

Comments on ‘Linguistic structure: A plausible theory’

Richard Hudson^a

^a University College London, r.hudson@ucl.ac.uk.

Paper received: 22 October 2016

Published online: 20 November 2016

DOI: 10.31885/lud.4.1.230

Abstract. This comment on Sydney Lamb’s article “Language structure: A plausible theory” explores the similarities and differences between Lamb’s theory and my own theory called Word Grammar, which was inspired by Lamb’s work in the 1960s. The two theories share Lamb’s view that language is a symbolic network, just like the rest of our knowledge. The note explains this claim, then picks out a number of differences between the theories, all of which centre on the distinction between types and tokens. In Word Grammar, tokens are represented as temporary nodes added to the permanent network, and allow the theory to use dependency structure rather than phrase structure, to include mental referents, to recognise the messiness of spreading activation and to include a monotonic theory of default inheritance.

Keywords: dependency, Word Grammar, spreading activation, default inheritance, tokens, network

Sydney Lamb’s ideas about networks inspired me in the 1960s, and provided the foundations for my own network-based theory, called ‘Word Grammar’ (Hudson 1984, 1990, 2007, 2010). So it’s perhaps not surprising that we agree on the most important issues:

1. Language is a network in which the nodes are just unstructured atoms and the labels are redundant.
2. This network is a symbolic network (Lamb’s ‘relational network’) in which each node corresponds to a separate mental unit (in my terms, to a concept).
3. There are no boundaries between this network and the rest of the network for general cognition.

For example, here’s the core of a network entry for the lexical item DOG. It uses Word-Grammar notation, but conceptually it could come from either theory.

Language Under Discussion, Vol. 4, Issue 1 (June 2016), pp. 38–43

Published by the [Language Under Discussion Society](#)



This work is licensed under a [Creative Commons Attribution 4.0 License](#).

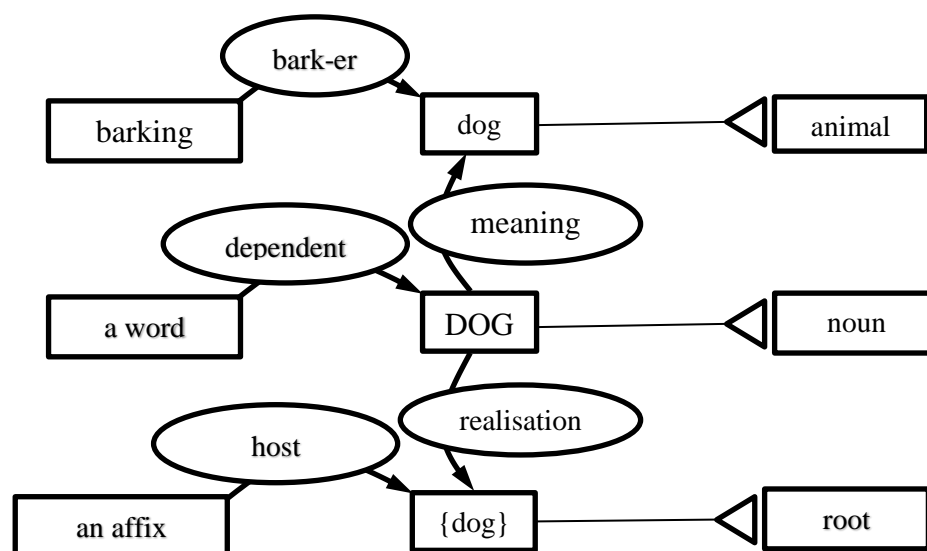


Figure 1. A conceptual network

In words, Figure 1 asserts that the lexeme DOG is a noun (the small triangle denotes the ‘isa’ relation), which means ‘dog’ (the general concept of a typical dog), which is an animal and which barks. DOG is typically dependent on another word, and is realised by the root morpheme {dog}, which may be the host for an affix. The horizontal arcs denote Lamb’s tactic patterns (how things combine with one another), the lines with triangles denote his unordered downward links, while the vertical arcs denote his other upward and downward links. The labels are redundant because each node has a unique position in the total network, and the only way to invoke a node is by linking directly to it; so a label such as ‘DOG’ never appears anywhere else in the network. The nodes are symbolic because each identifies a single concept, which is in turn only represented by that node. And there are no boundaries in the network between the ‘linguistic nodes’ in the bottom half of the network and the ‘non-linguistic nodes’ of ordinary cognition in the top half.

Claims #2 (symbolic network) and #3 (no boundaries) link us to the Cognitive Linguistics movement, which of course Lamb’s early work predated by a couple of decades. Maybe #1 (nothing but networks) distinguishes us from all other contemporary theories covering the whole of language; but we should both be pleased to note that our views are shared by an important theory of morphology, Network Morphology (Brown and Hippisley 2012). And, of course, Head-driven Phrase Structure Grammar presents a somewhat similar view of language as a rather limited kind of network, a Directed Acyclic Graph (Pollard and Sag 1994).

These general claims are easy to justify in terms of cognitive psychology, and Lamb (2016) offers a much better bridge to neuroscience than I could ever do. Lamb also makes the important point that the claims are easy to justify with evidence from linguistics, and that this evidence is an important contribution by linguistics to the general science of the mind. It is odd how reluctant other linguists are to make this point, especially given the importance of language in experimental cognitive psychology.

But having established these general claims, we then needed to work through their consequences for a detailed theory of language structure, the traditional target of linguistics. This is a project to which we have both dedicated most of our working lives, but working independently, so it is perhaps unsurprising that we have reached different conclusions in

various places. All the points of disagreement are issues for research rather than simply matters of personal taste, so we must hope that those who are attracted by the idea of 'network linguistics' will also engage in serious exploration of these questions.

To kick-start this process, here is a short list of points on which we at least appear to disagree. I offer it while recognising that it is all too easy to misunderstand a theory, so I may well have simply misunderstood.

Tokens. Word Grammar claims that token-concepts play an important part because we create a new node (and therefore a new concept) for every linguistic item that we process—for every word, sound segment, letter or whatever. The type-token distinction is of course a standard part of linguistic theory (Wetzel 2006), and figures outside linguistics in the 'type-token ratio' which is standardly used to measure the breadth of vocabulary in a text (Malvern et al. 2004); in terms of this distinction, the sentence *The cat sat on the mat* contains six tokens of five types (because the type *the* occurs twice). However, the conventional wisdom assigns types and tokens to two quite separate areas of theory, 'competence' for types and 'performance' for tokens. This theoretical contrast used to be supported by a similar contrast in cognitive psychology between long-term memory (competence) and short-term (or working) memory (performance), but many cognitive psychologists now reject this contrast in favour of a unified theory in which working memory is simply the part of long-term memory which is currently active (Ericsson and Kintsch 1995). Similarly, a network model of language treats the tokens of performance as temporary additions to the permanent network of types.

The claim in Word Grammar is both that we create temporary token nodes whether we are receiving a communication (as listeners or readers) or producing one, and that the same is true throughout cognition. Consequently, Word Grammar recognises the process of node-creation as one of a few very general mental operations. The structure of an utterance must be entirely composed of tokens, and not merely of activated stored types, because otherwise it would be impossible to recognise two tokens of the same type. In contrast, although I believe that Lamb's intention is to model production and perception, I don't see how token nodes can be created in his system.

Dependencies. One of the consequences of distinguishing token nodes from the types on which they are based is that a token word shows the influence of the words that depend on it. So in *French house*, the token word *house* is not simply a copy of the stored lexical item HOUSE, since its meaning is affected by the accompanying *French* so that this particular token of HOUSE actually means 'French house'. The Word-Grammar structure for an example phrase, *typical French houses*, is shown in Figure 2. The word tokens are labelled in italics, and the effect of 'modifying' a word by adding a dependent is shown by the '+' signs to distinguish one token from another. This example answers one of the most telling criticisms of dependency theory, that if each word has just one token node there is no node with the meaning 'French house' to which the other dependent, *typical*, applies (Dahl 1980).

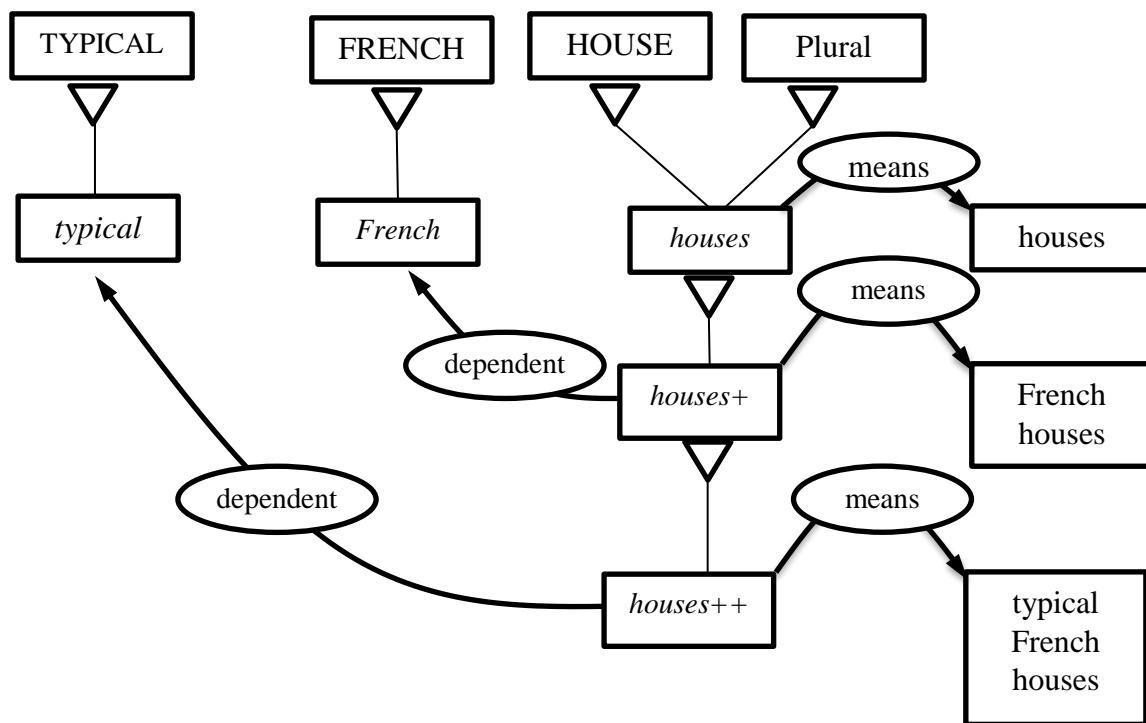


Figure 2. Dependency structure and meaning

The idea that a dependent creates a new token node removes the main motivation for phrase structure, and supports the much simpler approach of dependency structure in which the only units of syntax are single words; and one word may ‘govern’ another directly (e.g. the verb RELY selects the preposition ON, rather than a preposition-phrase headed by ON). However, Lamb’s mechanism for combining words appears to rely on the phrase-structure idea of a mother node which contains both the combined units as its parts; it doesn’t seem to allow for direct links between co-occurring units.

Referents. Another consequence of the earlier disagreement over tokens is that Word Grammar allows mental representations for experiences and objects in the world, just like the ones we have for tokens of words and other linguistic items; for instance, we can not only classify some bit of experience as a dog, but we can also represent this dog mentally by creating a temporary ‘token’ node for it. These item-bound, ad hoc and temporary concepts contrast both with the things in the world and with the stored categories to which we assign them. In contrast, Lamb (2016, 15) seems to say that there are no mental representations for objects in the world; but this means that we cannot identify one object as the same as another. For instance, if we hear a dog barking at 11 o’clock, we can classify it as a barking dog but we can’t recognise it as the same dog that we heard at 10 o’clock. This is a general weakness for any cognitive theory, but a critical weakness for a theory of language because it seems to rule out any theory of anaphora such as the one needed to explain the identity-of-reference anaphora in *It is barking again*.

Activation. Some of the strongest evidence for a network view of language comes from spreading activation (which Lamb 2016 doesn’t mention, although I’m sure he’s aware of it). This is responsible both for the priming effects found in experiments and for the speech errors we are all prone to; and in both cases the causal chain involves activation spilling over in an

uncontrolled and unintended way from a target node onto its neighbours. A standard example of priming is that we can retrieve the target word *nurse* more quickly shortly after hearing a related word such as *doctor*, *nursing* or *curse* than after hearing an unrelated word such as LORRY (Reisberg 2007, 257–59). Similarly in speech errors, the target word may be replaced by a related one (as when we mean ‘black’ but say *white*, or mean ‘eradicate’ but say *educate*) or by one which is already active in our planning (as when the famous Dr Spooner accused a student of tasting the whole worm when he meant ‘wasting the whole term’) (Harley 2006) Lamb’s much more orderly account of activation doesn’t seem to allow for activation spreading in this messy way if, as Lamb (2016, 7) says, activation is blocked where not all conditions are satisfied.

Exceptions. Both of us distinguish defaults from exceptions, but we do so in different ways. To make the discussion more concrete, suppose we are distinguishing the default pattern for English past-tense verbs (such as *baked*) from the exceptional form of an irregular verb such as *took*. How do we guarantee that the exception takes priority over the default? This is important because exceptions introduce serious uncertainty into any model of processing: if the processor can inherit the default property before the exceptional one, every property that’s predicted (‘inherited’) is just provisional because it may turn out later to be overridden; so for example if you’re generating the past tense of TAKE, and you first apply the default rule to produce *taked*, this will have to be revised when you find the exceptional rule for this particular verb. Because of this uncertainty, logicians often reject ‘default inheritance’ as ‘non-monotonic’ and unworkable, but it all depends on how we think default inheritance works, and Lamb and I both offer workable suggestions.

What Lamb (2016, 18) suggests is that exceptions actively block the default; so for example, the past tense of TAKE not only has the value *took*, but also has a blocking link to the default *-ed*. This is similar to a view I once espoused (Fraser and Hudson 1992), but I now think there is a better alternative, which once again involves the treatment of token nodes. In Word Grammar, default inheritance is part of the process of creating new token nodes, and only applies as part of this process; and as explained above, node-creation attaches the new node to an existing ‘type’ by means of an ‘isa’ link—the normal classification link shown by a small triangle in the figures. This means that the new token node sits at the very foot of an ‘isa hierarchy’ which links more particular concepts to more general concepts; for instance, in Figure 2 the token *houses++* (meaning ‘typical French houses’) isa *houses+* which isa *houses* which isa the types HOUSE and Plural, which in turn isa even more general concepts such as Common Noun, Noun and Word (not shown in the diagram). Now if the default inheritance mechanism always starts with the nearest ‘isa’ link and works cyclically up the isa hierarchy, it will automatically reach the exceptional value before the default value, so all we need is a general principle giving priority to any value which has already been inherited.

These research questions are really important because, as Lamb says, the study of language gives us an especially clear window into the human mind, so we linguists have the opportunity, and responsibility, to share our insights with our colleagues in other disciplines.

References

- Brown, Dunstan, and Andrew Hippisley. 2012. *Network Morphology. A Default-Based Theory of Word Structure*. Cambridge: Cambridge University Press.
- Dahl, Östen. 1980. "Some Arguments for Higher Nodes in Syntax: A Reply to Hudson's 'Constituency and Dependency'." *Linguistics* 18: 485–88.
- Ericsson, K.Anders, and W. Kintsch. 1995. "Long-Term Working-Memory." *Psychological Review* 102 (2): 211–45.
- Fraser, Norman, and Richard Hudson. 1992. "Inheritance in Word Grammar." *Computational Linguistics* 18: 133–58.
- Harley, Trevor. 2006. "Speech Errors: Psycholinguistic Approach." In *Speech Errors: Psycholinguistic Approach*, edited by Keith Brown, 739–45. Oxford: Elsevier.
- Hudson, Richard. 1984. *Word Grammar*. Oxford: Blackwell.
- Hudson, Richard. 1990. *English Word Grammar*. Oxford: Blackwell.
- Hudson, Richard. 2007. *Language Networks: The New Word Grammar*. Oxford: Oxford University Press.
- Hudson, Richard. 2010. *An Introduction to Word Grammar*. Cambridge: Cambridge University Press.
- Lamb, Sydney. 2016. "Linguistic Structure: A Plausible Theory." *Language Under Discussion* 4(1): 1–37.
- Malvern, David, Brian Richards, Ngoni Chipere, and Pilar Duran. 2004. *Lexical Diversity and Language Development: Quantification and Assessment*. Houndmills, UK: Palgrave Macmillan.
- Pollard, Carl, and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: Chicago University Press.
- Reisberg, Daniel. 2007. *Cognition. Exploring the Science of the Mind*. Third Media Edition. New York: Norton.
- Wetzel, Linda. 2006. "Type versus Token." In *Encyclopedia of Language and Linguistics, Second Edition*, edited by Keith Brown, 199–202. Oxford: Elsevier.