Hierarchies and networks in applied linguistics

Richard Watson Todd

King Mongkut's University of Technology Thonburi

Abstract

Methods of representing findings are rarely given careful consideration in applied linguistics. One example is the general preference for hierarchies rather than networks to represent relationships between concepts, largely because of the more easily understandable structure and seeming objectivity of hierarchies. However, the purported objectivity of hierarchies is open to dispute on several grounds. Networks, on the other hand, are the preferred way of representing relationships in much of cognitive psychology, mainly because they fit the research results well and are more empirically valid. Focusing on ways of representing organisation of concepts, this paper compares the use of hierarchies and networks as methods of representing relationships with an example of how each might be used to investigate topical relevance in aphasic discourse. It is concluded that, while hierarchies are clearer and easier to work with, networks provide a more valid method of representation.

Hierarchies and networks in applied linguistics

Language organises information, and the study of language necessarily entails the study of how information is organised, including how conceptual knowledge is organised and how concepts relate to each other. The study of organisation of concepts is also one of the main preoccupations of cognitive psychology. The approaches used to describe the organisation of concepts in linguistics and cognitive psychology, however, stand in contrast. Many proposed models of language processing especially at the discourse level do not reflect the cognitive structures that must underpin the processing. An example of this mismatch between models in applied linguistics and in cognitive psychology is the frequent use of hierarchies in applied linguistics compared to the emphasis placed on networks in cognitive psychology.

In this paper, I will compare the use of hierarchies and networks in describing the organisation of concepts, with the main focus being on how relationships between concepts can be represented. I hope to show that the more frequently used hierarchies are open to criticism on many grounds. An alternative way of representing the organisation of concepts, namely, networks, on the other hand, is a valid reflection of how people organise concepts in their minds, and thus networks may be preferable in many situations. Choosing between whether to use a hierarchy or a network to represent an organisation of knowledge, however, is no easy matter.

The nature of hierarchies

Perhaps the most familiar hierarchy is the taxonomic classification of living organisms originally proposed by Linnaeus. As with most hierarchies, the Linnaean classification is organised with a few very general categories at the top, such as the animal and plant kingdoms. Below these are a larger number of more specific categories, including the arthropod and mollusc phyla under the animal kingdom. This pattern of each level down the hierarchy having a greater number of more specific items is repeated throughout the Linnaean classification, so that, by the time we reach the bottom level of species, the five kingdoms at the top have been divided and subdivided into a vast number of very specific items.

Hierarchies, then, usually consist of a few (often only one) general categories at the top each of which is divided into several more specific categories or members. This hierarchical pattern of general to specific has also provided the basis for a lot of linguistic analysis. In semantics, for example, it is common to find words organised into hierarchical structures as one way of analysing meanings (see e.g. Lyons, 1977), and Figure 1 shows an example parallel to the Linnaean classification.

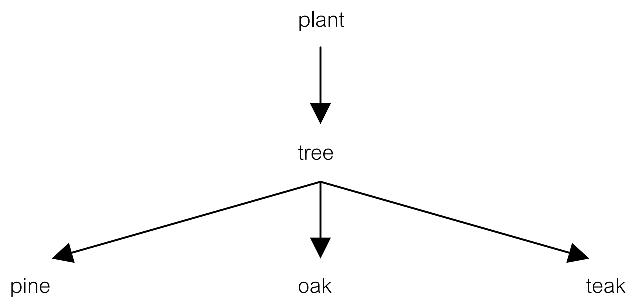

Figure 1 A typical hierarchy

The relationships in both the Linnaean classification and Figure 1 are hyponymic. In Figure 1, *pine* is a hyponym of *tree*, and conversely, *tree* is the superordinate of *pine*. Hyponymic relations can be defined through entailment, so for Figure 1 the statements given in (1) are true.

(1) If X is a pine, X is a tree.

If X is a tree, X is not necessarily a pine.

Hyponymy is not the only relation that can be used to build hierarchies. Since hierarchies represent differences in levels of specificity, any relation that distinguishes between general and specific can be used. Thus the hierarchical tree diagrams favoured by most models of syntax are based on meronymy (part-whole relations), where noun phrases and verb phrases are parts of a sentence. Other possible relations which can be used in hierarchies include entity-characteristic relations (e.g. elephant - big) and identifier - identified relations (e.g. robot - C3PO). There are, then, a variety of bases for constructing hierarchies, although hyponymy and meronymy predominate in linguistics.

The uses of hierarchies in applied linguistics

In addition to the hierarchies used in semantics and syntax discussed above, hierarchies have been applied in many other areas of linguistics, starting with Roget's lexicographic use of a hierarchy to organise his thesaurus.

Within discourse analysis, hierarchies have also been influential. A vast range of approaches has used hierarchies to explain the structure of discourse. For example, Sinclair and Coulthard's (1975) influential study of classroom discourse was predicated upon a hierarchical series of ranks in discourse with the lesson being the highest rank and the act the lowest. A classroom lesson analysed following Sinclair and Coulthard produces a hierarchy like Figure 2.

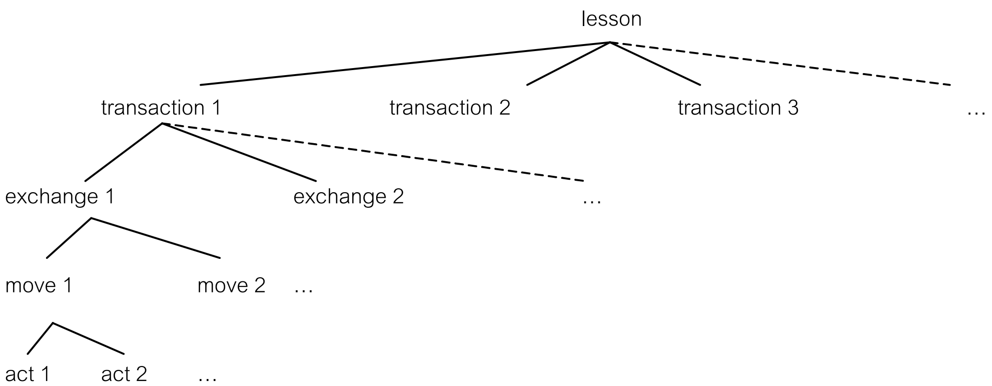

Figure 2 A hierarchical tree based on Sinclair and Coulthard (1975)

Other uses of hierarchies to analyse discourse include the following: texts have been described as a series of hierarchically-organised propositions (Grabe, 1984; Kintsch and Keenan, 1973; Tomlin *et al.*, 1997); coherence can be viewed as a hierarchy with local coherence subordinate to global coherence (Graesser *et al.*, 1994); topics, it has been argued, are best thought of as macro-propositions consisting of sub-topics and sub-sub-topics (Ellis, 1983; van Dijk and Kintsch, 1983); and both schemata (Long, 1989; Mann and Thompson, 1988; Slavin, 1994) and scripts (Abbott *et al*., 1985; Whitney, 1998) have been described using hierarchies.

Similarly, hierarchies have been very influential in educational theory, which together with linguistics provides much of the input into educational applied linguistics. Hierarchies have been used to describe the various educational disciplines (Bruner, 1960; Mohan, 1986), to describe syllabi (Woods, 1996), and to organise the teaching/learning process (Ausubel, 1963; Cole and Chan, 1987; Erickson, 1982). For example, Mohan (1986: 89) represents the structure of school-level mathematics using a hierarchy as shown in Figure 3.

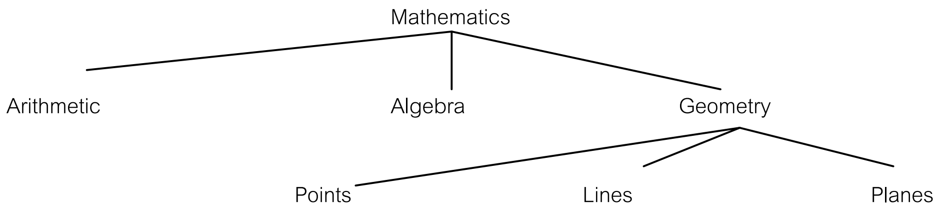

Figure 3 A hierarchical structure for mathematics (from Mohan, 1986: 89)

Criticisms of hierarchies

Given this massive wealth and range of uses of hierarchies, most of which have proved to be valuable and productive, why is there a need for a paper evaluating their value? It could be argued that the productiveness of the analyses using hierarchies already proves their value, and so no re-assessment of the use of hierarchies in applied linguistics is needed. However, there have also been several criticisms of the use of hierarchies which should give us pause to think more deeply about whether we should be using hierarchies as the basis of applied linguistic descriptions.

Some of these criticisms apply only to the use of hierarchies in specific areas. For example, we saw above that hierarchies have been used to describe topics, but Hudson (1980) argues that in situations where there is topic drift, in other words, where succeeding discourse moves seamlessly from one topic to another with no clearly identifiable boundaries, no hierarchical structure is evident.

A broader criticism of hierarchies is that they present a static picture of the world, whereas language use is dynamic. Thus, if we describe a text in terms of a hierarchy, there are problems in describing and explaining the sequence in which units in the hierarchy are introduced (van Lier, 1988).

The majority of criticisms, however, relate to the fact that hierarchies imply a very clear-cut view of knowledge that is not true in the messy realities of the world. For example, the hierarchy shown in Figure 1 appears clear-cut and indisputable. However, it is open to criticism. Firstly, the hierarchy implies that *plant* is the immediate superordinate of *tree*, but there may actually be several intermediate superordinates including *woody plant* and *flowering plant*. Secondly, if we add more concepts into Figure 1, such as *material used in making furniture*, we see that *pine*, *oak* and *teak* are co-hyponyms of two immediate superordinates, the relationship between which is unclear. Such cross-relationships, which hierarchies cannot cope with, are common in the real world (Schank, 1975; Strahan, 1989). Thirdly, there are frequently different principles available as the basis for organising a hierarchy. The taxonomic hierarchy in Figure 1 may differ from the hierarchy for trees conceptualised by forest park rangers who may categorise trees based on their health and fecundity rather than on species (Medin *et al*., 1997). These different bases for relations between concepts mean that the supposedly objective relations used to construct hierarchies frequently rely more on a researcher's subjective interpretations than on objectivity.

A further criticism of hierarchies is that a single hierarchy may include several different kinds of relations. For example, the hierarchic schema in Figure 4 taken from Slavin (1994: 196) is problematic.

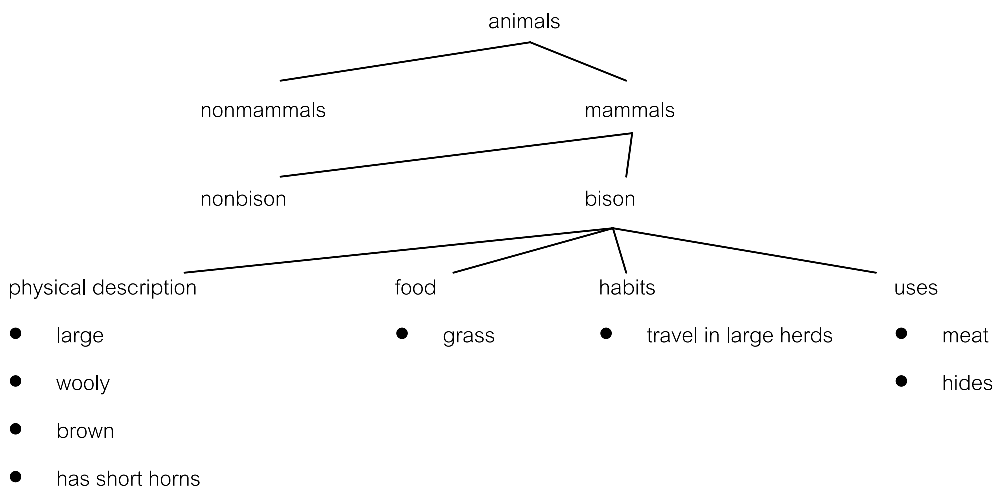

Figure 4 A schema for the word *bison* from Slavin (1994: 196)

In Figure 4, starting from the top, the first three levels are a taxonomic hierarchy based on hyponymic relations. The bottom level, however, is not based on hyponymy even though no diagrammatic distinction is made in the hierarchy to show any difference in relations between concepts. The relation between *bison* and, say, *grass* is more akin to one of entity and characteristic where eating grass is a characteristic of bison. Such mixing of different relations in hierarchies is quite common and clouds the supposedly neat picture that they present.

A further problem with Figure 4 concerns why *grass* should be placed at a lower level in the hierarchy than *bison*. We have seen that lower levels in hierarchies indicate items of a more specific nature, but is *grass* really more specific than *bison*? We can easily imagine someone looking at the schema in Figure 4 and deciding to construct a similar one for *grass* which might look like Figure 5.

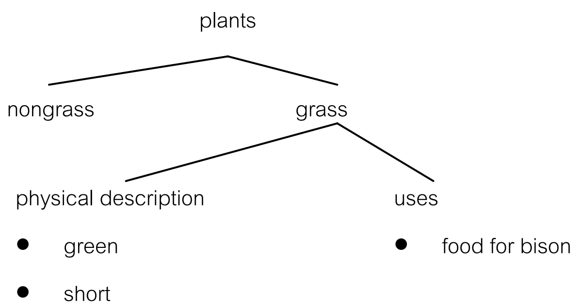

Figure 5 A schema for the word *grass*

In Figure 5, *bison* appears to be at a lower level in the hierarchy than *grass* in direct contrast to the relation shown in Figure 4. Problems such as these mean that the ostensible objectivity of a hierarchy may not actually be present and many hierarchies presented in the literature are suspect.

Reasons for using hierarchies

Having seen such a variety of criticisms of hierarchies, we are left with wondering why they are used so frequently in linguistics. As all the work from a Chomskyan perspective has shown, logic and the use of hierarchies can be powerful tools in analysing language.

The reason why Chomskyan syntax has been so successful rests on the fact that it provides theoretic, rather than empiric, explanations (Scribner, 1979). Most of the progress made in syntax has come from analyses of idealised language use rather than from analyses of messy real-world interaction. In applied linguistics, on the other hand, it is real rather than idealised language use that provides the focus for analysis, and thus applied linguistics seeks empiric explanations. Whether the logic and hierarchies used in Chomskyan approaches are relevant to the search for empiric explanations is unclear.

Similarly, the hierarchical tree diagrams used in syntax are valid by definition. If you define a subject as a component of a sentence, then the fact that a subject is a meronym of a sentence is self-evident by definition. The exact nature of the subject and the sentence are irrelevant as far as constructing meronymic hierarchies based on such meronymic definitions of syntactic components is concerned. In other words, such syntactic descriptions are internally valid but are restricted to descriptions of internalised I-language (Chomsky, 1988) divorced from context. In applied linguistics, on the other hand, analyses generally involve the state of the world, the context, and often speaker meanings and uses, as well as definitions of linguistic components. As such, applied linguistics is concerned with synthetic propositions rather than the analytic propositions of theoretical syntax (Pyles and Algeo, 1973). The placement of synthetic propositions into purportedly objective hierarchies can be fraught with problems. For example, statements whose meanings change depending on context cannot be easily assigned a place in a hyponymic semantic hierarchy. Because of problems like this, even if the use of hierarchies is valid in pure linguistics, we are still left searching for justifications of the use of hierarchies in applied linguistics.

The only justification of choosing a hierarchy to represent information organisation that I have found in the applied linguistics literature is given by Mohan (1986). In looking at ways to structure knowledge in content-based language teaching, Mohan compares the use of hierarchies and networks to represent information organisation. He chooses to use hierarchies in his analysis, arguing that "the main difference is the orderliness and precision of the [hierarchy] compared to the unruly proliferation of the [network]" (89). If the goal of an analysis, then, is to present a clear-cut and precise representation, hierarchies are preferable to networks.

Mohan's work, however, is just one of many in applied linguistics which have used hierarchies. The use of hierarchies in other studies, however, is not explicitly justified by researchers. This implies that the main reason for using hierarchies may be inertia. In other words, a hierarchy may be used in one study because a previous similar study used hierarchies. This lack of any rationale for using hierarchies demands that the use of hierarchies be evaluated, and such an evaluation must compare the use of hierarchies with the available alternatives.

The alternatives to hierarchies

A variety of ways of representing knowledge organisation within applied linguistics have been suggested. For organising concepts, these suggestions have included grids or matrices, flow charts, algorithms, hierarchical tree diagrams, and networks (Burgess, 1994; Graney, 1992; Mohan, 1986). Of these, flow charts and algorithms can be used to describe processes, and grids are useful for comparisons. For representing patterns of relationships in conceptual knowledge, then, we are left with a choice of hierarchies (and their mathematically equivalent alternatives such as Venn-Euler diagrams, see Lipschutz, 1964) and networks. Choosing between hierarchies and networks is a true choice in that they can frequently be substituted for each other in the same applications, albeit providing a different picture of the relationships. Some of the areas of discourse above which have been described using hierarchies can also be described using networks. For example, schemata have been presented as networks (e.g. Anderson and Pearson, 1984), and patterns of concepts in texts have also been described using networks (e.g. de Beaugrande and Dressler, 1981; Hoey, 1991).

Indeed, the fact that some patterns of relationships can be represented either through hierarchies or networks is such that, on occasion, it is unclear which is being used. For example, in analysing a large corpus of keywords, Scott (1997) found that keywords can be grouped into clumps which are best represented as networks. However, in a later article (Scott, 2000), using the same data and analysis, he argued that the different clumps produce a hierarchical tree of concepts, despite the fact that there is nothing in the data to indicate which of the clumps are specific and which general. Thus, Scott appears to be trying to force a network into a hierarchical structure. On the other hand, Halliday (1973), in presenting a hierarchical algorithm of options in language use, argues that the algorithm is a network. In this case, Halliday is doing the opposite of Scott and attempting to pass off a hierarchy as a network.

Given the criticisms of hierarchies I have presented above and the fact that hierarchies and networks may sometimes be used to describe the same patterns of relationships, let us now investigate the applicability of networks as an alternative to hierarchies.

A brief history of networks

While networks are used relatively infrequently in applied linguistics, they are almost de rigueur as a way of representing patterns of conceptual knowledge in cognitive psychology.

The use of networks to represent the organisation of concepts in cognitive psychology grew out of a dissatisfaction with hierarchies. In 1969, Collins and Quillian posited a mixture of hierarchies and networks to represent conceptual organisation. The basic structure was a hierarchical tree diagram, but at each of the nodes in the diagram other concepts were attached in the form of networks as shown in Figure 6. They further posited that response times taken to link two concepts would be proportional to the distance between the two concepts in the diagram. Their initial results suggested that this was true.

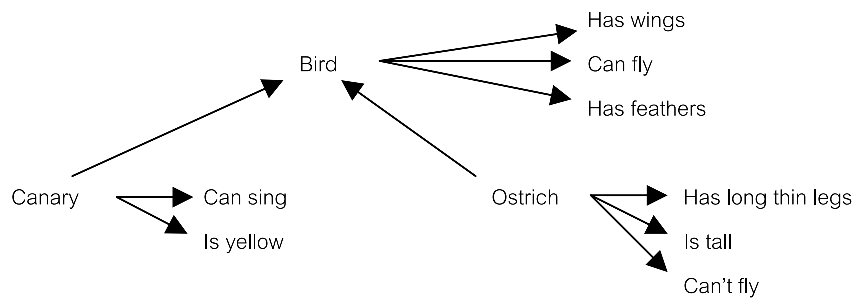

Figure 6 A mixed hierarchy and network (from Collins and Quillian, 1969)

However, Conrad (1972) took an alternative approach. Instead of predetermining the relationships between concepts as a hierarchy, he asked subjects to describe certain concepts and drew up networks of the concepts based on those descriptions. He then showed that these networks were better predictors of response times than Collins and Quillian's hierarchy. Since Conrad's research, networks have been more influential as ways of representing conceptual organisation in cognitive psychology than hierarchies.

There have, however, been several progressively more complicated variations on the theme of networks to represent the organisation of concepts. Initial models proposed a spreading activation between nodes in a network (e.g. Collins and Loftus, 1975). In other words, activation of nodes in a network spreads outwards along connections from one initially activated node to other nodes and then outwards again from these newly activated nodes.

Further refinements to this spreading activation model incorporated the strength of connections between nodes in the model (e.g. Anderson, 1980). In Collins and Quillian's (1969) model described above, the strength of a relationship between two concepts was posited as proportional to the distance between the two concepts in the diagram. In other words, related concepts are seen as existing close to each other in semantic space while unrelated concepts are distant (Van Dijk, 1977). This way of representing the extent of the relationship between two concepts as distance in semantic space can still be found in some network models (e.g. Hofstadter, 1996), but more usually, strengths of relationships are represented by weightings on connections. A strongly weighted connection indicates a close relationship or closeness in semantic space between the two concepts it links. Yet further refinements allowed the weightings of connections to be negative as well as positive (usually ranging from -1 to +1) where a negative weighting inhibits the activation of the node to which it is linked, and allowed activation to move backwards and forwards along connections rather than only spreading outwards.

A more recent development of network models produces somewhat less intuitive models. Instead of concepts being represented as nodes, the nodes themselves carry no meaning - instead, concepts are represented by combinations of connections between nodes (McClelland and Rumelhart, 1985). Such models are termed connectionist. Connectionist models are based on the metaphor of neurons in the brain and, as with the brain, can be fearsomely complicated. In the brain, each of the 20 million Purkinje cells can have up to 200,000 connections. Similarly, even a simple connectionist model can contain thousands of connections.

Network models in cognitive psychology, then, largely originated from dissatisfaction with a hierarchical model. The original simple network models have been continuously refined so that most present models are very complicated. In the next section, I will present an example of a basic network to illustrate most of the points discussed in this section. The model I will present is a spreading activation model with positive and negative weightings on connections and with backwards and forwards activation along connections. I will not attempt to present a connectionist model as the counter-intuitiveness and complexities of connectionism make such models difficult to explain and understand, but the example presented below could be reworked as a connectionist model relatively easily.

An example of a network

The example I will present here is very simple in that it consists of three nodes only. It aims to show how the hyponymic relations underpinning many hierarchies may be represented as a network using Schank's (1975) concept of *ISA* (see also Schank and Abelson, 1975, 1977). The *ISA* relation relates a hyponym to its superordinate, so from Figure 1 we can say *pine ISA tree*. Figure 7 shows a network with three nodes and weighted two-way connections.

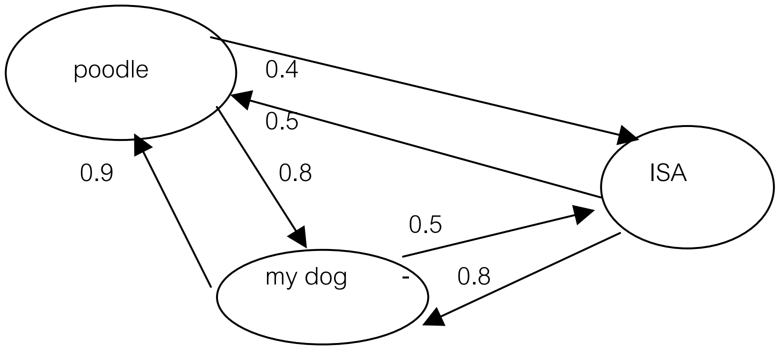

Figure 7 An example of a simple network

In Figure 7, we can see that each of the three nodes is connected to each of the others and that each connection has a number between -1 and +1 indicating its weighting. A high number (such as 0.9 from *my dog* to *poodle*) indicates a close relationship between the two concepts at the nodes. A negative weighting, on the other hand, means that activating the initial node (e.g. *ISA*) inhibits the activation of another node (e.g. *my dog*). A negative weighting is given from *ISA* to *my dog* since *my dog* is not a concept superordinate to any others.

To see how the network operates, let us suppose that I am thinking of *my dog*, in other words, *my dog* is activated (i.e. there is a value at the *my dog* node of 1). This activation will spread along the connections to other nodes giving an activation at the second node equivalent to the activation at the first node multiplied by the weighting of the connection from the first to the second node. Thus, *poodle* receives an activation of 1 (the level of activation of *my dog*) multiplied by 0.9 (the weighting on the connection from *my dog* to *poodle*) or 0.9, and *ISA* receives an activation of 0.5. The activations of *poodle* and *ISA* become new inputs into the spreading activation in addition to continued input from *my dog*. For example, *ISA*, having received an activation of 0.5 from the first 'round' of spreading activation originating from *my dog*, now receives further activation from both *my dog* and *poodle*. In this second 'round', *ISA* receives an activation of 0.5 from *my dog* and a further 0.36 (0.9 [the level of activation of *poodle*] x 0.4 [the weighting on the connection from *poodle* to *ISA*]) from *poodle*. In each round, the activation received by each node is:

Activation received at node A = $\sum$ ((the level of activation at each connected node) x (weighting on the connection from the connected node to node A))

(see Best, 1999 for details)

We can set a threshold for activation of, say, 0.8 (this threshold is necessary because following the feedback loops of continuously spreading and respreading activation will eventually lead to limits of 0 or infinity), which, once reached, means the node is activated. Thus, *my dog* is activated from the initial input of 1 and *poodle* is activated from the first 'round' of spreading activation when it reaches a value of 0.9. *ISA* is not activated in the first 'round' (when it only receives 0.5 from *my dog*), but in the second 'round', *ISA* receives 0.5 from *my dog* and 0.36 from *poodle* giving a total of 0.86 which is above the threshold level for activation.

The upshot of all this is that all three nodes are activated, meaning that every time *my dog* is activated, *ISA* and *poodle* are also activated. In other words, thinking of *my dog* always brings the realisation that *my dog ISA poodle*.

Arguments for and against networks

From even such a simple example as the three-node network in Figure 4, we can see that the use of networks is not straightforward. It takes a fair amount of work to find out that activating *my dog* activates *my dog ISA poodle*. In contrast, a simple hierarchy with *poodle* at the top and *my dog* at the bottom would show the hyponymic relation clearly at a glance. The implications and meanings residing in networks are generally not immediately apparent.

Furthermore, any sequencing of relations or concepts in a network like Figure 4 is unclear. I summarised Figure 4 as showing *my dog ISA poodle*, but there is no reason within the network why it could not be summarised as *ISA poodle my dog* or another variant. A network of concepts, then, does not automatically include a syntax for sequencing the concepts. It is possible to redraw the network to include such a sequence, but doing this would involve adding many extra nodes and connections, making the network even more convoluted and opaque.

A third problem with networks is how to construct them. Figure 4 is a purely imaginary network that was created to illustrate a point. Drawing up networks to describe and analyse real language data should not rely on intuition. However, it is difficult to see how networks, and especially the weightings of connections, could be constructed without relying on intuition. I will return to this point below.

Despite these problems with networks, they do have one massive advantage over hierarchies in representing relationships between concepts. Networks predict research results well and so are likely to be better representations of how people really organise knowledge than hierarchies. For example, there has been a lot of research using networks which has produced results matching people's reading performance and acquisition (e.g. Ans *et al.*, 1998; McClelland and Rumelhart, 1981; McEneaney, 1994; Seidenberg, 1992). Other aspects of linguistics which have been validly modelled by networks include child acquisition of German articles (McWhinney *et al.*, 1989), learning of regular and irregular forms of English verbs (Ellis and Schmidt, 1998), and ability to identify word boundaries (Christiansen *et al.*, 1998). If the goal of an analysis is to provide a valid and realistic picture of relationships among concepts, networks should be used instead of hierarchies.

One further consequence of the predictive validity of networks can be viewed as either an advantage or a disadvantage depending on perspective. Some of the research into describing aspects of language using networks (e.g. McWhinney *et al.*, 1989) has explicitly contrasted network models with rule-based models. In many networks, and especially in connectionist models, there is no explicit manipulation of symbol systems - in other words, there is no use of rules as they are usually understood in linguistics (Sokolik, 1990). The research contrasting network models and rule-based models has generally shown that networks explain human thinking better (e.g. Hunt, 1989; Ney and Pearson, 1990). If this is true and generalisable to describing organisation of concepts and other aspects of linguistics, then the goal of much linguistic research could change from a search for generalisable patterns expressible as rules to constructing networks that describe data well and following the implications and meanings of these networks. It should be stressed, however, that not all use of networks results in such an either-or choice between networks and rules.

Lastly, unlike hierarchies, networks are not restricted to organising concepts on the basis of relations. Hierarchies, as we have seen, can be built around a variety of general-specific relations, although to be logically valid these relations should not be mixed in a single hierarchy as is the case in the schematic hierarchy of Slavin (1994) discussed above. Networks can also be used to represent these relations albeit less clearly than hierarchies. In addition, networks can also represent associations which cannot be dealt with by hierarchies. For example, in a recent study Schmitt (1998) elicited association responses to the word *dark*. The most frequent responses were *light*, *night* and *fear*. These associations can easily be represented as a network by placing *dark* in a central node connected to three peripheral nodes of *light*, *night* and *fear*. The weighting of the connection from *dark* to each of these three concepts would be proportional to the frequency at which they were elicited as association responses. Further connections could be added to the network by eliciting association responses for the three peripheral concepts.

The associations between *dark*, on the one hand, and *light*, *night* and *fear*, on the other, cannot be represented in a single hierarchy. As opposites, *dark* and *light* are co-hyponyms of the same superordinate, say, *brightness*; *dark* is a characteristic of *night*; and *fear* is an characteristic of *dark*. *Dark*, then, is at a more specific level under two general concepts, *brightness* and *night*. Since there is no relation holding between these two more general concepts and their relations with *dark* are not of the same kind, we would need to separate them into two different hierarchies. So, while a set of associations can be dealt with easily, the amount of variation of types in most associative sets means that hierarchies are generally inappropriate as a way of representing associations. Given that both relations and associations can be represented in networks, we can conclude that networks can deal with a greater range of types of closeness between concepts than hierarchies.

An example of choosing between hierarchies and networks

So far in this paper, I have presented a range of arguments for and against both hierarchies and networks. Unsurprisingly then, the choice of whether to use a hierarchy or a network in a description is dependent on the particular analysis, its situation and its goals. To illustrate this, I will look at a very small data set and compare the use of hierarchies and networks to describe it.

The example concerns topical relevance in aphasic speakers. Some kinds of semantic pragmatic disorder result in aphasic speakers having difficulty keeping contributions relevant to a topic (Lesser and Milroy, 1993). For example, in (2) taken from Stubbs (1986: 185), the contribution of B, an aphasic patient, does not appear relevant to A's question.

(2) A: Were these children in it?
B: I saw a baby chick trying to fly.

The problem here is how to measure whether B's utterance is relevant to the topic of A's question. Relevance to a topic is manifested by succeeding concepts which are close in semantic space (Van Dijk, 1977; Watson Todd, 1998). In other words, two succeeding concepts separated by a great distance in semantic space can indicate a lack of topical relevance (although this is dependent on context, see Levinson, 1983; Widdowson, 1979). The problem of measuring relevance, then, frequently resolves to identifying distance in semantic space, and ways of representing closeness between concepts may be helpful in identifying this.

Let us start, then, by looking at how the concepts in (2) could be represented in a hierarchy. Keeping things very simple and ignoring the unclear *it* and the functional rather than conceptual *I*, the main concepts in (2) are *children*, *baby chick* and *try to fly*. Looking for relations between these three concepts, we may identify *try to fly* as an characteristic of *baby chick*, but it is difficult to identify any relation between *children* and the other two concepts. It should be noted that identifying relations like these is not truly objective but relies to a certain extent on the analyst's interpretation. Nevertheless, if we accept that *baby chick* and *try to fly* exhibit a entity-characteristic relation and *children* is unrelated, we can construct a hierarchy representing these relations as in Figure 8.

Figure 8 A hierarchical representation of (2)

In Figure 8, we can see a close general-specific relation for *baby chick* and *try to fly*, whereas *children* is not placed within this hierarchy but is in a separate hierarchy on its own. This suggests that *baby chick* and *try to fly* are close in semantic space, but *children* is distant from both these concepts. Applying these distances in semantic space to (2), B's utterance is internally topically relevant but is not relevant to A's question.

From this analysis, it appears that hierarchies can be used to aid analyses of topical relevance. However, it should be stressed that in analysing (2), we are looking at only three concepts. For longer stretches of data (and most stretches of data will be longer), the number of concepts involved will increase, and there will be a concurrent increase in the likelihood that valid hierarchies linking concepts on the basis of a single relation cannot be constructed.

Turning now to networks, if we attempt to make a network based solely on the data in (2), we end up with three connected nodes for *children*, *baby chick* and *try to fly*. The weightings on the connections must be high since the three concepts at the nodes are closely associated as they appear in the same short stretch of discourse. Any network based solely on a short stretch of discourse would show such high levels of connectedness, irregardless of whether the discourse exhibited topical relevance or not.

We therefore need to look at ways of constructing a network which are based on data from outside the extract. Such a network would act as a benchmark against which the extract could be compared. As we have seen, one way in which this could be done is to build a network based on word associations. However, this would require a large group of subjects from whom word associations could be elicited, and even then, there is no guarantee that any association between, say, *baby chick* and *try to fly* would be elicited.

An alternative way of constructing a benchmark network would be to create a network from the language data of a large corpus. Concepts which are close in semantic space should co-occur in a corpus relatively frequently. We could set a range of, say, 20 words as a span for counting co-occurrences. The proportion of co-occurrences as a percentage of all occurrences of the concepts would give us a measure of how closely two concepts are associated. Applying this method to *children*, *baby chick* and *try to fly* using *The COBUILD Bank of English* corpus, unfortunately we find that the phrases *baby chick* and *try* to fly* occur too infrequently to be of much use for analysis (6 and 87 occurrences respectively). I will therefore look at the number of co-occurrences within a span of 20 words of *children*, *chick* and *fly (v)*. From the corpus, we find that 0.003% of occurrences of *children* co-occur with *chick*, 0.062% of occurrences of *children* co-occur with *fly (v)*, and 0.308% of occurrences of *chick* co-occur with *fly (v)*. From these proportions, we could draw up a network converting the percentages of co-occurrences into weightings ranging from -1 to +1. Such a network would show that connections between *chick* and *fly (v)* are most strongly weighted, while connections between *children* and the other two concepts are more weakly, and perhaps even negatively, weighted. The outcome of inputting the data from (2) into such a network would show that B's response in (2) is internally topically relevant but not relevant to A's question. This finding matches the finding concerning topical relevance obtained by using a hierarchy, but may be preferable as it is founded on language data and avoids the perhaps subjective steps involved in identifying relations needed to construct a hierarchy.

Conclusion

In the previous section, I hope that I have shown that the organisation of concepts in language data can be represented as either a hierarchy or a network. Hierarchies are much easier to construct and understand, but their purported objectivity founded on the fact that relations like hyponymy are testable through logic is suspect. With the exception of tautologous theoretic explanations, a large amount of subjective interpretation may still be needed in constructing a hierarchy. Networks, on the other hand, although laborious to construct and work with, are more empirically valid as they can be based on real language data. For this reason, networks may present a truer picture of conceptual organisation than hierarchies. In addition, their empirical foundations mirror the recent emphasis in linguistic research on analyses of real language data as epitomised by corpus linguistics rather than analyses of made-up examples.

The choice between hierarchies and networks, therefore, seems to come down to a choice between providing a clear picture and providing a true picture. While this paper has focused on ways of representing organisation of concepts, I believe that the arguments also apply to investigations in other areas of applied linguistics. Since most research in applied linguistics aims to provide empirically valid explanations of language use, careful consideration needs to be given as to whether the more difficult but more valid network representations of data should be used in preference to the presently predominant hierarchies.

References

Abbott, V., J. B. Black, and E. E. Smith. 1985. 'The representation of scripts in memory.' In L. K. Komatsu (ed.): *Experimenting with the Mind: Readings in Cognitive Psychology*. Pacific Grove, CA: Brooks/Cole.

Anderson, J. 1980. *Cognitive Psychology and its Implications*. San Francisco: Freeman.

Anderson, R. C. and P. D. Pearson. 1988. 'A schema-theoretic view of basic processes in reading comprehension.' In P. Carrell, J. Devine, and D. Eskey (eds.): *Interactive Approaches to Second Language Reading*. Cambridge: Cambridge University Press.

Ans, B., S. Carbonnel and S.Valdois. 1998. 'A connectionist multiple-trace memory model for polysyllabic word reading.' *Psychological Review* 105: 678-723.

Ausubel, D. P. 1963. *The Psychology of Meaningful Verbal Learning*. New York: Grune and Stratton.

Best, J. B. 1999. *Cognitive Psychology*, 5th edition. Belmont, CA: Brooks/Cole.

Bruner, J. 1960. *The Process of Education*. Cambridge, MA: Harvard University Press.

Burgess, J. 1994. 'Ideational frameworks in integrated language learning.' *System* 22: 309-318.

Chomsky, N. 1988. *Language and Problems of Knowledge*. Cambridge, MA: MIT Press.

Christiansen, M. H., J. Allen and M. S. Seidenberg. 1998. 'Learning to segment speech using multiple cues: a connectionist model.' *Language and Cognitive Processes* 13: 221-268.

Cole, P. G. and L. K. S. Chan. 1987. *Teaching Principles and Practice*. Sydney: Prentice Hall.

Collins, A. M. and E. F. Loftus. 1975. 'A spreading activation theory of semantic processing.' *Psychological Review* 82: 407-428.

Collins, A. M. and M. R. Quillian. 1969. 'Retrieval time from semantic memory.' In L. K. Komatsu (ed.): *Experimenting with the Mind: Readings in Cognitive Psychology*. Pacific Grove, CA: Brooks/Cole.

Conrad, C. 1972. 'Cognitive economy in semantic memory.' Journal of *Experimental Psychology* 92: 149-154.

De Beaugrande, R. and W. Dressler. 1981. *Introduction to Text Linguistics*. London: Longman.

Ellis, D. G. 1983. 'Language, coherence and textuality.' In R. T. Craig and K. Tracy (eds.) *Conversational Coherence: Form, Structure and Strategy*. Beverley Hills, CA: Sage.

Ellis, N. C. and R. Schmidt. 1998. 'Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax.' *Language and Cognitive Processes* 13: 307-336.

Erickson, F. 1982. 'Classroom discourse as improvisation: relationships between academic task structure and social participation structure in lessons.' In L. W. Anderson (ed.): *The Effective Teacher: Study Guide and Readings*. New York: McGraw-Hill.

Grabe, W. 1984. 'Written discourse analysis.' *Annual Review of Applied Linguistics* 5: 101-123.

Graesser, A. C., M. Singer, and T. Trabasso. 1994. 'Constructing inferences during narrative text comprehension.' *Psychological Review* 101: 371-395.

Graney, J. M. 1992. 'A framework for using text graphing.' *System* 20: 161-167.

Halliday, M. A. K. 1973. *Explorations in the Functions of Language*. London: Edward Arnold.

Hoey, M. 1991. *Patterns of Lexis in Text*. Oxford: Oxford University Press.

Hofstadter, D. R. 1996. 'Analogy-making, fluid concepts, and brain mechanisms.' In A. Clark and P. J. R. Millicon (eds.): *Connectionism, Concepts and Folk Psychology: The Legacy of Alan Turing vol. 2*. Oxford: Clarendon Press.

Hudson, R. A. 1980. *Sociolinguistics*. Cambridge: Cambridge University Press.

Hunt, E. 1989. 'Connectionist and rule-based representations of expert knowledge.' *Behavior Research Methods, Instruments and Computers* 21: 88-95.

Kintsch, W. and J. Keenan. 1973. 'Reading rate and retention as a function of the number of propositions in the base structure of sentences.' *Cognitive Psychology* 5: 257-274.

Lesser, R. and L. Milroy. 1993. *Linguistics and Aphasia: Psycholinguistic and Pragmatic Aspects of Intervention*. London: Longman.

Levinson, S. C. (1983) *Pragmatics*. Cambridge: Cambridge University Press.

Lipschutz, S. (1964) *Schaum's Outline of Theory and Problems of Set Theory and Related Topics*. Singapore: McGraw-Hill.

Long, D. R. 1989. 'Second language listening comprehension: A schema-theoretic perspective.' *The Modern Language Journal* 73: 32-40.

Lyons, J. 1977. *Semantics Volume 1, Volume 2*. Cambridge: Cambridge University Press.

Mann, W. C. and S. A. Thompson. 1988. 'Rhetorical Structure Theory: toward a functional theory of text organization.' *Text* 8: 243-281.

McClelland, J. L. and D. E. Rumelhart. 1981. 'An interactive activation model of context effects in letter perception: part 1. An account of basic findings.' *Psychological Review* 88: 375-407.

McClelland, J. L. and D. E. Rumelhart. 1985. 'Distributed memory and the representation of general and specific information.' In L. K. Komatsu (ed.): *Experimenting with the Mind: Readings in Cognitive Psychology*. Pacific Grove, CA: Brooks/Cole.

McEneaney, J. E. 1994. 'Does Parallel Distributed Processing provide a plausible framework for modeling early reading acquisition?' Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, 4th-8th April 1994.

McWhinney, B., J. Leinbach, R. Taraban and J. McDonald. 1989. 'Language learning: cues or rules?' *Journal of Memory and Language* 28: 255-277.

Medin, D. L., E. B. Lynch, J. D. Coley and S. Atran. 1997. 'Categorization and reasoning among tree experts: do all roads lead to Rome?' *Cognitive Psychology* 32: 49-96.

Mohan, B. A. 1986. *Language and Content*. Reading, MA: Addison-Wesley.

Ney, J. W. and B. A. Pearson. 1990. 'Connectionism as a model of language learning: parallels in foreign language teaching.' *The Modern Language Journal* 74: 474-482.

Pyles, T. and J. Algeo 1973. *English: An Introduction to Language*. New York: Harcourt Brace Jovanovich.

Schank, R. C. 1975. 'The structure of episodes in memory.' In D. G. Bobrow and A. Collins (eds.): *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press.

Schank, R. C. and Abelson, R. P. (1975) 'Scripts, plans, and knowledge.' In P. N. Johnson-Laird and P. C. Wason (eds.) *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press.

Schank, R. C. and Abelson, R. P. (1977) *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*. Hillsdale, NJ: Lawrence Erlbaum.

Schmitt, N. 1998. 'Quantifying word association responses: what is native-like?' *System* 26: 389-401.

Scott, M. 1997. 'PC analysis of key words - and key key words.' *System* 25: 233-245.

Scott, M. 2000. 'Lexical upshots from a large keywords database, or getting at the Guardian's view of the world.' Paper presented at the University of Liverpool, 5th April 2000.

Scribner, S. 1979. 'Models of thinking and ways of speaking: culture and logic reconsidered.' In R. O. Freedle (ed.): *New Directions in Discourse Processing: Advances in Discourse Processes Volume II*. Norwood, NJ: Ablex.

Seidenberg, M. S. 1992. 'Connectionism without tears.' In S. Davis (ed.): *Connectionism: Theory and Practice*. New York: Oxford University Press.

Sinclair, J. McH. and R. M. Coulthard. 1975. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils*. Oxford: Oxford University Press.

Slavin, R. E. 1994. *Educational Psychology: Theory and Practice*, fourth edition. Boston, MA: Allyn and Bacon.

Sokolik, M. E. 1990. 'Learning without rules: PDP and a resolution of the adult language learning paradox.' *TESOL Quarterly* 24: 685-696.

Strahan, D. B. 1989. 'How experienced and novice teachers frame their views of instruction: An analysis of semantic ordered trees.' *Teaching & Teacher Education* 5: 53-67.

Stubbs, M. 1986. *Educational Linguistics*. Oxford: Blackwell.

Tomlin, R. S., L. Forrest, M. M. Pu, and M. H. Kim. 1997. 'Discourse semantics.' In T. A. van Dijk (ed.): *Discourse as Structure and Process: Discourse Studies: A Multidisciplinary Introduction Volume 1*. London: Sage.

Van Dijk, T. A. 1977. *Text and Context: Explorations in the Semantics and Pragmatics of Discourse*. London: Longman.

Van Dijk, T. A. and W. Kintsch. 1983. *Strategies of Discourse Comprehension*. New York: Academic Press.

Van Lier, L. 1988. *The Classroom and the Language Learner*. London: Longman.

Watson Todd, R. 1998. 'Topic-based analysis of classroom discourse.' *System* 26: 303-318.

Whitney, P. 1998. *The Psychology of Language*. Boston: Houghton Mifflin.

Widdowson, H. G. (1979) *Explorations in Applied Linguistics*. Oxford: Oxford University Press.

Woods, D. 1996. *Teacher Cognition in Language Teaching: Beliefs, Decision-Making and Classroom Practice*. Cambridge: Cambridge University Press.

Richard Watson Todd works in the Department of Applied Linguistics, King Mongkut's University of Technology Thonburi. He is the author of 'Classroom Teaching Strategies' (1997) and 'Ways of Learning English' (2000) as well as numerous articles in many areas of language teaching.