23

Dependency Grammar

Richard Hudson

1 The Dependency Tradition

Dependency grammar (DG) is the modern continuation of the European tradition of grammatical analysis stretching back to classical antiquity, in contrast with the much shorter American tradition of constituency-based grammar which started with Leonard Bloomfield and was formalized by Noam Chomsky as Phrase Structure Grammar. This European tradition is also the source of Word-and-Paradigm (WP) morphology, so DG is normally combined, simply as a matter of history, with WP, in contrast with the American tradition which sees morphology simply as syntax below the word.

For example, take sentence (1).

(1) Small babies cried.

In the European grammatical tradition, strongly influenced as it was by analyses of Latin and Greek, the basic unit of grammar within the clause is the word, so there are three units of grammar: the words *small*, *babies*, and *cried*. Syntax can relate one word directly to another, so these words are all that is needed for the syntax, which allows *small* to modify *babies*, and *babies* (rather than *small babies*) to act as subject of *cried*. The word also defined the domain of morphology, so the relation between *cried* and the morphs *cry* and *-ed* is handled by morphology (traditionally called "accidence"), which is sharply separated from syntax. This tradition effectively forces the recognition of morphological structure as a distinct linguistic level, contrasting with syntax as well as with morphology. So the European tradition, including DG, combines three different assumptions:

I should like to thank Igor Mel'cuk for comments on an earlier draft of this paper.

- Dependency structure (DS): units are related by asymmetrical dependency rather than by part-whole relations.
- Lexical integrity: the units of syntax are words, not morphs.
- Morphological structure: morphology defines a distinct level of analysis.

On the other hand, although these assumptions tend strongly to be combined for historical (and geographical) reasons, they are logically separate, and indeed we shall see below that some recent work does separate them, and that the traditional packaging may in fact be unhelpful.

Since the main defining characteristic of DG is an approach to syntax which challenges one of the main assumptions underlying the mainstream PS tradition, we must start with an introduction to DS. The history of the choice between DS and PS has been ably summarized as follows by Mel'cuk, one of the leading proponents of dependency-based analysis:

> Syntactic dependency has been used to talk about the structure of sentences from Antiquity and throughout the Middle Ages to modern times. All respectable pre-20th century grammatical traditions in syntax have been based on it, as has much language teaching. By the 8th century, Arab grammarians (e.g. Sibawaih, who died in 798) already distinguished the governor vs. the dependent in syntax and used this distinction when formulating rules for word order and inflection . . . . One finds dependency trees as a means of describing sentence structure in German syntax books from the 19th century . . . . In point of fact, constituency representation in syntax, i.e. what became known as phrase-structure, was first introduced . . . in the early 20th century. The dependency approach was properly presented for the first time in Tesnière 1959 . . .; this profound treatise made syntactic dependency available for serious theorizing. Yet, due to the dominance of Chomskian Transformational-Generative Grammar—which used, as its main syntactic tool, the phrase-structure representation (i.e. constituency)—the [dependency] approach did not become popular in modern linguistics until the beginning of the 1980s. (Mel'cuk 2003: 22)

The difference between PS and DS emerges from a very simple question about our earlier example: What is the subject of *cried*, and what is it the subject of? According to PS, the entire noun phrase *small babies* is the subject of the entire clause *small babies cried*, but for DS, the subject relation involves just two words: the noun *babies* and the verb *cried*. The DS analysis gives no "official" status to the noun phrase and the clause, although they are implicit in the analysis; conversely, the PS analysis gives no explicit recognition to the relations between the words *babies* and *cried*, although this is again implicit. Each system of analysis has its own standard diagramming system: the PS tree and the DS "stemma" illustrated in Figure 23.1.

In both analyses the essential relations are asymmetrical, and in both diagramming systems the vertical dimension is used, iconically, to represent this asymmetry. In PS the essential relation is the part-whole relation between a smaller "constituent" and the larger one of which it is a part,

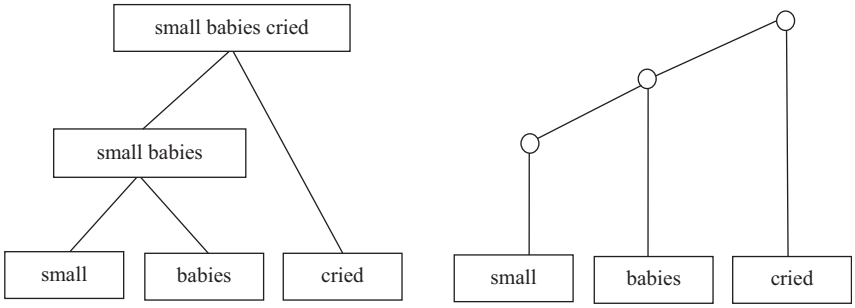

Figure 23.1. A PS tree and a DS stemma

whereas in DS it is the dependency relation between one word and the word on which it depends. (Unfortunately there is no generally accepted name for the latter word, so I shall use my own preferred term, "parent"; so in (1) *babies* is the parent of *small*, and *cried* is the parent of *babies*.)

However, it is important to recognize that DG is no more unified than PSG. The only unifying feature is, in fact, DS, and as we shall see below, at least one version of DG rejects the other two traditional partners of DS: lexical integrity and morphological structure. The DG tradition includes a large number of distinct theories (listed in §4), which are divided by deep theoretical issues. One of these concerns the notation for DS. One objection to the standard diagramming systems is that they place irrelevant limitations on our theoretical options. Since there is only one vertical dimension, each one only allows one asymmetrical relation. But what if there are good theoretical reasons for assuming more than one? This situation would arise in two cases:

- If we need to combine PS and DS analysis by recognizing a direct relation between two words in addition to some part-whole constituency relation.
- If we need to recognize mutual dependency between two words.

Although both PS and DS are widely assumed to preclude these possibilities, it may in fact be premature to rule them out—and indeed there is some evidence for both cases: evidence (discussed in §2) that Welsh mutation requires phrases to be recognized as such (Tallerman 2009) and evidence (reviewed briefly in §5) that extraction (in English as in other languages) with wh-pronouns requires mutual dependency (Hudson 2003b).

In short, we must abandon the vertical dimension for at least one of these relations, and my own preferred solution is to indicate asymmetrical dependency relations with arrows, as in Figure 23.2. These diagrams also show how a DS analysis may be enriched with functional labels such as "adjunct" or "subject". The diagram on the left is pure DS, while the one on the right combines DS and PS. Even if this combined analysis is wrong, its wrongness should be established by theoretical and empirical debate rather than by being ruled out a priori by the notation.

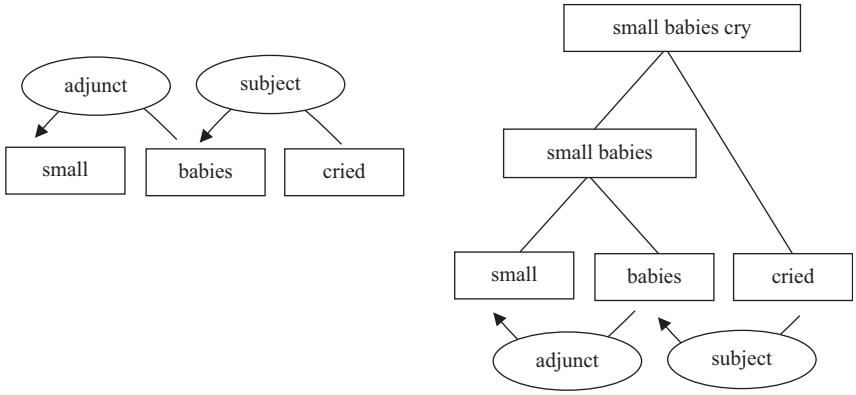

Figure 23.2. A dependency structure using arrows

The relative merits of PS and DS have hardly been debated in the theoretical literature; this is largely because of the social and numerical dominance of the PS community, which has by and large ignored the arguments for DS. This attitude was encouraged by some early articles which claimed that the two approaches were weakly equivalent (Hays 1964; Gaifman 1965; Robinson 1970)—a claim which is certainly true for some very limited versions of PS and DS, but which equally certainly does not apply to all possible versions. For example, a "non-projective" version of DS theory which allowed stemma lines to cross would certainly not be weakly equivalent to a standard version of PS theory. Similarly, and as suggested above, a version of DS might allow mutual dominance, which is completely impossible in PS. The following list summarizes some of the main arguments in favor of DS:

- Simplicity: A DS analysis requires just one node per word, so there are no phrasal nodes (unless, of course, DS is combined with PS).
- Word order: A DS analysis, as such, is neutral as to word order so it needs to be supplemented with some other system for fixing word order where needed; but this incompleteness is also a source of flexibility for accommodating different kinds of word-order systems. In contrast, word order is, at least in principle, built into PS, so some kinds of word-order phenomena such as "scrambling" are challenging even given the Generalized Phrase Structure Grammar separation of "immediate dominance" and "linear precedence" within a single phrase.
- Semantics: Syntactic dependencies are similar to the semantic roles with which they are paired, whereas PS requires part-whole relations to be mapped onto classified semantic roles such as "agent" or "instrument."
- Headedness: For DS, every construction is endocentric by definition, because every construction is defined by dependency relations between its head word and the other words that depend on it. In contrast, PS

accommodates exocentric constructions as easily as endocentric, and it is only by theoretical stipulation that X-bar syntax bans the former. (Indeed, X-bar syntax already combines PS with DS by importing the latter's asymmetry into a PS analysis.) Unlike word order, therefore, DS is more restrictive than PS in that it requires headedness.

- Valency: For DS, most of syntax (or even the whole of syntax) is controlled by the syntactic requirements of individual words (including the requirements they inherit from more general categories). In the DS tradition, these requirements are called "valency" (a term introduced by Tesnière and applied in the first instance just to verbs), but the notion is similar to traditional government and modern subcategorization or functional completeness. The advantage of DS in this area is that valency requirements follow dependencies—in fact, the two notions are mutually defining—whereas PS does not provide any direct structural link between a word and the words that it governs.
- Psychology: The basic assumption of PS is that part-whole relations are the only relations that are available for syntax. Taken as a psychological claim this is absurd, given the wide range of relations that we are capable of recognizing outside language (e.g., between people). But if we are mentally able to recognize dependencies between words, why not allow such relations in syntactic theory (Hudson 2003a)?

In short, the arguments in favour of DS rather than PS are strong, and deserve to be taken much more seriously than they have been. Unfortunately, in the absence of proper debate, the choice between the two approaches is currently made largely on social and geographical grounds, with PS dominating in the USA and other anglophone countries and DS much stronger in Europe—a distinction which is interestingly reminiscent of the split in morphology between the American 'Item and Arrangement' model and the European 'Word and Paradigm' model (Robins 1959).

2 Morphological Evidence for Phrase Structure in Syntax

On the other hand, there does appear to be some evidence from morphology that DS may not be enough on its own. This case is based on so-called "edge phenomena" in morphology, where some syntactic property of an entire phrase is signalled by a purely morphological phenomenon applied to a word which is located on the edge of the phrase. Clearly, if the phrase's edge is a crucial element in the analysis, then to that extent the phrase itself is represented, though it is an open question whether anything more than the edges is ever needed. One prediction of "pure" DS (i.e., DS without PS) is that such phenomena are impossible, but there is good evidence for edge phenomena in some languages.

The crucial characteristics of the phenomena concerned are these:

- Some word W undergoes a change of form when it is on the edge of a syntactically relevant phrase.
- W is not always the head of the phrase.
- The change to W cannot be analyzed as the result of combining it with a clitic that is syntactically outside the phrase (as with the English "group genitive" in *the King of England's daughter*).
- Nor can the change to W be explained in terms of regular phonological sandhi processes (such as assimilation or apocope).

According to these criteria, a number of languages offer clear evidence for edge phenomena (Anderson et al. 2006). One such language is Welsh, where the choice between PS and DS has been discussed in detail by Tallerman (2009).

The relevant phenomenon is soft mutation (e.g., the change from *llawn* 'full' to *lawn*, or from *dwy* 'two' to *ddwy*), which according to a widely accepted analysis called the "XP trigger hypothesis" applies to the first word of one complement phrase when it immediately follows another (Borsley et al. 2007). This immediately satisfies the first criterion for edge phenomena by applying to a word on the edge of a phrase. The second criterion is also satisfied because, although the mutated word usually is the head of this phrase, it need not be, as witness examples such as (2) and (3) (from Tallerman 2009). As in Tallerman's article, the mutated word is underlined and its unmutated form is shown in brackets after the sentence. The brackets show the boundaries of the mutated phrase.

(2) Dw i [lawn mor grac â chi] (llawn)
be. PRES.1S I [full as angry as you]
'I'm just as angry as you.'

(3) Rhoddodd Elen y delyn [ddwy droedfedd i ffrwdd] (dwy)
put. PAST.3S Elen the harp [two foot away]
'Elen put the harp two feet away.'

The significance of these examples is that it would be virtually impossible to justify an analysis in which the mutated word was the head of its phrase; this is why soft mutation passes the second criterion. The third criterion is easy: there is nothing at all additive about mutation, so once again there is nothing to be said for an analysis in terms of a clitic external to the mutated phrase. And finally, mutation is definitely not due to regular phonological processes because it is strongly influenced by the lexical and syntactic context.

In short, syntactic soft mutation in Welsh is a clear example of a morphological edge phenomenon. That being so, it is strong evidence that a grammar needs to be able to refer to the edge of a phrase so, at least in that sense, it calls for the apparatus of PS. However, this

conclusion needs to be interpreted with care. The evidence does not prove any of the following propositions:

- that DS is not needed. After all, the arguments in favour of DS still hold true, so if anything the evidence favors a mixed theory which allows both DS and PS to coexist in the same structure (as hinted at in Figure 23.2).
- that the whole of PS is needed. In its familiar form, PS is equivalent to a labeled bracketing. Edge phenomena show the need for brackets (or equivalent) but not for labeling—in other words, for phrasal categories as such; so it is easy to imagine a marriage of DS and PS to which the latter contributes nothing but the edge marking.
- that edge marking is needed in every language. It seems likely that although phrasal edges are relevant to a grammar of Welsh, they are irrelevant and therefore non-existent in a grammar for English; and, in psychological terms, that Welsh speakers define phrasal edges and pay attention to them, but English speakers do not.

This discussion of morphological edge phenomena shows how important morphology can be to discussions of syntax, and in particular, how important it can be for the choice between DS and PS. We now turn to influence working in the other direction: how the choice between DS and PS affects theoretical choices in morphology.

3 Morphology as a Distinct Level of Analysis

One of the most fundamental choices facing any theory of morphology is whether to recognise morphology as za distinct level of analysis between syntax and phonology (Chapters 8, 10). This involves two separate questions about morphological patterning: Is it distinct from syntactic patterning? And is it distinct from phonological patterning? The first of these questions is about the word: are words sufficiently similar to sentences to require the same kind of analytical apparatus?

According to PS, a sentence consists of "constituents" of various sizes, all related to their parts in the same way. In this hierarchical structure, words are included among the constituents, and since they appear to have parts—their constituent morphs[1]—it is reasonable to contine the analysis down inside the word. For example, in our earlier example *Small babies cry*, it is tempting to see the relation between the suffix *-s* and the word *babies* as an example of the same relation as the one between *babies* and *small babies*. In that analysis, words have no special status in the grammar; they are simply phrases, albeit small ones. Of course, it is also possible to reject this

[1] I use the term *morph* rather than *morpheme* in order to stress that these units are concrete rather than abstract. A morph is part of a word which has a position before or after other morphs and is realized phonologically.

conclusion, as many morphologists do, in favour of some version of the Word-and-Paradigm approach, so that PS applies just to sentence structure while some other model applies to word structure. But the point is that this decision is not required by the assumptions of PS.

In DS, on the other hand, words do have a very special status because (in the absence of phrases) they are the only units of grammar (Chapter 13). There is no hierarchy based on size in which words might be middle-sized units. DS is essentially a theory about how words are related to one another—a "bottom-up" theory about how words can be combined to form sentences, in contrast with the "top-down" PS theory of how sentences could be broken down into smaller parts. The special status of words in grammatical theory presumably reflects their special psychological status which allows even small children to insert word-spaces in their writing with very little teaching, and which also allowed words to be isolated very early in the history of writing (as in the word lists and verb paradigms of the Babylonian grammarians of four thousand years ago: Gragg (1994)). Indeed, one of the salient facts about the Western grammatical tradition is its lack of interest in morphological structure, in contrast with the attention lavished on sentence structure (Robins 1967: 56). Instead of dividing words into morphs, the Greek and Latin grammarians simply listed typical examples ("paradigms"), leaving learners to do their own analysis.

These assumptions about the place of words in DS lead logically to a single answer to our question: morphology is different from syntax. Of course, it is possible to ignore or deny this logic. In the PSG tradition the opposite position is occupied by Distributed Morphology (Chapter 15), and I shall explain in Section 4 that DS has sometimes been combined with the assumption that syntax in some sense includes morphology. However this assumption flies in the face of research evidence for "morphology-free syntax" (Chapter 9)—the idea that syntax is blind to purely morphological structure such as morphs (Zwicky 1992; Corbett 2009)—not to mention all the evidence in favour of the Word-and-Paradigm model.

One way to view the different implications of DS and PS would see DS as relevant to syntax, with PS as the model for morphology. After all, if *babies* is a unit of grammar where morphology and syntax meet, then one could see it as the whole of which the morphs {baby} and {z} are the parts, even if it is not itself a part of any larger unit. In this view, then, PS would be an appropriate model for the relation between words and morphs, with DS applying to the relations between words.

However, this simple view faces a serious objection: the part-whole relation requires the parts and the whole to exist on a similar level of abstraction. For example, a book can be viewed either as a physical object or in terms of its content. As a physical object, its parts are sheets of paper and a cover; but as content, its typical parts are a table of contents, a number of chapters, a list of references and an index. Each of these is a

different "representational redescription" of the same book (Karmiloff-Smith 1994), with its own vocabulary and properties. Returning to morphology, the idea that a morph could be part of a word is like treating a sheet of paper as part of a book's content. The Word-and-Paradigm approach assigns morphs and words to different levels of abstraction, with quite different properties. For example, whereas words are classified as nouns, verbs, and so on, morphs are classified as roots or affixes; and whereas words have meaning, morphs do not. Applied to English, this allows the {z} morph to occur, with completely different consequences, in both nouns and verbs. And in Latin, one of the languages for which WP was developed, it allows properties such as "first-person singular" to be expressed by many different indivisible morphs, such as the {o}, {am}, {i}, and {or} in *am-o* 'I love,' *am-āb-am* 'I used to love,' *am-av-ī* 'I have loved,' and *am-or* 'I am loved,' each of which expresses a number of other distinctions at the same time.

This argument leads to the conclusion that PS is not, in fact, a suitable model for relating words to morphs. Words and morphs exist on different levels of abstraction, so morphs cannot be parts of words. Instead, their relation must be an example of realization, the relation between one entity and another which makes it more "real"—less abstract (Stump 2001). However, this conclusion does not follow from the assumption of DS; as emphasized earlier, DS is independent of assumptions about lexical integrity and morphological structure.

On the other hand, there clearly is a sense in which {z} is part of a larger unit, even if this larger unit is morphological rather than a syntactic word. Indeed, one of the striking characteristics of morphological structure is its rigidity, which is just what we might expect from a whole which holds all its parts closely together. It seems, therefore, that we have a combination of relations: realization relating a word to its morphological structure, and part-whole relations between this structure and its parts. These part-whole relations call for an analysis in terms of PS, with larger units consisting of smaller ones. The structure of *babies* must therefore be something like Figure 23.3, where the curved arrows labelled "r" (or "realization") show realization and the numbered straight arrows show parts. The dotted lines represent the relation between the syntactic word "BABY, plural" (the plural of BABY) and the more general categories BABY and "plural"; these dotted lines anticipate the "is-a" relation of Figure 23.8.

Once again we have to ask whether PS and DS are mutually exclusive. Is there any evidence of a dependency relation between the morphs? At one time a popular answer was that every word has an internal dependency structure in which the right-most element was always the head—the Right-hand Head Rule (Williams 1981). The claim is supported in part by English morphology because we use suffixes rather than prefixes for changing word classification, so in *babies* the suffix shows that the word is plural and in

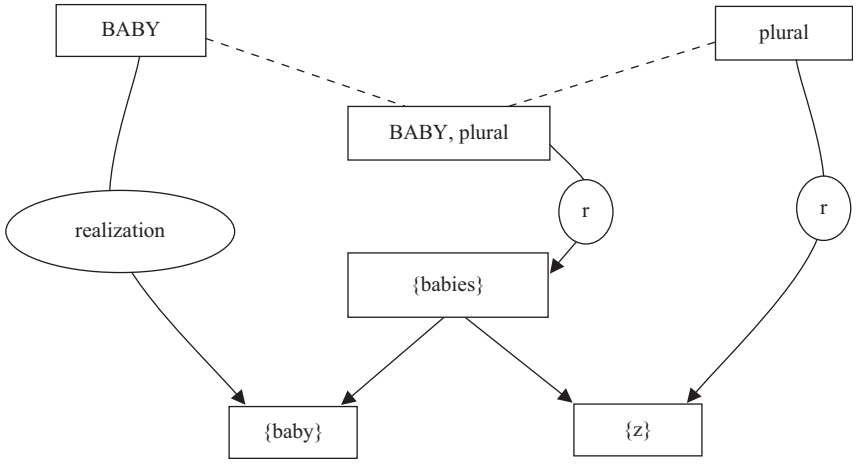

Figure 23.3. Syntax and morphology of *babies*

exploration the suffix indicates that it is a noun. If head-hood is merely a matter of word class and inflection, the claim is correct for English; but it faces a number of serious objections even when applied to English (Borer 1998). Perhaps the most important objection is that it confuses syntax with morphology. In purely morphological terms, the direction of dependency between a root and a suffix is clearly the reverse of the one claimed: the suffix must depend on the root. After all, that is precisely what "suffix" means—a secondary element added to the primary one, incapable of acting on its own, and with its position defined in relation to the primary one. The simple fact is that {baby} can occur without {z}, but {z} cannot occur without some lexical form to support it. The fact that {z} indicates a syntactic category is irrelevant to its status in morphology. After all, precisely the same is true of {baby}, which signals a different syntactic category (noun); and in an inflected verb such as *takes*, it is the lexeme realized by {take} that supplies all the syntactically vital information about valency.

This discussion already offers one argument for a dependency relation between affixes and roots, where the asymmetry is somewhat similar to that between a syntactic dependent and its parent in terms of autonomy, selection, and position. A part-whole analysis on its own does not show this asymmetry any more than PS did before the X-bar version incorporated asymmetry from DS. However, we can go further by exploring the selection relations among morphs. Take nominalization in English, for example. English has a number of alternative suffixes which are used for changing adjectives into nouns, as in *sincerity* and *gentleness* (Huddleston and Pullum 2002: 1702–3). These two words illustrate the two main alternatives: {ity} and {ness}. In a few words these are both possible (e.g., *clarity*, *clearness*), but in general they are selected by the morphological structure of the base to which they are added. In particular, there is a strong selection relation

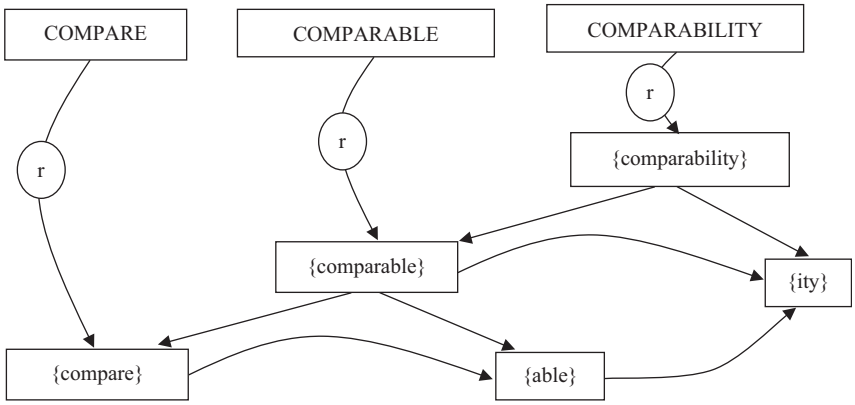

Figure 23.4. The morphological structure of *comparability*

between {able} and {ity}, as in *ability*, *capability*, *comparability*. Not only do we tend to favour {ity} after {able}, but we have a general morphophonological rule which realizes this pair of morphs as *ability*. The significance of these observations is that they conflict with a strictly PS analysis because the morph {able} is not a sister of {ity} in a word such as *capability*, where the sister of {ity} is the complex form {capable}; so in a word like this, {able} should not be able to influence {ity}. Consequently, the strong relation between these two morphs must be a dependency, coexisting with the PS structure. Figure 23.4 shows a possible analysis of the word *comparability*. The main point to note in this diagram is that it shows three dependencies, two of which point to {ity}; so this morph depends both on {comparable} and on {able}. This leads to a very general theoretical claim: that morphological structure requires both part-whole relations and dependency relations in order to reveal the complex relations among morphs, in addition to the realization relations between morphological patterns and words.

Another area of morphology where dependencies appear at first to be relevant is compounding, where single-word morphological compounds such as *postman* (meaning 'man who delivers post') are very similar in meaning and structure to two-word pairs such as *grocery man* ('man who delivers groceries'), which are clearly related by dependency. The need for dependencies is especially clear in English because of the very unclear boundary between syntactic pairs and morphological compounds. For example, Google displays a typical level of uncertainty when asked to define *soap box* (with a space) by offering a definition of *soapbox* (no space) which, for good measure, includes the expression *soap-box*. At the same time, Google is aware of the differences because it gives a different frequency of occurrence for each of the three alternatives *soap box*, *soapbox*, and *soap-box*. However, it is quite unclear what this uncertainty tells us about morphological structure. Even in *soap box,* it is the words, not the morphs, that are linked by dependency; the morphs {soap} and {box} are simply juxtaposed, like any other pair of morphs realizing adjacent words. This being so, there

is no reason to assume a morphological dependency in *soap-box* or *soapbox*, or more generally between the parts of any compound word.

4 Theories of Dependency Grammar

The discussion so far has focused on very general ideas about syntactic structure and how they impact on morphology. After the general outline of the differences between DS and PS in Section 1, Section 2 argued that even if DS is preferred to PS in general, there is morphological evidence that at least the edges of phrases are sometimes an essential part of a structural analysis, implying a theory of syntax based on DS supplemented with ideas from PS. Section 3 then argued that the choice of DS in syntax leads naturally to a theory in which morphology is a distinct level of analysis mapped onto syntactic words by realization rules. In contrast with syntax, morphology needs PS for the obvious part-whole relations, but once again we actually need both approaches because PS needs to be supplemented with dependencies between morphological elements.

Given these important links between the DS/PS choice in syntax and the theory of morphology, one might expect advocates of dependency syntax to have developed morphological theories to match their theories of syntax. Unfortunately this is not the case, and most work on DS syntax has failed to develop a matching theory of morphology. I know of only three exceptions:

- recent work by Gross and Osborne within Catena theory
- longstanding work by Mel'cuk and others within the Meaning-text Model
- longstanding work by Hudson and others with Word Grammar.

To contextualize this work, I start with a brief review of DS-based syntactic theories.

Just like PS, DS has been developed in a large number of different directions, so there are almost as many competing theories based on DS as there are on PS. Here is an indicative but incomplete list:

- Generative Dependency Grammar (Vater 1975; Diaconescu 2002)
- Case Grammar (Anderson 1977)
- Functional Generative Description (Sgall et al. 1986)
- Lexicase (Starosta 1988)
- Tree-adjoining Grammar (Joshi and Rambow 2003)
- Link Grammar (Sleator and Temperley 1993)
- Catena Theory (Osborne et al. 2012)
- Meaning Text Theory (Mel'cuk 1974, 1988, 2004, 2009)
- Word Grammar (Hudson 1984, 1990, 2007, 2010)

The theories listed here are general academic theories of language structure, but DS is also widely used in computational linguistics (Kunze 1975;

Debussmann 2006; Kübler et al. 2009; Chapter 29 in this volume) and in books for school-teaching (Heringer 1993).

The last three theories listed are the ones where we find the most extended discussion of morphology, but (as noted in Gross and Osborne 2013, footnote 4) it has often been suggested that the minimal units of syntax are in fact morphs, or at least that syntactic rules can refer directly to morphs (Chapter 8). As I mentioned earlier, this assumption sits much more comfortably with PS than with DS; indeed, it is notable that the very ancient tradition based on DS paid little attention to morphological structure as such (in contrast with the paradigms illustrating different morphological patterns). However, the influence of PS reaches even into discussions of DS, so this PS-based assumption is accepted explicitly in many recent discussions of DS, including the first one, Hays 1964. The same is true of Catena Theory.

Catena theory is based on the term catena, which in this context means a chain of units which are linked by dependency. In syntax, the units are words, and the theory develops various claims about the role of catenae in the analysis of idioms, ellipsis, and other patterns (e.g., Osborne et al. 2012). The theory's main claim about morphology is that morphs are related by dependencies in the same way as whole words are, with roots depending on affixes as in some PS-based analyses. The examples in Figure 23.5 are taken from Gross and Osborne 2013, but with their stemmas replaced by my horizontal arrows (pointing, as usual, towards the dependent). The third example shows how the morphological structure is simply part of the syntactic structure. (The discontinuities created by this analysis may allow the same raising analysis as purely syntactic discontinuities, but they are not discussed explicitly.)

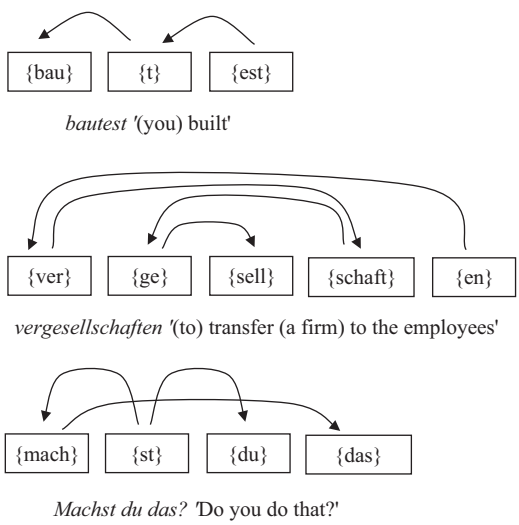

Figure 23.5. Morphological structure with catenae

As a theory of morphology, Catena Theory faces the same objections as similar theories based on PS (including the Right-hand Head Rule discussed earlier).

- The analysis of morphological structure is driven too strongly by the needs of syntax. Whenever the syntax needs a morphosyntactic category, this must be carrried by a separate morph, regardless of the actual morphological facts. For example, if German infinitives are marked as such by the presence of the morph {en}, what about English infinitives? Must we assume a zero morph, which is never realized phonologically? And what about nonconcatenative morphological patterns such as umlaut or mutation, or the multiplicity of morphological markers found in German plural nouns?
- Moreover, the choice of relevant syntactic properties is arbitrary. Although some syntactic properties of a word are reflected by its affixes, others are determined by its lexical root. This is particularly true of government, where notions such as valency and transitivity apply to the lexeme, not to the inflectional categories. For example, the German verb *bautest* 'you built' is not only past-tense, second-person singular (as indicated by its inflectional endings), but it is also transitive, so it allows an object. Why should the first property determine the word's internal structure, but not the second?

So far, the morphological sub-theory of Catena Theory has only been sketched in published form, so further work may address some of these objections.

In contrast, Mel'cuk has developed and described Meaning Text Theory (MTT) extremely thoroughly over several decades (Mel'cuk 1988, 1992–2000, 2003; Kahane 2004), and his work gives considerable attention to morphology (Mel'cuk 2008). One of the most impressive characteristics of this work is the range of languages and phenomena that it covers. I cannot do justice here to the depth and breadth of the work, so the focus will be on more general "architectural" features of the theory.

Meaning Text Theory assumes that a sentence's complete analysis can be broken down into seven distinct, but related, structures, including one for semantics, two for syntax, two for morphology, and two for phonology. In each pair, one is "deep" and the other is "surface," so unlike most theories of morphology, MTT provides two distinct structures, one oriented toward syntax and the other toward phonology. However, the teminology is misleading for outsiders, because word order is treated entirely as a matter of deep morphology, so MTT's morphology includes an important part of what is elsewhere called "syntax"—linear order, but no dependencies. The following example of deep morphology is based on Mel'cuk's analysis of his example *The people's support for the prime minister amazes Mr Bumbo-Yumbo* (Mel'cuk 2008: 10). Figure 23.6 shows Mel'cuk's deep morphological structure on the top line, and his surface morphological structure on the

Figure 23.6. Morphology in MTT

second line. It can be seen that deep morphology shows morphologically unanalyzed words, identified in terms of lexical and morphosyntactic categories, while surface morphology interprets these categories in terms of morphemes, some of which are zero morphemes. It is only the surface morphological representation that clearly counts as morphological representation in the usual sense.

MTT is relevant to the present chapter because the two syntactic levels are both organized in terms of dependencies. In contrast, neither of the morphological levels shows dependencies, which illustrates the profound differences described earlier between syntax and morphology. Indeed, Mel'cuk's notation for surface morphology assumes a part-whole relation between morphs and words, as suggested in Section 3. He also provides a rich specialized apparatus for changing one morphological pattern into another such as a rule for changing Tagalog nouns into adjectives by shifting stress (e.g. *álam* 'knowledge' becomes *alám* 'known'):

$$A_{PASSADJ} = \langle \langle \text{which underwent } \ldots \rangle ; / - \text{'} - / \Rightarrow / - - \text{'} / ; \sum = \text{applies to a noun, } \ldots \rangle$$

(Mel'cuk 2008: 304). These morphophonological rules are clearly unique to morphology, and in general MTT morphology is a clear example of a theory which combines DS in syntax with a distinct level of morphology.

On the other hand, MTT also recognizes "morphological dependency" as one of three main types of dependency: semantic, syntactic, and morphological (Mel'cuk 2003). Semantic dependencies exist on the level of semantics, and correspond to what others call semantic roles or argument structure. Similarly, syntactic dependencies apply to syntactic structures and are a mixture of universal and parochial syntactic relations such as "subjectival" and "determinative." Morphological dependencies, however, mix syntax, semantics, and morphology by linking the morphology of one word to another word that determines it by agreement, government, or "congruence"—the agreement between an anaphor and its antecedent. In most versions of DS, agreement and government are taken as evidence for syntactic dependency relations, but MTT distinguishes such relations from those called "syntactic." It is unclear how this terminological move improves the analysis; nor is it clear that agreement and government have anything in common with anaphor-antecedent relations, the third type of morphological dependency.

Returning to the MTT theory of morphology proper, its rich typological basis means that it offers a very useful summary of the phenomena that any theory needs to accommodate. It also offers a model of how to analyze morphology separately from syntax, and how this is encouraged by combining it with syntax based on DS. This separation of morphology from syntax is a strength. Arguable weaknesses of the theory include a lack of attention to morphological structure as such, beyond some rather simple part-whole structures, and a traditional lack of interest in similarities between morphology and anything outside language. Is morphology really totally *sui generis*? And in particular, is morphology similar to other things that we hold in our minds? To some extent of course it is a matter of personal taste whether or not to consider mental structures, but language is clearly a mental phenomenon so in the long run we must reconcile our theories of morphology with general pyschological theory. The next section argues that this is possible and that a certain amount has already been achieved.

5 Network-based Morphology: Word Grammar

One of the clearest findings that emerges from syntax is that syntactic structure is complex—much too complex for the simple versions of either PS or DS without massive enrichment of some kind. For example, "raising" structures like (4) contain one word which has two parents (or, in PS terms, two mothers).

(4) It stopped raining.

In this example the word *it* clearly depends (as subject) on *stopped*, but it must also depend (again as subject) on *raining*, because it satisfies the latter's need for *it* as subject. Another kind of complexity is mutual dependency, which is found (among other places) in examples like (5) and (6).

(5) Who came?
(6) I wonder who came.

The pronoun *who* is clearly the subject of *came* in both examples, but in (6) it is equally clearly dependent on *wonder*, a verb which requires some kind of question word as its complement. But if *who* depends on *wonder*, *came* must depend on *who*. In short, *who* and *came* depend on each other (Hudson 2007: 142).

This kind of complexity calls for a formal analytical apparatus which is much more flexible and powerful than the strictly hierarchical machinery of either standard PS or standard DS. These elementary models can be enriched in a number of ways, and one of the main items on the research agenda of theoretical linguistics has been the best way to enrich available models, with transformations and enriched feature structures among the

options considered. One particularly productive avenue of exploration has been the idea that syntactic structure is in fact a network rather than a hierarchy. This idea has generated interesting work in the branch of mathematics called graph theory (Ferrer i Cancho et al. 2004; Barabasi 2009; Solé et al. 2010), as well as in research on neural networks (Onnis et al. 2006). It also meshes well with the idea that we may use the same mental apparatus for syntax as we do for other areas of our mental life, given that we clearly use networks in general cognition. After all, one of the main conclusions of cognitive psychology is that memory is a network where each node is associated with many others (Reisberg 2007).

If our mental representations for syntax are indeed networks, then we can ask what the relations and units are without ruling out any possibilities on formal grounds. This idea provides the basis for two models of morphology: Network Morphology and Word Grammar. These two theories have a great many similarities (not least the central place they both give to the logic of default inheritance), but since Network Morphology is fully described in Chapter 18 in this volume I shall concentrate here on Word Grammar (WG: Hudson 1984, 1990, 2007, 2010; Gisborne 2010; Duran-Eppler 2011), which also has the distinction of combining a rich theory of morphology (presented in the previous references as well as in Rosta (1997), Creider and Hudson (1999), Camdzic and Hudson (2007), Hudson (forthcoming), Gisborne (forthcoming-a, forthcoming-b) with a DS-based theory of syntax. The following account of WG includes some changes to the previously published versions of the theory in the light of the earlier arguments.

The earlier sections of this article concluded that both syntax and morphology need to combine DS with PS, though in different proportions: whereas syntax needs DS supplemented with parts of PS, morphology needs PS supplemented from DS. The main reason why morphology needs PS is that a typical word needs a single morphological realization, a single unit of morphology related (by realization) to a single unit of syntax. This morphological unit is a word-form, so just as the lexeme CAT is realized by the morph {cat}, the combination of CAT with 'plural' (called CAT, plural) is realized by the word-form {cats}. (In WG notation, {. . .} is used for morphology just as /. . ./ is for phonology, so any unit of morphology is written between braces regardless of its internal complexity.) We now have two morphological units, {cat} and {cats}, but the latter clearly contains the former, alongside another unit written {z}. In short, {cats} contains two parts, so the analysis requires PS. But, in contrast with Distributed Morphology, the morphological structure involves the word-form {cats} while the syntactic structure involves the word "CAT,plural"; the two meet via the realization relation between CAT,plural and {cats}, but they are completely separate.

On the other hand, there is also an asymmetrical relation between these parts whereby the affix depends on the root; like any syntactic dependency,

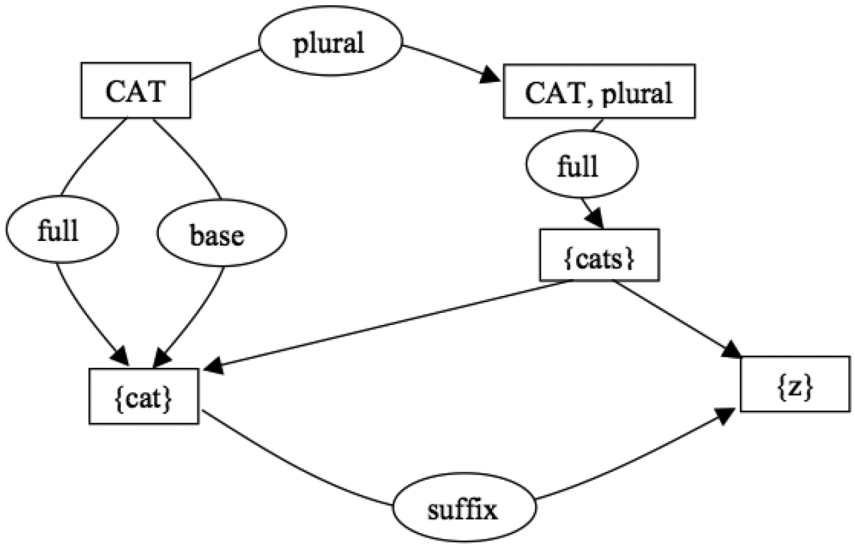

Figure 23.7. The morphology of *cats* in Word Grammar (1)

this morphological dependency brings together at least two properties, which in this case are:

- selection: {cat} selects {z} (in contrast with irregular noun roots such as {mouse} or {analysis})
- linear order: {z} is a suffix, so it follows {cat}.

Figure 23.7 shows this small area of the total morphological network for English, with simplifications that will be discussed below. The label "full" means "fully inflected form"—a particular kind of realization relation, contrasting (in this case) with the more familiar "base." In words, the plural of CAT is CAT,plural, whose fully inflected form {cats} has two parts: {cat} and its suffix, {z}; {cat} is also, by default, both the base of the lexeme CAT and its fully inflected form. Every link and every unit can be justified, so at least this part of the grammar must have a much richer structure than a mere hierarchy—in short, it must be a network.

But although this network is not itself organized hierarchically, it clearly coexists with a purely hierarchical structure which provides the generalizations that allow and explain the network. For example, CAT is an example of a noun, and CAT,plural is an example of a plural noun, with the default morphology of this category. Every unit is related to a more general category from which it inherits properties; and the same is true of every relation. For example, both "full and "base" are examples of the more general "realization" relation. This hierarchical organization is shown in Figure 23.8, where the small triangle signals an "is-a" relation; for example, a noun is a word, CAT is a noun, and so on. As in Network Morphology, the logic of generalization is default inheritance, so exceptions are handled with ease (Hudson forthcoming).

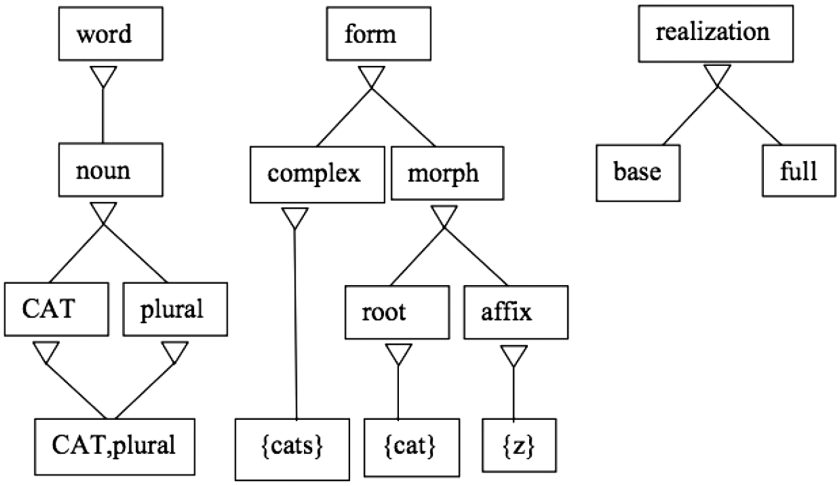

Figure 23.8. Three hierarchies in WG

This very simple example already illustrates most of the formal apparatus of WG—a hierarchy of units ranging from the most general (word) to the most particular (CAT,plural), where each unit is also part of a non-hierarchical network which relates it to other units, and where each relation (other than is-a) is hierarchically related to other more or less general relations. Moreover, the formal apparatus is paired with an equally simple model of processing and learning, including default inheritance, node-creation and binding, and a model of activation and how it spreads and affects these processes (Hudson 2010). These cognitive elements play an important part in the grammar, but the formal apparatus can be seen in its own right simply as a theory of how language is organized.

One of the potential attractions of WG for morphologists is that it is not simply a theory of morphology; in fact, it is primarily a theory about syntax, so the interface with syntax is already available for inspection. This is important because morphology and syntax obviously have to be reconciled sooner or later, and there are areas of research where they have to be reconciled immediately. The obvious example is the analysis of clitics, which I discuss briefly below. On the other hand, it has to be admitted that WG has no proper theory of phonology (though it may be compatible with some existing theories such as Dependency Phonology: van der Hulst 2006). Nor does it have the typological range of MTT, having been developed mainly on the basis of English and other western European languages (with occasional glances at Semitic and Cushitic). But perhaps the most important research gap is the lack of a computer implementation comparable with the DATR-based implementation available for Network Morphology. The potential interactions in a network really need to be tested by computer, and especially so if they include cognitive patterns such as activation.

Turning to clitics, the WG theory of clitics has evolved through a series of studies (Hudson 2001, forthcoming; Camdzic and Hudson 2007), so the following account reflects the latest theory rather than a single unchanging theory. The leading idea is that clitics are full syntactic words which, exceptionally, are realized by affixes (Chapter 2). Like other affixes, these need a "host," but in WG the host is a larger morphological structure which contains them rather than a syntactic word that they "lean" on (as suggested etymologically by the term *clitic*), so in (7) the suffix {ll} is part of the morphological complex {cats'll}.

(7) The cats'll eat it.

This is similar to the structure for *cats*, where {z} is part of {cats}, except that where {cats} is the realization of a syntactic word, {cats'll} is not. This is a purely morphological entity called a "hostform" which does not realize anything. Since the ordinary word order rules apply, all the grammar has to say about {ll} is that its host takes the realization of the previous word as its "anchor" (and as its first part). This structure is shown in Figure 23.9. which omits most of the morphological structure for *cats* as shown in Figure 23.7.

The syntactic structure is shown at the top of the diagram and the morphological at the bottom, so the two are mapped onto one another by the realization arrows labeled "f" (for "fully inflected form"). In the syntax, the root of the sentence is the finite verb WILL (subclassified here as "reduced") with THE as its subject and the infitive of EAT as its subject complement. These two dependents have further dependents of their own (CAT,plural as complement of THE, and IT as object of EAT,infin); but this is a raising structure, so THE is the subject of DO as well as of WILL. In the morphology, almost every word is realized by a full word-form, with the exception of WILL, which (because of its classification as "reduced") is realized by a mere affix{ll}. This brings

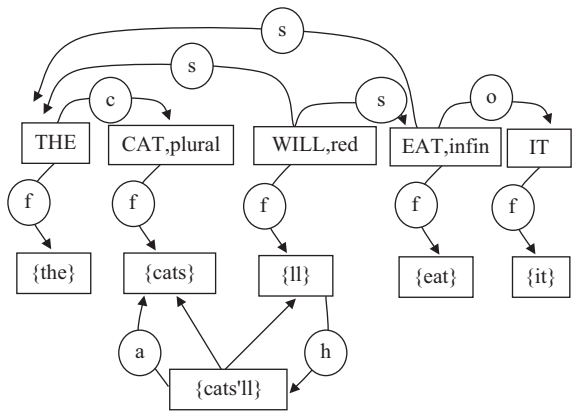

Figure 23.9. WG syntax and morphology for a simple clitic

with it a host, consisting of {ll} and the anchor, which is whatever happens to be the preceding word. The diagram omits some details, such as the distinction between tokens and types (which in fact plays an important part in WG theory) and all the apparatus for linear order.

The theory can be extended to special clitics such as the clitic clusters of French (Hudson forthcoming) and Serbo-Croatian (Camdzic and Hudson 2007). Take an example such as the French (8), compared with a syntactically similar sentence without clitics such as (9).

(8) Il la leur présentera
he her to-them will-introduce
'He will introduce her to them.'

(9) Jean présentera Marie à ses amis
John will-introduce Mary to his friends
'John will introduce Mary to his friends.'

In this case, each clitic is once again a full syntactic word, with ordinary syntactic dependencies but (exceptionally) only an affix as its realization. But unlike English clitics, French ones have a structurally complex host which provides a template of "slots," with the clitic concerned already assigned to one of these slots; for example, *il* occupies the "subject" slot, while *la* and *leur* occupy slots for "third-person direct object" and "third-person indirect object." These relatively abstract slots are then mapped onto slots for linear order which vary according to the type of verb, to take account of the differences between affirmative imperatives, as in (10), and other verbs as in (8) above.

(10) Présentez-la-leur!
introduce her to-them
Introduce her to them!

Figure 23.10 shows a partial structure for example (8). The top line shows the syntactic structure consisting of five words and their syntactic dependencies; for example, '3sg,m' is the pronoun *il*, and 'PRÉSENT,fut,3sg' is *présentera*. The middle line shows the four morphological forms that realise these (with a fused realisation of the preposition À and the pronoun '3pl' as {leur}). And at the bottom is the single complex morphological form that contains all four of these forms.

This figure makes a number of controversial assumptions. Most obviously, it assumes that the clitic pronouns occupy their expected positions after the verb in syntax, leaving the morphology to show their actual positions. (This follows the practice of Autolexical Syntax, Sadock 1991.) Another assumption is that {leur} is a fused realization of two words, both the preposition À and the pronoun called "3pl," in just the same way that {au} realizes À and the plural article *les* in *aux hommes* 'to the men.' The example thus illustrates two major mismatches between the syntax

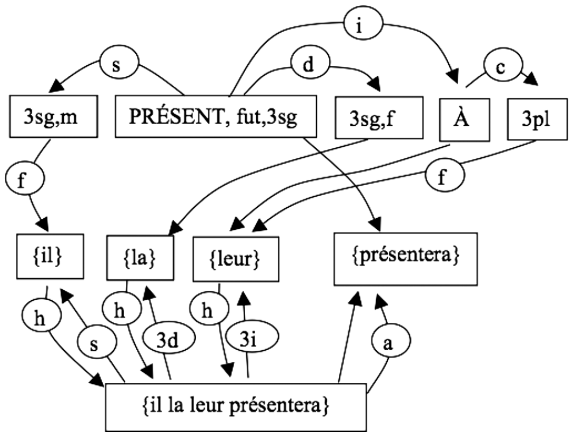

Figure 23.10. Syntax and semantics for special clitics

and the morphology, as well as the WG treatment of special clitics as parts of a complex hostform.

This analysis of French clitic pronouns illustrates a number of strengths in the WG theory of morphology. The associated theory of syntax takes the word as its basic unit, leaving the internal structure of words to the morphology (Chapter 10). The strict separation of morphology from syntax allows mismatches between the two structures, such as the two noted above. Another attraction involves the logic of default inheritance, which accommodates alternative patterns in a deeply explanatory way; for example, it turns out that the default pattern for French clitics is in fact the one found in affirmative imperatives such as (10), which in turn follows precisely the same order of elements as the regular syntactic pattern. In contrast, the exceptional, but more common, pattern found with other verbs can be explained in terms of the presence of two clitics, the subject and the negative (Hudson forthcoming). And finally, we can address the question of how clitics form clusters, given that each clitic introduces a new hostform. The WG answer builds on the theory of processing, which includes a process called "binding" that applies very generally and tends to identify compatible and highly active nodes. So if two clitics are both highly active at the same time, and both introduce a virtual hostform and anchor verb, the processor binds their hostforms and anchor verbs into one, creating a clitic cluster.

The discussion so far has concentrated on the WG treatment of morphosyntax, the interface between syntax and morphology, because this is the part of the theory which is most relevant to this chapter. The pairing with DS is less relevant to other topics in morphology such as allomorphy and syncretism, but WG morphology does provide for them as well (Hudson 2007: 63–116).

6 Conclusion

Although dependency grammar has so far contributed disappointingly few insights to the theory of morphology, it has great potential as the syntactic counterpart of a fully adequate theory of morphology. As argued in both MTT and WG, this is because dependency grammar generally takes the word as the basic unit of syntax, so syntax has to stop at the word and leave morphology to deal with the word's internal structure, using partially different machinery. At the risk of simplifying, the dominant relation is the word-word dependency in syntax but the part-whole relation between word-forms and their parts in morphology. So although lexical integrity is logically separate from DG, at least the existence of well developed versions of DG shows that it is possible to combine lexical integrity with a theory of syntax.

References

Anderson, John. 1977. *On Case Grammar: Prolegomena to a Theory of Grammatical Relations*. London: Croom Helm.

Anderson, Stephen; Lea Brown, Alice Gaby, and Jacqueline Lecarme. 2006. Life on the edge: There's morphology there after all!, *Lingue e Linguaggio* 5, 33–48.

Barabasi, Albert L. 2009. Scale-free networks: A decade and beyond, *Science* 325, 412–13.

Borer, Hagit. 1998. Morphology and syntax. In Andrew Spencer and Arnold Zwicky (eds.), *The Handbook of Morphology*, 151–90. Oxford: Blackwell.

Borsley, Robert; Maggie Tallerman, and David Willis. 2007. *The Syntax of Welsh*. Cambridge University Press.

Camdzic, Amela, and Richard Hudson. 2007. Serbo-Croat clitics and word grammar. *Research in Language (University of Lodz)* 4, 5–50.

Corbett, Greville. 2009. Morphology-free Syntax: Two potential counterexamples from Serbo-Croat. In Steven Franks, Vrinda Chidambaran, and Brian Joseph (eds.), *A Linguist's Linguist: Studies in South Slavic Linguistics in Honor of E. Wayles Browne*, 149–66. Bloomington, IN: Slavica.

Creider, Chet, and Hudson, Richard. 1999. Inflectional morphology in Word Grammar. *Lingua* 107, 163–87.

Debussmann, Ralph. 2006. Extensible Dependency Grammar: A Modular Grammar Formalism Based on Multigraph Description. PhD dissertation, Universität des Saarlandes, Saarbrücken.

Diaconescu, Stefan. 2002. A Generative Dependency Grammar. In Mitsuru Ishizuka and Abdul Sattar (eds.), *7th Pacific Rim International Conference on Artificial Intelligence Tokyo, Japan, August 18–22, 2002 Proceedings*, 605. Berlin: Springer.

Duran-Eppler, Eva. 2011. *Emigranto: The Syntax of German-English Code-switching*. Vienna: Braumüller.

Ferrer i Cancho, Ramon; Ricard Solé, and Reinhard Köhler. 2004. Patterns in syntactic dependency networks. *Physical Review E* 69, 1–8.

Gaifman, Haim. 1965. Dependency systems and phrase-structure systems. *Information and Control* 8, 304–37.

Gisborne, Nikolas. 2010. *The Event Structure of Perception Verbs*. Oxford University Press

Gisborne, Nikolas. forthcoming. Word grammar and new morphology. In Andrew Hippisley and Nikolas Gisborne (eds.), *Defaults in Morphological Theory*. Oxford University Press.

Gisborne, Nikolas. forthcoming. Word grammar morphology. In *Oxford Handbook of Morphology*.

Gragg, G. 1994. Babylonian grammatical texts. In Ronald Asher (ed.), *Encyclopedia of Language and Linguistics*, 296–8. Oxford: Pergamon.

Gross, Thomas, and Timothy Osborne. 2013. Katena und Konstruktion: Ein Vorschlag zu einer dependenziellen Konstruktionsgrammatik. *Zeitschrift Für Sprachwissenschaft* 32, 41–73.

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

Heringer, Hans J. 1993. Dependency syntax: Basic ideas and the classical model. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Venneman (eds.), *Syntax: An International Handbook of Contemporary Research*, Vol. 1, 298–316. Berlin: Walter de Gruyter.

Huddleston, Rodney, and Geoffrey Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press

Hudson, Richard. 1984. *Word Grammar*. Oxford: Blackwell.

Hudson, Richard. 1990. *English Word Grammar*. Oxford: Blackwell.

Hudson, Richard. 2001. Clitics in word grammar, *UCL Working Papers in Linguistics* 13, 243–94.

Hudson, Richard. 2003a. The psychological reality of syntactic dependency relations. In Sylvain Kahane and Alexis Nasr (eds.), *Proceedings of the First International Conference on Meaning-text Theory* 181–92. Paris: Ecole Normale Supérieure.

Hudson, Richard. 2003b. Trouble on the left periphery. *Lingua* 113, 607–42.

Hudson, Richard. 2007. *Language Networks: The New Word Grammar*. Oxford University Press

Hudson, Richard. 2010. *An Introduction to Word Grammar*. Cambridge University Press

Hudson, Richard. forthcoming. Default inheritance, word grammar morphology and French clitics. In Andrew Hippisley and Nikolas Gisborne (eds.), *Defaults in Morphological Theory*. Oxford University Press.

Joshi, Aravind, and Owen Rambow. 2003. A formalism for Dependency Grammar based on Tree Adjoining Grammar. In Sylvain Kahane and

Alexis Nasr (eds.), *Proceedings of the First International Conference on Meaning-text Theory*. Paris: École Normale Supérieure.

Kahane, Sylvain 2004. The Meaning-text Theory. In Vilmos Àgel, Ludwig Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans-Jürgen Heringer, and Henning Lobin (eds.), *Dependency and Valency: An International Handbook of Contemporary Research*. Berlin: De Gruyter.

Karmiloff-Smith, Annette. 1994. Precis of Beyond modularity: A developmental perspective on cognitive science. *Behavioral and Brain Sciences* 17, 693–745.

Kübler, Sandra; Ryan McDonald, and Joakim Nivre. 2009. Dependency parsing. *Synthesis Lectures on Human Language Technologies* 2, 1–127.

Kunze, Jürgen. 1975. *Abhängigkeitsgrammatik*. Berlin: Akademie-Verlag.

Mel'cuk, Igor. 1974. *Opyt teorii lingvisticeskix modelej "Smysl<=>Tekst"*. [Outline of a Theory of Linguistic Models of Meaning-text Type]. Moscow: Nauka.

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

Mel'cuk, Igor. 1992–2000. *Cours de morphologie générale*, 5 vols. Montreal: Les Presses de l'Université de Montréal.

Mel'cuk, Igor. 2003. Levels of dependency in linguistic description: Concepts and problems. In Vilmos Agel, Ludwig Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans Jürgen Heringer, and Henning Lobin (eds.), *Dependency and Valency: An International Handbook of Contemporary Research*, Vol. 1, 188–229. Berlin: de Gruyter.

Mel'cuk, Igor. 2004. Levels of dependency in linguistic description: Concepts and problems. In Vilmos Àge, Ludwig Eichinger, Hans-Werner Eroms, Peter Hellwig, Hans-Jürgen Heringer, and Henning Lobin (eds.), *Dependency and Valency: An International Handbook of Contemporary Research*, 188–229. Berlin: Walter de Gruyter.

Mel'cuk, Igor. 2008. *Aspects of the Theory of Morphology*. Berlin: De Gruyter Mouton.

Mel'cuk, Igor. 2009. Dependency in natural language. In Alain Polguère and Igor Mel'cuk (eds.), *Dependency in Linguistic Description*. Amsterdam: John Benjamins.

Onnis, Luca; Morten Christiansen, and Nick Chater. 2006. Human language processing: Connectionist models. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn., 401–9. Oxford: Elsevier.

Osborne, Timothy; Michael Putnam, and Thomas Gross. 2012. Catenae: Introducing a novel unit of syntactic analysis. *Syntax* 15, 354–96.

Reisberg, Daniel. 2007. *Cognition: Exploring the Science of the Mind*, 3rd edn. New York: Norton.

Robins, Robert. 1959. In sefence of WP. *Transactions of the Philological Society* 99, 114–44. (Reprinted in 2001).

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

Robinson, Jane. 1970. Dependency structure and transformational rules. *Language* 46, 259–85.

Rosta, Andrew. 1997. English Syntax and Word Grammar Theory. Ph.D. dissertation, UCL, London.

Sadock, Jerrold. 1991. *Autolexical Syntax: A theory of Parallel Grammatical Representations*. University of Chicago Press.

Sgall, Petr; Eva Hajicová, and Jarmila Panevova. 1986. *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*. Prague: Academia.

Sleator, Daniel D., and David Temperley. 1993. Parsing English with a link grammar. In *Proceedings of the Third International Workshop on Parsing Technologies*, 277–92. Tilburg.

Solé, Ricard; Bernat Corominas Murtra, Sergi Valverde, and Luc Steels. 2010. Language networks: their structure, function and evolution. *Complexity* 15, 20–27.

Starosta, Stanley. 1988. *The Case for Lexicase*. London: Pinter Publishers.

Stump, Gregory T. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge University Press.

Tallerman, Maggie. 2009. Phrase structure vs. dependency: The analysis of Welsh syntactic soft mutation. *Journal of Linguistics* 45, 167–201.

van der Hulst, Harry. 2006. Dependency Phonology. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn., 451–8. Oxford: Elsevier.

Vater, Heinz. 1975. Toward a generative dependency grammar. *Lingua* 36, 121–45.

Williams, Edwin. 1981. On the notions "lexically related" and "head of a word." *Linguistic Inquiry* 12, 245–74.

Zwicky, Arnold. 1992. Morphology: Morphology and syntax. In William Bright (ed.), *International Encyclopedia of Linguistics*, 10–12. Oxford University Press.