includes DS, including my own Word Grammar. Here too I found welcome similarities, not least the use of default inheritance in some versions of the theory. I shall argue below that inheritance offers a novel solution to one of the outstanding challenges for the dependency tradition.

The next section sets the historical scene. This is important because it's all too easy for students to get the impression (mentioned above) that PS is just default syntax, and maybe even the same as 'traditional grammar'. We shall see that grammar has a very long and rather complicated tradition in which the default is actually DS rather than PS. Later sections then address particular issues shared by HPSG and the dependency tradition.

2. Dependency and constituency in the

history of syntax

The relevant history of syntax starts more than two thousand years ago in Greece. (India may have started even earlier, but it seems to have had little effect on the European tradition.) Greek and Roman grammarians focused on the morphosyntactic properties of individual words, but since these included a rich case system, they were aware of the syntactic effects of verbs and prepositions governing particular cases. However, this didn't lead them to think about syntactic relations, as such; precisely because of the case distinctions, they could easily distinguish a verb's dependents in terms of their cases: 'its nominative', 'its accusative' and so on (Robins 1967, 29). Both the selecting verb or preposition and the item carrying the case inflection were single words, so the Latin grammar of Priscian, written about 500 AD and still in use a thousand years later, recognised no units larger than the word: 'his model of syntax was word-based – a dependency model rather than a constituency model' (Law 2003, 91). However, it was a dependency model without the notion of 'dependency' as a relation between words.

The dependency relation, as such, was first identified by the Arabic grammarian Sibawayh in the 8th century (Owens 1988; Kouloughli 1999). In the Arabic tradition which then developed, it was relatively easy for grammarians to recognise dependency because of two typological properties of Classical Arabic: three cases (which were extended to include three verbal moods) and fairly consistent head-initial word order. As a result, the dependency relation, including subjects as well as objects, was recognised not only as a bearer of case government but also as relevant to word order, and the metalanguage distinguished the governor-governed relation from the particular case selected. In fact, the relation abstracted in this way was solid enough to allow a major debate between the two centres of grammatical theory (Basra and Kufa) over whether mutual dependency was possible (Owens 1988, 52) – an issue to which I return in Section 5.3. But even though grammatical theory recognised dependency as a relation, it only recognised one kind of dependency: that between a governor and a governed noun or verb. This dependency included subjects as well as objects and other complements, but it excluded adjuncts; so no relation was recognised between a noun and its modifying adjectives.

In Europe, grammar teaching in schools was based on 'parsing' (in its original sense), an activity which resonates with HPSG but which was

formalised in the ninth century (Luhtala 1994). The activity of parsing was a sophisticated test of grammatical understanding which earned the central place in school work that it held for centuries – in fact, right up to the 1950s (when I did parsing at school). In HPSG terms, school children learned a standard list of attributes for words of different classes, and in parsing a particular word in a sentence their task was to provide the values for its attributes, including its grammatical function (which would explain its case). In the early centuries the language was Latin, but more recently it was the vernacular (in my case, English).

Alongside these purely grammatical analyses, the Ancient World had also recognised a logical one, due to Aristotle, in which the basic elements of a proposition (*logos*) are the logical subject (*onoma*) and the predicate (*rhēma*). To Aristotle it was obvious that a statement such as 'Socrates ran' requires the recognition both of the person Socrates and of the property of running, neither of which could constitute a statement on its own (Law 2003, 30–31). By the twelfth century, grammarians started to apply a similar analysis to sentences; but in recognition of the difference between logic and grammar they replaced the logicians' *subiectum* and *praedicatum* by *suppositum* and *appositum* – though the logical terms would creep into grammar by the late eighteenth century (Law 2003, 168). This logical analysis produced the first top-down analysis in which a larger unit (the logician's proposition or the grammarian's sentence) has parts, but the parts were still single words, so *onoma* and *rhēma* can now be translated as 'noun' and 'verb'. If the noun or verb was accompanied by other words, the older dependency analysis applied.

The result of this confusion of grammar with logic was a muddled hybrid analysis in the Latin/Greek tradition which persists even today in some school grammars, and which took centuries to sort out in grammatical theory. For the subject and verb, the prestige of Aristotle and logic supported a subject-verb division of the sentence (or clause) in which the subject noun and the verb were both equally essential – an analysis which even logicians have now abandoned in favour of a Fregean dependency analysis (where the subject is just one argument among many). Moreover the grammatical tradition even includes a surprising number of analyses in which the subject noun is the head of the construction, ranging from the modistic grammarians of the twelfth century (Robins 1967, 83), through Henry Sweet (Sweet 1891, 17), to no less a figure than Otto Jespersen in the twentieth (Jespersen 1937), who distinguished 'junction' (dependency) from 'nexus'

(predication) and treated the noun in both constructions as 'primary'.

The first grammarians to recognise a consistently dependency-based analysis for the rest of the sentence (but not for the subject and verb) were the French *encyclopédistes* of the eighteenth century (Kahane forthcoming), and by the nineteenth century much of Europe accepted a theory of sentence structure based on dependencies, but with the subject-predicate analysis as an exception – an analysis which by modern standards is muddled and complicated. Each of these units was a single word, not a phrase, and modern phrases were recognised only indirectly by allowing the subject and predicate to be expanded by dependents; so nobody ever suggested there might be such a thing as a 'noun phrase' until the late nineteenth century. Function words such as prepositions had no proper position, being treated typically as though they were case inflections.

The invention of syntactic diagrams in the nineteenth century made the inconsistency of the hybrid analysis obvious. The first such diagram was published in a German grammar of Latin for school children (Billroth 1832), and the nineteenth century saw a proliferation of diagramming systems[1], including the famous Reed-Kellogg diagrams which are still taught (under the simple name 'diagramming') in some American schools (Reed and Kellogg 1877); indeed, there is a website (Sentence Diagrammer, by 1aiway) which generates such diagrams, giving diagrams such as the one reproduced in Figure 2. The significant feature of this diagram is the special treatment given to the relation between the subject and predicate (with the verb *are* sitting uncomfortably between the two), with all the other words in the sentence linked by more or less straightforward dependencies. (The geometry of these diagrams also distinguishes

[1] See a small selection at http://dickhudson.com/sentence-diagramming/.

grammatical functions.)

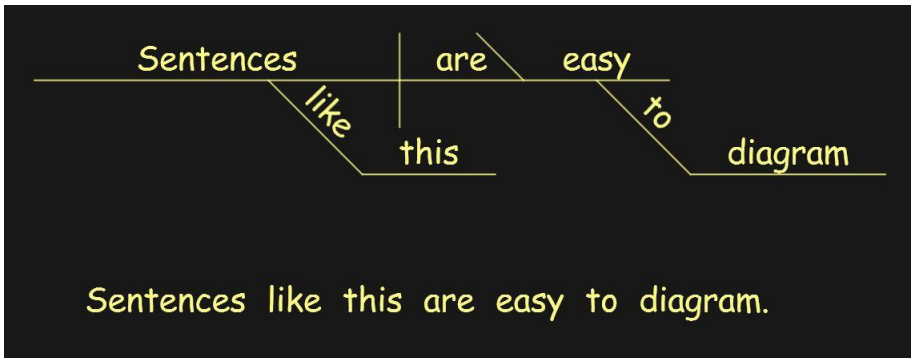


Figure 2: Reed and Kellogg diagram by Sentence Diagrammer

One particularly interesting (and relevant) fact about Reed and Kellogg is that they offer an analysis of *that old wooden house* in which each modifier creates a new unit to which the next modifier applies: *wooden house*, then *old wooden house* (Percival 1976, 18) – a clear hint at more modern structures (including the ones proposed in 4.1), albeit one that sits uncomfortably with plain-vanilla dependency structure.

However, even in the nineteenth century there were grammarians who questioned the hybrid tradition which combined the subject-predicate distinction with dependencies. Rather remarkably, three different grammarians seem at roughly the same time to have independently reached the same conclusion: hybrid structures can be replaced by a homogeneous structure if we take the finite verb as the root of the whole sentence, with the subject as one of its dependents. This idea seems to have been first proposed in print in 1873 by the Hungarian Sámuel Brassai (Imrényi 2013; Imrényi and Vladár forthcoming); in 1877 by the Russian Aleksej Dmitrievsky (Sériot 2004); and in 1884 by the German Alexander Kern (Kern 1884). Both Brassai and Kern used diagrams to present their analyses, and used precisely the same tree-structures which Lucien Lucien Tesnière in France called 'stemmas' nearly fifty years later (Tesnière 1959, 2015). The diagrams have both been redrawn here, with English translations of the grammatical terminology.

Brassai's proposal is contained in a school grammar of Latin, so the

example is also from Latin:

(5) Uxor amans flentem flens acrius ipsa tenebat,
wife loving crying crying more.bitterly herself was.hugging

imbre per indignas usque cadente genas. (Latin)
shower on unbecoming continuously falling cheeks

'The wife, herself even more bitterly crying, was hugging the crying one, while a shower [of tears] was falling on her unbecoming cheeks [i.e. cheeks to which tears are unbecoming].'

Brassai's diagram, including grammatical functions as translated by the authors (Imrényi and Vladár forthcoming), is in Figure 3. The awkward horizontal braces show the limitations of the analysis rather than a nod in the direction of PS, as is evident from the fact that the bracketed words are not even adjacent in the sentence analysed.

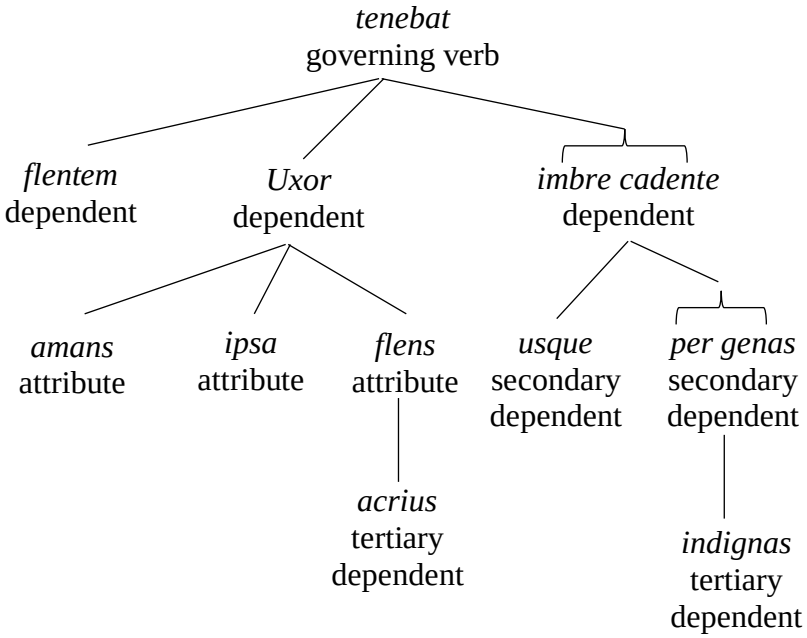


Figure 3: A verb-rooted tree from Brassai 1873

Kern's tree, on the other hand, is for a German sentence. Once again,

the original includes function terms which are translated here into English:

(6) Eine stolze Krähe schmückte sich mit den
a proud crow decorated himself with the

ausgefallenen Federn der Pfauen. (German)
fallen-out feathers of.the peacocks

'A proud crow decorated himself with the dropped feathers of the peacocks.'

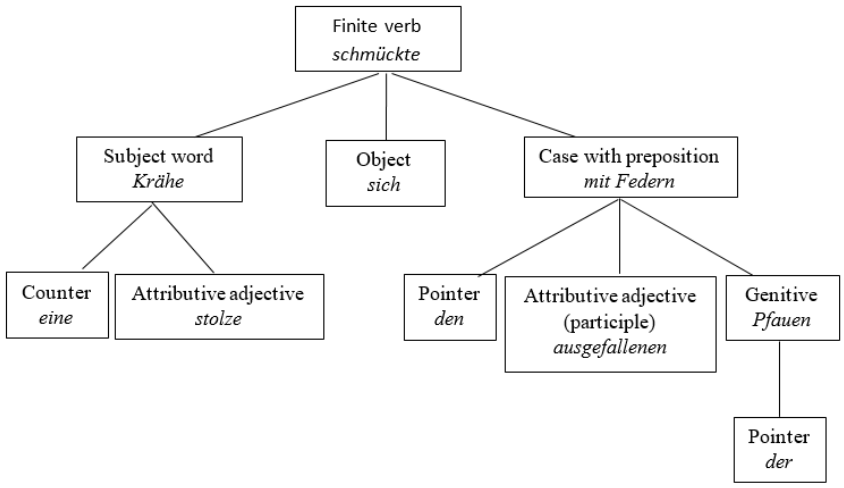


Figure 4: A verb-rooted tree from Kern (1884)

Once again the analysis gives up on prepositions, treating *mit Federn* as a single word, but Figure 4 is an impressive attempt at a coherent analysis which would have provided an excellent foundation for the explosion of syntax in the next century. In this approach,

> ... the sentence is not a basic grammatical unit, but merely results from combinations of words, and therefore ... the only truly basic grammatical unit is the word. A language, viewed from this perspective, is a collection of words and ways of using them in word-groups, i.e., expressions of varying length. (Percival 1976, 21)

If Brassai, Dmitrievsky and Kern had had the influence they deserved, it is reasonable to assume that modern theories would have been

predominantly based on DS.

But the vagaries of intellectual history and geography worked against this. When Leonard Bloomfield was looking for a theoretical basis for syntax, he could have built on what he had learned at school:

> … we do not know and may never know what system of grammatical analysis Bloomfield was exposed to as a schoolboy, but it is clear that some of the basic conceptual and terminological ingredients of the system that he was to present in his 1914 and 1933 books were already in use in school grammars of English current in the United States in the nineteenth century. Above all, the notion of sentence "analysis," whether diagramable or not, had been applied in those grammars. (Percival 1976, 18)

And when he visited Germany in 1913-14 he might have learned about Kern's ideas which were already influential there. But instead, he adopted the syntax of the German psychologist Wilhelm Wundt. Wundt's theory applied to meaning rather than syntax, and was based on a single idea: that every idea consists of a subject and a predicate. For example, a phrase meaning 'a sincerely thinking person' has two parts, one meaning 'a person' and the other 'thinks sincerely'; and the latter breaks down into 'thought' and 'is sincere' (Percival 1976).

For all its reliance on logic rather than grammar, the analysis is a clear precursor to neo-Bloomfieldian trees: it recognises a single consistent part-whole relationship (a 'partonomy') which applies recursively. This, then, is the beginning of the PS tradition: an analysis based purely on meaning and developed by a psychologist, not a grammarian – an unpromising start for a theory of syntax. However, Bloomfield's school experience presumably explains why he combined Wundt's partonomies with the hybrid structures of Reed-Kellogg diagrams in his classification of structures as endocentric (headed) or exocentric (headless). For him, exocentric constructions include the subject-predicate structure and preposition phrases, both of which were problematic in sentence analysis at school. Consequently, his Immediate Constituent Analysis (ICA) perpetuated the old hybrid mixture of headed and headless structures.

The DS elements of ICA are important in evaluating the history of PS, because they contradict the 'standard' view of history expressed here:

> Within the Bloomfieldian tradition, there was a fair degree of consensus regarding the application of syntactic methods as

> well as about the analyses associated with different classes of constructions. Some of the general features of IC analyses find an obvious reflex in subsequent models of analysis. Foremost among these is the idea that structure involves a part–whole relation between elements and a larger superordinate unit, *rather than* an asymmetrical dependency relation between elements at the same level. (Blevins and Sag 2013, 202–3, my italics)

This quotation implies, wrongly, that ICA discarded DS altogether.

What is most noticeable about the story so far is that even in the 1950s we still haven't seen an example of pure phrase structure. Every theory visited so far has recognised dependency relations in at least some constructions. Even Bloomfieldian ICA had a place for dependencies, though it introduced the idea that dependents might be phrases rather than single words and it rejected the traditional grammatical functions such as 'subject' and 'object'. Reacting against the latter gap, and presumably remembering their schoolroom training, some linguists developed syntactic theories which were based on constituent structure but which did have a place for grammatical functions, though not for dependency as such. The most famous of these theories are

- Tagmemics (Pike 1954)
- Functional Grammar (Dik 1989; Siewierska 1991)
- Systemic Functional Grammar (Halliday 1961, 1967)
- Relational Grammar (Perlmutter and Postal 1983; Blake 1990)

However, in spite of its very doubtful parentage and its very brief life, by the 1950s virtually every linguist in America seemed to accept without question the idea that syntactic structure was a partonomy.

This is the world in which Noam Chomsky introduced phrase structure, which he presented as a formalisation of ICA, arguing that "customarily, linguistic description on the syntactic level is formulated in terms of constituent analysis (parsing)" (Chomsky 1957, 26). But such analysis was only 'customary' among the Bloomfieldians, and was certainly not part of the classroom activity of parsing (Matthews 1993, 147).

Chomsky's phrase structure continued the drive towards homogeneity which had led to most of the developments in syntactic theory since the early nineteenth century. Unfortunately, Chomsky dismissed both

dependencies and grammatical functions as irrelevant clutter, leaving nothing but part-whole relations, continuity and sequential order, and category-labels.

Rather remarkably, the theory of phrase structure implied the (psychologically implausible) claim that 'sideways' relations such as dependencies between individual words are impossible in a syntactic tree. Less surprisingly, having defined PS in this way, he could easily prove that it was inadequate and needed to be greatly expanded beyond the plain-vanilla version. His solution was the introduction of transformations, but it was only thirteen years before he also recognised the need for some recognition of dependency structure in X-bar theory (Chomsky 1970). At the same time, others had objected to transformations and started to develop other ways of making PS adequate – by combining it with a functional structure as in LFG or by greatly enriching the categories as in GPSG and (later) HPSG.

Meanwhile, the European ideas about syntactic structure culminating in Kern's tree diagram developed rather more slowly. Lucien Tesnière in France wrote the first full theoretical discussion of DS in 1939 but it was not published till 1959 (Tesnière 1959, 2015), complete with 'stemmas' looking like the diagrams produced seventy years earlier by Brassai and Kern. Somewhat later, these ideas were built into theoretical packages in which DS was bundled with various other assumptions about levels and abstractness. Here the leading players were from Eastern Europe, where DS flourished: the Russian Igor Mel'čuk (Mel'cuk 1988), who combined DS with multiple analytical levels, and the Czech linguists Petr Sgall and Eva Hajičová (Sgall, Hajicová, and Panevova 1986) who included information structure. My own theory Word Grammar (developed, exceptionally, in the UK), also stems from the 1980s (Hudson 1984, 1990, 2007, 2010; Eppler 2004; Gisborne 1996). This is the theory which I compare below with HPSG, but it is important to remember that other DS theories would give very different answers to some of the questions that I raise.

DS certainly has a low profile in theoretical linguistics, and especially so in anglophone countries, but there is an area of linguistics where its profile is much higher: natural-language processing (Kübler, McDonald, and Nivre 2009). For example, at the time of writing (July 2018):

- the Wikipedia entry for 'Treebank' classifies 50 of its 101

treebanks as using dependency structure.
- The ‘Universal dependencies’ website lists more than 100 dependency-based treebanks for 60 languages.
- Google’s n-gram facility allows searches based on dependencies.
- The Stanford Parser (Chen and Manning 2014; de Marneffe et al. 2014) uses DS.

The attraction of DS in NLP is that the only units of analysis are words, so at least these units are given in the raw data and the overall analysis can immediately be broken down into a much simpler analysis for each word. This is as true for a computer scientist building a treebank as it was for a school teacher teaching children to parse words in a grammar lesson. Of course, as we all know, the analysis actually demands a global view of the entire sentence, but at least in simple examples a bottom-up word-based view will also give the right result.

To summarise this historical survey, PS is a recent arrival, and is not yet a hundred years old. Previous syntacticians had never considered the possibility of basing syntactic analysis on a partonomy. Instead, it had seemed obvious that syntax was literally about how words (not phrases) combined with one another.

3. HPSG and Word Grammar

The rest of this chapter considers a number of crucial issues that distinguish PS and DS by focusing specifically on how they distinguish two particular manifestations of these traditions, HPSG and Word Grammar (WG). The main question is, of course, how strong the evidence is for the PS basis of HPSG, and how easily this basis could be replaced by DS.

The comparison requires some understanding of WG, so what follows is a brief tutorial on the parts of the theory which will be relevant in the following discussion. Like HPSG, WG combines claims about syntactic relations with a number of other assumptions; but for WG, the main assumption is the Cognitive Principle:

(7) The Cognitive Principle

Language uses the same general cognitive processes and resources as general cognition, and has access to all of them.

This principle is of course merely a hypothesis which may turn out to

be wrong, but so far it seems correct (Müller 2018, 494), and it is more compatible with HPSG than the innatist ideas underlying Chomskyan linguistics. In WG, it plays an important part because it determines other parts of the theory.

On the one hand, cognitive psychologists tend to see knowledge as a network of related concepts (Reisberg 2007, 252), so WG also assumes that the whole of language, including grammar, is a conceptual network (Hudson 1984, 1, 2007, 1). One of the consequences is that the AVMs of HPSG are presented instead as labelled network links; for example, we can compare the elementary example of the HPSG entry for a German noun (Müller 2018, 264) with an exact translation using WG notation:

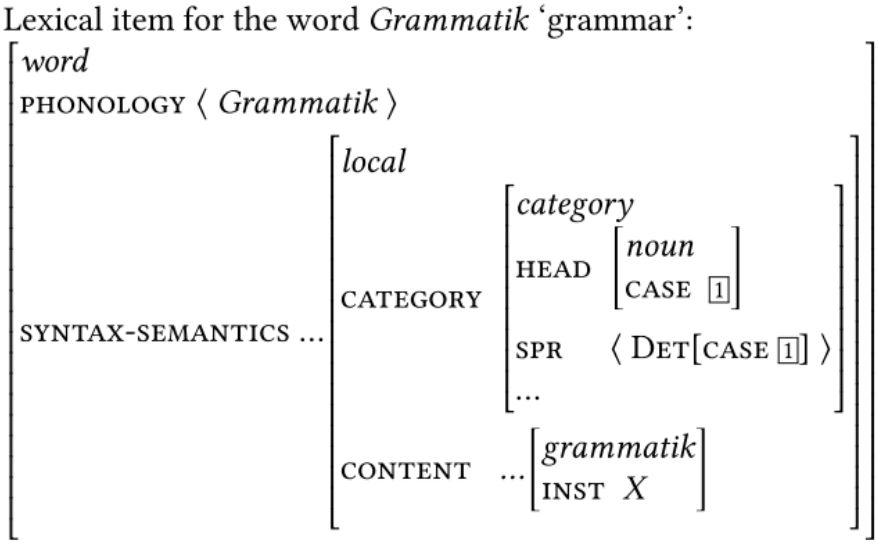


Figure 5: Grammatik *in HPSG notation (Müller 2018)*

Translating this AVM into network notation is straightforward but visually complicated so I take it in two steps. First I introduce the basic notation in Figure 6: a small triangle showing that the lexeme GRAMMATIK 'isa' word, and a headed arrow representing a labelled attribute (here, 'phonology') and pointing to its value. The names of entities and attributes are enclosed in rectangles and ellipses

respectively.

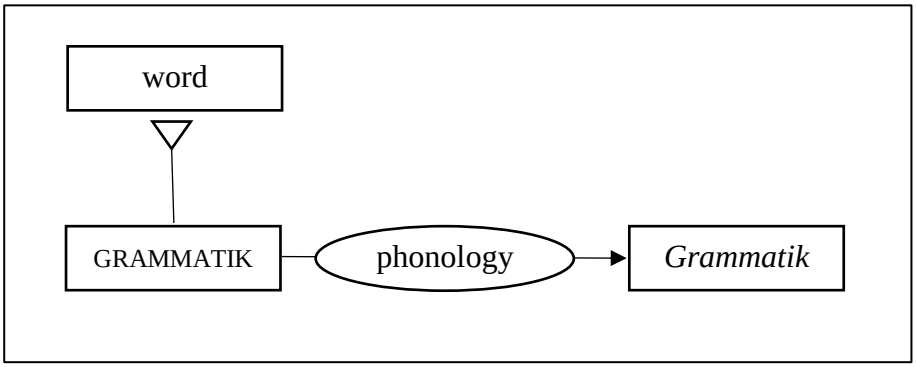


Figure 6: Grammatik *in a WG network 1*

The rest of the AVM translates quite smoothly (ignoring the potential list for SPR), giving Figure 7, though an actual WG analysis would be

rather different in ways that are irrelevant here.

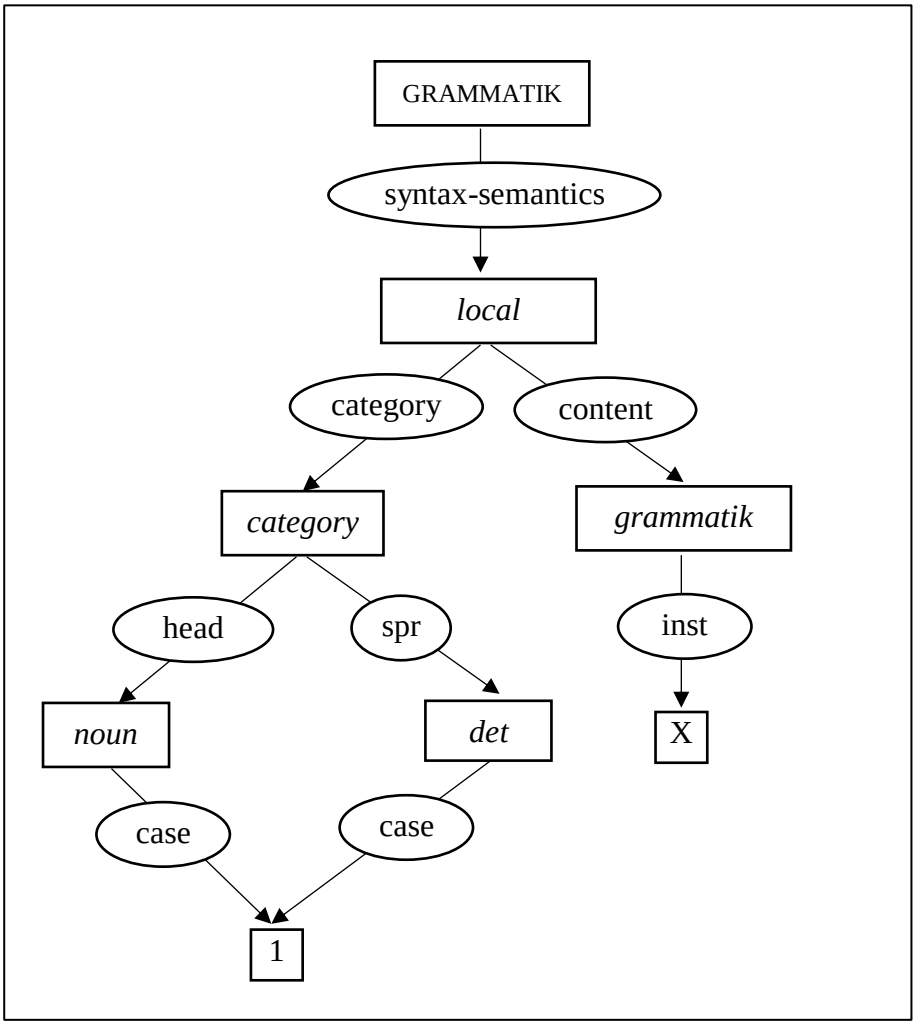


Figure 7: Grammatik *in a WG network 2*

The difference in notation may seem trivial, but it encourages very different ways of thinking about syntax, in terms of the boxes of a PS partonomy or the separate words of a DS analysis.

The other difference based on cognitive psychology between HPSG and WG is that many cognitive psychologists argue that concepts are built round prototypes (Rosch 1973; Taylor 1995), clear cases with a periphery of exceptional cases. This claim implies the logic of default inheritance (Briscoe, Copestake, and De Paiva 1993), which is popular in AI though less so in logic. In HPSG, default inheritance is accepted by some but not by others (Müller 2018, 403), whereas in WG it plays

a fundamental role, as I show in 4.1 below. WG uses the 'isa' relation to carry default inheritance, and avoids the problems of non-monotonic inheritance by restricting inheritance to node-creation (Hudson 2017). Once again, the difference is highly relevant to the comparison of PS and DS because one of the basic questions is whether syntactic structures involve partonomies (based on whole:part relations) or taxonomies (based on the 'isa' relation).

Another reason for discussing default inheritance and the isa relation is to explain that WG, just like HPSG, is a constraint-based theory. In HPSG, a sentence is grammatical if its structure unifies consistently with entries in the grammar; and in WG, it is grammatical if its word tokens can all be inherited, without overrides, from entries in the grammar.

This completes the tutorial on WG, so we are now ready to consider the issues that distinguish HPSG from this particular version of DS. In preparation for this discussion, we return to the four distinguishing assumptions in (1) to (4):

- Containment: in PS, but not in DS, if two items are directly related, one must contain the other.
- Continuity: therefore, in PS, but not in DS, all the items contained in a larger one must be adjacent.
- Asymmetry: in DS, but not in PS, a direct relation between two items must be asymmetrical, with one depending on the other (the head of the relation).
- Functions: therefore, DS, but not PS, recognises subtypes of dependency, viz the traditional grammatical functions (e.g.

'subject') as distinct relations.

These distinctions will provide the structure for the discussion:

- Containment and continuity:
  - Semantic phrasing
  - Coordination
  - Phrasal edges
  - Word order
  - Pied piping
- Asymmetry and functions:
  - Structure sharing and raising/lowering
  - Headless phrases
  - Complex dependency
  - Grammatical functions

4. Containment and continuity (PS but not DS)

4.1. Semantic phrasing

One apparent benefit of PS is what I call 'semantic phrasing' (Hudson 1990, 146–51), in which the effect of adding a dependent to a word modifies that word's meaning to produce a different meaning. For instance, the phrase *typical French house* does not mean 'house which is both typical and French', but rather 'French house which is typical (of French houses)' (Dahl 1980). In other words, even if the syntax doesn't need a node corresponding to the combination *French house*, the semantics does need one.

For HPSG, of course, this is not a problem because every dependent creates a new structure, semantic as well as syntactic; so the syntactic phrase *French house* has a 'content' which is 'French house'. But for DS theories, this is not generally possible because there is no syntactic node other than those for individual words – so, in this example, one node for *house* and one for *French* but none for *French house*.

Fortunately for DS, there is a solution: create extra word nodes but treat them as a taxonomy, not a partonomy (Hudson 2017). To appreciate the significance of this distinction, the connection between the concepts 'finger' and 'hand' is a partonomy, but that between 'index finger' and 'finger' is a taxonomy; a finger is part of a hand, but it is not a hand, and converely an index finger is a finger, but it is

not part of a finger.

In this analysis, then, the token of *house* in *typical French house* would be factored into three distinct nodes:

- *house*: an example of the lexeme HOUSE, with the inherited meaning ‘house’.
- *house+F*: the word *house* with *French* as its dependent, meaning ‘French house’.
- *house+t*: the word *house+F* with *typical* as its dependent, meaning ‘typical example of a French house’

These three nodes can be justified as distinct categories because each combines a syntactic fact with a semantic one: for instance, *house* doesn’t simply mean ‘French house’, but has that meaning because it has the dependent *French*. The alternative would be to add all the dependents and all the meanings to a single word node, thereby removing all the explanatory connections; this seems much less plausible psychologically. The proposed WG analysis of *typical French house* is shown in Figure 8, with the syntactic structure on the left and the semantics on the right.

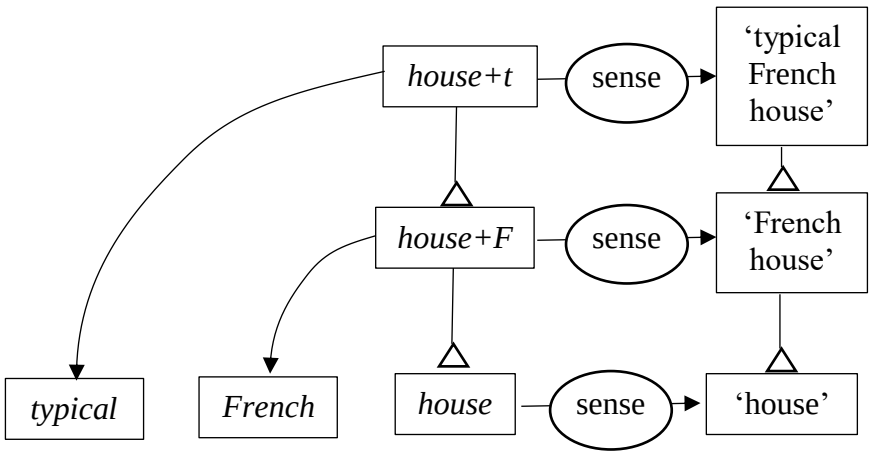


Figure 8: typical French house *in WG*

The number of syntactic nodes in this analysis is the same as in an HPSG analysis, but crucially these nodes are linked by the ‘isa’ relation, and not as parts to wholes – in other words, the hierarchy is a

taxonomy, not a partonomy.

4.2. Coordination

Another apparent argument for PS, and against DS, is based on coordination: coordination is a symmetrical relationship, not a dependency, and it coordinates phrases. For instance, in (8) the coordination clearly links *came in* to *sat down* and puts them on equal grammatical terms; and it is this equality that allows them to share the subject *Mary*.

(8) Mary came in and sat down.

But of course, *came in* and *sat down* are not syntactic items in a DS analysis, so we have a prima facie case against DS.

Fortunately, there is a solution: sets (Hudson 1990, 404–21). We know from the vast experimental literature (as well as from everyday experience) that the human mind is capable of holding ordered sets (strings) of words, so all we need to assume is that we can apply this ability in the case of coordination. The members of a set are all equal, so their relation is symmetrical; and the members may share properties (e.g. a person's children constitute a set united by their shared relation to that person and a multitude of other shared properties). Moreover, sets may be combined into supersets, so both conjuncts such as *came in* and *sat down* and coordinations (*came in and sat down*) are lists. According to this analsis, then, the two lists (*came, in*) and (*sat, down*) are united by their shared subject, *Mary*, and combine into the coordination ((*came, in*) (*sat, down*)). The precise status of the conjunction *and* remains to be determined. The proposed analysis is

shown in network notation in Figure 9.

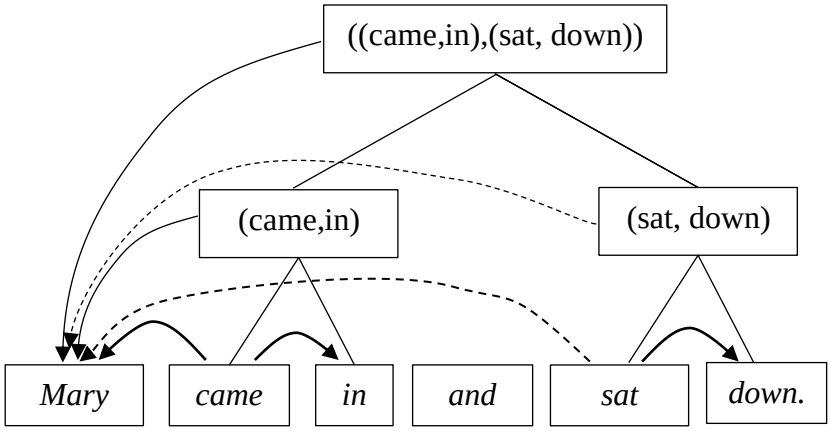


Figure 9: Coordination with sets

Once again, inheritance plays a role in generating this diagram, though the isa links have been omitted to avoid clutter: the dependency from *came* to *Mary* is inherited first by (came, in) and then by the whole coordination, from which it is inherited first by (sat, down) and then by *sat*.

The proposed analysis may seem to have adopted phrases in all but name, but this is not so because the analysis easily accommodates 'incomplete' conjuncts (Hudson 1982) precisely because there is no expectation that strings are complete phrases. This claim is born out by examples such as (9) (meaning '… and parties for foreign girls …').

(9) We hold parties for foreign boys on Tuesdays and girls on Wednesdays.

In this example, the first conjunct is the string (*boys, on, Tuesdays*), but the relevant phrases would be *parties for foreign boys* and *on Tuesdays*.

4.3. Phrasal edges

One of the differences between PS and DS is that PS formally recognises phrasal boundaries, and a PS tree can even be converted to a bracketed string where the phrase is represented by its boundaries. In contrast, standard DS implies phrases (since a phrase can be defined as a word and all the words depending on it either directly or

indirectly), but doesn't mark their boundaries.

This turns out to be problematic in dealing with Welsh soft mutation (Tallerman 2009). Tallerman's article is one of the few serious discussions by a PS advocate of the relative merits of PS and DS, so it deserves more consideration than space allows here. It discusses examples such as (10) and (11), where the underlined words are morphologically changed by 'soft mutation' in comparison with their underlying forms shown in brackets.

(10) Prynodd y ddynes delyn. (telyn)

buy.PAST.3S the woman harp

'The woman bought a harp.'

(11) Gwnaeth y ddynes [werthu telyn]. (gwerthu)

do.PAST.3S the woman sell.INF harp

'The woman sold a harp.'

Soft mutation is sensitive to syntax, so although 'harp' is the object of a preceding verb in both examples, it is mutated when this verb is finite (*prynodd*) and followed by a subject, but not when the verb has no subject because it is non-finite. Similarly, the non-finite verb 'sell' is mutated in example (11) because it follows a subject, in contrast with the finite verbs which have no mutation.

The standard PS explanation for such facts (and many more) is the 'XP Trigger Hypothesis': that soft mutation is triggered on a subject or complement (but not an adjunct) immediately after an XP boundary. The analysis contains two claims: that mutation affects the first word of an XP, and that it is triggered by the last word of another XP. The first claim seems beyond doubt: the mutated word is simply the first word, and not necessarily the head. Examples such as (12) are conclusive.

(12) Dw i [lawn mor **grac** â chi]. (llawn)

be.PRES.1S I full as angry as you

'I'm just as angry as you.'

The second claim is not beyond challenge; for instance, it relies on controversial assumptions about null subjects and traces in examples such as (13) and (14) (where *t* and *pro* stand for a trace and a null

subject respectively).

(13) Pwy brynodd t delyn? (telyn)

who buy.PAST.3S t harp

'Who bought a harp?'

(14) Prynodd pro delyn. (telyn)

buy.PAST.3S pro harp

'He/she bought a harp.'

But suppose both claims were true. What would this imply for DS? All it shows is that we need to be able to identify the first word in a phrase (the mutated word) and the last word in a phrase (the trigger). This is certainly not possible in WG as it stands, but the basic premiss of WG is that the whole of ordinary cognition is available to language, and it's very clear that ordinary cognition allows us to recognise beginnings and ends in other domains, so why not also in language? The Welsh data do not show that we need phrasal nodes complete with attributes and values. In other words, edge phenomena such as Welsh mutation show that DS needs to be expanded, but not that we need the full apparatus of PS. Exactly how to adapt WG is a matter for future research, not for this chapter.

4.4. Word order

In both WG and HPSG, dominance and linearity are separated, but this separation goes much further in WG. In basic HPSG, linearization rules apply only to sisters, and only determine whether the head precedes or follows its sisters (Müller 2018, 272). However this is obviously inadequate given the freedom of ordering found in many languages, so one proposal is an ordered list for each item of all its dependents(Müller 2018, 294). This proposal joins other analyses with dependent lists (Bouma, Malouf, and Sag 2001) which obviously bring HPSG much nearer to DS, though these dependent lists still contain no function labels.

WG takes the separation of linearity from dominance a step further by introducing two new syntactic relations dedicated to word order: 'position' and 'landmark', each of which points to a node in the overall network (Hudson 2017). As its name suggests, a word's landmark is the word from which it takes its position, and is normally

the word on which it depends (as in the HPSG list of dependents); this is what by default holds phrases together, because dependents keep as close to their landmarks as possible and a general principle bans intersecting landmark relations. Moreover, the word's 'position' relative to its landmark may either be free or defined as either before or after.

However, this default pattern allows exceptions, and because 'position' and 'landmark' are properties, they are subject to default inheritance such as raising and extraction (discussed in 5.1) and pied piping (in the next section). To give an idea of the flexibility allowed by these relations, we start with the very easy English example in Figure 10, where 'lm' and 'psn' stand for 'landmark' and 'position', and '<' and '>' mean 'before' and 'after'.

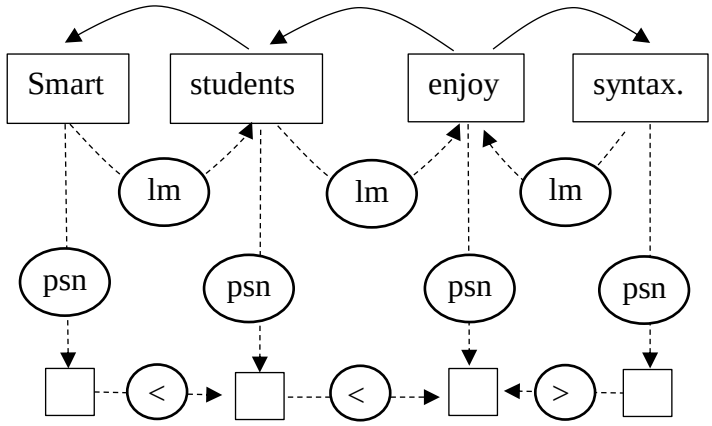


Figure 10: Easy word order in English

It could be objected that this is a lot of formal machinery for such a simple matter as word order. However, it is important to recognise that the conventional left-right ordering of writing is just a written convention, and that a mental network (which is what we are trying to model in WG) has no left-right ordering. Ordering a series of objects (such as words) is a complex mental operation, which experimental subjects often get wrong, so complex machinery is appropriate.

Moreover, any syntactician knows that language offers a multiplicity of complex relations between dependency structure and word order. To take an extreme example, non-configurational languages pose

problems for standard versions of HPSG, as illustrated by this Wambaya sentence (Bender 2008):

(15) Ngaragana-nguja ngiy-a gujinganjanga-ni
having-grog.ACC 3sg.past mother

jiyawu ngabulu.
give milk.ACC

'(His) mother gave (him) milk with grog in it.'

The simplified literal gloss shows that both 'having-grog' and 'milk' are marked as accusative, which is enough to allow the former to modify the latter in spite of their separation. The word order is typical of many Australian non-configurational languages: totally free within the clause except that the auxiliary verb (glossed here as '3sg.past') comes second (after one dependent word or phrase). Such freedom of order is easily accommodated if landmarks are independent of dependencies: the auxiliary verb is the root of the clause's dependency structure (as in English), and also the landmark for every word that depends on it, whether directly or (crucially) indirectly. Its second position is due to a rule which requires it to precede all these words by default, but to have just one 'preceder'. A simplified structure for this sentence (with Wambaya words replaced by English glosses) is shown in Figure 11, with dotted arrows below the words again showing landmark and position relations.

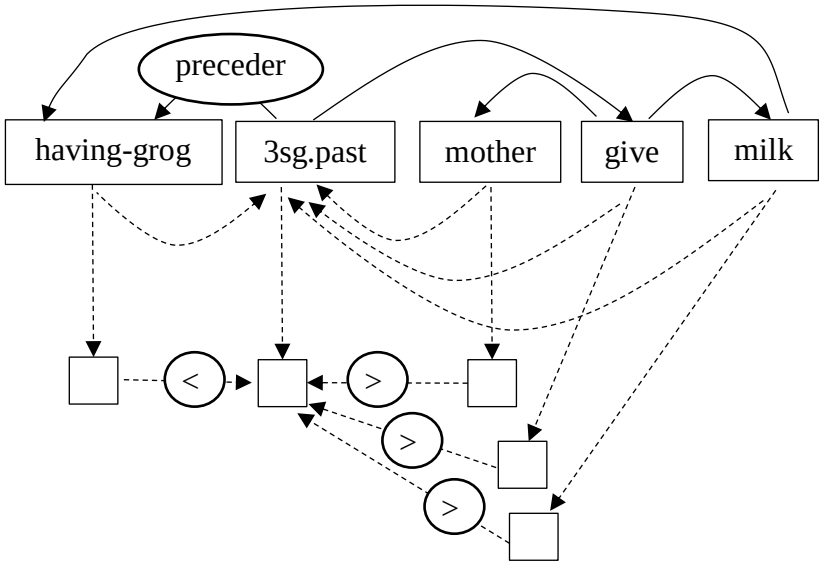


Figure 11: A non-configurational structure

Later sections will discuss word order, and will reinforce the claims of this subsection: that plain-vanilla versions of either PS or DS are woefully inadequate and need to be supplemented in some way.

4.5. Pied piping

The construction called ‘pied piping’ is unusual in that the position of a phrase is determined by a word inside it which is not its head. For example, in (16) the phrase *in which* is in the extracted position reserved for relative pronouns, but its head is the preposition *in*, not the relative pronoun *which*. In the metaphor based on the legend of the Pied-piper of Hamlyn, *which* has ‘pied-piped’ (or ‘lured’) the pronoun to accompany it to the front.

(16) This is the house from which she escaped.

The standard HPSG analysis of pied piping (Müller 2018, 302–3) uses feature-percolation whereby the ‘relative’ property of the relative pronoun percolates up to the phrase containing it. This has the unfortunate consequence of classifying *from which* as a relative pronoun, and, since one of the properties of relative pronouns is that they are linked anaphorically to the antecedent noun, this property should also percolate up to *from which*. But this is semantically

wrong, because *house* is the antecedent of *which*, not *from which*.

The WG analysis (Hudson 2017), in contrast, assigns exactly the same dependency structure to (16) as it does to (17).

(17) This is the house which she escaped from.

Since the dependencies guide the semantics, this also guarantees that the two examples have the same semantics. The crucial difference is merely a matter of word order, explained in terms of the landmark and position of *from* and *which*. The analysis is enabled by an extra relation 'pipee' which links *which* (as 'piper') to *from*, the word which takes over its expected landmark and position.

This completes the discussion of 'containment' and 'continuity', the characteristics of PS which are missing in DS. We have seen that the continuity guaranteed by PS is also provided by default in WG by a general ban on intersecting landmark relations; but, thanks to default inheritance, exceptions abound. Moreover, WG offers a great deal of flexibility in other relations: a word may be part of a string (as in coordination), its phrase's edges may need to be recognised structurally (as in Welsh mutation), and the word order may be mediated by ad hoc relations such as 'pipee'.

5. Asymmetry and functions

5.1. Structure sharing and raising/lowering

This section considers the characteristics of DS which are missing from PS: asymmetry and functions. Does syntactic theory need these notions? The notion of asymmetry is easy to defend in a discussion of HPSG, because the asymmetry of the head and its dependents is fundamental to the theory. But what about functions? We start with the functions involved in extraction, handled in both WG and HPSG by means of structure-sharing, described as 'the most important means of expression in HPSG' (Müller 2018, 267). For example, consider example (18), in which *where* is extracted from its expected position (as complement) after *are*.

(18) Where are you?

Figure 12 sketches two analyses both of which involve structure-

sharing, one in the spirit of HPSG and the other in a simplified version of WG. In the HPSG analysis, the structure sharing is shown by coindexing *what* either to a trace in its expected position, or to the relevant entry in the COMPLEMENTS list, while the WG analysis shows it as a double dependency, with *what* depending both on *did* and on *see*. (The function labels can be read as 'extractee', 'subject' and 'extractee and complement'.)

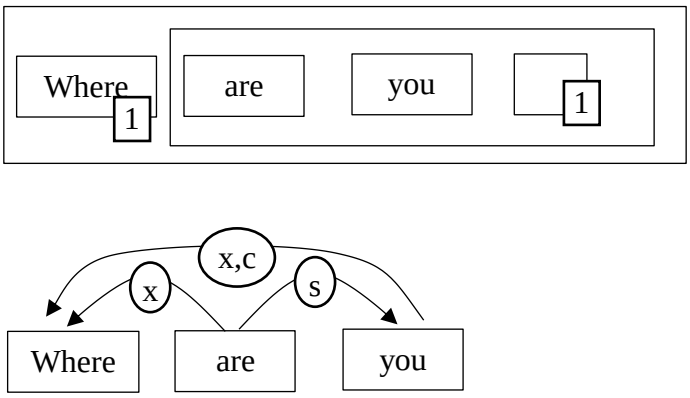


Figure 12: Extraction in HPSG and WG

In HPSG, *where* is linked to two positions but one is just a trace or just virtual so it's clearly not the surface word. But why must the overt item be in the extracted position? Since the two positions are co-indexed, in a sense they are the same and interchangeable. A similar question arises in WG, where there's just one word, *where*, and no clear reason why this should take its position before *are* rather than after it (as expected given that it is the complement of *are*). Why not **Are you where*?

The solution lies, once again, in default inheritance, but in this case the solution rests on the functional difference between complements and extractees. It also involves the multiple word tokens discussed in connection with semantic phrasing in 4.1: we recognise two distinct tokens of *where*, one functioning as complement of *are* (*where-c*) and one functioning as extractee (*where-x*); this is exactly like an HPSG analysis in which *where-c* is a virtual complement and *where-x* is the extracted item. Of these, the one that defines the expected, default, position is *where-c*, whereas *where-x* has an exceptional position. As with semantic phrasing, the analysis has to distinguish the word tokens because each brings together two properties: *where-c* combines

postverbal position with the complement function, but *where-x* combines preverbal position with the extractee function. If the two were simply collapsed into a single item these connections would be lost.

This analysis is shown in Figure 13, where *where-x* isa *where-c*. Since *where-x* is in position B relative to C, this overrides position A, and the surface position of *where* is explained.

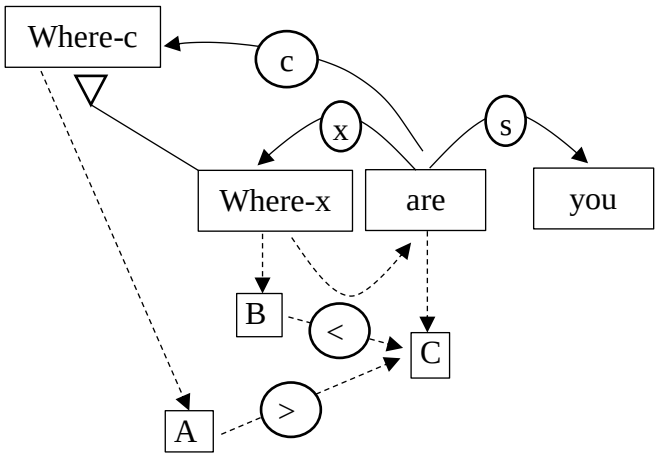


Figure 13: Extraction with default inheritance

This ‘raising’ configuration, in which basic dependencies provide default positions which is overridden by others (such as ‘extractee’), is part of the grammar, and indeed this ‘raising’ seems to be a default for grammars across languages, but default inheritance allows exceptions, and we do indeed seem to find exceptional cases of lowering such as German Partial VP Fronting (Müller 2018, 198; Hudson 2007, 143–44).

5.2. Headless phrases

Bloomfield assumed that phrases could be either headed (endocentric) or not (exocentric). According to WG (and other DS theories), there are no headless phrases. Admittedly, utterances may contain unstructured lists (e.g. *one two three four …*), and quotations may be unstructured strings, as in (19), but presumably no-one would be

tempted to call such strings ‘phrases’.

(19) He said “One, two, three, testing, testing, testing”

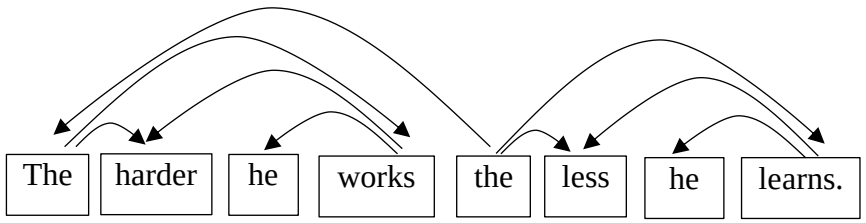
Is this claim tenable? There are a number of potential counterexamples including the following (Müller 2018, 403–5; Jackendoff 2008; Arnold and Borsley 2014):

(20) The rich get richer.
(21) The harder he works, the less he learns.
(22) In they came, student after student.
(23) However smart the students, a lecture needs to be clear.

Starting with (20), *the rich* is allowed by *the*, which has a special sub-case which allows a single adjective as its complement, meaning ‘generic people’; this is not possible with any other determiner. In a determiner-headed analysis this is unproblematic, and the head is *the*.

The comparative correlative in (21) is clearly a combination of a subordinate clause followed by a main clause (Culicover and Jackendoff 1999), but what are the heads of the two clauses? The obvious dependency links the first *the* with the second (hence ‘correlative’), so it is at least worth considering an analysis in which this dependency is the basis of the construction and, once again, the head is *the*. Figure 14 outlines a possible analysis, though it should be noted that the dependency structures are complex. The next section discusses such complexities.

Figure 14: A WG sketch of the comparative correlative



Example (22) is offered by Jackendoff as a clear case of headlessness, but there is an equally obvious headed analysis of *student after student* in which the structure is the same as in commonplace N P N examples like *box of matches*. The only peculiarity of Jackendoff’s example is the lexical repetition, which is beyond most theories of syntax. For WG, however, the solution is easy: default inheritance. This example

illustrates an idiomatic but generalisable version of the N P N pattern in which the second N isa the first and the meaning is special; as expected, the pattern is recursive.

The 'exhaustive conditional' or 'unconditional' in (23) clearly has two parts: *however smart* and *the students*, but which is the head? A verb could be added, giving *however smart the students are*, so if we assumed a covert verb that would provide a head, but without a verb it is unclear – and indeed this is precisely the kind of subject-predicate structure that stood in the way of dependency analysis for nearly two thousand years.

However, there are good reasons for rejecting covert verbs in general. For instance, in Arabic a predicate adjective or nominal is in different cases according to whether 'be' is overt: accusative when it is overt, nominative when it is covert. Moreover, the word order is different in the two constructions: the verb normally precedes the subject, but the verbless predicate follows it. In Arabic, therefore, a covert verb would simply complicate the analysis; but if an analysis without a covert verb is possible for Arabic, it is also possible in English.

Moreover, even English offers an easy alternative to the covert verb based on the structure where the verb BE is overt. It is reasonably uncontroversial to assume a raising analysis for examples such as (24) and (25), so (26) invites a similar analysis.

(24) He keeps talking.
(25) He is talking.
(26) He is cold.

But a raising analysis implies a headed structure for *he cold* in which *he* depends (as subject) on *cold*. Given this analysis, the same must be true even where there is no verb, as in example (23) *however smart the students* or (27).

(27) What, him smart? You're joking!

Comfortingly, the facts of exhaustive conditionals support this analysis because the subject is optional, confirming that the predicate is head:

(28) However smart, students have to be motivated as well.

In short, where there is just a subject and a predicate, without a verb,

then the predicate is the head.

Clearly it is impossible to prove the non-existence of headless phrases, but the examples considered have been offered as plausible examples, so if even they allow a well-motivated headed analysis, it seems reasonable to hypothesise that all phrases have heads.

5.3. Complex dependency

How complex can dependencies be? Is there a theoretical limit such that some geometrical patterns can be ruled out as impossible? Two particular questions arise:

- Can a word depend on more than one other word? This is of course precisely what structure sharing allows, but this only allows 'raising' or 'lowering' within a single chain of dependencies. Is any other kind of 'double motherhood' possible?
- Is mutual dependency possible?

The answer to both questions is yes for WG, but is less clear for HPSG.

Consider the dependency structure for an example such as

(29) I met Mary, who lives nearby.

Of interest are three dependencies:

- *who* depends on *Mary* because *who* needs a previous word as its antecedent; so *who* is an adjunct of *Mary*.
- *who* also depends on *lives* because it is the subject of *lives*.
- *lives* depends on *who*, because relative pronouns need a following finite verb, so *lives* is the complement of *who*; moreover, it is through *who* that *lives* is related to *Mary* and the rest of the sentence.

Each of these dependencies is quite obvious and uncontroversial when considered in isolation. The problem, of course, is that they combine in an unexpectedly complicated way; in fact, this one example illustrates both the complex conditions defined above: *who* depends on two words which are not otherwise syntactically connected (*Mary* and *lives*), and *who* and *lives* are mutually dependent. A WG analysis of the relevant dependencies is sketched in Figure 15 (where 'a', 's'

and ‘c’ are ‘adjunct’, ‘subject’ and ‘complement’).

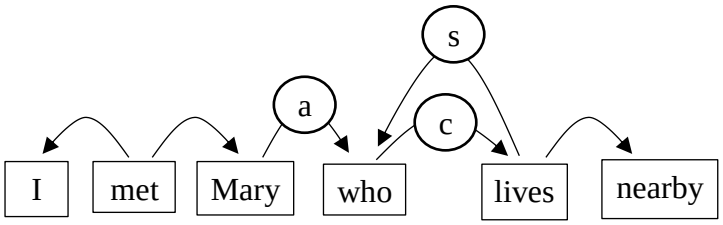


Figure 15: Complex dependencies in a relative clause

It might be argued that the mutual dependency can be avoided by invoking the idea of sub-tokens introduced earlier. But this is not in fact so because of a general principle that a word’s dependents are all fully specified – i.e. they are the most specific sub-tokens available. Thus the most fully specified tokens of *lives* and *who* are *lives+s* (as modified by its subject *who*) and *who+c* (as modified by its complement *lives*), so *who+c* and *lives+s* are mutually dependent.

The complexities of relative clauses have evoked a great deal of discussion in the HPSG community and at least some analyses seem to recognise that the relative pronoun depends on the antecedent as well as on a word inside the relative clause; for instance, Sag’s 1997 analysis of his example (33) coindexes it with the former and extracts it from the latter (Sag 1997). As for mutual dependency, it may be relevant to mention the mutual selection of determiner and noun (Müller 2018, 331).

And so we finish this review of complex dependencies by answering the question that exercised the minds of the Arabic grammarians in the Abbasid Caliphate: is mutual dependency possible? For plain-vanilla PS and DS the answer has to be no for the simple reason that the notation uses the vertical dimension to represent dominance, and geometry doesn’t allow two nodes each to be higher than the other. The arrow notation of WG removes this constraint, and allows the answer yes, but it would be much harder to use the boxes of HPSG to show mutual dependency.

5.4. Grammatical functions

As we have seen, WG, like other DS-based theories, can easily accommodate grammatical functions in the same way that it accommodates any relations: as relational categories, comparable in

general cognition with social relations and spatial relations. This being so, we can define a taxonomy of functions such as the one sketched in Figure 16 (which ends with 'the second complement of *from*', as in *from London to Edinburgh*, given as an example of an extremely specific function).

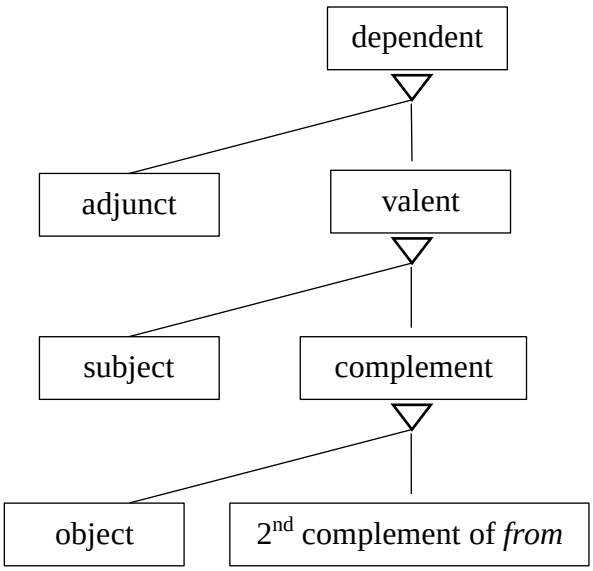


Figure 16: A taxonomy of grammatical functions

HPSG, however, shows its strong roots in the PS tradition by rejecting explicit grammatical functions in principle, although various versions of HPSG do recognise a number of attributes dedicated to functional information, and with potentially intersecting lists of values (Müller 2018, 266–67):

- SPECIFIER, which may include subjects in languages such as English.
- COMPLEMENTS, which may include subjects in languages such as German.
- ARGUMENT STRUCTURE, which includes all the arguments in the SPECIFIER and COMPLEMENTS lists.
- MODIFIED, containing the head item which can be, or is, modified by an adjunct.
- DEPENDENTS, the special list mentioned in 4.4 which

contains all the dependents of an item, including its adjuncts.

For more precise functional classification within a list, however, HPSG uses the category of the dependent combined with its position in the list; so, for example, the direct object in English might be the first NP in the COMPLEMENTS list.

The treatment of grammatical functions is strikingly different in the two theories:

- WG provides a rich and open-ended taxonomy of relational categories (for instance, 'direct object' isa 'object' isa 'complement' isa 'dependent') which are quite separate from the taxonomy of word classes.
- HPSG just distinguishes very general types of function, while finer distinctions depend on part-of-speech categories and position in a list.

Which is right? In particular, do we need to separate function from class? And do we need to separate it from order?

A classic challenge for any theory that merges function and class is the coordination of unlike categories, as in (30) to **Error! Reference source not found.**.

(30) They're very tired and in a bad mood. (AP + PP)
(31) He looks a nice guy and very friendly. (NP + AP)
(32) He told me he was rich and stuff like that. ('S' + NP)
(33) I'll put mine the other side or in the bowl. (NP + PP)

The standard explanation for such examples is that the functions concerned (e.g. predicative complement of BE) are available to different categories, so these categories may combine as a coordinated realisation of a single function; in other words, what these coordinated items share is their function, and not their classification. The problem is very familiar to HPSG users, as witness the 2006 survey by Chaves (Chaves 2006). One popular HPSG solution (supported by Chaves) is to assume that the coordination is actually sentential but with shared material removed by ellipsis, but unlike categories can be coordinated even when the semantics forbids sentential coordination, as in (34).

(34) We have to choose between here and the other side.

Another familiar argument for distinct functions is that some complements must be defined in terms of their meaning rather than in terms of their syntactic classification. For example, the verbs PUT and

BEHAVE both require a semantically defined complement, defined as a location for PUT and a manner for BEHAVE.

(35) I put it *(on the shelf / there / somewhere else / where I'd be able to find it easily).

(36) He behaved *(badly / in his usual way / as he usually does / the way he usually does).

Purely semantic selection fits poorly into a system which only mentions syntactic class-membership.

Finally, the policy of leaving functions implicit seems to prevent important generalisations. For example, every English finite verb has a subject (which arguably may be covert if the verb is non-finite or imperative). How can this generalisation be expressed if the subject is simply the first item in an ordered list? In contrast, a WG analysis allows each function to be handled by an appropriate generalisation. Figure 17 shows how subjects and objects are inherited, one from the 'verb' category and the other from the particular lexeme. (The '1' in a square bracket is an abbreviation for an additional link from the node to its numerosity, so '1' means 'is obligatory'.)

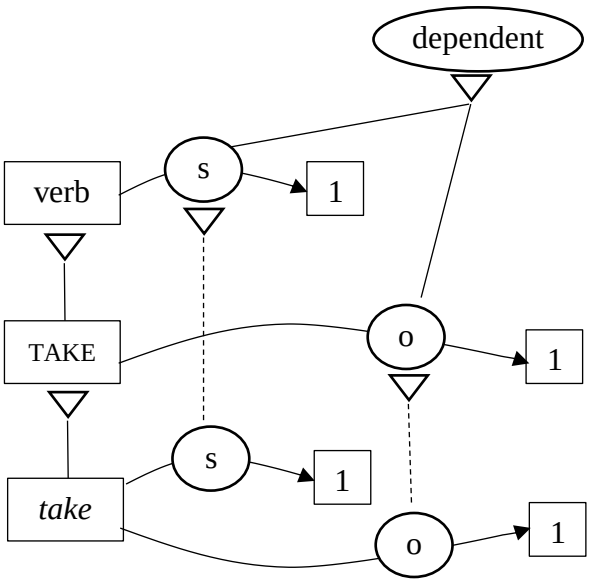


Figure 17: How subjects and objects are inherited in WG

The question of ordering arguments is more difficult, as this is an area

where DS lags behind. In 4.1 I argued that 'semantic phrasing' was needed and that this could be provided by creating a taxonomy of 'sub-tokens' each defined by one dependency; but that discussion only looked at nouns with multiple adjuncts such as *typical French house*. If this kind of analysis is extended to verbs, we need to know how to order the sub-tokens created by subjects, direct objects and indirect objects (not to mention a multiplicity of other types of dependent). For instance, in (37) we recognise at least three sub-tokens: *bought+M*, *bought+F* and *bought+s*, but how are these arranged in a taxonomy?

(37) Mary bought Fred a scarf.

There do seem to be strong reasons (such as the lack of idioms with a fixed subject and variable object) for treating the subject as the 'external argument', the one which modifies the results of all the other arguments. This ranking is expressed in the HPSG ARGUMENTS list, but DS theories have no comparable mechanism.

However, an ordered list of arguments is problematic because it's unclear how the other arguments should be ordered (Müller 2018, 284). Should the next argument in the list (after the subject) be the indirect object (*Fred*), because it comes next in the sentence and is easier to passivize, or the direct object (*a scarf*), because it's more likely to form idioms with the verb and more accessible to processes such as relativization? Given this uncertainty, we have to leave future research to sort out how arguments (and other dependents) should be ranked.

6. HPSG without PS?

This chapter raises a fundamental question for HPSG: does it really need PS? Introductory textbooks present PS as an obvious and established approach to syntax, but the historical sketch showed very clearly that nearly two thousand years of syntactic theory assumed DS, not PS, with one exception: the subject-predicate analysis of the proposition (later taken to be the sentence). Even when PS was invented by Bloomfield, it was combined with elements of DS, and Chomsky's PS, purified of all DS elements, only survived from 1957 to 1970.

The challenge for HPSG, then, is to explain why PS is a better basis than DS. The debate has hardly started, but suppose the debate favoured DS; would that be the end of HPSG? Far from it. It could

survive almost intact, with just two major changes.

The first would be in the treatment of grammatical functions. All the objections in 5.4 could be answered by treating each argument-type as a separate attribute, though there may also need to be two lists: one containing all adjuncts, and the other like the extra list proposed earlier in which all the dependents are listed in order.

The other change would be the replacement of phrasal boxes by a single list of words. Here is a list for the example with which we started (with round and curly brackets for ordered and unordered sets, and a number of sub-tokens for each word):

(38) ({*smart, smart+h*} {*students, students+a*}, {*enjoy, enjoy+o, enjoy+s*}, *syntax*)

Each word in the analysis stands for a whole box of attributes which include isa links between tokens and sub-tokens, and also syntactic dependency links to other words in the set. The internal structure of the boxes would otherwise look very much like standard HPSG, as in the following schematic neo-HPSG structure (where the brackets for sets are replaced by the left-right and top-down dimensions of the diagram):

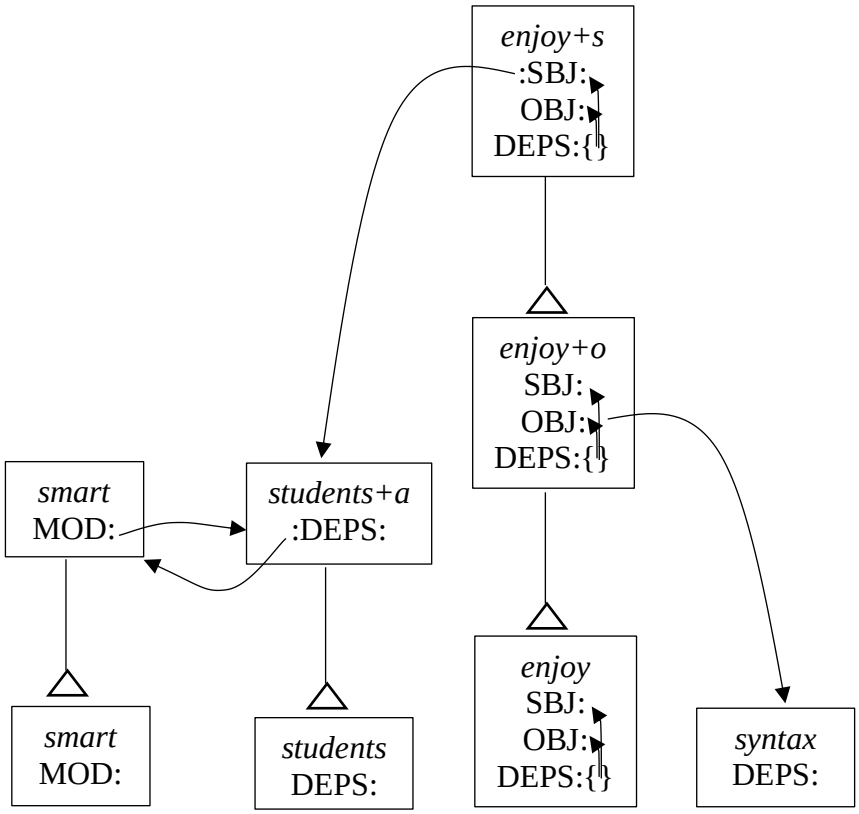


Figure 18: A neo-HPSG analysis

References

Arnold, Douglas, and Robert Borsley. 2014. “Arnold, D. and R. D. Borsley (2014), On the Analysis of English Exhaustive Conditionals.” In *Proceedings of the HPSG14 Conference, Stanford*, edited by Stefan Müller, 27–47. CSLI Publications.

Bender, Emily. 2008. “Radical Non-Configurationality without Shuffle Operators: An Analysis of Wambaya.” In *Proceedings of the 15th International Conference on Head-Driven Phrase Structure Grammar*, edited by Stefan Müller, 6–24. CSL! Publications.

Billroth, Johann. 1832. *Lateinische Syntax Für Die Obern Klassen*

Gelehrter Schulen. Leipzig: Weidmann.

Blake, Barry. 1990. *Relational Grammar*. London: Croom Helm.

Blevins, James, and Ivan Sag. 2013. “Phrase Structure Grammar.” In *The Cambridge Handbook of Generative Syntax*, edited by Marcel den Dikken, 202–25.

Bouma, Gosse, Robert Malouf, and Ivan Sag. 2001. “Satisfying Constraints on Extraction and Adjunction.” *Natural Language and Linguistic Theory* 19: 1–65.

Briscoe, Ted, Ann Copestake, and Valeria De Paiva. 1993. *Inheritance, Defaults, and the Lexicon*. Studies in Natural Language Processing. Cambridge England: Cambridge University Press.

Chaves, Rui. 2006. “Coordination of Unlikes without Unlike Categories.” In *Proceedings of the 13th International Conference on Head-Driven Phrase Structure Grammar, Held in Varna, Bulgaria*, edited by Stefan Müller, 102–22. Stanford: CSLI Publications.

Chen, Danqi, and Christopher Manning. 2014. “A Fast and Accurate Dependency Parser Using Neural Networks.” In *Proceedings of EMNLP (Empirical Methods on Natural Language Processing) 2014*, edited by Alessandro Moschitti, Bo Pang, and Walter Daelemans.

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

———. 1970. “Remarks on Nominalizations.” In *Readings in Transformational Grammar*, edited by R Jacobs and Peter Rosenbaum, 184–221. Waltham, MA: Ginn.

Culicover, Peter W., and Ray Jackendoff. 1999. “The View from the Periphery: The English Comparative Correlative.” *Linguistic Inquiry* 30 (4): 543–71. https://doi.org/10.1162/002438999554200.

Dahl, Östen. 1980. “Some Arguments for Higher Nodes in Syntax: A Reply to Hudson’s ‘Constituency and Dependency’.” *Linguistics* 18: 485–88.

Dik, Simon. 1989. *The Theory of Functional Grammar: The Structure of the Clause*. Dordrecht: Foris.

Eppler, Eva. 2004. “The Syntax of German-English Code-Switching.”

UCL.

Gisborne, Nikolas. 1996. “English Perception Verbs.” UCL, London.

Halliday, Michael. 1961. “Categories of the Theory of Grammar.” *Word* 17: 241–92.

———. 1967. “Notes on Transitivity and Theme in English. Part 2.” *Journal of Linguistics* 3: 199–244.

Hudson, Richard. 1982. “Incomplete Conjuncts.” *Linguistic Inquiry* 13: 547–50.

———. 1984. *Word Grammar*. Oxford: Blackwell.

———. 1990. *English Word Grammar*. Oxford: Blackwell.

———. 2007. *Language Networks: The New Word Grammar*. Oxford: Oxford University Press.

———. 2010. *An Introduction to Word Grammar*. Cambridge: Cambridge University Press.

———. 2017. “Pied Piping in Cognition.” *Journal of Linguistics Online*, 1–54. https://doi.org/10.1017/S0022226717000056.

Imrényi, András. 2013. “Constituency or Dependency? Notes on Sámuel Brassai’s Syntactic Model of Hungarian.” In *VLlxx. Papers Presented to László Varga on His 70th Birthday.*, by Péter Szigetvári, 167–82. Budapest: Tinta.

Imrényi, András, and Zsuzsa Vladár. forthcoming. “Sámuel Brassai in the History of Dependency Grammar András Imrényi –.” In *The Origins of Dependency Syntax. A Historical Survey from Panini to Tesnière.*, edited by András Imrényi and Nicolas Mazziotta. John Benjamin.

Jackendoff, Ray. 2008. “Construction after Construction and Its Theoretical Challenges.” *Language* 84: 8–28.

Jespersen, Otto. 1937. *Analytic Syntax*. Chicago: University of Chicago Press.

Kahane, Sylvain. forthcoming. “How Dependency Syntax Appeared in the French Encyclopedia: From Buffier (1709) to Beauzée (1765).” In *The Origins of Dependency Syntax. A Historical Survey from Panini to Tesnière.*, edited by András Imrényi and Nicolas Mazziotta.

Amsterdam: John Benjamins.

Kern, Franz. 1884. *Grundriss Der Deutschen Satzlehre.* Berlin: Nicolaische Verlags-Buchhandlung.

Kouloughli, Djamel-Eddine. 1999. “Y A-t-Il Une Syntaxe Dans La Tradition Arabe.” *Histoire Épistémologie Langage* 21: 45–64.

Kübler, Sandra, Ryan McDonald, and Joakim Nivre. 2009. “Dependency Parsing.” *Synthesis Lectures on Human Language Technologies* 2: 1–127.

Law, Vivien. 2003. *The History of Linguistics in Europe: From Plato to 1600*. Cambridge: Cambridge University Press.

Luhtala, A. 1994. “Grammar, Early Mediaeval.” In *Encyclopedia of Language and Linguistics*, edited by Ronald Asher, 1460–68. Oxford: Pergamon.

Marneffe, Marie-Catherine de, Timothy Dozat, Natalia Silveira, Katri Haverinen, Filip Ginter, Joakim Nivre, and Christopher Manning. 2014. “Universal Stanford Dependencies: A Cross-Linguistic Typology.” In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC).*, 4585–92.

Matthews, Peter. 1993. *Grammatical Theory in the United States from Bloomfield to Chomsky*. Cambridge: Cambridge University Press.

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

Müller, Stefan. 2018. *Grammatical Theory: From Transformational Grammar to Constraint-Based Approaches. Second Revised and Extended Edition.* Textbooks in Language Sciences 1. Berlin: Language Science Press.

Owens, Jonathan. 1988. *The Foundations of Grammar: An Introduction to Mediaeval Arabic Grammatical Theory*. Amsterdam: Benjamins.

Percival, Keith. 1976. “On the Historical Source of Immediate Constituent Analysis.” In *Notes from the Linguistic Underground.*, edited by James McCawley, 229–42. London: Academic Press.

Perlmutter, David, and Paul Postal. 1983. “Some Proposed Laws of Basic Clause Structure.” In *Studies in Relational Grammar 1*, 81–128.

Chicago: University of Chicago Press.

Pike, Kenneth. 1954. *Language in Relation to a Unified Theory of the Structure of Human Behavior.* Glendale, CA: Summer Institute of Linguistics.

Reed, Alonzo, and Brainerd Kellogg. 1877. *Higher Lessons in English. A Work on English Grammar and Composition, in Which the Science of the Language Is Made Tributary to the Art of Expression. A Course of Practical Lessons Carefully Graded, and Adapted to Every Day Use in the School-Room. (1890 Edition)*. New York: Clark and Maynard.

Reisberg, Daniel. 2007. *Cognition. Exploring the Science of the Mind. Third Media Edition*. New York: Norton.

Robins, Robert. 1967. *A Short History of Linguistics*. London: Longman.

Rosch, Eleanor. 1973. "Natural Categories." *Cognitive Psychology* 4: 328–50.

Sag, Ivan. 1997. "English Relative Clause Constructions." *Journal of Linguistics* 33 (2): 431–83.

Sériot, Patrick. 2004. "L'affaire Du Petit Drame : Filiation Franco-Russe Ou Communauté de Pensée? (Tesnière et Dmitrievskij)." *Slavica Occitania, Toulouse*, 93-118, .

Sgall, Petr, Eva Hajicová, and Jarmila Panevova. 1986. *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Pro Functional-Generative Description. Prague: Academia.

Siewierska, Anna. 1991. *Functional Grammar*. London: Routledge.

Sweet, Henry. 1891. *A New English Grammar Logical and Historical. 2 Vol.* Oxford: Clarendon Press.

Tallerman, Maggie. 2009. "Phrase Structure vs. Dependency: The Analysis of Welsh Syntactic Soft Mutation." *Journal of Linguistics* 45: 167–201.

Taylor, John. 1995. *Linguistic Categorization: Prototypes in Linguistic Theory.* 2nd ed. Oxford: Clarendon.

Tesnière, Lucien. 1959. *Éléments de Syntaxe Structurale*. Paris:

Klincksieck.

———. 2015. *Elements of Structural Syntax*. Translated by Timothy Osborne and Sylvain Kahane. Benjamins.

Zwicky, Arnold. 1985. “The Case against Plain Vanilla Syntax.” *Studies in the Linguistic Sciences* 15: 1–21.