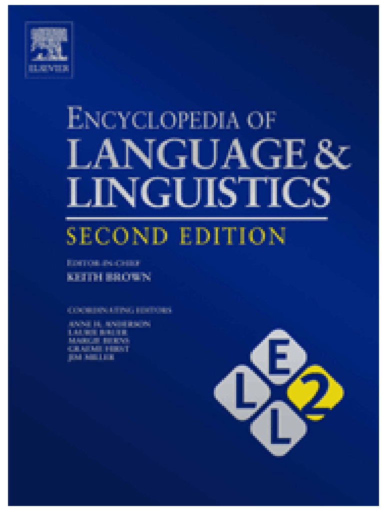

Hudson R (2006), Word Grammar. In: Keith Brown, (Editor-in-Chief) *Encyclopedia of Language & Linguistics, Second Edition*, volume 13, pp. 633-642. Oxford: Elsevier.

Bibliography

Adams V (1973). *An introduction to modern English word-formation.* London: Longman.

Aronoff M (1976). *Word formation in generative grammar.* Cambridge MA: MIT Press.

Bauer L (1983). *English word-formation.* Cambridge: Cambridge University Press.

Booij G, Lehmann C & Mugdan J (eds.) (2000). *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung (Morphology: An international handbook on inflection and word-formation)* (vol. 1). Berlin and New York: de Gruyter.

Collins-Ahlgren M (1990). 'Word formation processes in New Zealand Sign Language.' In Fischer S D & Siple P (eds.) *Theoretical issues in sign language research. vol. 1: linguistics.* (International Conference on Theoretical Issues in Sign Language Research.) Chicago and London: University of Chicago Press. 279–312.

Fleischer W (1975). *Wortbildung der deutschen Gegenwartssprache* (4th edn.). Tübingen: Niemeyer.

Henzen W (1965). *Deutsche Wortbildung* (3rd edn.). Tübingen.

Lang M F (1990). *Spanish word formation.* London and New York: Routledge.

Marchand H (1969). *The categories and types of present-day English word-formation* (2nd edn.). Munich: Beck.

Miranda A J (1994). *La formación de palabras en español.* Salamanca: Ediciones Colegio de España.

Plag I (2002). *Word-formation in English.* Cambridge: Cambridge University Press.

Rainer F (1993). *Spanische Wortbildungslehre.* Tübingen: Niemeyer.

Seymour R K (1968). *A bibliography of word formation in the Germanic languages.* Durham, NC: Duke University Press.

Stein G (1973). *English word-formation over two centuries.* Tübingen: Narr.

Stekauer P & Szymanek B (2002). *A bibliography of English word-formation (1960–2000).* Presov: Acta Facultatis Philosophicae Universitatis Presoviensis.

Worth D S (1977). *A bibliography of Russian word-formation.* Columbus, OH: Slavica.

Word Grammar

R Hudson, University College London, London, UK

Overview

'Word grammar' (**WG**) is the name of a theory of grammar in which words play a specially central role as the main unit of syntax; for example, the syntactic structure of (1) recognizes just four units, one per word, and phrases play no explicit part in the grammar:

(1) Small children often cry.

This rejection of phrase structure was the main distinctive feature of WG at the time when it received its name (Hudson, 1984), but the focus of the theory has broadened since then to include the outlines of a general theory of conceptual knowledge, with language as a particular subcase. The psychological assumptions are, by and large, conservative and uncontentious, but the linguistic theory that results is rather different from most mainstream theories and belongs firmly in the recent tradition of 'cognitive linguistics' (e.g., Barlow and Kemmer, 2000). This article will present some of the psychological assumptions (in the sections 'Inheritance Networks' and 'How to use a Network'), but will mostly be concerned with their consequences for the theory of language structure.

As far as language is concerned, the leading idea is (2):

(2) Language is a network.

This was already explicit from the early days (Hudson, 1984: 1) but its implications are much clearer now. 'Inheritance Networks' will spell out the nature of this network in more detail, and will highlight the importance of is-a relations (as in 'John is-a linguist' or 'BUT is-a conjunction') and its related logic, default inheritance. This logic is more fully developed in WG than in any other theory of language structure (Hudson, 2003b); for example, it is default inheritance that explains the difference:

- between regular and irregular morphology
- between basic and special word order
- between a morpheme's basic shape and its allomorphs
- between sense and referent
- between types and tokens.

In fact, it is tempting to conclude that default inheritance is responsible for all the complexity (and flexibility) in language. However, it is important also to recognize the important role played by another very general principle of processing, variable binding, which will be explained in 'How to Use a Network.' Both default inheritance and variable binding are heavily dependent on spreading activation.

Language structure, as such, is completely declarative – as indeed it has to be if it is a network. It is like a map, which may be used as a guide for action but which is not, in itself, either an active machine or even a set of instructions for actions. Language is simply a (very large) collection of nodes and their (even larger) set of interconnecting links. However, the network is highly structured through a rich classification of the nodes and links, so (unlike distributed connectionist networks) this theory is easy to compare with more conventional theories based on rules and lexical items. For example, entities in language are of a small number of basic types that define the traditional levels of language – semantics, syntax, morphology, and phonology. This structure will be outlined in 'Levels of Language.' Similarly, links fall into a fairly traditional range of categories that are specialized for particular places in the total network – e.g., for linking a word to a meaning, an affix to a stem, or one word to another. However, although the categories used for classifying links are mostly quite traditional, WG is the only formal theory of language (or knowledge) that allows links to be freely classified in just the same way as entities.

If language is a network as described here, and if this network is a subnetwork of general knowledge, then language cannot be modular in the very strong sense of this word (Fodor, 1983). Neither language, nor any of its parts (e.g., syntax), is informationally encapsulated so that elements inside it cannot communicate directly with elements outside it. On the contrary, it is quite typical for elements of language to be closely related to nonlinguistic categories such as speaker types and situations – e.g., people only use *cookie* if they are American and *gee-gee* when speaking to children. However, networks do allow more limited kinds of modularity, including subnetworks whose elements have relatively dense internal links and relatively sparse external ones.

The network theory also favors explanations of language structure in functional rather than genetic terms. If language has the same formal properties as the rest of knowledge, its formal properties need no special explanation, genetic or otherwise. Moreover, even the substantive properties – the particular categories and relations – are rather easy to learn by induction from experience (Tomasello, 2000), so the contents of the network can be explained by a theory of learning and we can conclude that language is not an 'instinct' (Pinker, 1994).

Inheritance Networks

If language is a network, then a major research question is how it fits into a general typology of networks: what kind of network is it? This question can be approached either quantitatively or qualitatively. Graph theory takes a quantitative approach by applying various kinds of mathematical measure to a network. For example, what is the average topological distance – i.e., the average number of links in the shortest path – between nodes? Many complex human networks, such as society and the internet, turn out to have quite small distance averages, so they are called small worlds (Barabási, 2003); research so far indicates that the same may be true of language (Ferrer i Canchol and Solé, 2001; Motter *et al.*, 2002). Another quantitative question is whether the number of links per node has a normal distribution (like that of people's heights) or whether it follows a power law, with a small number of extremely richly connected nodes and a long tail of nodes with very few connections. The second kind of network, called scale-free, also turns out to be common among complex human networks; and once again language too seems to have this property (Ferrer i Canchol and Solé, 2001).

Linguistics, unlike graph theory, is concerned with the qualitative characteristics of the language network. WG is not the only linguistic theory that has concerned itself with the question. The whole of synchronic linguistics in the 20th century was an attempt to flesh out Saussure's idea that "language is a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others" (Saussure, 1959: 114 – in short, a network of terms. However, given the enormous size of a language, it is easy to lose sight of the total network in order to focus on some small area such as phonology or syntax, so most theories say very little about the overall structure of the network of relations. Indeed, this methodological preference has turned into the factual claim that language really is modular (Fodor, 1983), with distinct submodules that have quite different organizing principles – "a modularization of the grammar reminiscent of the modularity of the mind" (Smith, 1999: 68). The first fully articulated theory to stand back and consider the total network of language was stratification grammar (Lamb, 1966) and its more recent manifestation called neurocognitive linguistics (Lamb, 1998). Lamb's ideas played a major part in the development of WG (in spite of the negative comment in Hudson, 1990: 15, but WG's claims about the language network are different from Lamb's.

According to WG, language, like the rest of conceptual knowledge, is an inheritance network. What this means is that:

- the classification relation is-a has a special status, because it allows generalizations by default inheritance (discussed more fully below).
- every node is part of at least one inheritance hierarchy united by is-a; for instance: TAKE is-a Verb is-a Word. In other words, all nodes are classified.
- every link is also part of at least one inheritance hierarchy, such as Object is-a Complement is-a Dependent. Thus, classification applies not only to nodes but also to links.

Because of its privileged status, the is-a relation has a special notation: a small triangle whose base is next to the super-category. Figure 1 shows an elementary WG diagram for the examples just given.

Choosing to model language as a network is not simply a choice of notation. As we shall see in 'How to Use a Network,' it has important implications for understanding how we use language, but it also affects the purely linguistic analysis of the language system. The main effect is that categories have to be justified carefully because of the following general principle:

(3) Every category is defined, and distinguished from other categories, solely by its links to other categories.

The first consequence of this principle is that strictly speaking, the labels that we give to nodes and links carry no information at all, so a network would still carry exactly the same information even if we removed all labels (Lamb, 1998: 59). Universal categories identified only by their name fit very badly into WG theory. Another consequence is that labels should not be used, as in most other theories, to show class membership. For example, if something is a noun, this fact should be shown by an is-a link to Noun, rather than by labeling it 'noun.' Labels should not be used in order to smuggle classificatory information into the network.

Nodes are simply nodes – points where links meet. The function of a node, in fact, is merely to show a correlation between two or more links. In this sense, then, a node defines some 'property'; for example, an Object link from TAKE represents the property of taking an object. Each link can in fact be seen as defining a property for mathematical sense: a relation that takes an argument and yields a unique value for it (compare the mathematical function +2 which gives a unique value for any number to which it is applied). As in mathematics, the value varies with the argument; for example, the value of the object function varies from verb to verb and from token to token, so at the level where generalizations are made about objects of all verbs, the value must be a variable. In WG notation, variables are shown either as a simple dot or (perhaps more helpfully) as a question-mark: '?'. Clearly (3) requires that this labeling convention must duplicate some kind of relation difference, but it is not yet clear what this difference is. Variables will play an important role in the WG theory of processing outlined in 'How to Use a Network.'

Figure 2 shows a simple network including the dependencies and word types distinguished in Figure 1. It shows that a word's complement is normally impossible: its quantity ('#') is 0. This general ban on complements explains why, unlike adjuncts, they are selected by specific lexical items, which can override the general ban as explained in the next paragraph; for example, TAKE has an obligatory object (whose quantity is 1). Figure 2 also shows that objects are nouns, so this property also applies by inheritance to the object of TAKE.

As explained earlier, the main characteristic of inheritance networks is the possibility of generalization by default inheritance. This seems to be the basic logic

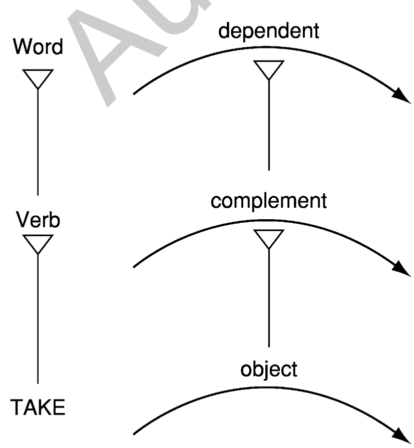

Figure 1 Inheritance hierarchies for nodes and links.

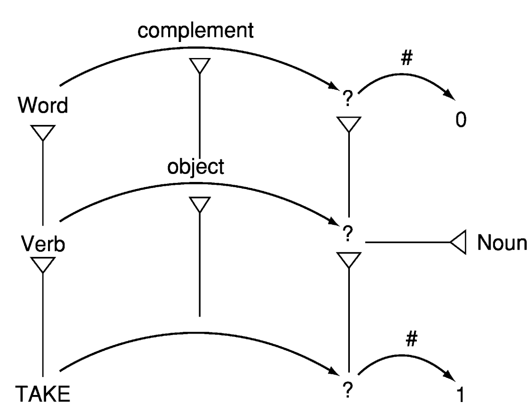

Figure 2 A network for TAKE and its object.

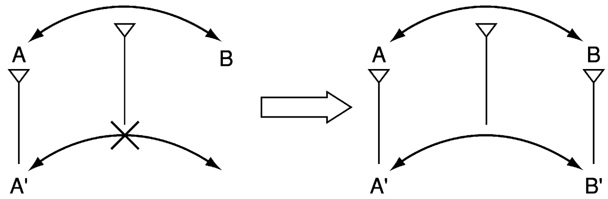

Figure 3 Default inheritance.

of everyday reasoning, in which we generalize from the typical case in order to guess information that we cannot observe. For example, if we recognize something as a cat we assume it has four legs, like the typical cat; in other words, this cat inherits (in our minds) the properties of the typical cat by default, in the absence of information to the contrary. However, if we already know that it has only three legs this information overrides the default so we can continued to classify it as a cat in spite of the exceptional leg number. The same logic applies to linguistic reasoning:

- by default, a dependent is optional,
- but if a dependent is (more specifically) a complement, it is impossible,
- but if a complement is (more specifically) the object of TAKE, it is obligatory.

In each case, a fact located lower in the inheritance hierarchy of entities or relations takes priority over one located above it. The logic is presented more formally in Figure 3, where the double-headed arrow means an arrow going in either direction and the crossed lines mean that the link concerned is not already present.

Default inheritance is controversial among logicians because it is nonmonotonic, meaning that it allows us to draw conclusions which later turn out to be false. It is also controversial in computational linguistics because the system spends a great deal of time checking for possible overriders. However, a network model reduces both objections to triviality because the search space for competitors is very small.

Levels of Language

We turn now to the specifics of language structure. WG distinguishes the traditional levels of language through its classification of entities and relations rather than by trying to divide all the data of linguistics into discrete components, which has the advantage of avoiding boundary disputes. The relevant node types are as follows:

- words of all degrees of generality from word tokens (e.g., *tokens* in this sentence) through lexemes (e.g., TOKEN) to inflectional categories (Plural) and word classes (Noun) and even the category Word itself,
- forms of all sizes from morphemes (e.g., {s}) to word forms (e.g., {tokens}), and of all degrees of generality from tokens to form classes (e.g., Suffix),
- sound segments of all sizes and generalities,
- letters and other units of written language,
- any kind of entity or relation may act as the meaning (see ‘Semantics and Sociolinguistics’ section) of a word.

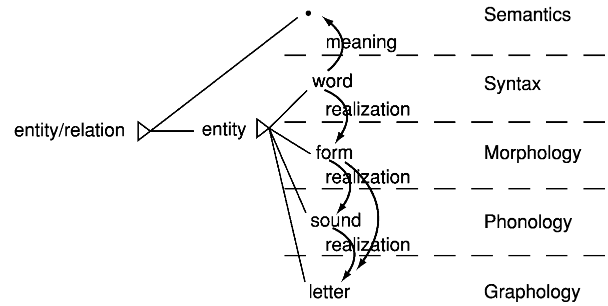

Figure 4 The architecture of language.

Various kinds of sets of these entities are also needed; for example, word strings are used in coordination (see ‘Syntax’ section).

As for relations, they fall into a range of types that include the following, all of which are relevant to language; the terms in brackets are specific subcases of the more general relation which heads the paragraph:

- Meaning (or referent) links a word to its meaning.
- Realization (base or fully inflected form) links a word to its component morphemes, or larger form, and also links a form to its phonological structure. (Written structures may realize forms directly or indirectly through phonology.) Realization may even be considered the converse of Meaning, in which case it is also found between words and their meanings.
- Co-occurrent (dependent, next, etc.) links a word to words that co-occur with it. It also handles the combinational patterns of co-occurring morphemes, sounds, and letters, and indeed may be found in semantics (e.g., as semantic roles such as Actor).
- Part (part 1, last part, etc.) links a larger form to the smaller forms within it, and is also found in many other areas of knowledge. Part may also include Member, which relates a set to its parts.

The overall architecture of language is summarised in Figure 4. One of the noteworthy features of this diagram is the absence of a separate lexicon. In WG,

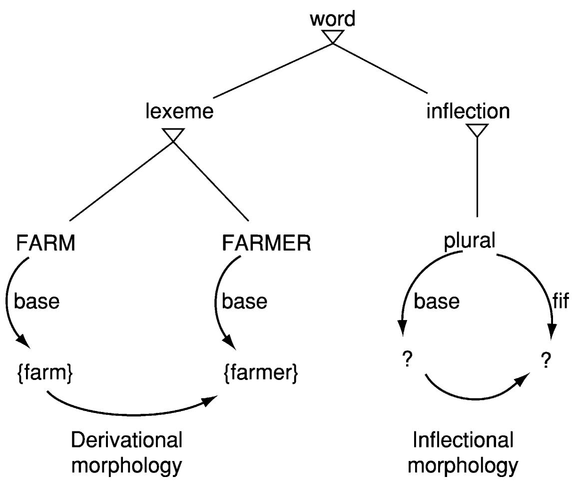

Figure 5 Derivational and inflectional morphology.

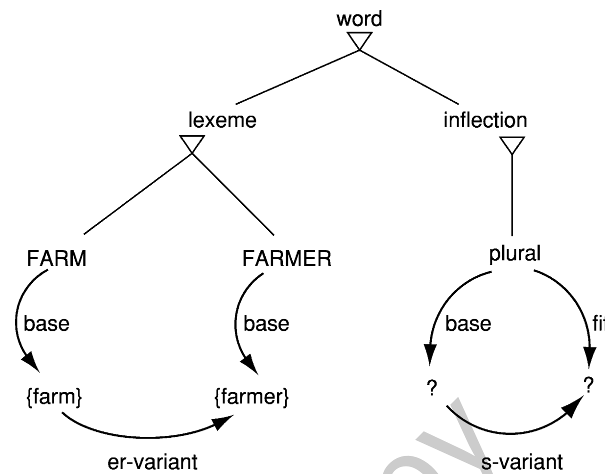

Figure 6 Morphosyntactic relations between words and variants.

as in other cognitive theories (Langacker, 2000), the difference between lexical and more general facts is merely a matter of degree; for example, the valency facts about TAKE are shown in the same network as those for more general categories such as Verb and Word (Figure 2).

Morphology

The only level of language in Figure 4 that needs justification is morphology, the level whose units are forms, which are distinguished in notation by the use of the traditional morpheme boundary markers, {. . .}. For example, the form {dog} is distinguished from the word DOG as well as from the phonological structure /dog/. Forms are morphomic (Aronoff, 1994), so:

- unlike words they have no meaning or syntactic categorization
- and unlike phonological units they have no pronunciation – indeed, they correspond just as easily to written structures.

The relation between a form and its corresponding word is Realization, not Part, i.e., abstractness, not size. The size relation applies to different forms, so we find complex forms that contain smaller ones; e.g., {dogs} contains {dog} and {s}, and {farmers} contains {farmer}, which in turn contains {farm} and {er}.

The WG treatment of morphology rests on the familiar distinction between two kinds of word: lexemes (e.g., DOG) and inflections (e.g., Plural). Thanks to default inheritance, unmarked inflections can be treated simply as the default, so DOG is by default singular in contrast with its plural, DOG:-plural. Only the marked form needs any mention in morphology. This distinction between lexemes and inflections allows a distinction between two kinds of Realization relation: Base and Fully inflected form (abbreviated to 'fif'). A word's base is determined by its lexeme and its fif by its inflection. By default the base and fif are the same, but marked inflections distinguish them in some way. This is the province of inflectional morphology, while derivational morphology deals with the relations between the bases of different lexemes, e.g., between those of FARM and FARMER. The difference between derivational and inflectional morphology is shown schematically in Figure 5.

Although derivational and inflectional morphology are very different at the level of words, at the level of form they offer the same range of possible morphological differences. These differences are each defined by a morphological function called a variant, so for example, the plural form is the s-variant of the base and the agent noun is the er-variant of the source base. This treatment of morphological structure is an important (and controversial) part of WG theory because it establishes a clear separation between morphological structure and its syntactic and semantic correlates. The morphosyntactic realization relations of inflectional and derivational morphology map lexemes and inflections onto variants, while the morphophonological relations map variants onto morphological and (eventually) phonological structures. The morphosyntactic mapping is shown in Figure 6, in which the variants have replaced the unlabeled relations of Figure 5.

One of the immediate benefits of this separation of morphosyntax from morphophonology lies in the treatment of syncretism, cases where the same morphological pattern realizes more than one word type. To take an obvious example, the s-variant pattern is used not only for plural nouns (e.g., {dogs}), but also for present tense singular verbs (e.g., {eats}). More interestingly, the passive and past participles of English verbs are always the same however irregular the verb may be. This can be explained by saying that both inflections use the same variant of the base, the en-variant (Hudson, 1990: 90).

Variants are the bridge from words to morphological structure. In simple cases, the variant of a form consists of a copy of that form combined with an affix ({farm} + {er} or {farmer} + {s}). These patterns are handled in terms of parts, with Part1 and Part2 occurring in that order. This kind of pattern is illustrated in Figure 7 for the er-variant, of which the relation between {farm} and {farmer} is an example.

Exceptions are easily accommodated thanks to default inheritance, as shown in Figure 8. This illustrates two degrees of irregularity: suppletion ({person}, {people}), where the entire base is replaced, and vowel-change ({goose}, {geese}), where only the vowel is changed. Of course, morphology can be a great deal more complicated than the examples discussed here, but WG has already been shown to be able to accommodate various kinds of complexity such as multiple affixation and cliticization (Camdzic and Hudson, 2002; Creider and Hudson, 1999; Hudson, 2001).

Syntax

The most distinctive characteristic of syntax in WG is the rejection of phrase structure in favor of

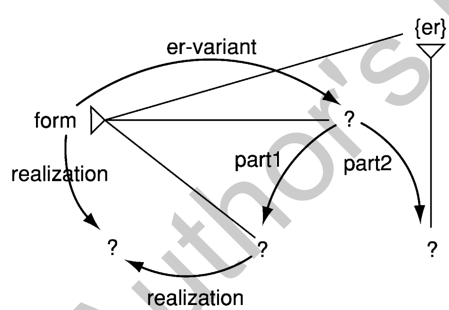

Figure 7 Variation by affixation.

dependency structure, which follows from the network assumption (2). The question is not whether to show dependencies between words, but how. Phrase structure identifies dependencies indirectly via part relations to phrases, whereas dependency analysis recognizes them directly. Consider this very simple example:

(4) Gas smells.

Linguistics would generally accept the traditional view that *gas* is related to *smells* in a way that explains, for example, the {s} on {smells}. The question is how to include such word–word relations in a syntactic structure. Phrase structure recognizes a number of phrase nodes, including (at least) an NP node, some kind of clause or sentence node, and (probably) a VP node. The relation between *smells* and *gas* is merely implied by phrase structure (such as the rather conservative one in Figure 9) but is explicit in a dependency structure.

The dependency analysis fits well in a general network-based theory of language, where all relations are shown explicitly. In contrast, standard phrase structure is incompatible with network theory because word–word relations are merely implicit. Indeed, it is not at all obvious that a phrase really is more than the sum of its parts, so the basis for WG syntax is not only that a phrase really is more than the sum of its parts, so the basis for WG syntax is not only that word–word dependencies must be part of the network because they are psychologically real, but also that phrases should be excluded from the network because they are psychologically real, but also that phrases should be excluded from the network because they are not real. A good analogy outside language is social structure. Interpersonal relations are certainly real, but groups such as families and communities are much more nebulous and hard to define; for example, where are the boundaries of ‘my family’ and – even worse – ‘my community’? At best, the case for such groups needs to be made,

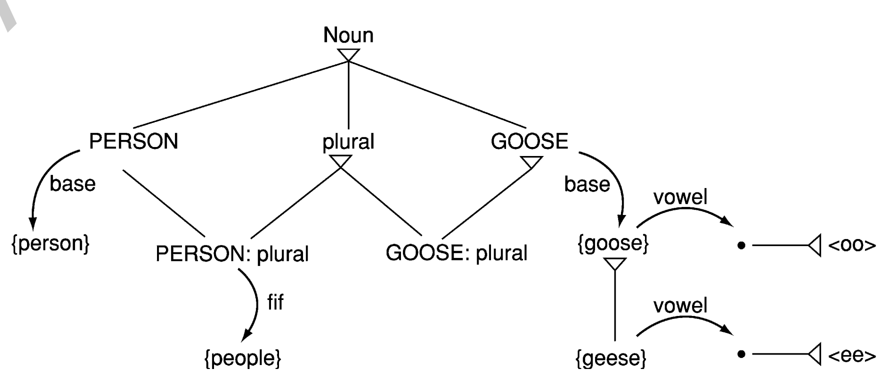

Figure 8 Suppletion and vowel change.

and similarly for phrases in syntax; but all the standard arguments for phrases simply ignore the possibility of using word–word dependencies instead. However, even if phrases really are indispensable, this will merely show that phrases are needed in addition to dependencies, as in some proposed mixed theories (Hudson, 1976). Meanwhile, the WG answer is that every generalization that can be stated in terms of phrases can be stated at least as well in terms of dependencies between single words (Hudson, 1990: 104).

Suppose that linguistic knowledge is a network. It then follows that sentence structure is also a network with the same flexibilty as other networks For example, there is nothing in the network idea to prevent one word from depending on more than one other word; and in fact all the familiar evidence for structures such as raising and extraction shows that this possibility is reality. The structure in Figure 10 illustrates the kind of richness that is possible in a quite ordinary sentence:

(5) What do you think we should wait for?

Of course, there are limits to possible dependency structures and, as in any other theory of syntax, our aim is to find and explain them. For example, the dependencies in Figure 10 follow a very general principle:

(6) The dependencies in a sentence must include a tangle-free substructure, i.e., one in which the implied phrases do not overlap.

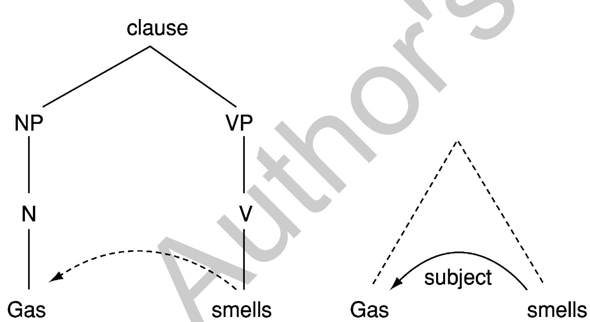

Figure 9 Phrase structure or dependency structure?

This is the substructure shown above the words. In contrast, sentence (7) is impossible because its dependencies do not include a tangle-free substructure.

(7) *I complicated like sentences.

The structure for (7) is shown in Figure 11, which is ruled out by (6). It is easy to imagine functional explanations for the restriction in (6); for example, it helps the hearer to identify word–word relations, and speaking is easier if one phrase is finished before the next one is started.

As mentioned earlier, the syntax of coordination is handled in terms of word strings, which are mere unstructured strings of words, like the strings that are possible as object of SAY:

(8) He said, "One, two, three, testing, testing."

A coordination is itself a word string consisting of two or more smaller strings (conjuncts) linked by a conjunction. However, the ordinary rules of syntax also apply to these strings, subject only to one extra constraint:

(9) Any dependency between a word inside one of the conjuncts and a word outside the entire coordination must also apply to one word inside each of the other conjuncts.

In other words, conjunction-reduction must apply across the board. Word strings are indicated in the notation by brackets, as illustrated in Figure 12. It is important to notice how this example shows that word strings are not phrases: for example, *London on Tuesday* is not a phrase in terms of any phrase-structure analysis.

Syntax is the level of language at which WG has been applied most intensively, so further details are easily available in the literature (Hudson, 1984, 1990, 1998, 1999, 2003a, 2003c).

Semantics and Sociolinguistics

The similarities between semantics and sociolinguistics are at least as striking as their differences. Semantics handles (at least) referential meaning, such as the

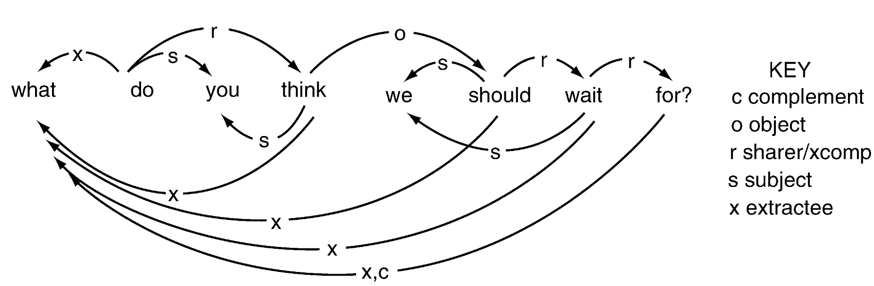

Figure 10 A syntactic structure.

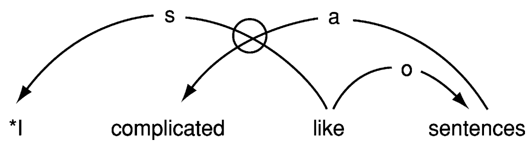

Figure 11 An ungrammatical structure.

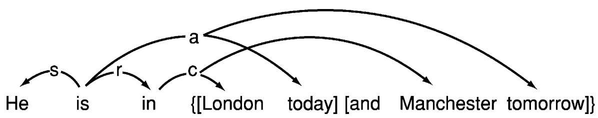

Figure 12 Coordination of non-phrases.

difference between apples and pears or between singular and plural, while sociolinguistics deals (at least) with the social meaning of words such as *hello* and *ain't*. However, as the terms imply, in both cases we are concerned with a kind of meaning, information about the world that follows from the choice of words. In both cases, this information can be inferred by anyone who is aware of the correlation between the word and the thought; for example, between the word *apple* and the concept Apple, and between *hello* and Meeting. In terms of spreading activation (which I discuss more fully in the next section), the word activates the thought, and both participants know that the same is true for both of them. In short, by using words "we can shape events in each other's brains with exquisite precision" (Pinker, 1994: 15). This is equally true of referential or social information. Moreover, there is a great deal of overlap between the two kinds of meaning; for example, the difference in (10) and (11) between polite and blunt imperatives could be classifed in either way:

(10) Do come in!

(11) Come in!

WG distinguishes the two kinds of meaning in terms of relations. Referential meaning is a property of the referent (the person or thing being referred to), whereas social meaning generally belongs to the speaker or the addressee (or both). In both cases, the particular individual is classified in terms of some general category: the sense for the referent, and some general social category for the speaker and addressee. We can illustrate both kinds of meaning in the name MOMMY, which refers to the mother of a child who is either speaking or being addressed, as in:

(12) Mommy will be back soon.

The advantages of a network analysis are especially obvious in a case like this where linguistic categories are related to non-linguistic ones. The WG analysis is shown in Figure 13, which treats Mother as a relation between two variables.

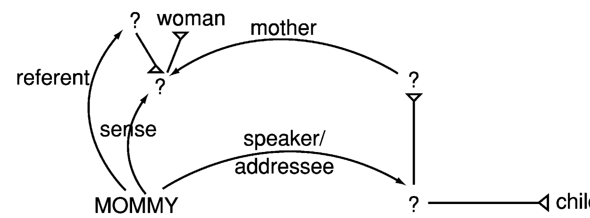

Figure 13 Referential and social meaning of MOMMY.

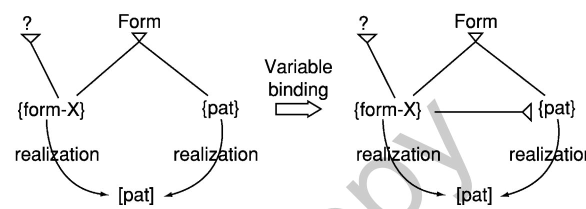

Figure 14 Classification by variable binding.

As might be expected in a highly lexicalized theory, WG has been applied quite extensively to problems of lexical semantics (Hudson and Holmes, 2000, Hudson, 2005), but one of the great advantages of dependency analysis in syntax is the relatively simple relation between syntactic dependencies and semantic relations. The published WG literature includes serious proposals for the semantic structure of plurals and quantifiers (Hudson, 1990: 139–146), past tenses and imperatives (Hudson, 1990: 222–223), and the semantic phrasing of arguments (Hudson, 1990: 146–151). However it has to be admitted that all these proposals need to be fleshed out.

Social meaning also benefits greatly from the network approach. In this case, one of the strengths of the approach is that it allows links to have different strengths; these are an essential ingredient of the model for spreading activation and are highly relevant to the quantitative work that has been so successful in recent sociolinguistics (Hudson, 1996: Chap. 7, 1997).

How to Use a Network

One of the attractions of a network model of language is the bridge that it offers between theoretical linguistics and psychology, where models of long-term memory are standardly framed in terms of networks (Reisberg, 1997: Chap. 7). Our knowledge of language, according to network models, is an example of long-term memory, but recent theories of working memory suggest that this too can be modeled in the long-term network: "Working memory is closely linked to LTM, and its contents consist primarily of currently activated LTM representations" (Miyake and Shah, 1999: 450). If these theories are right, then linguistic processing depends crucially on activity in the language network.

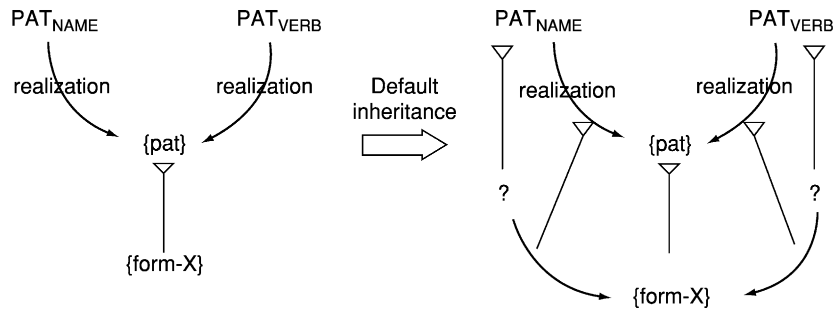

Figure 15 Enrichment by default inheritance.

WG now includes the basis for a theory of how processing works. It assumes, as already explained, that all long-term memory for language is held in a declarative network, but it also assumes that the links and nodes of the network have levels of activity that can be represented as numbers. This activity spreads in a completely mechanical way from active nodes to their neighbors, and the aim of processing is to allow nodes to compete for selection as the target meaning (in understanding) or as the target pronunciation (in planning). This far the model is similar to many others (e.g., Levelt *et al.*, 1999), but there are two crucial differences:

- The WG processor can create new nodes and links, which is not generally allowed in other theories of processing. A new node is created for each new token of experience, and for each property inherited by default inheritance (see Figure 3).
- WG processing is oriented toward the processor's current goals and interests, defined by highly active nodes and links; for example, a hearer's goal of finding meanings is reflected in the high activation of the Referent relation, and the hearer's interpretations are heavily biased by the general-knowledge nodes that are currently active.

The processor uses a small number of very general processes:

- Node and link creation, as described above, which creates new variable nodes (by default inheritance) and new constant nodes (for tokens).
- Variable binding, in which any highly active variable node is bound to the most active node that has the same super-category. This is how the processor assures the best fit between the token and the stored knowledge (Hudson, 1990: 46–52).
- Default inheritance, which again is controlled by activity so that it only applies to highly active variable nodes. This restriction prevents pointless inheritance that leads nowhere.

In principle this theory applies equally to productive and receptive processing, but for simplicity we can consider just one direction, reception (i.e., understanding). Suppose we hear the sounds [pat], and we have the form {pat} in our vocabulary. Hearing the sounds activates their respective nodes, which in turn activates the form {pat}, which contains all of them (in the right order). However, that activation does not itself tell us that we have just heard {pat}; to reach this conclusion we must first create a node for the unidentified form, which we can call {form-X}, together with an is-a link to the general category Form and a variable node '?' for its super-category. The latter node is highly active (thanks to both newness and the links to the active sounds [pat]), so variable binding automatically binds '?' to the most active node which is-a Form, namely {pat}. The effect of variable binding can be seen in Figure 14.

Once {form-X} is classified as an example of {pat}, it can be enriched by inheriting the latter's realization relations to words such as the name PAT and the verb PAT; this change is shown in Figure 15.

Activation then spreads blindly along these inherited realization relations to the word nodes, where it may meet activation flowing from nodes that are already active, such as the one for a person called Pat. Any highly active variable nodes will in turn inherit from their super-category, e.g., the name PAT, and so on until the word's referent is found, the processor loses interest, and activation dwindles.

This theory of processing is currently (2004) being modeled computationally in a program called WGNet++. Information about this and other developments in WG can be found at the WG site.

See also: Architecture of Grammar; Cognitive Linguistics; Cognitive Semantics; Constituent Structure; Construction Grammar; Declarative Morphology; Dependency Grammar; Grammatical Relations and Arc-Pair Grammar; Head-Driven Phrase Structure Grammar; Meaning: Cognitive Dependency of Lexical Meaning; Nonmonotonic Inference; Paradigm Function Morphology; Prototype Semantics; Relational Network Grammar; Right Node Raising; Unification, Classical and Default.

Bibliography

Aronoff M (1994). *Morphology by itself. Stems and inflectional classes.* Cambridge: MIT Press.

Barabási A -L (2003). *Linked: how everything is connected to everything else and what it means for business, science and everyday life.* London: Penguin.

Barlow M & Kemmer s (2000). *Usage based models of language.* Stanford: CSLI.

Camdzic A & Hudson R (2002). 'Serbo-Croat-Bosnian clitics and word grammar.' *UCL Working Papers in Linguistics 14,* 321–354.

Creider C & Hudson R (1999). 'Inflectional morphology in word grammar.' *Lingua 107,* 163–187.

Ferrer i Canchol R & Solé R (2001). 'The small world of human language.' *Proceedings of the Royal Society Series B 268,* 2261–2265.

Fodor J (1983). *The modularity of the mind.* Cambridge: MIT Press.

Hudson R (2005). 'Buying and selling in word grammar.' In Andor J & Pelyvás P (eds.) *Empirical, cognitive-based studies in the semantics-pragmatics interface.* Oxford: Elsevier Science.

Hudson R (1976). *Arguments for a non-transformational grammar.* Chicago: Chicago University Press.

Hudson R (1984). *Word grammar.* Oxford: Blackwell.

Hudson R (1990). *English word grammar.* Oxford: Blackwell.

Hudson R (1996). *Sociolinguistics* (2nd edition.). Cambridge: Cambridge University Press.

Hudson R (1997). 'Inherent variability and linguistic theory.' *Cognitive Linguistics 8,* 73–108.

Hudson R (1998). *English grammar.* London: Routledge.

Hudson R (1999). Subject-verb agreement in English. *English Language and Linguistics 3,* 173–207.

Hudson R (2001). 'Clitics in word grammar.' *UCL Working Papers in Linguistics 13,* 243–294.

Hudson R (2003a). 'Gerunds without phrase structure.' *Natural Language & Linguistic Theory 21, 579–615.*

Hudson R (2003b). 'Mismatches in default inheritance.' In Francis E & Michaelis L (eds.) *Mismatch: form-function incongruity and the architecture of grammar.* Stanford: CSLI. 269–317.

Hudson R (2003c). 'Trouble on the left periphery.' *Lingua 113,* 607–642.

Hudson R & Holmes J (2000). 'Re-cycling in the encyclopedia.' In Peeters B (ed.) *The lexicon/encyclopedia interface.* Amsterdam: Elsevier. 259–290.

Lamb S (1966). *Outline of stratificational grammar.* Washington, DC: Georgetown University Press.

Lamb S (1998). *Pathways of the brain. The neurocognitive basis of language.* Amsterdam: Benjamins.

Langacker R (2000). 'A dynamic usage-based model.' In Barlow M & Kemmer S (eds.) *Usage-based models of language.* Stanford: CSLI. 1–63.

Levelt W J M, Roelofs A & Meyer A S (1999). 'A theory of lexical access in speech production.' *Behavioral and Brain Sciences 22,* 1–45.

Miyake A & Shah P (1999). 'Toward unified theories of working memory. Emerging general consensus, unresolved theoretical issues and future research directions.' In Miyake A & Shah P (eds.) *Models of working memory. Mechanisms of active maintenance and executive control.* Cambridge: Cambridge University Press. 442–481.

Motter A, de Moura A, Lai Y & Dasgupta P (2002). 'Topology of the conceptual network of language.' *Physical Review E 65,* 102.

Pinker S (1994). *The language instinct.* London: Penguin.

Reisberg D (1997). *Cognition. Exploring the science of the mind.* New York: Norton.

Saussure F de (1959). *Course in general linguistics* (translated by W. Baskin; French edition 1916). Lausanne: Payot.

Smith N V (1999). *Chomsky. Ideas and ideals.* Cambridge: Cambridge University Press.

Tomasello M (2000). 'First steps toward a usage-based theory of language acquisition.' *Cognitive Linguistics 11,* 61–82.

Relevant Website

www.phon.ucl.ac.uk – WG website.

Word Order and Linearization

A Siewierska

This article is reproduced from the previous edition, volume 9, pp. 4993–4999, © 1994, Elsevier Ltd.

Word order and linearization are major areas of linguistic inquiry within all linguistic traditions: descriptive, theoretical, and typological. This article concentrates on the typological perspective centering on the notion of 'basic order.' It begins with a consideration of the notion of basic order, then proceeds to review the distribution of word order patterns on a cross-linguistic basis, and next presents a brief account of the host of interrelated factors deemed to underlie the recurring linearizations. The article closes with a specification of the major problems posed for theory specific accounts of universal word order phenomena.